Material perception and action

The role of material properties in object handling

Ingvarsdottir, Kristin

2021

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Material Perception and Action

The role of material properties in object handling

Kristín Ósk Ingvarsdóttir

DOCTORAL DISSERTATION

by due permission of the Faculty of Humanities, Lund University, Sweden.
To be defended at SOL, room C126, Lund, on November 19, 2021 at 13:00.

Faculty opponent
Roland W. Fleming

Organization: LUND UNIVERSITY. Cognitive Science. Department of Philosophy

Author: Kristín Ósk Ingvarsdóttir

Document name: Doctoral dissertation

Date of issue: 2021-10-18

Sponsoring organization

Title and subtitle: Material perception and action: the role of material properties in object handling

Abstract: This dissertation is about visual perception of material properties and their role in preparation for object handling. Usually before an object is touched or picked-up we estimate its size and shape based on visual features to plan the grip size of our hand. After we have touched the object, the grip size is adjusted according to the provided haptic feedback and the object is handled safely. Similarly, we anticipate the required grip force to handle the object without slippage, based on its visual features and prior experience with similar objects. Previous studies on object handling have mostly examined object characteristics that are typical for object recognition, e.g., size, shape, weight, but in the recent years there has been a growing interest in object characteristics that are more typical to the type of material the object is made from. That said, in a series of studies we investigated the role of perceived material properties in decision-making and object handling, in which both digitally rendered materials and real objects made of different types of materials were presented to human subjects and a humanoid robot. Paper I is a reach-to-grasp study where human subjects were examined using motion capture technology. In this study, participants grasped and lifted paper cups that varied in appearance (i.e., matte vs. glossy) and weight. Here we were interested in both the temporal and spatial components of prehension to examine the role of material properties in grip preparation, and how visual features contribute to inferred hardness before haptic feedback has become available. We found the temporal and spatial components were not exclusively governed by the expected weight of the paper cups, instead glossiness and expected hardness has a significant role as well. In paper II, which is a follow-up on Paper I, we investigated the grip force component of prehension using the same experimental stimuli as used in paper I. In a similar experimental set up, using force sensors we examined the early grip force magnitudes applied by human subjects when grasping and lifting the same paper cups as used in Paper I. Here we found that early grip force scaling was not only guided by the object weight, but the visual characteristics of the material (i.e., matte vs. glossy) had a role as well. Moreover, the results suggest that grip force scaling during the initial object lifts is guided by expected hardness that is to some extend based on visual material properties. Paper III is a visual judgment task where psychophysical measurements were used to examine how the material properties, roughness and glossiness, influence perceived bounce height and consequently perceived hardness. In a paired-comparison task, human subjects observed a bouncing ball bounce on various surface planes and judged their bounce height. Here we investigated, what combination of surface properties, i.e., roughness or glossiness, makes a surface plane to be perceived bounceable. The results demonstrate that surface planes with rough properties are believed to afford higher bounce heights for the bouncing ball, compared to surface planes with smooth properties. Interestingly, adding shiny properties to the rough and smooth surface planes, reduced the judged difference, as if surface planes with gloss are believed to afford higher bounce heights irrespective of how smooth or rough the surface plane is beneath. This suggests that perceived bounce height involves not only the physical elements of the bounce height, but also the visual characteristics of the material properties of the surface planes the ball bounces on. In paper IV we investigated the development of material knowledge using a robotic system. A humanoid robot explored real objects made of different types of materials, using both camera and haptic systems. The objects varied in visual appearances (e.g., texture, color, shape, size), weight, and hardness, and in two experiments, the robot picked up and placed the experimental objects several times using its arm. Here we used the haptic signals from the servos controlling the arm and the shoulder of the robot, to obtain measurements of the weight and hardness of the objects, and the camera system to collect data on the visual features of the objects. After the robot had repeatedly explored the objects, an associative learning model was created based on the training data to demonstrate how the robotic system could produce multi-modal mapping between the visual and haptic features of the objects. In sum, in this thesis we show that visual material properties and prior knowledge of how materials *look like* and *behave like* has a significant role in action planning.

Key words: material perception, visual properties, motor planning, cognition, humanoid robot

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language: English

ISSN and key title: 1101-8453 Lund University Cognitive Studies 182

ISBN: 978-91-89213-97-5

Recipient's notes

Number of pages: 135

Price

Security classification

Material Perception and Action

The role of material properties in object handling

Kristín Ósk Ingvarsdóttir

LUND
UNIVERSITY

Faculties of Humanities and Theology
Department of Philosophy
Cognitive Science

ISBN 978-91-89213-97-5 (print)
ISBN 978-91-89213-98-2 (digital)
ISSN 1101-8453

Printed in Sweden by Media-Tryck, Lund University
Lund 2021

Printed matter 3041 0903

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

In loving memory of my parents

Table of Contents

Acknowledgements **9**

List of original papers **11**

Paper I 11

Paper II 11

Paper III 11

Paper IV 11

Introduction **13**

Living in a material world 13

The scope of the thesis 14

Summary of papers **17**

Paper I: The Visual Perception of Material Properties Affects Motor Planning in Prehension: An Analysis of Temporal and Spatial Components of Lifting Cups. 17

Paper II: Coffee to go! Examination of grip force anticipation when lifting paper cups. 18

Paper III: Material properties and visual judgment of bounce height 19

Paper IV: Learning material properties from visual and proprioceptic features in a humanoid robot 20

What is material perception? **21**

Behavioral and biological evidence 21

Visual material properties 27

How do we perceive *Stuff?* **35**

Conceptual spaces of materials 36

Mathematical models 39

Material perception and action **41**

Developmental studies and learning physical properties 41

We as active perceivers 44

References **47**

Acknowledgements

Where should I begin? For a while I felt like I would never see the end of my journey as a doctoral student, as sudden hurdles continued to appear in front of me, causing the finishing line to move farther away from its original place. Finally, I have crossed the finishing line and my doctoral thesis gets to see daylight and get the recognition it deserves. For that I thank my darling Emil, who encouraged me during difficult times of grief and inspired me to think and write creatively. This thesis would not have been completed without the excellent supervision by my supervisor, professor Christian Balkenius, who always had great feedbacks available when I encountered new challenges. Christian provided a positive and thriving work environment, and his enthusiasm and passion for knowledge and research was inspiring for a young researcher like me. I truly hope I have adopted these characteristics and paid them forward to my students.

I am very thankful to my family, my brother Hartmann Ingvarsson and his girlfriend Annika Simonssen, for being supportive and wonderful people, and my cousin Sigríður Inga Þorkelsdóttir Svarfdal and her boyfriend Hannes Gunnarsson, for being encouraging and cheerful. I thank my parents in law, Ásdís Ósk Jóelsdóttir and Adolf Hólm Petersen, for all the support throughout the years and for all the help with my children. I also want to thank my lovely children, Ronja Áskatla Petersen and Thrándur Alvar Petersen, who have now reached the age where they ask questions about how the brain works (and then create their own funny explanations). Finally, I would especially like to thank my parents, Áslaug Hartmannsdóttir and Ingvar Stefánsson, who supported me from a far and encouraged me to pursue my research interests, despite difficult times in our lives. This thesis is dedicated to them. I am grateful to my friends. Especially, my *breakfast club*, Kristín Rut Kristjánsdóttir and Jess Chia-Sui Hsu. Our breakfast meetings in our early years as doctoral students were inspirational and a great platform for discussing our research, work, and other unrelated matters. I also want to thank Snædís Góa Guðmundsdóttir and Oddsteinn Guðjónsson for many years of great friendship and support. Last, but definitely not the least, I want to thank Erik Anderberg for being a great friend during our times at LUCS, first as master students and then later as employees.

Along the road there are many people who I am very thankful to. First, and foremost I thank the pillars of our institute: Ingela Byström, Anna Cagnan Enhörning, Fredrik Eriksson, Tobias Hansson Wahlberg, Martin Jönsson, Tomas Persson, Eva

Sjöstrand, Annah Smedberg-Eivers, Ylva von Gerber, and Anna Östberg. I thank my fellow LUCS members from my time in Lund, especially the Cognitive Zoology group that I shared office with at LUX: Katarzyna Bobrowicz for all the wonderful discussions, statistic advices and cookies, Ivo Jacobs, Can Kabadayi, Elainie Madsen, Helena Osvath, Mathias Osvath, Stephan Reber, and Claudia Zeiträg, for interesting talks about animal cognition. Moreover, I want to especially thank professor Peter Gärdenfors for instructive conversations about his theory of conceptual spaces, and Birger Johansson for his help and collaboration on paper IV. I also want to thank all the other *koggarna*, i.e., all the talented professors, researchers, lecturers, and doctoral students that either are *free range* like myself or are part of some of the amazing research groups we have here at LUCS: Andrey Anikin, Rasmus Bååth, Kerstin Gidlöf, Agneta Gulz,, Magnus Haake, Lars Hall, Jana Holsanova, Petter Johansson, Magnus Johnsson, Andreas Lind, Jens Nirme, Manuel Oliva, Philip Pärnamets, Maximilian Roszko, Björn Sjödén, Thomas Strandberg, Trond Arild Tjøstheim, Betty Tärning, Gabriel Vogel, and Annika Wallin. Outside of LUCS I want to thank Joost van de Weijer at the LU Humanities Lab for all the statistic courses and discussions, and professor Geoffrey Patching at the department of psychology at LU for his insights in the project of bounce height estimation. Finally, I want to thank *Drakamöllans Vänner för Kultur och Vetenskap* for the scholarship to stay at their residency for one week to work on my thesis.

I thank you all for the wonderful times in Lund, first in Kungshuset and then later in LUX!

Kristín Ósk Ingvarsdóttir

List of original papers

Paper I

Ingvarsdóttir, K.Ó., and Balkenius, C. (2020). The Visual Perception of Material Properties Affects Motor Planning in Prehension: An Analysis of Temporal and Spatial Components of Lifting Cups. *Front. Psychol. 11*:215. DOI: 10.3389/fpsyg.2020.00215.

Paper II

Ingvarsdóttir, K.Ó., and Balkenius, C. (in review). Coffee to go! Examination of grip force anticipation when lifting paper cups. © The Authors 2021.

Paper III

Ingvarsdóttir, K.Ó., and Balkenius, C. (submitted). Material properties of ground affects visual judgment of bounce height. © The Authors 2021.

Paper IV

Ingvarsdóttir, K.Ó., Johansson, B., and Balkenius, C. (manuscript ready for submission). Learning material properties from visual and proprioceptic features in a humanoid robot. © The Authors 2021.

Introduction

Living in a material world

The most important part of my daily morning routine at home is to make coffee. The preparation process is almost automatic, in which I take the coffee pot to the kitchen sink to get water, and without relying on any visual measuring lines on the glass pot, I stop the running tab water when I have precisely enough water to make four cups of coffee. The water level estimation is perfect almost every time thanks to collections of informative sensory information such as the haptic feedback of the increasing weight of the pot, the visual feedback of the rising water level in the pot, the duration time of the running water, and the changing pitch coming from the decreasing air volume in the pot. More interestingly, the estimation skill relies not only on the provided sensory information, but also on my stored knowledge of how water and glass behave as material entities in a similar scenario. In other words, memories of previous experiences with running water, the knowledge of how the weight of the glass pot can be changed with the increasing water level, and the understanding of how fragile glass can be when handled. Together, this collection of sensory information and memories create a meaningful pattern, that is then learned via the repetitive routine of making that same amount of coffee every morning. For that reason, a morning routine like this requires a knowledge of material properties.

Material categories and their properties are everywhere and interactions with them are unavoidable. For instance, we dress ourselves using textiles that vary in stiffness, e.g., cotton, wool, denim, and jersey, and the surface beneath our feet as we move around is comprised of materials that vary in roughness and friction, e.g., grass, gravel, hardwood, and carpet. Moreover, in scenarios when we cannot interact with physical objects directly using our hands, feet, or body, we use tools with certain physical properties to interact with them instead. For instance, when taking out a hot tray from the oven we normally use silicone oven mitts or dry kitchen towel to prevent our skin from burning. Additionally, we approach food using various kinds of tools and utensils depending on the physical property of the food and the tools, and what we want to do with them (e.g., stir, knead, pour, cut). Food that is different in terms of density and stiffness, such as root vegetables and sponge cakes, can both be cut using a sharp knife made of stainless steel, however, only the material properties of the sponge cake allow cutting using a plastic spoon or knife. We know that the cake can be cut using a breakable plastic utensil because the cake is soft, and we know that the utensil might break if we tried the same procedure, using the same tool, on a carrot. All because we have the ability to both recognize materials, and to learn about their properties and what they afford, consequently we have no problem with selecting the right tool for the task. In a broader context, material knowledge provides an important survival value. For instance, spoiled food has certain surface characteristics that can be easily detected visually before any consumption takes place. Fruits turn brown, bread gets covered with fuzzy mold, and water with unsafe pathogens in it has brown and cloudy appearance, that signals that consumption should be avoided to prevent sickness. As the forgoing examples demonstrate, we are constantly interacting with objects that come in different shapes, sizes, and colors. More importantly the examples demonstrate that objects are made of various materials that have characteristic appearances and intrinsic properties, e.g., roughness, glossiness, friction, stiffness, and viscosity, that we rely on in our daily activities. Regardless, research on material perception has been a small field in comparison to research on object recognition. Fortunately, this has changed, as material perception research has been a fast-growing research field over the last two decades.

The scope of the thesis

This doctoral dissertation examines the role of materials and their characteristics in decision-making and productions of motor actions when handling objects. While the main research focus is visual perception of material properties, the thesis also features tactile perception of materials and stored knowledge of material properties, since materials are rarely experienced passively via single sensory modality in the real world. That said, the theoretical discussion, both in the introduction chapters of the dissertation and in discussion sections of the research papers, assumes multi-modality when describing the nature of material perception. The dissertation intends to achieve the two following aims:

1. To examine how changes in visual material properties result in changes in motor behavior when handling objects.
2. To theorize about internal models describing the relationship between visual and intrinsic material properties.

Through the repeated experience of observing, lifting, weighing, smelling, listening, and walking on different types of materials, we learn about the physical world we live in and develop internal models that reflect our material knowledge. These models comprise learned associations (long-term priors) between object properties, such as object weight and object size (Gordon, et al. 1991; Flanagan and Beltzner, 2000; Flanagan et al., 2008), and object weight and material type (e.g., Buckingham et al., 2009; Baugh et al.,2012), and are formed to enable predictions as well as planning of future actions when encountering novel materials. The models bear a resemblance to the internal models of motor movements that predict the outcome of motor commands before tactile feedback becomes available, thus generating a desired movement in a forward manner (Johansson and Westling, 1988; Gordon, et al., 1993; Gordon et al., 1994; Nowak et al., 2007).

Materials are handled in various ways, in which action requirements are distinctive depending on categorization of the material and its properties. Some materials afford bouncing and sliding, while others afford folding, and splashing. Oftentimes, we instantly recognize the material category and its properties because materials have typical visual characteristics that we take advantage of. For example, the material category *wood* is easily recognized by the brown-colored streaks on its surface (i.e., growth rings), and subsequently, *hardness*, a typical material property of wood is identified. In comparison, some materials are so visually similar that their distinct action requirements, and end-goals of required actions, are hard to recognize instantly. For instance, natural yoghurt and face cream are both white colored substances with smooth texture that afford to be scooped or smeared, but they differ in required end-goals, in which the yoghurt is edible while the face cream is not (well, normally not). Without the packaging or labels, the two substances are difficult to quickly tell a part visually, but with more experience of using the two substances, the material knowledge expands and the task of telling yoghurt and face cream apart becomes easier (Adelson, 2001).

Early on in my research I felt the perception of action requirements had a reminiscent of *affordance*, a term created by Gibson (1966, 1979) to describe perceived action possibilities of objects based on shapes and forms. Gibson's ecological approach to visual perception proposes that action is always involved in one's perception, as we are active and explorative perceivers in our perceptual world. According to his approach, the structure of the perceiver's environment is meaningful, in which all objects have a range of action possibilities that we instantly pick up to know how to interact with those objects. For example, a chair affords to be sat on, but depending on the context and the end-goal of our intended action, the chair might also afford to be moved or to support weight while standing on it. For this reason, and since no strong scientific consensus of how material properties are recognized was present in the beginning of my research, I adopted Gibson's term affordance but extended it to perceived action possibilities based on material properties, as a way of describing an active perception of action possibilities based on materials rather than simply shapes. Nonetheless, Gibson's ecological approach argues against that perception involves learning, which goes against the theme of this dissertation. I therefore kept the approach as a reference framework when describing how action requirements of materials are visually perceived, rather than as a theory-based principles. In recent years, theories of visual perception of materials have shifted away from a purely bottom-up processing mechanism that involves no learning, in which the visual system makes inferences based on measurable low and mid-level features. Such approach is more suitable for research on visual perception from a computer vision perspective. The shift has been towards goal-directed and semantic driven high-level vision theories allowing the inclusion of cognitive constructions of materials (Fleming, 2017). I therefore include a discussion of two additional theoretical and mathematical approaches in the section *How do we perceive Stuff*? that are both suitable for explaining how material knowledge is obtained and stored. The first one is Gärdenfors's (2000) theory of conceptual spaces, which is particularly useful for understanding how we make a meaning of the stuff in the world as it assumes concept learning and organization based on dimensions of qualities. The second approach is Fleming's (2014, 2017; see also Fleming and Storrs, 2019) statistical appearance model, which specifically explains how materials are visually perceived and has become the most accepted theory of material perception within the research field.

The visual environment changes constantly as we move around within it. Changes in viewpoints and light source positions result in changes in surface appearances, and tactile properties are learned when objects are picked up. Consequently, our knowledge of material properties improves constantly when materials are viewed from various viewpoints and illumination and new associations between visual and tactile properties are learned. That said, this doctoral dissertation seeks to approach material perception as a continuous cycle between the active perceiver and the real-world, in which material properties are studied as natural properties that afford to be interacted with instead of static images that are passively viewed. For this reason, I used real-world items in the experiments (Paper I, II, and IV), with the exception of Paper III, which consists of dynamic computer-generated animations to obtain high-degree experimental control to study visually inferred hardness. The research papers include evidence for internal models that are specifically for material categorization and action control. The models conclude associations between visual characteristic and intrinsic properties of materials, and allow for anticipating action requirements for future handlings, both in humans (Paper I, II, III) and in a humanoid robot (Paper IV). Specifically, the papers included in this dissertation demonstrate: findings for visually inferred hardness (Paper III); that visual material properties are incorporated early in action- and grip force production when lifting objects (Paper I and II); and learned associations between visual and physical properties of materials in a humanoid robot (Paper IV).

Summary of papers

Paper I: The Visual Perception of Material Properties Affects Motor Planning in Prehension: An Analysis of Temporal and Spatial Components of Lifting Cups.

Research question

In this reach-to-grasp study, we examined the participants' expectations of how materials behave, to determine the role of visual material properties in early prehension control. Specifically, we were interested in surface gloss and object weight, to see how those properties contribute to the process of inferring hardness before tactile feedback becomes available.

Procedure

The results are based on data collection from fourteen participants, in which motion capture technology was used to obtain both temporal and spatial measurements of prehension when the participants lifted paper cups that either had original appearance or were altered to appear glossy. To vary the object weight, we presented half of the paper cups with added weight while the other half of the cups were kept empty. In addition, to examine the effect of content visibility on prehension control, half of the cups had their content hidden using a thin lid, while the other half had their content visible. The expected and perceived material properties were assessed using a seven-points semantic differential scale, in which the participants rated the cups' properties in terms of heaviness, glossiness, and hardness, both before and after lifting the cups.

Results and conclusion

Overall, we found the temporal and spatial components of prehension were not exclusively governed by the weight of the cups as previous studies have stated. Instead, the results revealed that material appearance and expected hardness have significant roles in prehension as well. For the initial object lifts, the prehension was guided by how the participants anticipated the functional properties of the cups, based on the visual characteristics of the cups and their prior experience with materials having similar visual structure. Whereas, after repeated trials of lifting the cups, prehension is characterized by the participants' perception of the physical properties of the cups, as object weight dominates the prehension control of the hand.

Paper II: Coffee to go! Examination of grip force anticipation when lifting paper cups.

Research question

This study is a follow-up on the previous reach-to-grasp study (Ingvarsdóttir and Balkenius, 2020). Here, we further examined the role of visual material properties in prehension by examining the early grip force control in object lifting. In particular, we measured the grip force magnitudes applied by each of the five digits (fingers) individually, while the participants lifted paper cups that varied in surface gloss and object weight.

Procedure

This study is based on data collection from sixteen participants. The same experimental procedure and stimuli were used as in Paper I, but with the addition of grip force collection during the reach-to-grasp task. The force data was collected using five touch sensors that were individually attached to each digit of the hand.

Results and conclusion

The results support our previous findings and highlights the significance of visual material properties in prehension control. We found that early grip force scaling was not only guided by the weight of the cups, but also by the surface gloss and whether the participants could visually infer the weight of the cups based on their content or not, i.e., the presence of another material. Comparisons between the grip force magnitudes guided by expected vs. perceived properties suggest that we have predetermined expectations of visually inferred hardness that we rely on in grip force scaling during initial object lifts.

Paper III: Material properties and visual judgment of bounce height.

Research question

In this visual judgment task, we examined how the visual material properties, roughness and glossiness and their association with hardness, are incorporated into the visual assessment of bounce height. In general, it is believed that learned associations based on previous experiences of how materials *look* and *feel* help us to anticipate the nature of a new material. For instance, it is important for us to anticipate the *walkability* of a surface plane in front of us to prevent us from falling over. Here, we asked what visual properties make surface planes *bounceable*? That is, what visual properties are judged to afford high bounce heights as a consequence of being associated with hard qualities.

Procedure

In two paired-comparison experiments, 60 participants (30 in each experiment) were asked to observe a computer-generated ball bouncing on surface planes with either smooth or rough characteristics that were presented with either matte or shiny finishing. The purpose was to see if the participants paired higher bounce heights with surface planes with certain types of visual characteristics. Participants were also asked to rate the surface properties of the planes using three semantic-differential scales: roughness, hardness, and glossiness, to obtain data on perceived properties and visually inferred hardness.

Results and conclusion

Our results demonstrate context effects in judged bounce height. We found that the participants judged physically high bounce heights to be higher than physically low bounce heights, however, the judged bounce height was also affected by the visual characteristics of the surface planes. The participants judged surface planes with rough characteristics to provide higher bounce heights compared to surface planes with smooth visual properties. Moreover, in the second experiment we found adding shiny properties to the rough and smooth surface planes reduced the previously judged difference in bounce height. The participants judged bounce heights on shiny surface planes to be high, irrespective of the textured characteristics of the surface planes. The results demonstrate that perceived bounce height incorporates not only the physical elements of the ball and the bounce, but the visual material properties of the plane as well. Here we provide evidence for learned associations between the visual material properties, roughness and glossiness, and visually inferred surface hardness.

Paper IV: Learning material properties from visual and proprioceptic features in a humanoid robot

Research question

In this experiment, we examined the development of material knowledge in a humanoid robot, named Epi, to further investigate the role of material properties in skilled object manipulations. Here we were especially interested in two physical properties that normally are perceived using touch, namely weight and hardness. Moreover, we were also interested to see if the robotic system could categorize the experimental objects visually, and if multi-modal representations could be produced, based on the information provided by the haptic- and the camera systems. The experiment was also set out to see if the haptic feedback coming from the servos that control the arm and shoulder of the robot could be used to measure weight and hardness. Using feedback signals from the servos, rather than signals coming from tactile sensors built within the hand, reduces the likelihood of broken fragile parts in the robot fingers when lifting heavy objects.

Procedure

In two experiments, Epi repeatedly picked up and placed experimental objects that varied in type of materials, visual appearance, weight and hardness. For the first experiment we tested eight different objects, including three types of weight, to see if the robot could measure weight and hardness using feedback signals from the servos controlling its arm. In the second experiment, same experimental procedures were conducted as in the first experiment, but the number of experimental objects was extended to twenty-five objects, and the weight used for the training data for the weight learning ranged in 50g steps from 50g to 500g. After training, an associative learning mechanism was created to show how the Epi the robot can learn the mapping between the two modalities, vision and haptic, and to create multi-modal representations of the different types of materials that constitute the experimental objects.

Results and conclusion

Using inputs from the visual and the haptic systems, Epi learned to estimate the weight and hardness of the experimental objects using feedback signals provided by the servos that control the arm and the shoulder of the robot, and to furthermore map those material properties to the visual features of the experimental objects. Like an infant who learns about the material properties of objects through active exploring using sight and touch, Epi learned to map the weight and hardness of the experimental objects to their visual features. In the future, this modelling mechanism could be used as a part of an affordance learning system.

What is material perception?

Behavioral and biological evidence

Material perception is recognition of material categories and their properties, and it is a multi-modal process in the sense that real-world materials are perceived via sight, touch, hearing, smell, and taste (see discussion on multimodality in material perception in Fleming et al., 2015; and Fujisaki et al., 2014). Nevertheless, *visual characteristics* of materials and visual perception of material properties has dominated the field of material perception research. The great interest in visual material properties does not come as a surprise, considering the technical advances in computer graphics and algorithms over the years, which allow for photo-realistic graphics of how materials appear as well as how they behave to the untrained eye. Similarly, research on material recognition, together with advances of artificial intelligence and neural networks, has assisted the development of smarter computer vision software (Adelson, 2001; Cui et al., 2018).

The visual process of material properties has been thoroughly outlined in a primer written by Barton L. Anderson (2011), which also includes an exceptional overview of the significant conducted early in the field. The following examples briefly describe the visual process of materials. To begin with, the light travels and interacts with objects in the physical world, the light hits the surface of the objects and bounces back off and into our eyes carrying important information about the geometrical structure of the physical world. At the same time, the structured light entering our eyes tells us also about the materials in our visual environment, as objects vary not only in structure but also in surface characteristics that are specific for the material that the object is made of. For instance, while a polished metal and a sheet of printing paper both have smooth surface characteristics, the light interacts differently with the two objects as they are made from different types of materials and have different properties as a result of that. For instance, the light hitting the metal gets reflected into one direction, creating highlights and shiny appearances, whereas the light hitting the paper sheet scatters at various angles when it exits the object, creating diffused reflection and matte appearance. The purpose of visual material perception is then to encode the retinal images (i.e., proximal stimulus) comprised of low-level features, organize the provided visual information into representations of the real-world materials (i.e., distal stimulus) and apply semantic meaning to that visual information, to achieve material recognition at a higher cognitive level.

It is remarkable that humans can effortlessly recognize materials based on seemingly meaningless surface structures, considering there are similarities in visual characteristics across different groups of materials (see figure 1), and variations in visual features within a single group of materials (see figure 2). To clarify, the two types of vases seen in figure 1 have similar reflection spots, i.e., the white highlights on their surfaces, and could therefore be classified into the same group of reflective material.

Figure 1. An example of similarities in surface appearance across different types of material categories. The two vases have similar visual characteristics in terms of reflectance, despite belonging to separate material categories, glass and ceramic.

However, despite the similarities in visual features between the two vases, we can easily classify them into separate material groups, i.e., a transparent glass and a glossy ceramic. Likewise, although the two types of foliage seen in figure 2 have distinct visual characteristics, as one leaf has *hairy* texture whereas the other leaf has smooth and glossy surface features, we have no problem classifying them into the same material category. That said, there is a lot of visual ambiguity that needs to be sorted out in order for materials to be recognized.

Figure 2. An example of variation in visual properties within a single material category. Here we see that foliage comes in various forms and texture, in which one leaf has smooth and glossy surface appearance, while another leaf has matte surface with *hairy* texture. Nevertheless, the two leaves are classified as belonging to the same material category, despite their difference in visual properties.

Despite such visual ambiguity, where similarities in visual appearance is found across different types of materials, and variations in visual appearance is found within a single type of material, it has been demonstrated that visual recognition of materials is a very quick process when tested using high-level material categories and real-world images (Sharan, 2009; Sharan, et al., 2009; and Sharan, et al., 2014). Sharan (2009) used images of real-world materials and asked participants to categorize the images into material categories, and measured their timing of visual selection using rapid serial visual presentation paradigm. The results did not only demonstrate that material recognition is extremely fast and accurate, but also that the participants could even recognize material categories from images that were briefly displayed in just 40 milliseconds, irrespective of the visual diversity present within each material category. Later, Wiebel, et al., (2013) confirmed that categorization of materials is both fast and accurate but argued that it was less accurate than categorization of simple objects. However, the remarkable performance found in the studies by Sharan (2009) and Sharan, et al. (2014) was found to be equally accurate as scene- and object categorization. Moreover, the performance could neither be explained by simple categorization of shapes, as no overall shapes were present in the image stimuli, nor could the performance be explained by low-level visual features such as color or contrasts, as the images were manipulated to discard those features. Wolfe and Myers (2010) have also supported that the visual system does not rely on pre-attentive attributes when rapidly classifying high-level material categories based on material properties, as they found that material properties were inefficient to guide selective attention in a visual search task. Suggesting that material properties are not pre-attentive features in nature, unlike color and line orientation that are processed early in the visual system.

Although both material properties and shapes are fundamental parts of object's identity, evidence have shown that visual recognition based on material properties, on one hand, and shapes on the other hand, are two separate cognitive processes that rely on distinct features in the distal world. In other words, material recognition is simply not a conventional object recognition. This is unsurprising considering that objects from the same object category can be made of completely different materials, and the same material can be used to create distinct objects. For instance, the object category *drinking cup* is comprised of cups made of all kinds of materials, such as plastic and metal with applied enamel, which are all classified as an object to drink liquid out of despite their differences in material properties. Similarly, *plastic* is a very versatile material as it is not only used for creating drinking cups but also toys, bicycle helmets, packaging, etc. Let alone, it is also imprinted in our language that objects are *things* that can be counted (e.g., one pillow, two pillows, etc.), whereas materials are *stuff* that cannot be counted in the same way as objects are. For instance, the tactile property of a pillow, *softness*, is not counted in the same sense as the pillow itself, unless the softness property would be turned into a noun (e.g., *one softie, two softies, etc.*). Instead, the softness property is measured in terms of quantity, using either measurement units such as how much the fabric bends or subjective rating scales (Peirce, 1930; see also Kilinc-Balci, 2011, for overview of studies on comfort properties of textiles). Findings by Sawayama and Nishida (2018), on intensity gradient information and material perception, confirm that the visual system relies on separate image characteristics for visual judgment of materials and shapes. According to their findings, glossiness judgment is greatly affected by image manipulations in which intensity gradient magnitude information were changed (think increased intensity of white patches, or highlights, that are indication of a glossier surface), whereas shape estimation was unaffected by them. In turn, shape estimation was heavily affected by image manipulations in which intensity gradient order information was changed (think white patches that are usually positioned on top of the highest points of a surface, indicating the shape of the object).

Many have argued that the visual system relies on low-level visual features when categorizing materials, and often those studies emerge from the field of computer vision, in which image statistics (e.g., luminance histograms) are required for image analysis (e.g., Pont and Koenderink, 2005; Varma and Zisserman, 2009; Sharan, et al., 2013). It is difficult, however, to argue that visual material perception is simply based on image statistics and estimation of physical parameters like the Bidirectional Reflectance Distribution Function (BRDF), which describes the reflection characteristics of the surface (for more on image statistics see Nishida, 2019), as similar appearances can be found for different types of materials as seen in Figure 1. The two materials, glass and ceramic, produce similar retinal images when the images are cropped to reduce the impact of other object properties like shape. Therefore, it would be problematic to just rely on low-level retinal image features to accurately categorize those two materials into separate categories. Instead, it has been suggested that categorization of materials involves stored knowledge and is to some extent a semantic process of visual input. In two experiments, Fleming, et al. (2013) asked participants to judge the material properties from both images of materials and from verbal names of materials and found systematic relations between the judged material properties and the presented material categories, in which judged material properties clustered into distinct feature spaces (i.e., material categories). Moreover, they found correspondence between the two tasks and concluded that the visual judgment of material properties and the judgment of memory-based material categories emerged from the same stored knowledge of materials.

In recent years there has been a growing interest in the neural mechanism underlying material perception, in which, functional magnetic resonance imaging measures (fMRI) have been used for theorizing how materials are perceived and organized in the brain. In terms of visual perception, neuroimaging findings suggest that visual material recognition is a hierarchical structure, ranging from recognizing simple image features in the early visual areas, specifically in the primary and secondary visual cortexes, V1 and V2 (Baumgartner and Gegenfurtner, 2016), to classifying material categories in higher order areas in the parahippocampal gyrus (Jacobs, et al., 2014), fusiform gyrus and collateral sulcus (Goda, et al., 2014; Hiramatsu, et al., 2011; Komatsu and Goda, 2018). Furthermore, electroencephalography measures (EEG) have been practical for examining the time course of material categorization, in which event-related potentials data (ERP) has demonstrated relatively fast neural responses (approximately 140 ms after stimuli onset), suggesting an early differential visual processing of materials, at least for simple image features (Wiebel, et al., 2014). Neuroimaging techniques have also confirmed that material perception is cross-modal in nature, where every sensory modality is involved in perception of materials. In terms of cross-modal interactions, increased neural activity has been found in the visual cortex together with an increased activity in the somatosensory cortex during a haptic judgment of surface textures, suggesting visual-haptic interactions (Podrebarac, et al., 2014; see also review in Komatsu and Goda, 2018). Moreover, cross-modal interaction between vision and hearing have been demonstrated as increased activity has been found in the ventral visual cortex when listening to sounds of materials, whereas sounds of voice did not elect such activity (Arnott et al., 2008). Behavioral studies have supported the cross-modal aspect of material perception as well, for instance has surface appearance been found to affect the haptic perception of weight, as has been demonstrated with the *material-weight illusion,* which is an illusion that occurs when lifting and comparing the weight of two objects with equal mass but difference in material appearance (e.g., stone vs. Styrofoam). When the two objects are lifted, the Styrofoam (i.e., the lighter-looking object) is judged to be heavier than the stone (i.e., heavier-looking object), possibly because expected weight based on previous experience with stones and Styrofoams is violated, resulting in an overestimated weight of the lighter looking object (Buckhingham et al., 2009). Surface friction (slipperiness) has been found to affect visual judgment of surface gloss, in which perceived slippery objects are judged glossier (Adams et al., 2016). Materials can also be recognized based on their contact sound (Klatzky, et al., 2000), but matching visual appearance of one material with a contact sound of another material results in a perception of a third material, a perceptual phenomenon resembling the *McGurk effect*. The effect is a multisensory illusion in which incongruent pairing of acoustic speech (*ba*) with visual speech (*ga*) results in a perception of a third sound (*da*) (Fujisaki et al., 2014; McGurk and MacDonald, 1976).

Material perception is not unique to humans. Studies on animal cognition suggest that some animals have a cognitive mechanism to perceive properties that are not necessarily attached to the size or the shape of an object. For instance, Capuchin monkeys can identify the functional property, such as rigidity (a material property related to hardness), of a tool and select appropriate tool for the task requirement by simply observing an experimenter manipulating it (Manrique, et al., 2011). Additionally, Capuchin monkeys have also been found to use sequential tools (i.e., use a tool to get to another tool to then access the reward) that differed in rigidity (Sabbatini, et al., 2013). Furthermore, New Caledonian crows are known to be able to infer object weight based on external forces, i.e., they can infer the weight of light and heavy objects based on whether the objects move or remain stationary when getting breezed by an electric fan (Jelbert et al., 2019).

All things considered, we currently know that visual perception of materials is a remarkably quick and accurate cognitive process that is distinct to object recognition, has strong correlations with the neural mechanism in the cortical areas in the ventral visual pathway, is cross-modal, and is not unique to humans.

Visual material properties

The four research papers included in this dissertation aim to examine visual material properties and our behavioral reactions to them. Overall, the papers comprise of assessments of visual perception of *glossiness* and *roughness*, and their associations with inferred *hardness*, for the purpose of examining the role of inferred material properties in grip planning when handling everyday objects. The three material properties, *roughness, glossiness,* and *hardness* were of special interest due to several reasons as explained in this subsection.

Firstly, *glossiness* and *roughness* are two fundamental visual properties that we commonly use, besides color, when describing material appearances. We use glossiness (or shininess) when describing the reflective properties of the material, in which we position our description somewhere on the range between *matte surface appearance* and *shiny surface appearance*. Whereas, we use roughness when describing textured surface appearance, in which our description is positioned between *smooth surface appearance* and *rough surface appearance*. Given the common usages of these two descriptions when describing subjective experiences of materials, it is important that research on material perception includes assessments of perceived glossiness and roughness, and the interplay between them. Besides, both glossiness and roughness can be easily digitally rendered, using today's computer graphics, which allows for reproduction of specific surface characteristics without the cost of losing experimental control. That way, the perceptual effects of each individual property can be examined separately at various intensities, or jointly with other properties to obtain knowledge of their interacting perceptual effects. As a result, many findings on material recognition and categorization have been derived from studies on computer vision and texture analysis (see review in Adelson, 2001; and Thompson, et al., 2011), in which the most studied material properties in visual perception of materials are often *glossiness* (see review in Fleming et al., 2015) and *roughness* (e.g., Bergmann Tiest, 2010; Bergmann Tiest and Kappers, 2007).

Secondly, materials are rarely perceived passively and without being touched, which is why I believe it is important to examine material knowledge as it is obtained through continuous interactions between the perceiver and the physical world. Research on material knowledge, should therefore include material properties that describe not only the visual appearance of materials, but also the intrinsic and non-visual qualities of materials. In particular, the intrinsic properties are useful to anticipate in order for the perceiver to achieve successful interactions with the physical world. Therefore, *hardness* is included in the papers as a fundamental material property to anticipate when interacting with materials, because it informs us about the required grip force when handling objects with care. In this dissertation, glossiness and roughness are both defined as optical properties that describe how light interacts with surfaces, whereas *hardness* is defined as a mechanical property that describes how materials react to applied force (e.g., poking soft bread vs. hard apple). Although, glossiness is normally perceived visually, whereas roughness and hardness are usually perceived via touch. Evidence show that mechanical properties can be visually inferred from glossiness and roughness, via learning (see, e.g., Schmid and Doerschner, 2018), which is why the interplay between glossiness and roughness as visual properties, and their possible involvement in inferred hardness for soft and hard materials, were of special interest in this dissertation.

Glossiness describes the shininess of a surface. In other words, it describes the perceived surface reflectance, which is how much light (and specular direction) is reflected of a surface. Glossiness has been extensively studied, both within computer vision science and human psychophysics (Thompson, et al., 2011). The visual distinction between matte and shiny materials are specular highlights that are characterized by white colored patches, in which the intensity of perceived glossiness is achieved by increasing the size and brightness of the specular highlights, consequently expanding the specular coverage of the surface area as a result (Beck and Prazdny, 1981). Gloss perception has also been found to be affected by additional surface characteristics, such as the sharpness and luminance contrast of the highlights, including the location of the highlights on the three-dimensional object, as highlights must correspond to the shading profile of the object (Anderson and Kim, 2009; Kim, et al., 2011 and 2012; Marlow, et al., 2012; Marlow and Anderson, 2013; see review in Chadwick and Kentridge, 2015). Furthermore, in recent years psychophysical studies have revealed distinct kinematic information for matte and shiny materials (i.e., material-from-motion). These findings have shown that light travels differently across surface depending on the specular property of the material, creating distinct optical flow, which the visual system detects. For instance, a static computer-generated matte object painted with highlights corresponding the 3-D shape of that object, is judged equally glossy as static shiny object rendered with specular highlights. However, when the objects are rotated, the shiny object is instantly judged glossy whereas the matte object is not. The rotation results in distinct optical flow for the two types of materials, in which the shiny object has light running across the surface, while the matte object has no specular flow and appears to have highlights *painted* on it (Doerschner et al., 2011;

Hartung and Kersten, 2002; Kam, et al., 2015). In turn, specular flow has been found to influence judgment of rotation motion of an object (Doerschner et al., 2013). Clearly, reflectance is a complex material property that needs to be studied in a broader experimental setting than simple passive viewing of static images, considering that perceived glossiness depends on the available information at that moment.

Material recognition based on reflective properties alone, has been challenging to explain due to the complex visual structure of specular highlights and their interplay with the environment the material is presented in. For instance, glossy materials that are highly reflective (e.g., glass vase) have specular highlights that display distorted images of the surrounding environment, that are a *part* of the external environment rather than the material itself. As seen in figure 1, the distorted images of windows are not painted decorations on the vase itself, but rather a reflection of the living room the vase is positioned in. Somehow, we easily recognize these distorted images as specular reflectance rather than decorations, both for computer-generated glossy surfaces (Nishida and Shinya, 1998), and real-world images (Fleming, et al., 2003) as long as the illumination is realistic and not atypical. This suggests we have stored knowledge of typical structures of reflectance.

Others have argued that visual material perception is merely a perception of low-level image statistics (e.g., Nishida and Shinya, 1998; Sawayama, et al., 2017). For instance, the intensity of the reflectance property of a material can be represented by a bidirectional reflectance distribution function (BRDF), which is a function that is based on the amount of incident light hitting the material surface, and the direction of the light reflected from the surface, given that reflectance is directional (Nicodemus et al., 1977). Using image statistics as visual cues to particular materials is understandably a logical approach to understand visual perception of materials, as such an approach offers a measurable connection between the physical world (the distal stimuli), the image statistics (proximal stimuli) and material perception (see review in Nishida, 2019). Regardless, psychophysical data suggest that some assumptions are required together with the image statistics in order to recognize material properties, as perceived material properties can have perceptual effects on one another. For instance, perceived glossiness is not only influenced by the illumination strength, but also by the surface relief (bumpiness), and the three-dimensional shape of the material (Fleming, et al., 2003; Ho, et al., 2008; Marlow et al., 2012; Nishida and Shinya, 1998; Olkkonen and Brainard, 2011; Vangorp, et al., 2007; see also Fleming, 2014). Figure 3 demonstrates how adding reflective features to rough surfaces can dramatically alter their appearances.

Although surface gloss is perceived visually it can convey tactile information, such as friction (e.g., Paulun et al., 2016). Interestingly, there are studies that have shown that surface gloss is associated to more qualities than surface friction. Bennet and Porteus (1961) found that people could predict moisture by estimating the number of areas on the surface that had high luminance (i.e., white highlights), and Schaller and Park (2011) demonstrated a close interplay between glossiness and the functioning of the behavioral immune system (BIS), a concept developed by Schaller (2006) to describe the psychological- and behavioral aspect of the physiological immune system (PIS). When pathogens, toxins, or allergen, enter the body, PIS becomes activate and starts to eliminate the invading stimuli to protect the body and prevent sickness. Since PIS is merely *reactive*, the role of BIS is to compensate the immune system by detecting the visual cues of pathogens and other uninvited guests before they enter the body, and consequently initiate an avoidance behavior. In a recent study, Iwasa et al. (2020), demonstrated a presence of a visual perceptual mechanism for detecting moisture within BIS. Iwasa et al. (2020) presented images of dough with moisture at various scales to participants in a pairwise manner and asked the participants to compare them according to perceived moisture as well as to rate them in terms of disgust evoked by them (e.g., how likely they were to touch/avoid the dough), in order to assess the BIS response. Their findings showed that both moisture magnitude and the number of areas with high luminance (i.e., highlights) accurately predicted the visual detection of moisture. Moreover, they found that BIS response peaked for images of dough having moderate amount of moisture, indicating an association between glossiness and moisture. Certainly, moisture is related to friction and is therefore perceived as slipperiness, moist surfaces (e.g., wet floors) are usually slippery. Nevertheless, glossiness could possibly be associated with other qualities such as hardness. For instance, some hard materials have lustre or shiny appearance due to their nature (e.g., gold, steel), whereas other have dull appearance (e.g., paperboard, wood) but are often coated with varnish or other glossy material to ensure durability.

Roughness describes the height differences of a surface relief (i.e., bumpiness) that provide haptic impressions which can be utilized for surface categorization, both in humans **(**Bergmann Tiest and Kappers, 2007; Bergmann Tiest, 2010**)**, and in biologically inspired self-organizing robot systems (Johnsson and Balkenius, 2008 and 2010). Even though roughness is typically studied within the haptic system (and often using sandpapers), its salient visual features have proven to be highly informative attributes for visual recognition of materials as well (e.g., Adelson, 2001; Anderson, 2011; Bergmann Tiest and Kappers, 2007). After all, surface roughness (and smoothness) can be found anywhere, from rocky stone terrain to polished concrete floor, where their surface characteristics are first perceived visually before being touched or walked on, which makes their visual characteristics equally important as their haptic impressions. Consequently, the consensus is that roughness is an important material property for both the haptic and the visual system to perceive, and that the systems are equally good at perceiving roughness (Bergmann Tiest and Kappers, 2007). Figure 4 illustrates examples of computer-generated surfaces that vary in roughness.

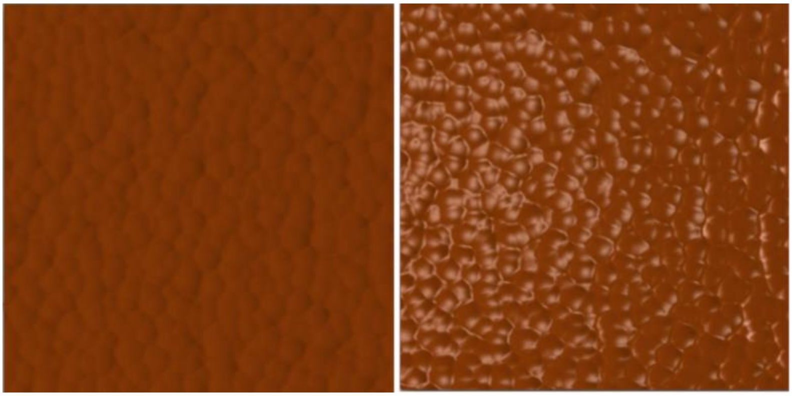

Figure 3. Examples of identical surface textures, presented with and without reflectance properties to demonstrade the qualitative difference the interpay between roughness and glossiness has on material appearances.

Figure 4. Examples of computer-generated surfaces resembling an avocado skin. Here, surface roughness is reduced by increasing the noise level of the Voronoi tessellation texture map, creating larger tessellation tiles and consequently smoother appearance.

In material perception, roughness and glossiness are considered to operate on the same scale to describe surfaces (microscale). Other scales being mesoscale (e.g., visible textures and shapes like wood appearance or folded cloth), and megascale (e.g., overall object shape) (Koenderink and van Doorn, 1996). As a result, a perceptual interplay in judgment of materials has been found between roughness and glossiness, in which negative correlations have been found between visually judged roughness and glossiness (Fleming et al., 2013). Surfaces with rough texture are judged less glossy compared to surfaces with smooth texture, which is perhaps unsurprising considering how light interacts with surfaces. In particular, the incidental light reflects differently depending on the surface structure of the material. Smooth surfaces tend to reflect the light into one direction not far away from the initial entry-location, creating a perfect reflection of the environment (think mirrors) or alternatively very bright highlights. In turn, rough surfaces scatter the light into many directions, creating diffused reflection and matter appearance (e.g., Komatsu and Goda, 2018).

Hardness, sometimes known as compliance or stiffness, is a physical property providing haptic information and is frequently used for describing the elasticity or deformability of an object/material, e.g., the glass table is hard whereas the cleaning sponge is soft. Therefore, the perception of hardness has been almost exclusively studied within the haptic system, in which judged hardness is based on surface deformation and force/finger-displacement (Bergmann Tiest and Kappers, 2009; Bergmann Tiest, 2010). Nevertheless, we often have clear assumptions of how hard or soft materials are before touching them because of an association mechanism that we have, which is based on prior experiences with similar materials (Baumgartner et al., 2013; Bergmann Tiest, 2010; Paulun et al., 2017). Although perceived hardness without tactile feedback is often based on the visual characteristics of the material, studies have shown that perceived hardness can be based on auditory information as well, in which contact sounds provide information about the hardness/softness of the material (Arnott, et al., 2008; Giordano and McAdams, 2006; Vickers, 1981; and Vickers, 1987).

Oftentimes, visual perception of hardness is based on shape cues (e.g., folds in textiles, visual deformation depth) or how shape changes over time (Fakhoury, et al., 2015; Paulun et al., 2017; Schmid and Doerschner, 2018; Schmidt, et al., 2017; Zoeller et al., 2019). Interestingly, correlations have been found between perceived weight and hardness, in which rated soft materials tend to be rated lighter, and rated hard materials tend to be rated heavier, which is perhaps unsurprising considering that light objects are often less dense and tend to be softer in the sense they can easily deform (think textiles) (Ashby, 2013). Furthermore, certain visual features appear to be correlated with increased rated weight and hardness, such as surfaces with *spiky* appearances that tend to be rated both heavier and harder than less spiky surfaces (Schmid and Doerschner, 2018; Schmidt, et al., 2020).

Certainly, there are plenty of other material properties worth mentioning, although they were not particularly examined in this doctoral dissertation. Color, for instance, has been briefly mentioned in the context of material recognition. Although color has been found to facilitate the recognition of materials it is not crucial for categorizing materials (Yoonessi and Zaidi, 2010; Zaidi, 2011), thus color was not directly manipulated in the research papers presented in this dissertation. In the recent years there has been a growing interest in perception and estimation of mechanical properties of materials, specifically in the mechanical difference between *rigid* and *non-rigid* materials. Liquids (non-rigid material) and visual estimation of viscosity of liquids has been of special interest, where it has been shown that we are fairly good at estimating viscosity based on how the liquids flow (Kawabe, et al., 2015; Van Assen, et al., 2018) and shape cues (Paulun et al., 2015). Textiles are also non-rigid materials which are often studied in relations to perceived softness/comfortability (e.g., Xiao, et al., 2016). Also, visual perception of complex material properties such as transparency and translucency has become increasingly popular in the recent years, due to the increased interest in non-rigid materials and progress in computer graphics (Fleming, et al., 2011; Motoyoshi, 2010).

How do we perceive *Stuff*?

It is remarkable that we manage to visually categorize materials into clusters and categories such as *ceramics, papers, glass,* and *liquids,* considering that there are variations in visual characteristics within a single material category, and similarities between different types of material categories. For instance, you can easily recognize that the paperback dissertation, that you are currently holding in front of you, is made from paper. The same goes for the takeaway coffee cup you got at *Pressbyrån* this morning, you assign that disposable cup to the material category *paper,* without any hesitations. More interestingly, although both items, the printed copy of this dissertation and the paper cup, are both considered to belong to the same material category and therefore categorized as such, they are by no means identical in appearance. Rather, their optical images differ in shape, illumination, and wear due to external forces. The printed pages in the dissertation are thin and easily folded or teared, whereas a greater force is required to make at least a small dent on the paper cup, what then to tear it apart. The two items appear qualitatively different and afford different actions but are nevertheless judged to be made from the same material. Thus, recognizing materials is a complex process when considering that members of the same material category can appear differently. Moreover, members of different material categories can appear more similar than members that belong within the same material category. As figure 5 illustrates beautifully, quesadillas may look like the skin of giraffes, and blueberry muffins may look like Chihuahua, in a way that can be hard to visually distinguish at a first glance.

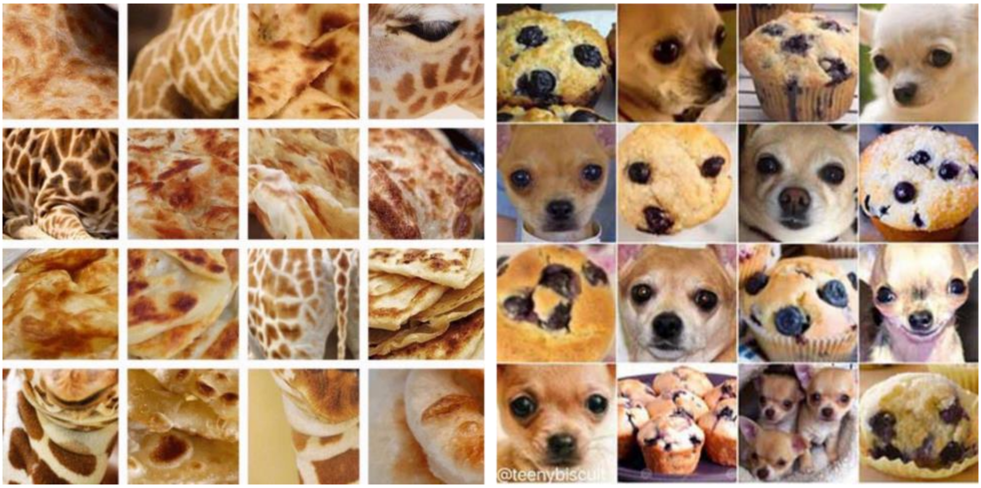

Figure 5. Quesadilla or giraffe? Muffin or chihuahua? Image courtesy of Karen Zack @teenybiscuit.

Conventionally, feed-forward models have been used when describing visual perception, in which the structured light enters the eyes after it has interacted with objects in the external environment which, in turn, creates retinal images of the environment comprised of low-level features (e.g., edges, contrasts, color, and motion). A detected edge (*proximal stimuli*) in the retinal image is, however, not particularly informative about visual environment as it could correspond to multiple different properties in the external environment (*distal stimuli*). Consequently, the visual system begins its analysis by encoding these local image properties and transforms them into representations of surfaces and materials. Here is when interpretation comes in, which is the never-ending struggle of mid-level vision theories, as well as high-level vision theories. It is challenging to explain the mechanism of perceptual organization since the mapping between the distal and the proximal stimuli is ambiguous (the inverse problem in optics) (see, e.g., Anderson, 2011; Fleming, 2017). In this section, two practical frameworks that explain the organization of materials are introduced.

Conceptual spaces of materials

One way to understand how the cognitive system recognizes material categories from diverse groups of materials, and how material properties are connected with the categories, is to define perceived materials as learned concepts rather than estimations of physical parameters. That way, we can utilize category-based inference models to explain the organization of perceived *stuff*. The theory of *Conceptual Spaces* proposed by Gärdenfors (2000) is an explanatory framework that uses geometric psychological space to describe the organization of concept knowledge and concept learning using *quality dimensions* and *domains.* According to the framework, our perceptions of real-world objects, such as a drinking glass, are represented as concepts that are comprised of *quality dimensions*, the qualities of the objects that either are obtained via the sensory modalities (e.g., weight, color, transparency, hardness and glossiness) or are abstract attributes assigned to the perceived objects (e.g., fragility, pleasantness and consolation). A *domain* is a cluster of quality dimensions, which represents qualities that are separable from other clusters of dimensions representing that same real-world object. Figure 6, a reprint from Gärdenfors and Warglien (2012) illustrates the definition of quality dimensions and domains using a color spindle. The color spindle describes all possible colors perceived by humans using only three quality dimensions that each of which contain color qualities that are separable from one another, namely the hue (the color itself), the saturation (the intensity of the color), and the brightness (ranging from black to white) of colors. Together, these three dimensions, including the variation within each dimension, make up the *color* domain in color perception, which is separable from other domains.

For example, does the domain for *roughness* comprise of quality dimensions representing surface properties (e.g., the height of the surface relief and the frequency of the spatial bumpiness), and those dimensions contain values that cannot exist as values in the dimensions found within the color domain, or in the *glossiness* domain. Ultimately, a concept of a real-world object is represented as a *region* within the conceptual spaces, using either single or multiple domains that are correlated.[1]

[1] Concepts are not necessarily comprised of collections of correlated domains, as real-world objects that are simple in nature can also be described using only a single domain. A good example of concepts made of a single domain are simple stimuli used for psychophysical experiments examining reaction times and choices to infer about the otherwise unobservable computation mechanism during a cognitive task. Reaction time studies often require using simple stimuli made of a single property such as color, line, or orientation.

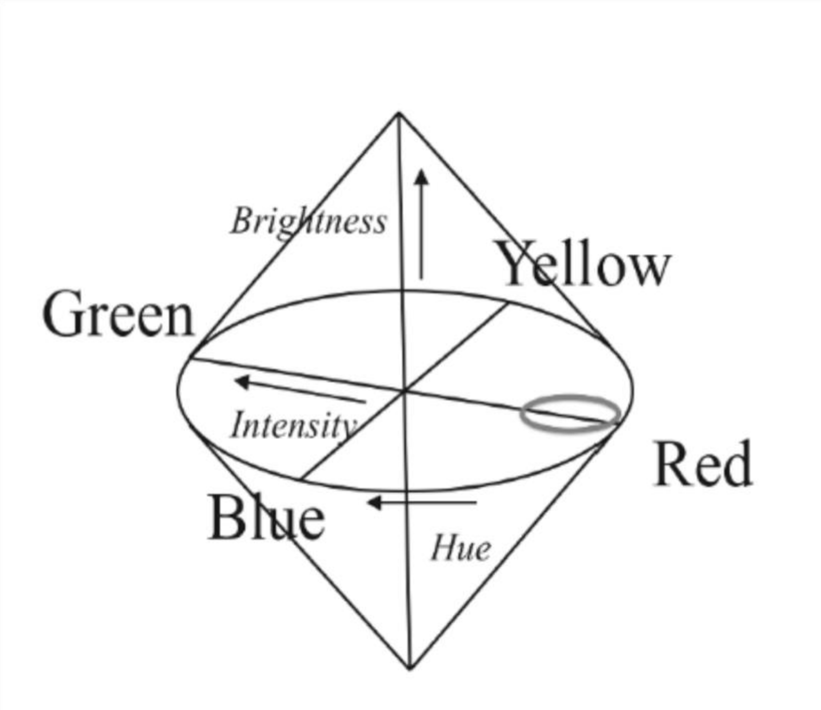

Figure 6. An example of a color spindle representing the color domain in human color perception, comprised of three quality dimensions that are dissimiliar, brightness, intensity, and hue (reprinted from Gärdenfors and Warglien, 2012, pg. 490).

Concepts are represented as *convex regions,* which are geometrical representations of the properties of the quality dimensions and are found in domains. The regions are attached together in a form of a Voronoi tessellation that changes its formation when new concepts are encountered. Concept learning occurs when the formation of *regions* within the conceptual space is altered. When new concepts are constructed, new data points are added to the quality dimensions within the domains, altering the convex regions and consequently expanding the size and shapes of the tessellation tiles and ultimately expanding the conceptual spaces. Figure 7 is a reprint from Gärdenfors (2000) and illustrates an example of a Voronoi tessellation. The number of data points within the dimension describes the perceived variation of the property of that dimension, and the average of all the data points within the same region defines the prototypical reference point of that convex region. The prototypical points are constantly changing and shifting to new coordinates when new information is learned and consequently new data points are added to the quality dimensions, changing the formation of the convex regions. The geometric distances (*metric*) between these prototypical points within the geometric space describe the judged similarities between the properties and the concepts. The shorter the distance the more similar they are. That said, the cognitive system only needs to rely on the size of the tessellation tiles, that is, the prototypical points and the number of points it is based on, when learning a new concept, instead of remembering the exact categorization of all data points (Gärdenfors, 2008). The theory of *Conceptual spaces* has inspired research on the mathematical structure of category-based inference models (Osta-Vélez and Gärdenfors, 2020), and preceded modelling of neural representations of psychological spaces (Balkenius and Gärdenfors, 2016), including a mapping of cognitive spaces in the hippocampus specifically (Bellmund, et al., 2018).

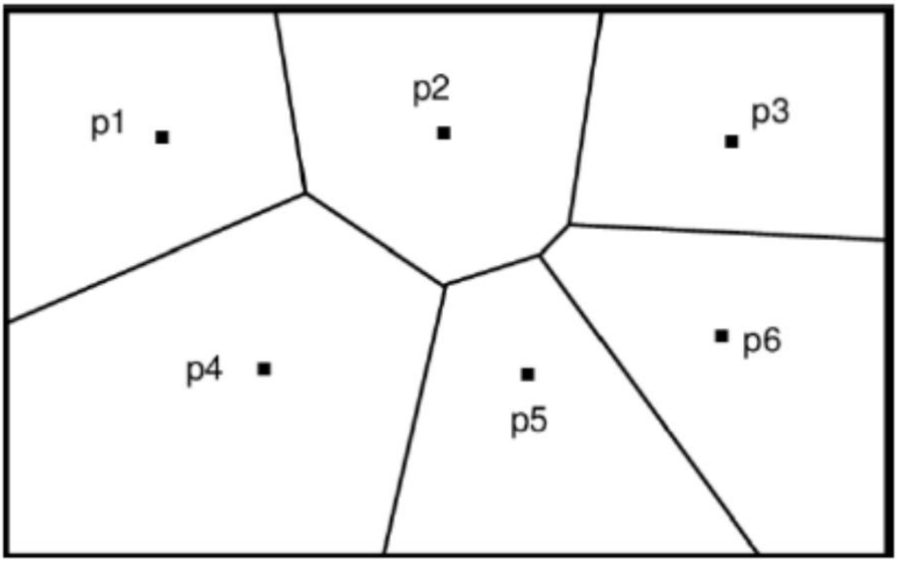

Figure 7. Illustration of a Voronoi tessellation of convex regions containing prototypical points (from Gärdenfors, 2000).

Mathematical models

We move around in the physical world learning about its properties, however, the sensory information we receive is often too ambiguous, which creates uncertainty for the visual system. For instance, a spilled face cream on a dining table might look like a spilled yoghurt, as both are white colored with smooth texture and similar density. Thankfully, our previous experience with these two substances and the provided context the substances are presented in (i.e., the dining table), we are much more likely to assume that the stain on the dining table is a spilled yoghurt rather than face cream. Lately, mathematical models, in particular Bayesian inference models, have been considered promising for explaining the cognitive process behind visual recognition (e.g., Kersten, et al., 2004). Bayesian inference models can account for the uncertainty that derives from ambiguous visual information, by using probabilities to describe how likely an event *A* has taken place given certain conditions, where the probability of the event *A* is updated as more information becomes available to us. Furthermore, the probability of event *A* may be based on more than the sensory information provided, since the probability of event *A* may also be based on prior knowledge and previous experience of similar events (Nour and Nour, 2015).

The *statistical appearance model* proposed by Fleming (2014, 2017; see also Fleming and Storrs, 2019) is another promising framework for explaining how material perception is achieved. Similar to Gärdenfors's theory of conceptual spaces, this statistical appearance model is a framework that assumes the presence of a high-dimensional feature space. However, unlike the theory of conceptual spaces, the statistical appearance model involves generative models that are specifically designed for encoding and predicting materials. According to Fleming (20xx), the perceiver learns the variation of the informative attributes that represent materials, and their exemplars, and develops generative models of the *typical appearances* of each material. These are based on learned differences and similarities both within and between materials, which allows for material recognition to occur. For instance, the typical attributes for a high-gloss material, such as a glossy plastic laminate used on interiors, are reflective highlights. The informative highlights differ in contrast and size depending on how glossy the material is, especially in comparison to matte materials like sanded wood that has low-gloss attributes. Rather than estimating the true physical parameters of surface properties (e.g., surface reflectance), the statistical appearance model relies on feature dimensions in a high-dimensional space and unsupervised learning rules when organizing perceived materials. In other words, instead of learning the ground truth physical principles of the external environment, in which pre-defined physical parameters of material properties are estimated, we learn to encode the systematic variations, both within and between materials. Recently Fleming and Storrs (2019) successfully implemented the statistical appearance model using deep-learning networks and unsupervised learning rules, which is generally a more realistic description of how we learn about our surroundings in the real-world, i.e., without pre-determined labelled training data. It is much more feasible and less time consuming to learn regularities in data, rather than to match instances and stimuli with stored sets of pre-labelled data.

Material perception and action

At the kitchen table, the three-year old Ronja bangs her fork made of stainless steel on several objects laid out in front of her on the table. Rhythmically banging on each item on the table and the table itself, she creates sounds that vary in pitch depending on what type of material each object is made of. As an experiment I ask her "is this hard or soft?" and point at the kitchen paper roll. Without hesitation Ronja replies "it is soft" and starts poking the paper roll with her fork to confirm (and to show me how soft it is), next, she moves on to the paper roll holder (made of stainless steel), and while hitting it with her fork she yells "and this is not soft, I know it!", and continues to drum on it because it makes far more exiting noise than the paper roll. – *a memory from everyday life.*

Developmental studies and learning physical properties

As we grow, our knowledge of the physical properties of the world expands. Through repeated encounters of seeing, grasping, and handling objects made of various materials, infants and toddlers learn that sharp corners can hurt little fingers while round edges are safer. Moreover, they learn that objects made from light materials are easier to handle and play with, compared to objects made of heavier material. A large part of our scientific understanding of material perception is, however, based on findings on adults' ability to recognize and categorize materials, while research on the developmental aspect of material knowledge has only newly begun. Yang, et al. (2013a) have shown that the ability to visually perceive and represent surface gloss emerges at an early age, and it is sufficiently developed at the age of 7 months old. Using a preferential looking task, Yang et al. (2013a) demonstrated that infants at the age of 7-8 months old could use specular properties to distinguish matte materials from shiny material when observing yellow and gold objects, whereas 5-6 months old infants lacked that ability. Consequently, this confirms their previous finding that representation of glossiness emerges around the age of 7 months old and not earlier (Yang, et al., 2011). Although the ability to discriminate colors and lightness has been shown to emerge as early as 4 months old (Yang, et al., 2013b), the age of 7 months appears to be the "*golden age*" of the emergence of the ability to discriminate visual attributes, such as telling shapes from shadings, (Imura et al., 2008; Tsuruhara, et al., 2009), and motions from shadows

(Yonas and Granrud, 2006; Imura et al., 2006). In addition, Paulus and Hauf (2011) have demonstrated the developmental progression of material knowledge in 9-, 11-, and 13- months old children using a preferential reaching task which showed that 11 months old infants were able to visually recognize previously preferred light object over a heavy object, and 13 months old infants were able to generalize that association on to a new object made of same material as the previously explored object (thus, based on their visual features). Furthermore, accurate visual recognition fails often for children under the age of two, especially when categorizing instances presented in visually crowded scenes (Farzin, et al., 2010). Even though the children are judging instances from known categories, such as *dogs*, they tend to over- and undergeneralize the instances, such as categorizing all four-legged furry animals as dogs. Interestingly, the children have developed the adult-like ability to infer about variations within a known category based on one learned example, already at the age of 2 years old. At that age they rarely put all four-legged furry animals into the same animal category, but rather they can infer that some four-legged animals are actually *cats*, even after meeting or seeing only one cat (see review in Smith and Slone, 2017). With age the children develop better motor skills and become more flexible to move around and explore the physical world. As a result of that development, children experience more variations of visual images and learn that categories are comprised of several variations of instances. That said, children learn about objects (and materials) through self-generated visual experiences.

In a recent study, Balas (2017) studied the progression of material knowledge and examined material categorization as a function of age and type of material, to demonstrate how material categorization at a higher level than discrimination of visual attributes develops during childhood. In his study, he compared the ability to categorize and discriminate materials between adults and five-to-ten years old children. An age range typically considered the developmental years for learning visual features for categorization (e.g., Pascalis et al., 2011). Balas (2017) found that the ability to use visual summary statistics to discriminate materials was already well developed at a young age, as young children showed adult-like abilities to encode materials based on summary statistics alone. However, it was much more challenging for the young children to assign labels to the material categories based on those visual summary statistics, which suggests that material perception develops slowly throughout the middle childhood. Later, using the same age-range as before, Balas et al. (2020) showed that the visual processing of materials shifted from relying on local visual information to relying on global visual features, and that shift occurred slowly during the middle childhood.

Figure 8. Epi, a humanoid robot, preparing to lift various objects.

In terms of material recognition and motor control, several developmental studies have demonstrated that the control of the temporal span between peak velocity in reaching and the end of the reaching behavior (i.e., *adjustment time*) is an adaptive mechanism, which arises early in the perceptual motor development and is learned through repeated experience (Berthier et al., 1999; Berthier and Keen, 2005; Karl and Whishaw, 2013; Newman, Atkinson, and Braddick, 2001; Pryde et al., 1998; Rocha et al., 2013). Studies have shown that infants initially lack the detailed motor control in reaching, therefore instead the infant reaches in a ballistic manner without any deceleration period to prepare the grip. Later, around the age of 5-6 months, the deceleration phase develops when infants start to be able to slow down the hand when reaching for objects and they continue to improve with age. Through repeated experiences of interacting with objects in various shapes and sizes, we learn the association between object properties and how to successfully interact with the objects. Furthermore, and in a similar way to children learning that a large object is usually heavier than a smaller one, we learn that objects made from different materials require different handling and that we need to scale our reach and grasp accordingly to achieve successful interaction. Moreover, we need to estimate the required force to lift objects that are made of materials that are durable (e.g., a golf ball) vs. other objects are made of a more fragile material and therefore require a careful approach to be handled without damage (e.g., an uncooked egg).

An alternative approach is to examine material perception using epigenetic robotics (see paper IV). Epi (see figure 8) is a humanoid robot controlled by a cognitive infrastructure, Ikaros, that uses system-level brain modelling for implementing and experimenting cognitive theories (Balkenius, et al., 2010). Epi's "eyes", i.e., the cameras, detect the object's visual appearance and send the information to sensory modules that perform a low-level visual processing on the visual input, which produce vectors of visual features that are then forwarded to an association module. Furthermore, Epi has two arms that move five degrees of freedom each, making it possible for the robot to manipulate objects in its hands, and the servos controlling the arms and the shoulders produce feedback signals that are read by the robotic system. The haptic feedback signals can then be mapped to the previously detected visual features of objects using associative learning modules. After repeated interactions with the objects, i.e., lifting, weighing, squeezing, Epi learns to associate the visual properties to the non-visual haptic features, producing multi-modal representation of the objects. Ultimately, Epi will be able to accurately plan the required hand movements to handle the objects before touching the objects. Epi's cognitive infrastructure is ideal for studying the developmental mechanism underlying the learning of a new skill, such as sorting materials, in which the robot learns to associate visual features of real-world objects to haptic object properties, such as weight and hardness.

We as active perceivers

We are active perceivers moving around within the physical world, interacting with objects made of different types of materials that vary in visual and mechanical characteristics. In many scenarios, the materials even undergo changes in visual characteristics due to external influences that need to be recognized and accounted for when moving around in order to prevent erroneous decisions or motor actions. As pedestrians we walk on flat and stable concrete surfaces with certain frictional properties that are easily influenced by external forces, like rainy weather. The frictional property of the concrete surface changes dramatically in the winter months when the sidewalk turns icy and slippery due to significant decrease in temperature. Thankfully, the changes in surface friction, due to the weather, are accompanied by changes in surface appearance, in which the dry and matte concrete surface is replaced by a sparkling and reflective surface. The "new" visual surface appearance is then utilized to infer the changes in frictional properties, and consequently motor strategies are reconfigured in order to reduce the risk of falling on the slippery sidewalk.

In this doctoral dissertation I set out to examine material perception as a perception-action cycle, in which predictions about the physical world are made based on past experiences with similar scenes, and actions are taken based on those predictions. The fundamental characteristics of the perception-action cycle is the continuous feedback loop between the perception and the action, which allows predictions being updated when a mismatch is present between the predictions and the perceived outcome. This results in constant expansion of material knowledge. To achieve this, I examined motor actions while participants interacted with everyday life objects made of various materials (Paper I, II, IV). I also examined changes in material perception while participants observed animations of interacting materials to address context effects in visual perception of materials (Paper III). Clearly, visual material properties have a complex but informative role in material perception. Our knowledge of materials, i.e., how they look and *behave*, and how objects should be handled based on the *stuff* they are made of, is constantly being updated throughout the childhood years, and continues to develop into our adult years as we learn from experience (Fleming and Storrs, 2019). This makes research on material recognition and its daily purpose even more interesting.

Our knowledge of materials is constantly expanding as we develop from novices to experts in recognizing and distinguishing materials. A small child that learns to handle their toys that are made from different materials, to a textile designer that is an expert in recognizing and analyzing significant elements in textiles. All materials have typical visual appearances and functional roles (e.g., slippery, bounceable, deformable) that we have learned to recognize and utilize for guiding our motor actions, as the features (or perceived affordance) determine how we should interact with them. Material properties are multidimensional as they are perceived by all senses (e.g., Castiello, et al (2010) on contact sound and grasping), although visual properties have attained most attention in the research field. Oftentimes, the functional properties of materials evoke responses that are affective, in which feelings are associated with certain types of materials (e.g., the baby’s plush toy is not only *soft* and *fluffy*, but it also provides a baby with a feeling of *warmth and security,* another example, the favorite t-shirt that often is associated with luck, happiness and comfort). In a broader perspective, material knowledge is much more than recognition of material properties. It serves a fundamental survival role, such as detecting contaminated drinking water and spoiled food based on visual information, in which a brown and cloudy water could be an indication that unsafe pathogens are present in the water, and spoiled food is indicated by changes in color and surface texture (e.g, fuzzy mold). Material knowledge is also comprised of attributes that are far more subjective in nature than material properties such as glossiness and roughness. For instance, the Japanese concept *Shitsukan* (質感, translation: the sense of quality) describes material perception as not only perception of material properties, but it also describes the physical state of materials (e.g., soft, battered, used or new), including the subjective impression of the materials (e.g., comfortable, cozy) (see review in Komatsu and Goda, 2018). Research on material perception is still young, but has grown exponentially in the recent years, as it is an important research topic for understanding how we go about our daily lives.

References

Adams, W.J., Kerrigan, I.S., and Graf, E.W. (2016). Touch influences perceived gloss. *Scientific Reports, 6*, 21866. DOI: 10.1038/srep21866.

Adelson, E. H. (2001). On seeing stuff: The perception of materials by humans and machines. In B. E. Rogowitz, and T. N. Pappas (Eds.), *SPIE Proceedings, Human Vision and Electronic Imaging VI*, *4299*, 1–12.

Anderson, B. L., and Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness *Journal of Vision, 9*(11), 10. DOI: 10.1167/9.11.10.

Anderson, B. L. (2011): Visual perception of materials and surfaces. *Current Biology, 21*(24), 978-983. DOI: 10.1016/j.cub.2011.11.022.

Arnott, S.R., Cant, J.S., Dutton, G.N., and Goodale, M.A. (2008). Crinkling and crumpling: an auditory fMRI study of material properties. *NeuroImage, 43,* 368-378. DOI:10.1016/j.neuroimage.2008.07.033.

Ashby, M. F. (2013). *Materials and the environment: eco-informed material choice.* Elsevier, Amsterdam, Netherlands. DOI: 10.1016/C2010-0-66554-0.

Balas, B. (2017). Children's use of visual summary statistics for material categorization. *Journal of vision*, *17*(12), 22. DOI: 10.1167/17.12.22.

Balas, B., Auen, A., Thrash, J., and Lammers, S. (2020). Children's use of local and global visual features for material perception. *Journal of Vision, 20*(2), 10. DOI: 10.1167/jov.20.2.10.

Balkenius, C., Morén, J., Johansson, B., and Johnsson, M. (2010). Ikaros: Building cognitive models for robots. *Advanced Engineering Informatics, 24,* 40-48. DOI: 10.1016/j.aei.2009.08.003.

Balkenius, C., and Gärdenfors, P. (2016) Spaces in the Brain: From Neurons to Meanings. *Frontiers of Psychology, 7,* 1820. DOI: 10.3389/fpsyg.2016.01820.

Baugh, L.A., Kao, M., Johansson, R.S., and Flanagan, J.R. (2012). Material evidence: interaction of well-learned priors and sensorimotor memory when lifting objects. *Journal of Neurophysiology, 108*(5), 1262–1269. DOI: 10.1152/jn.00263.2012.

Baumgartner, E., Wiebel, C., and Gegenfurtner, K. (2013). Visual and Haptic Representations of Material Qualities. *Multisensory research, 26*, 429-55. DOI: 10.1163/22134808-00002429.

Baumgartner, E., and Gegenfurtner, K. R. (2016). Image statistics and the representation of material properties in the visual cortex. *Frontiers in Psychology, 7,* 1185. DOI: 10.3389/fpsyg.2016.01185.

Beck, J., and Prazdny, S. (1981). Highlights and the perception of glossiness. *Attention, Perception, & Psychophysics*, *30*(4), 407-410.

Bellmund, J., Gärdenfors, P., Moser, E., and Doeller, C. (2018). Navigating cognition: Spatial codes for human thinking. *Science, 362*(6415). DOI: 10.1126/science.aat6766.

Bennett, H. E., and Porteus, J. O. (1961). Relation between surface roughness and specular reflectance at normal incidence. *Journal of the Optical. Society of. America. 51*, 123–129. DOI: 10.1364/JOSA.51.000123.

Bergmann Tiest, W. M., and 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.

Bergmann Tiest, M. W., and Kappers, A. M. L. (2009). Cues for haptic perception of compliance. *IEEE Transaction on Haptics, 2*(4), 189–199. DOI: 10.1109/TOH.2009.16.

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

Castiello, U., Giordano, B. L., Begliomini, C., Ansuini, C., and Grassi, M. (2010). When ears drive hands: The influence of contact sound on reaching to grasp. *PLOS ONE, 5*(8), 1-9.

Chadwick, A.C., and Kentridge, R.W. (2015). The perception of gloss: A review. *Vision Research*, *109*(Part B), 221-235. DOI: 10.1016/j.visres.2014.10.026.

Cui, Q., Mcintosh, S., and Sun, H. (2018). Identifying materials of photographic images and photorealistic computer-generated graphics based on deep CNNs. *Computers, Materials and Continua, 55*(2), 229–241. DOI: 10.3970/cmc.2018.01693.

Doerschner, K., Fleming, R.W., Yilmaz, O., Schrater, P. R., Hartung, B., and Kersten, D. (2011). Visual motion and the perception of surface material *Current Biology, 21*(23), 1–7. DOI: 10.1016/j.cub.2011.10.036.

Doerschner, K., Yilmaz, O., Kucukoglu, G., and Fleming, R. W. (2013). Effects of surface reflectance and 3D shape on perceived rotation axis. *Journal of Vision, 13*(11), 8. DOI:10.1167/13.11.8.

Fakhoury, E., Culmer, P. R., and Henson, B. (2015). *The effect of indentation force and displacement on visual perception of compliance*. In Colgate, J. E. (Ed.), IEEE World Haptics Conference 2015: Evanston, Il., USA, 22-26, June 2015 (pp. 88–93). Piscataway, NJ: IEEE, DOI: 10.1109/WHC.2015.7177696.

Farzin, F., Rivera, S. M., and Whitney, D. (2010). Spatial resolution of conscious visual perception in infants. *Psychological science, 21*(10), 1502–1509. DOI: 10.1177/0956797610382787.

Flanagan, J.R, Bittner, J.P., and Johansson, R.S. (2008). Experience can change distinct size-weight priors engaged in lifting objects and judging their weights. *Current Biology, 18*, 1742–1747. DOI: 10.1016/j.cub.2008.09.042.

Flanagan, J. R., and Beltzner, M. (2000) Independence of perceptual and sensorimotor predictions in the size-weight illusion. *Nature Neuroscience, 3*, 737–774. DOI: 10.1038/76701.

Fleming, R. W., Dror, R. O., and Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3*(5), 347 – 368. DOI: 10.1167/3.5.3.

Fleming, R. W., Jäkel, F., and Maloney, L. T. (2011). Visual perception of thick transparent materials. *Psychological Science, 22*(6), 812–820. DOI: 10.1177/0956797611408734.

Fleming, R. W., Wiebel, C., and Gegenfurtner, K. (2013). Perceptual qualities and material classes. *Journal of Vision, 13*(8), 9. DOI: 10.1167/13.8.9.

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

Fleming, R.W., Gegenfurtner, K. R., and Nishida, S. (2015). Visual perception of materials: The science of stuff. *Vision Research, 109*(Part B), 123-124. DOI: 10.1016/j.visres.2015.01.014.

Fleming, R. W. (2017). Material perception. *Annual Review of Vision Science, 3*(1), 365–388. DOI: 10.1146/annurev-vision-102016-061429.

Fleming, R. W., and Storrs, K. R. (2019). Learning to see stuff. *Current Opinion in Behavioral Sciences*, *30*, 100-108. DOI: 10.1016/j.cobeha.2019.07.004.

Fujisaki, W., Goda, N., Motoyoshi, I., Komatsu, H., and Nishida, S. (2014). Audiovisual integration in the human perception of materials. *Journal of Vision, 14*(4), 12. DOI: 10.1167/14.4.12.

Gärdenfors, P. (2000). *Conceptual Spaces: On the Geometry of Thought*. Cambridge: MIT press.

Gärdenfors, P. (2008). Concept learning. In P. Gärdenfors and A. Walling (Eds.), *A Smorgasbord of Cognitive Science* (pp. 165–181). Nora: Nya Doxa.

Gärdenfors, P., and Warglien, M. (2012). Using Conceptual Spaces to Model Actions and Events. *Journal of Semantics. 29*, 487-519. DOI: 10.1093/jos/ffs007.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton-Mifflin.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Boston: Houghton-Mifflin.

Giordano, B., and Mcadams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America, 119*, 1171-1181.

Goda, N., Tachibana, A., Okazawa, G., and Komatsu, H. (2014). Representation of the material properties of object in the visual cortex of nonhuman primates. *The Journal of Neuroscience*, *34*(7), 2660–2673. DOI: 10.1523/JNEUROSCI.2593-13.2014.

Gordon, A. M., Forssberg, H., Johansson, R. S., and Westling, G. (1991). Visual size cues in the programming of manipulative forces during precision grip. *Experimental Brain Research. 83*(3), 447–482. DOI: 10.1007/bf00229824.

Gordon, A., Westling, G., Cole, K., and Johansson, R. (1993). Memory representations underlying motor commands used during manipulation of common and novel objects. *Journal of Neurophysiology, 69*(6), 1789-1796. DOI: 10.1152/jn.1993.69.6.1789.

Gordon, A., Forssberg, H., and Iwasaki, N. (1994). Formation and lateralization of internal representations underlying motor commands during precision grip. *Neuropsychologia*, *32*(5), 555-568. DOI: 10.1016/0028-3932(94)90144-9.

Hartung, C. B., and Kersten, D. (2002). Distinguishing shiny from matte *Journal of Vision, 2*, 551–551.

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

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

Imura, T., Yamaguchi, M.K., Kanazawa, S., Shirai, N., Otsuka, Y., Tomonaga, M., and Yagi, A. (2006). Perception of motion trajectory of object from the moving cast shadow in infants. *Vision Research, 46*(5), 652-7.

Imura, T., Yamaguchi, M. K., Kanazawa, S., Shirai, N., Otsuka, Y., et al. (2008). Infants' sensitivity to shading and line junctions. *Vision Research, 48*, 1420–1426.

Iwasa, K., Komatsu, T., Kitamura, A., and Sakamoto, Y. (2020). Visual perception of moisture is a pathogen detection mechanism of the behavioral immune system. *Frontiers in Psychology, 11*, 170: DOI: 10.3389/fpsyg.2020.00170.

Jacobs, R. H., Baumgartner, E., and Gegenfurtner, K. R. (2014). The representation of material categories in the brain. *Frontiers in Psychology, 5*, 146. DOI: 0.3389/fpsyg.2014.00146.

Jelbert, S. A., Miller, R., Schiestl, M., Boeckle, M., Cheke, L. G., Gray, R. D., Taylor, A. H., and Clayton, N. S. (2019). New Caledonian crows infer the weight of objects from observing their movements in a breeze. *Proceedings of the Royal Society B: Biological Sciences, 286* (1894), 1-8. DOI: 10.1098/rspb.2018.2332.

Johansson, R.S., and Westling, G. (1998). Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Experimental Brain Research, 71*, 59–71. DOI:10.1007/BF00247522.

Johnsson, M., and Balkenius, C. (2008). Recognizing Texture and Hardness by Touch. *Proceedings of the IEEE International Conference on Intelligent Robots and Systems, France,* 482-487. DOI: 10.1109/IROS.2008.4650676.

Johnsson, M., and Balkenius, C. (2010). Haptic Perception with Self-Organizing ANNs and an Anthropomorphic Robot Hand. *Journal of Robotics*, [860790]. DOI: 10.1155/2010/860790.

Kawabe, T, Maruya, K., Fleming, R. W., and Nishida, S. (2015). Seeing liquids from visual motion. *Vision Research, 109*(part B), 125–138. DOI: 10.1016/j.visres.2014.07.003.

Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*, 271-304. DOI:10.1146/annurev.psych.55.090902.142005.

Kilinc-Balci, F.S. (2011). Testing, analyzing and predicting the comfort properties of textiles. In G. Song (Ed.), *Improving comfort in clothing*: *Woodhead Publishing Series in Textiles* (pp. 138-158). Woodhead Publishing: Cambridge, UK. DOI: 10.1533/9780857090645.1.138.

Kim, J., Marlow, P.J., and Anderson, B.L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision, 11*(9), 4. DOI: 10.1167/11.9.4.

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

Klatzky, R. L., Pai, D. K., and Krotkov, E. P. (2000). Perception of material from contact sounds. *Presence: Teleoperators and Virtual Environments*, *9*(4), 399-410.

Koenderink, J. J., and van Doorn, A. J. (1996). Illuminance texture due to surface mesostructure. *Journal of the Optical Society of America A*, *13*(3), 452. DOI: 10.1364/JOSAA.13.000452.

Komatsu, H., and Goda, N. (2018). Neural mechanism of material perception: quest on Shitsukan. *Neuroscience*, *392*, 329-347. DOI: 10.1016/j.neuroscience.2018.09.001.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748. doi:10.1038/264746a0. PMID 1012311.

Manrique, H. M., Sabbatini, G., Call, J., and Visalberghi, E. (2011). Tool choice on the basis of rigidity in capuchin monkeys. *Animal Cognition, 14*, 775–786.

Marlow, P. J., Kim, J., and Anderson, B. L. (2012). The perception and misperception of specular surface reflectance. *Current Biology, 22*(20), 1909–1913. DOI: 10.1016/j.cub.2012.08.009.

Marlow, P. J., and Anderson, B. L. (2013). Generative constraints on image cues for perceived gloss. *Journal of Vision*, *13*(14), 2.

Motoyoshi, I. (2010). Highlight-shading relationship as a cue for the perception of translucent and transparent materials. *Journal of Vision*, *10*(9), 6. DOI:10.1167/10.9.6.

Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W., and Limperis, T. (1977). *Geometrical considerations and nomenclature for reflectance.* NBS Monograph 160. National Bureau of Standards: Washington, DC.

Nishida, S., and Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A: Opt. Image Sci. Vis.*, *15*, 2951-2965.

Nishida, S. (2019). Image statistics for material perception. *Current Opinion in Behavioral Sciences, 30*, 94–99. DOI: 10.1016/j.cobeha.2019.07.003.

Nowak, D.A., Topka, H., Timmann, D., Boecker, H., and Hermsdörfer, J. (2007). The role of the cerebellum for predictive control of grasping. *Cerebellum 6,* 7. DOI:10.1080/14734220600776379.

Okamoto, S., Nagano, H., and Yamada Y. (2013). Psychophysical dimensions of tactile perception of textures. *IEEE Transactions on Haptics, 6*(1), 81-93.

Olkkonen, M., and Brainard, D.H. (2011). Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Perception*, *2*, 1014-1034.

Osta-Vélez, M., and Gärdenfors, P. (2020). Category-Based Induction in Conceptual Spaces. *Journal of Mathematical Psychology, 96*. DOI: 10.1016/j.jmp.2020.102357.

Pascalis, O., de Martin de Viviés, X., Anzures, G., Quinn, P. C., Slater, A. M., Tanaka, J. W., and Lee, K. (2011). Development of face processing. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(6). 666-675.

Paulun, V. C., Kawabe, T., Nishida, S., and Fleming, R. W. (2015). Seeing liquids from static snapshots. *Vision Research, 115*(Part B), 163–174. DOI: 10.1016/j.visres.2015.01.023.

Paulun, V. C., Gegenfurtner, K. R., Goodale, M. A., and Fleming, R. W. (2016). Effects of material properties and object orientation on precision grip kinematics. *Experimental Brain Research, 234*(8), 2253-2265. DOI: 10.1007/s00221-016-4631-7.

Paulun, V. C., Schmidt, F., van Assen, J. J. R., and Fleming, R. W. (2017). Shape, motion, and optical cues to stiffness of elastic objects. *Journal of Vision, 17*(1), 20. DOI: 10.1167/17.1.20.

Peirce, F.T. (1930). The handle of cloth as a measurable quantity. *Journal of the Textile Institute, 21*, T377.

Podrebarac, S.K., Goodale, M.A., and Snow J.C. (2014). Are visual texture-selective areas recruited during haptic texture discrimination? *NeuroImage, 94*, 129-137.

Pont, S.C., and Koenderink, J.J. (2005). Bidirectional texture contrast function. *International Journal of Computer Vision, 62*, 17-34. DOI: 10.1023/B:VISI.0000046587.42611.2c.

Sabbatini, G., Manrique, H. M., Trapanese, C., De Bortoli Vizioli, A., Call, J., and Visalberghi, E. (2013). Sequential use of rigid and pliable tools in tufted capuchin monkeys (Sapajus spp.). *Animal Behaviour, 87,* 213 – 220.

Sawayama, M., Adelson, E.H., and Nishida, S. (2017). Visual wetness perception based on image color statistics. *Journal of Vision, 17*(5), 7. DOI: 10.1167/17.5.7.

Sawayama, M., and Nishida, S. (2018) Material and shape perception based on two types of intensity gradient information. *PLOS Computational Biology 14*(4), e1006061. DOI: 10.1371/journal.pcbi.1006061.

Schaller, M. (2006). Parasites, behavioral defenses, and the social psychological mechanisms through which cultures are evoked. *Psychological Inquiry, 17* (2), 96–101. DOI:10.1207/s15327965pli1702_2.

Schaller, M., and Park, J.H. (2011). The behavioral immune system (and why it matters). *Current Directions in Psychological Science, 20*(2), 99–103. DOI:10.1177/0963721411402596.

Schmid, A. C., and Doerschner, K. (2018). Shatter and splatter: The contribution of mechanical and optical properties to the perception of soft and hard breaking materials. *Journal of Vision, 18*(1), 14. DOI: 10.1167/18.1.14.

Schmidt, F., Paulun, V. C., van Assen, J. J. R., and Fleming, R. W. (2017). Inferring the stiffness of unfamiliar objects from optical, shape, and motion cues. *Journal of Vision, 17*(3), 18. DOI:10.1167/17.3.18.

Schmidt, F., Roland, W. F., and Valsecchi, M. (2020). Softness and weight from shape: Material properties inferred from local shape features. *Journal of Vision, 20*(6), 2. DOI: 10.1167/jov.20.6.2.

Sharan, L. (2009). *The perception of material qualities in real-world images*. Unpublished Ph.D. Thesis. Massachussets Institute of Technology, Cambridge, MA.

Sharan, L., Rosenholtz, R., and Adelson, E. H. (2009). What can you see in a brief glance? *Journal of Vision, 9*(8), 784.

Sharan, L., Liu, C., Rosenholtz, R., and Adelson, E.H. (2013). Recognizing Materials using Perceptually Inspired Features. *International Journal of Computer Vision, 103*(3). 348-371. DOI: 10.1007/s11263-013-0609-0.

Sharan, L., Rosenholtz, R., and Adelson, E. H. (2014). Accuracy and speed of material categorization in real-world images. *Journal of Vision, 14*(9), 12. DOI: 10.1167/14.9.12.

Smith, L. B., and Slone, L. K. (2017). A Developmental Approach to Machine Learning? *Frontiers in Psychology, 8,* 2124. DOI: 10.3389/fpsyg.2017.02124.

Thompson, W., Fleming, R., Creem-Regehr, S., and Stefanucci, J. (2011). *Visual Perception from a Computer Graphics Perspective.* Boca Raton, FL: Taylor and Francis/CRC Press. DOI: 10.1201/b10927.

Tsuruhara, A., Sawada, T., Kanazawa, S., Yamaguchi, M.K., and Yonas, A. (2009) Infant’s ability to form a common representation of an object’s shape from different pictorial depth cues: a transfer-across-cues study. *Infant Behavior and Development 32,* 468–475.

Van Assen, J. J. R., Barla, P., and Fleming, R. W. (2018). Visual Features in the Perception of Liquids. *Current Biology: CB, 28*(3), 452–458. DOI: 10.1016/j.cub.2017.12.037.

Vangorp, P., Laurijssen, J., and Dutré, P. (2007). The influence of shape on the perception of material reflectance. *ACM Trans. Graph., 26*, 1-9.

Varma, M., and Zisserman, A.A. (2009). Statistical approach to material classification using image patch exemplars. *TPAMI, 31*(11), 2032–2047.

Wiebel, C., Valsecchi, M., and Gegenfurtner, K. R. (2013). The speed and accuracy of material recognition in natural images. *Attention, Perception, & Psychophysics, 75*(5), 954 – 966.

Wiebel, C. B., Valsecchi, M., and Gegenfurtner, K. R. (2014). Early differential processing of material images: evidence from ERP classification. *Journal of Vision, 14*(7), 10. DOI: 10.1167/14.7.10.

Wolfe, J. M., and Myers, L. (2010). Fur in the midst of the waters: visual search for material type is inefficient. *Journal of vision, 10*(9), 8. DOI:10.1167/10.9.8.

Xiao, B., Bi, W., Jia, X., Wei, H., and Adelson, E. H. (2016). Can you see what you feel? Color and folding properties affect visual–tactile material discrimination of fabrics. *Journal of Vision, 16*(3):34. doi: 10.1167/16.3.34.

Yang, J., Otsuka, Y., Kanazawa, S., Yamaguchi, M. K., and Motoyoshi, I. (2011) Perception of surface glossiness by infants aged 5 to 8 months. *Perception, 40,* 1491–1502.

Yang, J., Kanazawa, S., and Yamaguchi, M.K. (2013a) Can Infants Tell the Difference between Gold and Yellow? *PLoS One, 8*(6), e67535. DOI: 10.1371/journal.pone.0067535.

Yang, J., Kanazawa, S., Yamaguchi, M. K., and Kuriki, I. (2013b) Investigation of color constancy in 4.5-month-old infants under a strict control of luminance contrast for individual subjects. *Journal of Experimental Child Psychology*, *115*, 126–136.

Yonas, A., and Granrud, C.E. (2006). Infants' perception of depth from cast shadows. *Perception & Psychophysics*, *68*(1), 154-60.

Yoonessi, A., and Zaidi, Q. (2010). The role of color in recognizing material changes. *Ophthalmic & physiological optics: the journal of the British College of Ophthalmic Opticians (Optometrists), 30*(5), 626–631. doi:10.1111/j.1475-1313.2010.00722.

Zaidi, Q. (2011): Visual inferences of material changes: color as clue and distraction. *WIREs Cognitive Science, 2,* 686-700. DOI: doi.org/10.1002/wcs.148.

Zoeller, A. C., Lezkan, A., Paulun, V. C., Fleming, R. W., and Drewing, K. (2019). Integration of prior knowledge during haptic exploration depends on information type. *Journal of Vision, 19*(4), 1–15, DOI: 10.1167/19.4.20.

Paper I

frontiers in Psychology

ORIGINAL RESEARCH
published: 18 February 2020
doi: 10.3389/fpsyg.2020.00215

The Visual Perception of Material Properties Affects Motor Planning in Prehension: An Analysis of Temporal and Spatial Components of Lifting Cups

Kristín Ósk Ingvarsdóttir and Christian Balkenius*

Lund University Cognitive Science, Lund University, Lund, Sweden

The current study examined the role of visually perceived material properties in motor planning, where we analyzed the temporal and spatial components of motor movements during a seated reaching task. We recorded hand movements of 14 participants in three dimensions while they lifted and transported paper cups that differed in weight and glossiness. Kinematic- and spatial analysis revealed speed-accuracy trade-offs to depend on visual material properties of the objects, in which participants reached slower and grabbed closer to the center of mass for stimuli that required to be handled with greater precision. We found grasp-preparation during the first encounters with the cups was not only governed by the anticipated weight of the cups, but also by their visual material properties, namely glossiness. After a series of object lifting, the execution of reaching, the grip position, and the transportation of the cups from one location to another were preeminently guided by the object weight. We also found the planning phase in reaching to be guided by the expectation of hardness and surface gloss. The findings promote the role of general knowledge of material properties in reach-to-grasp movements, in which visual material properties are incorporated in the spatio-temporal components.

Keywords: motor planning, material properties, motion capture, center of mass, expectations

OPEN ACCESS

Edited by:
Yann Coello,
Université Lille Nord de France,
France

Reviewed by:
Christophe Bourdin,
Aix-Marseille Université, France
Francois Quesque,
INSERM U1171 Troubles Cognitifs
Dégénératifs et Vasculaires, France

****Correspondence:***
Kristín Ósk Ingvarsdóttir
kristin.ingvarsdottir@lucs.lu.se

Specialty section:
This article was submitted to Cognition, a section of the journal Frontiers in Psychology

Received: *09 August 2019*
Accepted: *30 January 2020*
Published: *18 February 2020*

Citation:
Ingvarsdóttir KÓ and Balkenius C (2020) The Visual Perception of Material Properties Affects Motor Planning in Prehension: An Analysis of Temporal and Spatial Components of Lifting Cups. Front. Psychol. 11:215. doi: 10.3389/fpsyg.2020.00215

INTRODUCTION

Previous research on material perception has largely focused on passive viewing of either real-world images or digitally produced images of material properties and classes, in which it has been found that we are relatively fast at recognizing and categorizing materials based on visual appearance alone (Sharan et al., 2009, 2014; Fleming et al., 2013; Fleming, 2014). However, as we move around in our environment actively exploring, we rarely perceive materials in static and passive scenes (Gibson, 1979). Nevertheless, despite the growing interest in material perception as an active process, less is known about how the visual appearance of material properties are incorporated into motor planning. Especially when conducting a seemingly effortless everyday task using familiar object, such as reaching for a cup of coffee in the morning.

One of many challenges in research on visual material perception is to understand how material properties that are intrinsic to an object, and usually perceived through touch, are visually estimated and anticipated. Before lifting an object, we anticipate the required kinematics and force to lift the object using the provided visual information and memory of previous sensorimotor experience of interacting with similar objects. Such anticipation is interesting considering that some material properties are almost exclusively perceived haptically (e.g., weight and hardness) or visually (e.g., color and glossiness) while other properties, such as size and shape, are readily accessible via both sensory systems (Johansson, 1996; Woods and Newell, 2004). Despite the complexity of the provided sensory information we seem to anticipate them effortlessly.

A growing number of studies (Gordon et al., 1991a,b; Johansson, 1996; Flanagan and Beltzner, 2000; Mon-Williams and Murray, 2000; Flanagan et al., 2006, 2008; Cole, 2008; Buckingham et al., 2009, 2016; Baugh et al., 2012) have shown that we are fairly good at anticipating appropriate application of grip-force when reaching for objects, as we are equipped with internal models that allow the grasping to be planned and adjusted accordingly to avoid slippage. These models are long-term priors formed through time with repeated experience of objects having certain visual appearances, weight, and density. Such as a size-weight association, that proposes that larger objects are normally heavier than smaller objects, and a material-density association, that proposes that objects made of materials with high density are heavier than objects with low density.

Due to its tender age as a research field, material perception has been largely studied using passive viewing of images of material properties and classes. Such approach disregards the impact of materials in a real-world setting. In the present study, we therefore examined the role of visual material properties in motor planning when handling familiar objects. In particular, we examined how changes in material surface appearance, and object weight, resulted in changes in motor movements. A seated reaching task was conducted, in which thin-walled paper cups, that varied in weight and glossiness, were grabbed and transported. We used a motion capture technology to accurately record the movement speed and movement position in three dimensions to achieve both temporal and spatial data for our analysis.

The role of object weight has been well-established in reach-to-grasp studies, whereas less is known about how such tactile information is anticipated based on the visual characteristics of the material category. Although few studies have addressed the noteworthy role of perceived surface texture and surface friction in prehension (Fikes et al., 1994; Flatters et al., 2012; Paulun et al., 2016), the investigation of the role of visual material properties in prehension is remarkably little, especially when using familiar objects. Conceptually, surface gloss is perceived visually but it can convey tactile information, such as friction. Adams et al. (2016) demonstrated such cross-modal relationship between surface gloss and perceived friction. Using discrimination paradigm, they presented computer-generated objects to their participants, either visually on a screen or haptically using a Phantom force feedback device, and found that slipperiness had a positive relationship with objects with shiny surfaces. Additionally, longer approaching times have been reported when reaching for slippery objects compared to objects with high friction (e.g., Fikes et al., 1994; Paulun et al., 2016).

On that note, it would be interesting to see if cross-modal relationships can be found for other material properties. In particular if hardness can be anticipated using the object's surface gloss, given that some objects with glossy appearances happen to be made of hard materials (e.g., plastic mobile cases, cutlery, and laminated paper). Although, Adams et al. (2016) found no correlation between perceived glossiness and physical compliance, and no relationship between physical surface gloss and perceived compliance. It would be interesting to examine if this is the case for real objects made of materials that are familiar to the observer, instead of phantom objects. Considering that both material category and context have plausible roles in anticipating the direction of the relationship between surface gloss and hardness. For instance, melted butter has glossy appearance and soft quality, whereas a glossy snooker ball has hard qualities. In the current study, we wanted to examine if glossiness could convey information about the hardness of the object when interacting with real objects. We used familiar objects made of similar materials, except some of them were altered with layers of varnish to create a glossy appearance and that way a different material appearance. Besides our interest in the role of surface gloss and object weight in prehension, we were also interested in the subjective impression of hardness based on the visual appearance of the objects alone as an additional measurement, to provide new and exciting data about material perception in motor control.

The kinematics of prehension are comprised of reaching and grasping. In reaching the hand is guided to the targeted location, whereas in grasping the observer prepares the grip by opening and closing the hand according to the targeted object's properties (Jeannerod, 1981, 1984). Given that prehension is planned with the purpose of the action in mind (Grafton, 2010; Flatters et al., 2012; Schubotz et al., 2014), we anticipated changes in both the temporal- and spatial components of reaching and grasping due to changes in material appearance of cups, in which we expected the participants to trade speed for precision when interacting with cups that required cautious manipulation due to their disposition (Fitts, 1954; Battaglia and Schrater, 2007; Petrenel et al., 2017). Besides recording maximum velocity (peak velocity) in reaching and transportation movements, we were especially interested in the adjustment time period in reaching, that is the deceleration period between the peak velocity of the reach and the moment of object contact. For the reason that such temporal phase reflects movement planning based on the object's identity rather than its location alone (Pryde et al., 1998; Egmose and Køppe, 2018; the term adjustment time is originally used by Rocha et al., 2013). We were also interested in the size of the maximum grip aperture (MGA) during reaching, and how early the MGA appeared for the different types of material properties when reaching for the cups. Moreover, we examined if the same material properties examined for the temporal data, influenced the spatial components of prehension, in particular the selection of grip position during the object contact. Grasping thin-walled deformable objects like paper cups requires a careful movement by the hand and appropriate grip position, where one needs to anticipate not only the weight of the cup but also its local stiffness points and its COM. Such anticipation is important, because applying too much grip force when grabbing a cup will dent it and cause a spillage of its content, while applying too little grip force will result in rotation and failing to hold the cup. Moreover, pouring a liquid into the cup alters its COM, and increases its density, which consequently creates a harder substance. We therefore expect grip preparation and anticipated hardness to be reflected in both the collected temporal- and the spatial data.

To summarize, we performed a seated reaching task to examine the contribution of weight of a material and surface gloss in reaching, grasping, and transporting paper cups. We were especially interested in the spatio-temporal data of the first encounters with the objects during the experiment, as they reflect expectations based visual properties and general knowledge of materials, rather than the direct tactile information. Specifically, we were interested in knowing if the motor movements and selection of grip position involved visual perception of hardness. That is, if a relationship exists between perceived glossiness and anticipated hardness. If the spatio-temporal components of the hand movements are affected by the plausible relationship between surface gloss and rated hardness, it would indicate a strategy, in which an association between what a material looks like and what it feels like, is anticipated in order to facilitate prehension. On the other hand, if no such relationship is found, it would indicate that the two material properties are processed independently.

MATERIALS AND METHODS

Participants

Twenty right-handed participants (seven females) with average age of 28 ± 6 years were recruited on Lund University campus to participate in the study. All participants gave their consent to participate in the study. Due to technical error when collecting the data, few of the participants had missing data in several trials because the cameras could not detect the reflective markers due to their small size and proximity to one another. Six participants were therefore removed from the analysis, and instead we report an analysis for 14 right-handed participants (five females) with the average age of 29 ± 8 years. Considering the smaller sample size than originally planned, *a posteriori* sensitivity power analysis was conducted to determine the minimum detectible effect size for our sample size in terms of Cohen's f, using the power analysis software G*Power 3 (Faul et al., 2007). The power analysis for our sample size ($N = 14$) allowed us to identify a medium effect size of 0.27 Cohen's f, with 80% power, and alpha of 0.05.

Stimuli

The experimental stimuli were eight drinking cups made of white paper lined with wax. The cups were presented either in their original appearances, matte ($N = 4$) or were altered to have shiny appearances ($N = 4$), by first applying two layers of glossy varnish using a soft foam roller to achieve smooth texture, then a permanent clear lacquer spray with high-gloss finish. Examples of the stimuli are seen in **Figure 1**. Source Four® jr™ ellipsoidal reflector spotlight with 575W high performance lamp with 50° beam angle (Electronic Theater Controls, Inc.) was positioned on the left of the participant, casting the light diagonally to make the shiny appearance of the stimuli more apparent (see **Figure 2**). The distance between the spotlight and the stimulus when located on position A was 140 and 180 cm when on position B. The stimuli had the shape of a conical frustum and had the capacity for 35 cl, and measured 114 mm in height and 57 mm outer diameter at the bottom of the cup and 87 mm at the top, and weighed 11 g. Each appearance, matte and shiny, had two empty cups and two cups with added weight (liquid), in which one of each had a lid on, a white printing paper, to hide the content

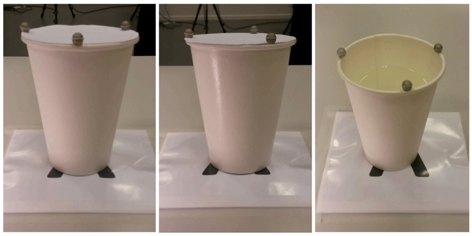

FIGURE 1 | Examples of the experimental stimuli: matte stimulus with hidden content **(left)**, shiny stimulus with hidden content **(center)**, and matte stimulus with visible content **(right)**.

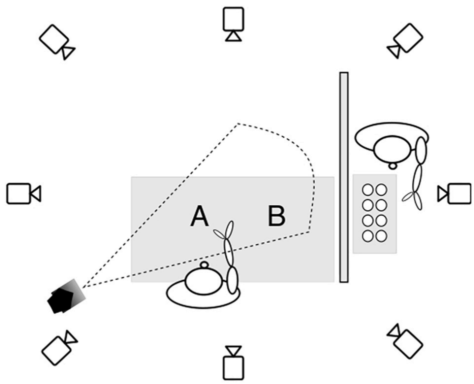

FIGURE 2 | The experimental setup in the motion capture studio. The cups were kept hidden behind a divider and presented by the experimenter one at a time to the participant on location A. Then the participant reached for the cup on location A and transported it to location B. The horizontal distance between locations A and B is 40 cm, and the vertical distance between location A and the start position is 40 cm. The spotlight is positioned on the left of the participant with a distance of 140 cm from location A, casting light diagonally.

of the cup. The added weight was a liquid at room temperature, and weighed together with the cup 252 g. With emphasis on our test exemplar, and according to Alt et al. (2015) model of grasp planning for deformable objects, there are three possible surface areas that afford secure object contact when lifting deformable cups, namely the base and the top due to their attached supportive rings, and the cup's COM. Holding the cups close to the COM is desirable as it prevents object rotation, and ensures a surface space above it during an upward lifting for repositioning the hand in case the cup starts to slide downward due to its weight. We asked our participants to grab the cups as they would normally do when reaching for a cup, that is to position their grip somewhere on the surface between the base and the top of the cup, using their whole hand.

Motion Capture System

The Qualisys motion capture system (Qualisys AB, Gothenburg, Sweden) operated with eight high-speed infrared cameras (Oqus 5 + series) was used to capture the motion data in three dimensions. The infrared cameras together with Qualisys track manager (QTM), a tracking software, recorded and located the position of each reflective marker in a three-dimensional space and in real-time. The motion capture system was also operated with one video camera (Oqus 210c), which was used for post-processing. The infrared cameras recorded at the sampling rate of 100 Hz. Five spherical markers made of reflective polystyrene (7 mm in diameter) were attached on the fingernails of each finger, and three markers were attached on the top of each experimental stimulus ($N = 8$). Marker positions were tracked in three dimensions, and coordinate values X (front–back), Y (left–right), Z (up–down), were exported into MATLAB 8.1.0.604 and R (R2013a) for further evaluation. The motion capture system was calibrated for each participant using a Wand calibration system, which consists of two calibration objects. A stationary L-shaped reference object (200 mm × 350 mm at size), which indicated the orientation and origin of the coordinate system (X-, Y-, and Z-axes), and a T-shaped wand (600 mm distance between the two reflective markers on the horizontal line), which was moved through the volume defining the experimental space. Calibration with average tracking residuals per camera below 0.4 mm was considered suitable for the experiment.

Design and Procedures

Each participant was given chance to get familiar with the experimental task, in which they went through one hypothetical trial of reaching and transporting an object from location A to location B. Next, each participant went through five blocks of test trials, where each block comprised of eight trials, in which each stimulus was presented once, a total of 40 trials per subject. All trials were randomized and counterbalanced across participants. The eight stimuli were four matte and four shiny cups, in which half of them contained liquid and the other half contained nothing. Then, half of the cups from each surface appearance were presented with a lid on to hide their content, and the other half with visible content (no lid). A seven-points semantic differential scale was used to collect subjective ratings of the stimuli in the study, where the participants verbally rated the visual appearance of the eight stimuli using the three bipolar dimensions; heaviness, hardness, and glossiness, both before and after interacting with the stimuli. The scales were as follow. *Heaviness:* if you were to lift the object how heavy would it feel? Low values represent light (light as a feather) and high values represent heavy (heavy as a book), ranging from (1) Very light, (2) Light, (3) Somewhat light, (4) Neutral, (5) Somewhat heavy, (6) Heavy, (7) Very heavy. *Hardness:* if you were to touch the object, would you be able to change the shape of the object using your hand? Low values represent soft (easily with a small force) and high values represent hard (difficult with a small force), ranging from (1) Very soft, (2) Soft, (3) Somewhat soft, (4) Neutral, (5) Somewhat hard, (6) Hard, (7) Very hard. *Glossiness: how shiny is the object?* Low values represent matte (matte as a writing paper) and high values represent shiny (shiny as a mobile phone screen), ranging from (1) Very matte, (2) Matte, (3) Somewhat matte, (4) Neutral, (5) Somewhat shiny, (6) Shiny, (7) Very shiny. The feather used as a reference was a rook feather that weighed less than a gram that are usually found around the university campus, and the book weighed 406 g (height 210 mm, length 150 mm, and width 13 mm).

The total runtime was approximately 45 min, and the procedures were the following. The participant sat in front of a table with his/her eyes closed, left hand rested under the table and right hand rested on a marked start/end position on the table. The participant was told to always keep their hand on that position at the beginning and at the end of each trial. Each experimental trial started after the experimenter placed a new stimulus on location A, taken from a pool of stimuli hidden behind a cardboard, and signaled the participant to open the eyes and start. During the experiment, the stimuli were placed on thin soft coasters to avoid any loud sound feedback when contacting the experimental table, which could give away the weight of the cup. Then, the participant reached for the stimulus on location A and moved it to location B on the right, then released the hand and returned the hand back to start/end position. The participants were allowed to approach the cups using the whole hand (all fingers), to reach for the cups and transport them as quickly and accurately they could, and were asked to position the grip somewhere between the base and the top of the cup during the object lifting. Before the first block of trials, the participants were asked to verbally rate the experimental stimuli and again in the final block after they transported them. When the participant had completed the experiment, he/she was debriefed and thanked.

Data Analysis

The data was analyzed using the statistical software R (R Core Team, 2018) and MATLAB 8.1.0.604 using the MoCap Toolbox (Burger and Toiviainen, 2013). Linear mixed effect models (LMMs) were fitted for the spatio-temporal data using the *lme4* package in R (Bates et al., 2015). All models were assigned a random intercept for each participant to account for individual differences. Deviations from normality were determined by visual inspection using Q–Q probability plots. The results were analyzed in the following manner.

First, the motion capture recordings were coded for the eight stimuli and for each finger using the Qualisys software. For current study we were only interested in the position of the grip center on the stimuli rather than the spread of the hand when grasping the cups, thus we only included the recordings for three fingers, thumb, index- and middle finger for our analysis. Then, the Euclidean norms were used to calculate the magnitudes of velocities for the two selected motor sequences, reaching and transporting, using the MoCap Toolbox in Matlab. This was done to create one informative vector based on the length of the three-dimensional velocity vector, instead of working with three vectors, one for each direction (X, Y, Z). Afterward, the data was exported to R for further statistical analysis. The relevant motor sequences, reaching and transporting, were then visually identified and coded while the remaining ones were discarded. An onset of a reaching movement was determined when the participant moved the hand at the velocity speed above 30 mm/s over a minimum of 10 successive samples, from the starting position to reach for the stimulus in location A, and the grip (object contact) was determined when the velocity speed returned to below 30 mm/s over a minimum of 10 successive samples. The transportation movement consisted of lifting the stimulus from location A and transporting it to location B, using the same onset and offset time thresholds as used for the reaching movements. Adjustment time in reaching was measured as the duration of the temporal span between the peak velocity in reaching and the end of the reaching sequence, measured in milliseconds and peak velocity was defined as the maximum speed measured in velocity (mm/s) for each of the motor sequences: reaching and transporting. The temporal variables, adjustment time in reaching, peak velocity in reaching, and peak velocity in transporting, were analyzed using repeated measures ANOVAs and linear mixed model regression analysis (LMM).

Second, a spatial analysis was performed to examine the effect of material properties on the MGA and the selection of grip position. The COM of the cups was calculated for the two different weights, empty- and filled cups, by calculating the geometric centroid for a conical frustum (formula 1) and COM that included the object weight (formula 2), as seen in the following formulas.

$$COM = \frac{h\left(R_1^2 + 2R_1R_2 + 3R_2^2\right)}{4\left(R_1^2 + R_1R_2 + R_2^2\right)} \tag{1}$$

$$COMweight = \frac{L_w COM + C_w COM}{L_w + C_w} \tag{2}$$

In formula 1, the centroid for the geometric shape of the cups and the liquid along the z-axis is calculated separately, in which h stands for respective vertical height of either the cup or the liquid in mm. R_1 represents the initial radius in mm for the smaller surface area at the base of the cup and the liquid, whereas R_2 represents the final radius in mm for the larger surface area at the top of the cup and the liquid. In formula 2, weight of the cup and the liquid are incorporated to get the COM for empty and filled cups, based on both the geometric shape and the weight of the cups and the liquid. Here COM represents the geometric centroid for the cups and the liquid respectively, L_w stands for the weight of the liquid in grams, and C_w for the weight of the cups in grams. An empty cup had its COM, measured along the vertical axis from the initial radius at the base of the cup to the final radius at the top, at 65 mm, whereas a cup with liquid had its COM at 46 mm. Next, we calculated the maximum distance (mm) between the thumb and the index finger during the reaching movement before the object contact to get the MGA, in which we were interested in both the size of the MGA for the different types of material properties and its timing. After, we examined the position coordinates of the grip during the object contact. First, the vertical position of the grip center (z-dimension) based on the average position coordinates of the thumb, the index finger, and the middle finger was found. Second, the degree of grip deviation from the COM, measured as the grip center's vertical distance from each object's COM in mm, in which positive values represented position coordinates above the COM and negative values represented position coordinates below the COM. The size and the position of MGA were analyzed using repeated measures ANOVA, whereas the selection of grip position, and data from the first block of trials, were analyzed using linear mixed effects regression models.

Finally, the relationship between the object properties (surface and content) and the rated properties, heaviness, hardness, and glossiness, both as pre-reaching ratings and post-transportation ratings, were examined using cumulative link mixed models (CLMMs) using the *ordinal* package in R (Christensen, 2018). Separate regression models were then conducted to examine the spatio-temporal data from the first block of trials, to see if the expected properties based on the pre-reaching ratings were reflected in the motor movements of the hand, or in the selection of grip position.

RESULTS

We expected distinct temporal- and spatial patterns in motor movements, MGA, and grip position for the different types of cups due to learned associations between material appearance and intrinsic properties. Furthermore, we wanted to explore if the motor movements, MGA, as well as the position of the hand on the cups, were guided by the expected properties, especially visually perceived glossiness and hardness.

The First Impression of the Material Properties

We were interested to know in what way the object properties influenced the subjective ratings, specifically before the first interaction with the cups. Thus, a cumulative link mixed regression analysis (CLMM) was conducted to examine the relationship between the object properties (surface appearance and content) and the rated properties (heaviness, hardness, and glossiness) based on the pre-reaching ratings. The expected properties represented as rated heaviness, hardness, and glossiness were computed by averaging over the seven items of the scale for each of the stimuli, separately for the pre-reaching ratings and the post-transportation ratings. The mean ratings of the scales for the eight stimuli are shown in **Figure 3**.

The CLMM analysis revealed statistically significant relations between the physical- and the rated properties. We found *rated*

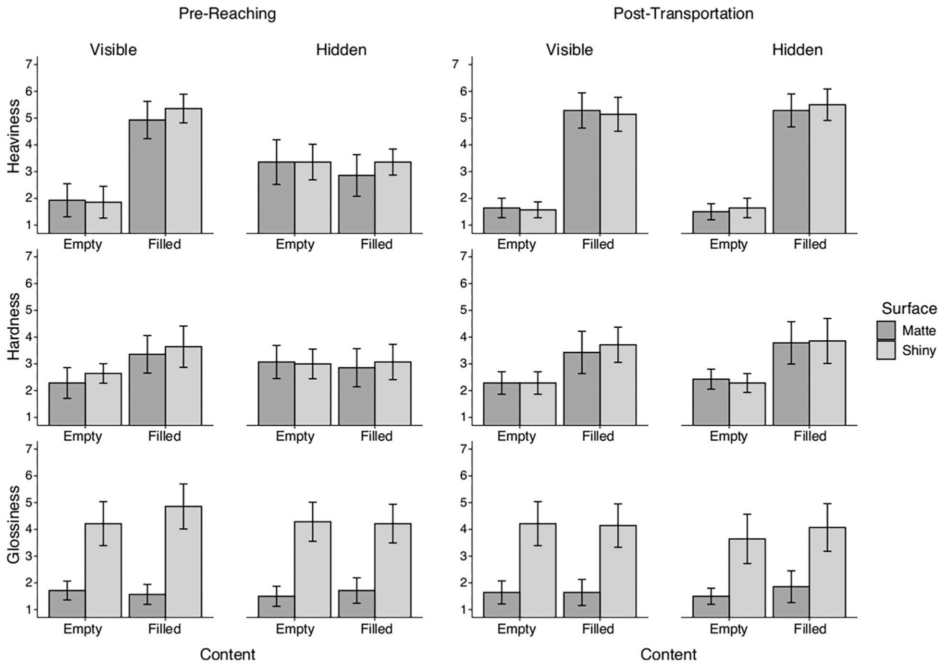

FIGURE 3 | Mean rated heaviness, hardness, and glossiness for the eight stimuli. **(Left)** Ratings collected before reaching for the stimuli. **(Right)** Ratings collected after transporting the stimuli. Error bars are 95% confidence intervals. The two graphs suggest that the pre-reaching ratings are based on expectations whereas the post-transportation ratings are based on sensorimotor information. The figure also suggests strong categorical impressions of the properties, heaviness, and glossiness (bimodal values) whereas the impression of hardness is less defined.

heaviness to be largely affected by the content of the cups, that is whether the cups contained liquid or not, $b = 1.98$, $SE = 0.12$, $\chi^2(1) = 299.94$, $p < 0.001$. According to the pre-reaching ratings, the participants rated the filled cups to be heavier ($M = 5.14$, $SD = 1.08$) than the empty cups when the content was visible to them ($M = 1.89$, $SD = 1.03$), whereas rated heaviness was similar for the two object weights when the content of the cup was hidden using a lid (filled: $M = 3.11$, $SD = 1.13$; empty: $M = 3.36$, $SD = 1.28$). We also found the surface appearance, that is whether the cups had matte or varnished appearance, influenced the rated heaviness, $b = 0.36$, $SE = 0.11$, $\chi^2(1) = 11.22$, $p < 0.001$, in which the shiny cups ($M = 3.61$, $SD = 2.02$) got rated heavier than matte cups ($M = 3.42$, $SD = 1.89$), both when the content was visible and hidden (shiny: $M = 3.36$, $SD = 0.99$; matte: $M = 3.10$, $SD = 1.40$). Intriguingly, we found *rated hardness* to be influenced by both the object weight and the surface appearance, in which filled cups got rated harder ($M = 3.50$, $SD = 1.26$) than empty cups when the content of the cups was visible ($M = 2.46$, $SD = 0.84$), $b = 1.41$, $SE = 0.14$, $\chi^2(1) = 115.86$, $p < 0.001$, whereas when the content was hidden the filled and empty cups were rated similar in terms of hardness (filled: $M = 2.96$, $SD = 1.17$; empty: $M = 3.04$, $SD = 1.00$). Looking into the role of surface appearance in rated hardness we found when the content was visible, the participants rated the shiny cups ($M = 3.14$, $SD = 1.15$) to be harder than the matte cups ($M = 2.82$, $SD = 1.22$), $b = 0.63$, $SE = 0.13$, $\chi^2(1) = 24.78$, $p < 0.001$. Whereas, the participants rated the two types of appearances to be similar in hardness when content was hidden (shiny: $M = 3.04$, $SD = 1.04$; matte: $M = 2.96$, $SD = 1.14$). For *rated glossiness* we found surface appearance largely influenced the impression of surface gloss, $b = 6.89$, $SE = 0.28$, $\chi^2(1) = 1421.9$, $p < 0.001$, in which shiny cups got rated shinier than matte cups, both when the content was visible (shiny: $M = 4.54$, $SD = 1.45$; matte: $M = 1.64$, $SD = 0.62$), and when the content was hidden (shiny: $M = 4.25$, $SD = 1.24$; matte: $M = 1.61$, $SD = 0.74$). No significant effect of object weight (i.e., content) was found on rated glossiness, $b = 0.13$, $SE = 0.11$, $\chi^2(1) = 0.23$, p = n.s.

Velocity Speed and Adjustment Time in Reaching

Overall, the three experimental conditions, (1) the content of the cups (Content: Empty vs. Filled), (2) the appearance of the surface (Surface: Matte vs. Shiny), and (3) the visibility of the content (Feedback: Visible vs. Hidden) had no effect on the peak velocity (mm/s) in reaching. A repeated measures within-subjects ANOVA revealed no significant effect for content, surface appearance, or feedback on peak velocity in reaching, over the five blocks of trials: *content* (Empty: M = 1039.89, SD = 267.26, Filled: M = 1011.67, SD = 236.26), $F(1,13) = 2.30$, $p = 0.15$.; *surface appearance* (Matte: M = 1022.33, SD = 268.59, Shiny: M = 1029.22, SD = 235.54), $F(1,13) = 0.52$, $p = 0.49$.; and *feedback* (Visible: M = 1018.38, SD = 249.12, Hidden: M = 1033.17, SD = 255.89), $F(1,13) = 0.61$, $p = 0.45$. All two-way and three-way interactions were non-significant as well (all ps = n.s.).

Looking into the planning phase when the participants approached the cups, a repeated measures within-subjects ANOVA with content, surface appearance, and feedback as predictive variables and adjustment time (ms) in reaching as the dependent variable, revealed significant effect for content on the adjustment time over all five blocks of trials, $F(1,13) = 71.76$, $p < 0.001$, $\eta^2G = 0.10$. Overall the participants had longer adjustment times when reaching for filled cups (M = 1024 ms, SD = 212 ms), compared to empty cups (M = 881 ms, SD = 139 ms). No other main effect or interaction effects were found for the five blocks of trials together (all ps = n.s.). A Kruskal–Wallis test revealed the adjustment time distributions to be significantly non-identical across the five blocks of trials, $\chi^2(4) = 14.54$, $p < 0.001$, which opted for a further examination of the adjustment times, block by block.

Comparison of Adjustment Times in Reaching

To examine the role of expectations of material properties in reaching, we conducted a linear mixed model regression analysis on the adjustment times for each of the five blocks of trials, with particular interest in the first and the last block of trials for comparison purposes. Each model consisted of three predictive variables as fixed factors, which were the object weight (content), surface appearance, and the visibility of the content (feedback), and a random intercept for each participant. The LMM analysis for the first block of trials revealed significant effect of content, in which adjustment time increased when reaching for filled cups (M = 1192.32, SD = 421.50 ms) compared to empty cups (M = 939.46 ms, SD = 299.88 ms), b = 252.86, SE = 52.08, $\chi^2(1) = 23.68$, $p < 0.001$. The analysis also showed the adjustment time to significantly increase when the participants reached for the cups with the matte appearance (M = 1120.54 ms, SD = 415.02 ms), compared to when the participants reached for the cups with the applied varnish (M = 1011.25 ms, SD = 349.03 ms), b = 109.29, SE = 52.08, $\chi^2(1) = 4.40$, $p < 0.05$. The visibility of the content had, however, no effect on the deceleration phase in reaching, as no significant main effect was found for the variable feedback on adjustment time for the first block of trials, b = 70.71, SE = 52.08, $\chi^2(1) = 1.84$, p = n.s.

For comparison purposes, we conducted a LMM analysis for the last block of trials (block 5), which revealed both content and surface appearance to have significant effect on the adjustment time in reaching, although the effect of surface appearance on the adjustment time in the last block of trials was much smaller compared to the adjustment time in the first block of trials. The adjustment time increased when reaching for the filled cups (M = 1006.79 ms, SD = 287.53 ms), compared to the empty cups (M = 850.36 ms, SD = 237.95 ms), b = 156.43, SE = 38.75, $\chi^2(1) = 16.30$, $p < 0.001$, and increased when reaching for the cups with the matte appearance (M = 967.14 ms, SD = 273.17 ms), compared to cups with the varnished appearance (M = 890.00 ms, SD = 272.21 ms), b = 77.14, SE = 38.75, $\chi^2(1) = 3.96$, $p < 0.05$.

For the remaining blocks (2, 3, and 4), the LMM regression analysis revealed no significant effect of surface appearance (matte vs. shiny) or feedback (visible vs. hidden content) on the adjustment times in reaching (all ps = n.s.). Only the content of the cups, that is the object weight, influenced the adjustment time for those blocks of trials. For the second block of trials, the filled cups had significantly longer adjustment times (M = 1015.71 ms, SD = 291.67 ms) compared to empty cups (M = 895.36 ms, SD = 252.63 ms), b = 120.36, SE = 32.66, $\chi^2(1) = 13.58$, $p < 0.001$. Similarly, filled cups had longer adjustment times (M = 922.50 ms, SD = 243.22 ms) compared to empty cups (M = 824.69 ms, SD = 235.12), b = 92.86, SE = 21.84, $\chi^2(1) = 18.08$, $p < 0.001$ in the third block of trials. As well as in the fourth block of trials, in which filled cups (M = 982.86 ms, SD = 270.54 ms) had longer adjustment time compared to empty cups (M = 892.50 ms, SD = 256.81), b = 90.36, SE = 28.46, $\chi^2(1) = 10.08$, $p < 0.001$. The results are illustrated in **Figure 4**.

In order to examine if the varnished surface appearance had enhanced role in motor planning when the participants could not estimate the weight of the cups based on their content (cups with lids). We further analyzed the reaching movement for cups with hidden content only. Using LMM regression analysis on adjustment times when reaching for cups with lids on, we found no difference in adjustment times for the two types of surface appearances (matte vs. shiny). Both when examining all blocks of trials together (Matte: M = 955.64 ms, SD = 180.06 ms; Shiny: M = 936.86 ms, SD = 188.06 ms), b = −18.79, SE = 33.18, $\chi^2(1) = 0.32$, p = n.s., and when examining only the first block of trials (Matte: M = 1161.07 ms, SD = 447.88 ms; Shiny: M = 1041.43 ms, SD = 388.30ms), b = −22.21, SE = 31.84, $\chi^2(1) = 0.49$, p = n.s. In fact, we found no significant effect of visibility (feedback: visible vs. hidden) on the adjustment times in reaching, neither for all blocks of trials together, b = −13.00, SE = 22.96, $\chi^2(1) = 0.32$, p = n.s., nor for the first block of trials, b = 70.71, SE = 59.05, $\chi^2(1) = 1.43$, p = n.s.

Grip Aperture and Grip Position

Before we examined the selection of grip position for the different types of material properties, we looked into the grip preparation moment in reaching, in which we examined the size of the MGA and the timing of it. First, we measured the MGA before object contact, defined as the maximum metric distance (mm) between the thumb and the index finger, for each stimulus per participant. Using repeated measures ANOVA, we found no significant effect of content, surface appearance, or feedback, on the size of the MGA when analyzing all blocks of trials (all ps = n.s.). Moreover, LMM regression analysis for the first block of trials revealed no significant effect of the three predictive variables on MGA

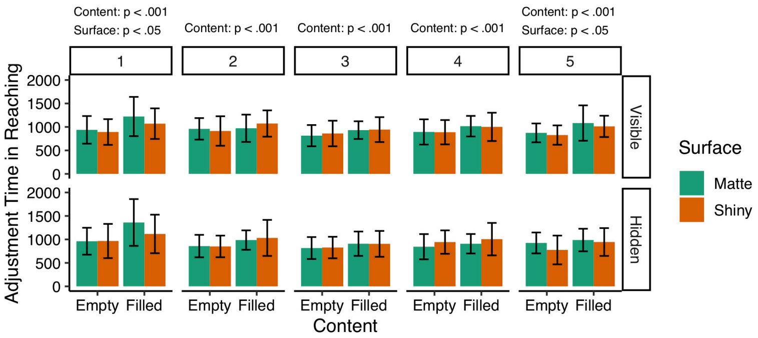

FIGURE 4 | Average adjustment times (ms) in reaching for each block of trials (*N* = 5) based on the cups content and surface appearance (material). Both variables, the weight of the cups and the surface gloss, had significant effect on the adjustment times during the first block of trials, and the last block of trials. The participants adopted longer adjustment times when reaching for cups with either matte appearance, or were filled with liquid (heavier). No significant effect was found for the feedback condition, indicating no difference in adjustment times when reaching for cups with or without lid.

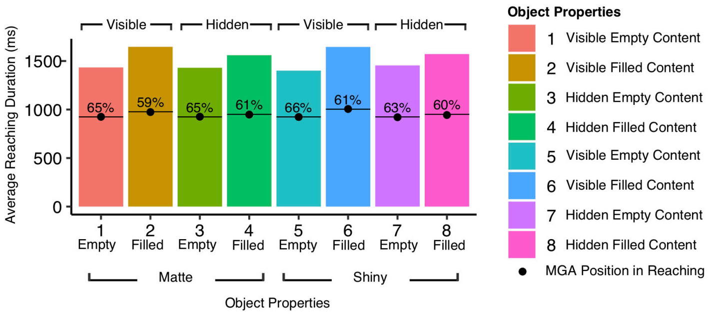

FIGURE 5 | Average duration times in reaching (ms) for the different types of cups presented with respective average timing of the maximum grip aperture (MGA). The figure demonstrates MGA to occur earlier in time when reaching for cups with filled content (*M* = 61.67% after movement start) compared to empty and consequently lighter cups (*M* = 66.78% after movement start). This effect was present when the timing of MGA was examined in relation to the reaching duration times, but was absent when the timing of MGA was examined as the absolute time from movement start.

size either (all *ps* = n.s.). Overall, the average size of MGA was 119.61 mm (*SD* = 12.98 mm). Next, we looked into the timing of the MGA as measured from the movement start to the moment the aperture reached its maximum value. Using repeated measures ANOVA on all blocks of trials, we found the absolute timing of the MGA was not guided by the object properties or their visibility. Furthermore, LMM regression analysis on the first block of trials revealed no significant effect for the same variables on the absolute timing of MGA (all *ps* = n.s.). On the average, the MGA for the different types of cups was at 945.86 ms (*SD* = 296.09 ms) after the movement start.

We found, however, variations in the time point of MGA in proportion to the duration time of reaching per condition. As seen in **Figure 5**, the duration times of reaching differs for the different types of cups depending on their properties, with average duration time ranging from 1401.14 ms (*SD* = 528.67 ms) to 1646.43 ms (*SD* = 699.55 ms). The figure also shows that the timing of MGA varies in proportion to the different presented duration times. Further examination using a repeated measures ANOVA revealed significant effect of object weight (content: filled vs. empty) on the time position of MGA in relations to the reaching duration time for the different types of cups. The MGA occurred earlier in time when reaching for filled cups ($M = 61.67\%$, $SD = 6.32\%$) compared to empty cups ($M = 66.78\%$, $SD = 7.53\%$), $F(1,103) = 14.57$, $p < 0.001$. A similar ANOVA analysis on the first block of trials revealed object weight had a significant effect on the timing of MGA, as we found MGA occurred earlier when the participants reached for filled cups ($M = 58.14\%$, $SD = 9.07\%$) compared to empty cups ($M = 65.86\%$, $SD = 10.52\%$), $F(1,103) = 16.71$, $p < 0.001$. No other main effects or interactions were found significant (all *ps* = n.s.).

The selection of grip position was examined using the vertical spatial position of the grip center of the hand, calculated as the average horizontal distance between the thumb, index-, and middle finger, using LMM. We were both interested in the grip position measured as the distance from the base of the cup in mm, and if the position deviated from the cups' center of the mass. When examining all blocks of trials together, the comparison of a full model to a nested model, revealed that the removal of *content* (empty, filled), $\chi^2(2) = 66.58$, $p < 0.001$, as main effect significantly reduced the model fit, whereas the removal of *surface* (matte, shiny) did not significantly affect the model fit, $\chi^2(2) = 0.19$, p = n.s. Further regression analysis on the selection of grip position using a model with only content as fixed effect revealed grip position to be guided by the content of the cups, in which filled cups were grasped lower than empty cups, $b = -4.11$, $SE = 0.49$, $\chi^2(1) = 70.73$, $p < 0.001$. Although the cups were on the average grasped at their respective COM, a further examination on the direction of grip deviations from the two types of COM, revealed an upward grip deviation for the empty cups, whereas we found a downward deviation for the filled cups with heavy properties, $b = -13.90$, $SE = 0.68$, $\chi^2(2) = 516.5$, $p < 0.001$. Moreover, a LMM analysis on the selection of grip position for the first block of trials, using a model with only content as the fixed effect, revealed grip position to be affected by the content of the cups, in which filled cups were grasped significantly lower than empty cups, $b = -2.26$, $SE = 0.81$, $\chi^2(1) = 7.72$, $p < 0.01$. **Figure 6** shows the relationship between the grip position and the COM of the cups, in which we found an upward grip deviation from COM for empty cups and downward deviation for filled cups, $b = -16.57$, $SE = 1.10$, $\chi^2(2) = 160.11$, $p < 0.001$.

Comparison of Transportation Velocities

Overall, a repeated measures ANOVA revealed the participants adopted slower transportation speed when reaching for filled cups ($M = 657.46$ mm/s, $SD = 112.68$ mm/s), compared to empty cups ($M = 834.01$ mm/s, $SD = 171.62$ mm/s), $F(1,13) = 73.77$, $p < 0.001$, $\eta^2G = 0.36$. No other main effect or interaction effects were found based on the overall performance (all *p*s = n.s.). A Kruskal–Wallis test revealed transportation peak velocity distributions to be significantly non-identical across the five blocks of trials, $\chi^2(4) = 16.31$, $p < 0.001$, which called for further examination.

Figure 7 shows the comparison between the five blocks of trials for the velocity measurements in transportation, in which an LMM regression analysis revealed a dominant effect of object weight (content) on the transportation speed when moving the cups. For the first block of trials, the transportation speed (peak velocity) was significantly slower for filled cups ($M = 600.48$ ms, $SD = 108.65$ ms) compared to empty cups ($M = 798.52$ ms, $SD = 191.55$ ms), $b = -198.03$, $SE = 18.30$, $\chi^2(1) = 117.00$, $p < 0.01$. As well as for block 2 (Filled: $M = 686.18$ ms, $SD = 111.36$ ms; Empty: $M = 845.79$ ms, $SD = 686.18$ ms), $b = -159.61$, $SE = 18.67$, $\chi^2(1) = 73.12$, $p < 0.01$. For block 3 (Filled: $M = 663.76$ ms, $SD = 102.77$ ms; Empty: $M = 853.93$ ms, $SD = 187.68$ ms), $b = -190.16$, $SE = 15.73$, $\chi^2(1) = 146.10$, $p < 0.01$. For block 4 (Filled: $M = 694.92$ ms, $SD = 99.44$ ms; Empty: $M = 854.77$ ms, $SD = 143.87$ ms), $b = -159.85$, $SE = 12.19$, $\chi^2(1) = 172.00$, $p < 0.01$. For block 5 (Filled: $M = 641.95$ ms, $SD = 117.86$ ms; Empty: $M = 817.06$ ms, $SD = 154.19$ ms), $b = -175.11$, $SE = 16.91$, $\chi^2(1) = 107.23$, $p < 0.01$. No significant effect was found for surface appearance (material) or for the visibility of the content (feedback), all *p*s = n.s.

Expected Material Properties in Prehension and Transportation

Figure 8 reveals the relationship between the rated properties and the spatio-temporal data. A LMM regression analysis on the effect of pre-reaching ratings on adjustment times for the first block of trials, revealed reaching to be significantly affected by the expectation of glossiness and hardness. We found rated glossiness, $b = -144.58$, $SE = 60.51$, $\chi^2(1) = 4.23$, $p < 0.05$, and the interaction between rated hardness and glossiness, $b = 39.43$, $SE = 19.36$, $\chi^2(1) = 4.14$, $p < 0.05$, to influence the adjustment time, whereas rated heaviness was found to have no effect on the adjustment time in reaching for the first block of trials, $b = 105.68$, $SE = 55.93$, $\chi^2(1) = 3.48$, $p = 0.06$. The figure shows longer planning phase, characterized by the longer adjustment times in reaching for cups with hard properties, which got longer the shinier the cups got rated. In comparison, the planning phase in reaching was shorter for cups with rated soft properties, and it got shorter the shinier the cups got rated. Indicating a distinct categorical perception for the two types of relationship directions. Moreover, when examining the relationship between the pre-reaching ratings and the selection of grip position, in which we found grip position to be affected by the rated glossiness, $b = 2.36$, $SE = 0.83$, $\chi^2(1) = 7.77$, $p < 0.01$, and by the interaction between glossiness and hardness, $b = -0.58$, $SE = 0.27$, $\chi^2(1) = 4.59$, $p < 0.05$, indicating that the spatio-temporal components of reaching and grasping are to some extend based on the visual impression of the material properties.

When examining the peak velocity for transporting the cups, no effect was found for the rated glossiness and hardness, neither when examining the pre-reaching ratings nor the post-transportation ratings (all ps = n.s.). A significant effect, however, was found for perceived heaviness based on the post-transportation ratings collected after the last block of trials, $\chi^2(1) = 14.49$, $p < 0.001$. The results suggest that the expected properties as based on the material appearance of the cups has a role in guiding reaching and grasping for the first encounters, but after the object lifting and during the transportation of the cup, the sensorimotor information based on the object weight becomes known, and consequently next physical interactions with similar objects are characterized by that updated knowledge.

DISCUSSION

Visual perception of material properties has largely been studied via passive viewing of either static or dynamic images of materials (e.g., Sharan et al., 2014). Consequently, less is known about the function of material perception in real environments when interacting with objects made of familiar materials. Here, we examined the role of visual perception of surface gloss and object weight in reach-to-grasp planning when handling familiar objects. Furthermore, we were interested to know if the visual perception of hardness was reflected in the spatio-temporal measurements of prehension. We therefore additionally

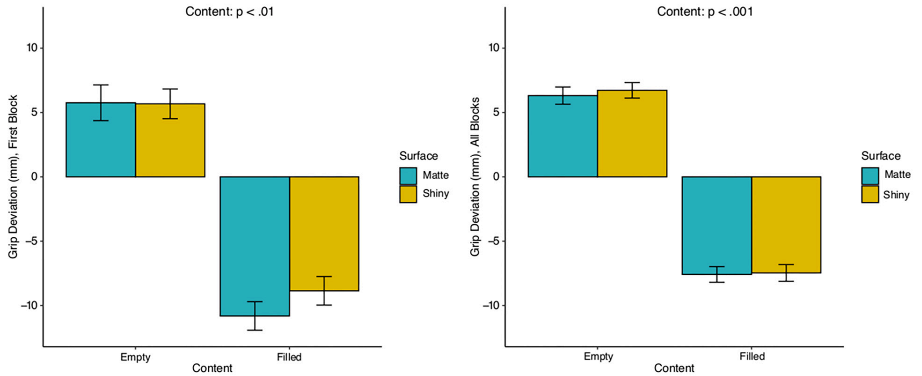

FIGURE 6 | Grip deviations from the center of mass (COM) in mm as a function of manipulated properties, for the first block of trials **(left)**, and across all blocks of trials **(right)**. The COM is measured from the base of the cups and is positioned at 65 mm for empty cups, and 46 mm for filled cups and is represented respectively as a zero on the y-scale. The error bars are standard errors. Significant difference was found between empty and filled cups. Surface is shown for comparison purposes.

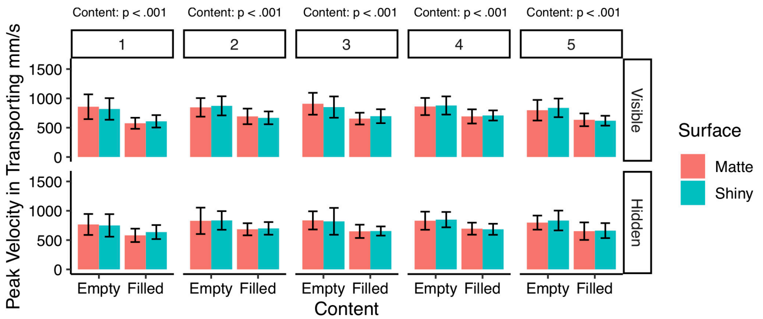

FIGURE 7 | Peak velocity (mm/s) and standard errors when transporting the cups from location A to location B. The graphs show that the filled cups are moved significantly slower than empty cups (all ps < 0.001), whereas no significant effect was found for surface appearance or visibility of the content on peak velocity in transporting.

examined if perceived hardness had a positive relationship with surface gloss, and if such relationship was incorporated in the handling of the objects.

According to the dual visuomotor channel theory in motor planning (Jeannerod, 1981, 1984), prehension is comprised of two movement components, reaching and grasping, encoded by two distinct visuomotor pathways that project from the visual cortex to the motor cortex. Due to the anatomical distinction, the two components are controlled separately, in which reaching is solely guided by the extrinsic properties such as the distance between the hand and the targeted object or the object's orientation, whereas grasping is based on the object's intrinsic properties such as its size. However, later studies have revealed that both intrinsic and extrinsic properties of objects can influence the motor control in reaching, thereby expanding the role of intrinsic properties in prehension (e.g., Paulignan et al., 1991a,b).

In our study, we found adjustment times in reaching were influenced by not only the object weight but also by the properties that described the materials the cups were made of, that is whether the cups had varnished appearance or not. In comparison, we found peak velocity in reaching to be unaffected by those same properties. This suggests that the beginning of reaching movements is for guiding the hand to the targeted location, whereas the deceleration period between the peak velocity and the object contact is for grasp preparation and therefore guided by the precision requirements based on the object properties. A meta-analysis based on 39 reach-to-grasp studies supports in fact such a division. Overall, Egmose and Køppe (2018) found extrinsic properties, such

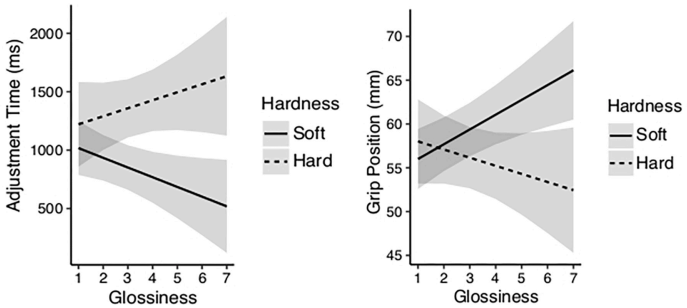

FIGURE 8 | (left) For the first block of trials, the estimated adjustment time based on LMM analysis for the interaction between the properties glossiness and hardness as based on the pre-reaching ratings. The figure shows cups assigned hard properties to have longer adjustment time than cups assigned soft properties, and that distinction to increase the shinier the cups got rated. **(right)** Estimated grip position based on LMM analysis for the same rated properties. The figure shows grip position on rated soft cups to move upwards in vertical direction the shinier they got rated, and move downwards for rated hard cups the shinier they got rated.

as the distance between the hand and the object, commonly guided the acceleration phase at the start of the reach, whereas the deceleration phase at the end of the movement was usually a preparation period for the grasp and therefore either guided by the object properties or by the abstract end-state goal, such as grasp-to-lift versus grasp-to-throw goals. Recent neurophysiological and neuroimaging studies support such earlier involvement of object properties in reaching as well (Grafton, 2010; Touvet et al., 2014; Turella and Lingnau, 2014; Milner, 2017; Freud et al., 2018).

Spatio-Temporal Evidence for Object Weight and Surface Gloss

We found the different object properties led to changes in reaching and grasping. The temporal components of reaching were initially guided by the material properties based on the visual appearance of the object and the object weight, while the timing of the MGA, and the selection of grip position were guided by the weight of the object alone.

Overall, we found object weight had a significant influence on adjustment time in reaching, both during the first block of trials, and across all blocks of trials, in which our participants took longer time to reach for heavier cups compared to lighter cups. Moreover, we found stimulus weight and rated heaviness based on the post-transportation ratings, influence the peak velocity during the transportation phase, in which filled cups (i.e., heavier) were moved significantly slower than empty (i.e., lighter) cups. In this respect, our results are in line with previous findings demonstrating longer planning phase of movements for objects made of heavy materials (e.g., brass) (Weir et al., 1991; Fleming et al., 2002; Paulun et al., 2014, 2016). The filled cups were heavier than the empty cups because they comprised of an additional material, water, which is heavier than the paper cups alone. The longer adjustment times when reaching for filled cups could therefore be explained by how the object weight was manipulated. Certainly, the precision requirement is greater for the filled cups, since their weight was created by pouring fluid in them. As a consequence, the participants had to anticipate the material properties of the water based on their knowledge of liquids, to prevent spillage when handling the cups. Interestingly, securing the opening of the cups with lids did not affect the adjustment time in reaching, even though such changes reduced the risk of spilling. Moreover, we found no overall effect of whether the cups had a lid on or not on the temporal components of reaching, which highlights the role of object weight in prehension, although other properties of water might contribute as well. Nonetheless, we found reaching movements are guided by the expected precision requirements, in which objects that require greater precision to be handled without errors are approached slower due to their disposition (Fitts, 1954).

The results also revealed a role of surface appearance in prehension. For the first block of trials, we found adjustment time was significantly shorter for cups with applied varnish, indicating a shorter grasp preparation time needed for such material appearance, compared to the longer planning phase, which we found for cups with matte appearance. Similarly, the regression analysis on the temporal data using the pre-reaching ratings as fixed effects revealed shorter adjustment time the shinier the cups got rated. We found, however, surface appearance to have no effect on the velocity speed when transporting the cups. Interestingly, the kinematic analysis revealed significant effect of object weight on adjustment time for all of the five blocks of trials individually. Whereas we only found significant effect of surface appearance (matte vs. shiny) on adjustment time for the first and the last block of trials, blocks 1 and 5, and not for the central blocks of trials, blocks 2, 3, and 4. It is likely that the participants were more aware of the object properties in the first and the last block of trials, compared to the central blocks due to the experimental task bound in them. In experimental blocks 1 and 5 the participants were asked to rate the cups using the provided scales for heaviness, hardness, and glossiness, whereas the central blocks of trials were conducted without any rating sessions. In block 1 the participants first rated the cups before reaching for them and relocating them, while in block 5 the participants knew they were going to rate the cups after they had relocated them. In other words, the instruction to form subjective impressions of the cups' properties might have enhanced the changes in the kinematic data.

The results demonstrate that object weight does not have a monopoly when it comes to planning and executing reaching movements. Instead visuomotor planning of reaching involves, to some extent, the expectations of material properties based on the surface appearance of the cup, in our case the glossiness of the cups. Then, after repeated object lifting the motor mechanism is updated by the provided sensorimotor information and the subsequent motor movements are adapted according to the object weight provided by the haptic feedback. Stimulus weight was found to have a prevailing effect on reaching movement after repeated trials, and continued to have an effect after the object contact by either reducing or increasing the transportation speed, depending on the object weight.

Further examination on the grip preparation before object contact, revealed MGA in reaching was not affected by the object properties, weight or surface gloss, as we found no significant effects on MGA size for the different types of cups. Instead, we found the duration times in reaching varied for the different types of properties and further examination revealed object weight (content: filled vs. empty) had a significant effect on the timing of the MGA in relations to the duration times for the different types of cups. The participants reached MGA earlier when reaching for filled cups compared to the empty cups, indicating a longer planning phase required after the MGA for those properties. Surprisingly, this effect on the MGA timing was not found when examining the absolute timing of MGA as measured as milliseconds from the movement start, which suggest the object properties have a role in motor planning after the MGA is reached. Similar findings for aperture in reaching have been reported before by Paulun et al. (2016), who found no effect on the size of MGA in reaching for their tested materials (Styrofoam, Wood, Brass, Slippery Brass), and no effect on its timing, measured as the duration time between the onset of reaching and MGA value. Instead, they found the timing of MGA varied in relations to the duration times, in which MGA occurred earlier when approaching objects made of Brass or Brass with slippery properties, compared to Styrofoam.

When examining the spatial components of the selection of grip position during the object contact, we found no effect of surface gloss. Instead, our results showed a clear effect of object weight, in which the participants positioned their grip at the cup's COM, rather than simply positioning the grip at the mere center of the cups. Supporting previous findings on the role of perceived mass in prehension (Weir et al., 1991; Brouwer et al., 2006; Eastough and Edwards, 2007; see Smeets and Brenner, 1999 for a review). While the results indicated no influence of surface gloss on the position of the grip during the object contact, it indicated that the participants coordinated their grip position according to the object's perceived material class. The participants grabbed the cups at their assigned COM, based on the inferred weight, in which they anticipated the liquid filled cups to have lower COM than the empty cups, and positioned their grip accordingly. Interestingly, one would not have expected any grip deviations for liquid filled cups, as the cups are heavy and require precision in grip position to ensure stableness and prevent torque. Regardless, we found the participants positioned their grip center either at the axis of the cups' COM, or they positioned it below the absolute COM when reaching for the filled cups. A similar effect found by Weir et al. (1991) and Paulun et al. (2016) when reaching for heavy objects. Our introspection suggests that the downward grip deviation from the COM is a strategy to secure a space above the grip in case the cups were heavier than expected and started to slip downward. That way the participants had space to reposition the grip if needed.

Rated Hardness and Glossiness

Hardness has been extensively studied in haptic perception (e.g., Han and Choi, 2010; Baumgartner et al., 2013), however in the context of visual material perception, little is known about hardness anticipation based on visual features when grasping objects made of materials that vary in compliance. In our study we hypothesized that the expected hardness of the cups would be reflected in the spatio-temporal measurements when reaching for the cups. Moreover, we hypothesized that rated hardness would have a positive relationship with the surface appearance of the cups, that is whether the cups had applied varnish or not, which would as well be reflected in the collected data.

In our study it is likely that the hardness ratings represented the expected density of the cups, as the ratings appeared to be influenced by the presence of the liquid in the cups. The filled cups were perceived harder (less deformable) compared to the empty cups, likely due to the liquidized content of the cups. Filling the cups with liquid creates a force against the inside walls of the cups, and results in perceived harder quality compared to cups without content. In terms of surface gloss, we found our participants rated the cups with the applied varnish to have harder quality compared to the ratings assigned to the matte cups. Moreover, CLMM analysis revealed significant relations between not only rated hardness and the content of the cups, but also between rated hardness and the surface gloss, which favors our assumption that there is a positive relationship between glossiness and hardness. When investigating the relationship between rated properties and the temporal data, we found adjustment time in reaching to be affected by the interaction between the expected glossiness and hardness, based on the pre-reaching data, however, the role of glossiness depended on how soft or hard the cups got rated. The participants reached slower for cups with rated hard qualities, compared to cups with rated soft qualities, and that distinction became larger the shinier the cups got rated. In other words, the planning phase in reaching was long for cups with rated hard properties and got longer the shinier the surface got rated, whereas the planning phase was shorter for cups with rated soft properties, and got shorter the shinier the surface got rated, suggesting a distinct categorical perception for hard and soft surfaces with shiny appearance.

Looking into the spatial data, we did not find surface appearance as an object property to have any effect on the selection of grip position. Regardless, we found cups with rated shiny and hard properties to be grabbed lower than cups with rated shiny and soft properties. We argue that cups with rated hard properties are expected to be denser, and due to the prior knowledge that dense materials tend to be heavy and that glossy surfaces are often slippery, the participants approach those cups with more caution. In comparison, cups with rated shiny and soft properties are expected to be lighter and with lower expected torque, which allows the participant to approach the cups with less precision and higher speed.

It is probable that we rely on stored associations based on material appearances to estimate hardness before interacting with an object. Associations that are based on previous experiences of interacting with objects with similar appearances and certain qualities, that are constantly being updated after every object contact. Fleming (2017) among others have noted that the estimation of hardness of objects before touching them is an indistinct task, as any visual information other than the shape deformation itself are ambiguous. As he cleverly clarifies with an example of an object with wooden appearance that is estimated to have hard qualities, but then when that same object is deformed when touched, it is perceived to have soft qualities despite the appearance usually associated with hardness. Clearly, the shape alteration of the object is a much stronger cue for hardness than the visual surface texture of the object, but some textured information led us to think that the object had hard quality to begin with.

For future research it would be interesting to examine expected hardness based on material appearance in relations to grasp force anticipation. The current consensus is that grasp force increases with object weight (Zatsiorsky and Latash, 2008; Marneweck et al., 2016), but less is known to what extend we rely on visual material appearances during force regulations when reaching for objects made of materials that require distinct approach due to their disposition. For instance, what optical characteristics segregate perceived soft objects from perceived hard objects, and do we rely on those characteristics to anticipate the required force to successfully lift the objects? Future research will hopefully clarify these questions.

We propose that the motor control of reaching involves a mechanism that has stored associations between visual material properties and their intrinsic properties, similar to what we have for size and weight (e.g., Baugh et al., 2012). A mechanism that is constantly being updated when new information becomes available when exploring the physical world. Developmental studies on motor control have demonstrated, in fact such adaptation in reaching movements. They have shown that the control of adjustment time in reaching is an adaptive mechanism, which arises early in the perceptual motor development and is updated through repeated experiences of objects in various shapes and with various weights (Pryde et al., 1998; Berthier et al., 1999; Newman et al., 2001; Berthier and Keen, 2005; Rocha et al., 2013).

In sum, we found the temporal characteristics of the hand movements during a reaching task to be influenced by not only the weight of the cups but also by their glossy appearance and expected hardness, extending previous findings on the role of material perception in prehension. Here, we demonstrated that the deceleration period of reaching is a grasp-preparation period guided by both the expectation of the object's weight and its glossy surface appearance. Moreover, we demonstrated that the selection of grip position before object lifting is guided by the expected material properties, in which the cups were held at their respective COM.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KI developed the idea for the research, created the experimental stimuli and set up, carried out the experiment and performed all computations, and wrote the manuscript with input from CB. CB verified the analytical methods and provided theoretical and methodological discussion during the preparation, conduction, and interpretation of the experiment. Both authors discussed the results and contributed to the final manuscript.

ACKNOWLEDGMENTS

The authors are thankful to Lund University Humanities Lab for the use of their motion capture studio.

REFERENCES

Adams, W. J., Kerrigan, I. S., and Graf, E. W. (2016). Touch influences perceived gloss. *Sci. Rep.* 6:21866. doi: 10.1038/srep21866

Alt, N., Xu, J., and Steinbach, E. (2015). "Grasp planning for thin-walled deformable objects," in *Proceedings of the Robotic Hands, Grasping, and Manipulation (ICRA Workshop)*, Seattle, WA.

Bates, D. M., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Battaglia, P. W., and Schrater, P. R. (2007). Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *J. Neurosci.* 27, 6984–6994. doi: 10.1523/jneurosci.1309-07.2007

Baugh, L. A., Kao, M., Johansson, R. S., and Flanagan, J. R. (2012). Material evidence: interaction of well-learned priors and sensorimotor memory when lifting objects. *J. Neurophysiol.* 108, 1262–1269. doi: 10.1152/jn.00263.2012

Baumgartner, E., Wiebel, C. B., and Gegenfurtner, K. R. (2013). Visual and haptic representations of material properties. *Multisens. Res.* 26, 429–455. doi: 10.1163/22134808-00002429

Berthier, N. E., Clifton, R. K., McCall, D. D., and Robin, D. J. (1999). Proximodistal structure of early reaching in human infants. *Exp. Brain Res.* 127, 259–269. doi: 10.1007/s002210050795

Berthier, N. E., and Keen, R. (2005). Development of reaching in infancy. *Exp. Brain Res.* 169, 507–518. doi: 10.1007/s00221-005-0169-9

Brouwer, A. M., Georgiou, I., Glover, S., and Castiello, U. (2006). Adjusting reach to lift movements to sudden visible changes in target's weight. *Exp. Brain Res.* 173, 629–636. doi: 10.1007/s00221-006-0406-x

Buckingham, G., Cant, J. S., and Goodale, M. A. (2009). Living in a material world: how visual cues to material properties affect the way that we lift objects and perceive their weight. *J. Neurophysiol.* 102, 3111–3118. doi: 10.1152/jn.00515.2009

Buckingham, G., Goodale, M. A., White, J. A., and Westwood, D. A. (2016). Equal-magnitude size-weight illusions experienced within and between object categories. *J. Vis.* 16:25. doi: 10.1167/16.3.25

Burger, B., and Toiviainen, P. (2013). "MoCap toolbox – a Matlab toolbox for computational analysis of movement data," in *Proceedings of the 10th Sound and Music Computing Conference, (SMC)*, ed. R. Bresin, (Stockholm: KTH Royal Institute of Technology).

Christensen, R. H. B. (2018). *Ordinal –Regression Models for Ordinal Data.* R Package Version 2018.8-25. http://www.cran.r-project.org/package=ordinal/ (accessed May 07, 2019).

Cole, K. J. (2008). Lifting a familiar object: visual size analysis, not memory for object weight, scales lift force. *Exp. Brain Res.* 188, 551–557. doi: 10.1007/s00221-008-1392-y

Eastough, D., and Edwards, M. G. (2007). Movement kinematics in prehension are affected by grasping objects of different mass. *Exp. Brain Res.* 176, 193–198. doi: 10.1007/s00221-006-0749-3

Egmose, I., and Køppe, S. (2018). Shaping of reach-to-grasp kinematics by intentions: a meta-analysis. *J. Mot. Behav.* 50, 155–165. doi: 10.1080/00222895.2017.1327407

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fikes, T. G., Klatzky, R. L., and Lederman, S. J. (1994). Effects of object texture on precontact movement time in human prehension. *J. Mot. Behav.* 26, 325–332. doi: 10.1080/00222895.1994.9941688

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47, 381–391. doi: 10.1037/h0055392

Flanagan, J. R., and Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size-weight illusion. *Nat. Neurosci.* 3, 737–741. doi: 10.1038/76701

Flanagan, J. R., Bittner, J. P., and Johansson, R. S. (2008). Experience can change distinct size-weight priors engaged in lifting objects and judging their weights. *Curr. Biol.* 18, 1742–1747. doi: 10.1016/j.cub.2008.09.042

Flanagan, J. R., Bowman, M. C., and Johansson, R. S. (2006). Control strategies in object manipulation tasks. *Curr. Opin. Neurobiol.* 16, 650–659. doi: 10.1016/j.conb.2006.10.005

Flatters, I. J., Otten, L., Witvliet, A., Henson, B., Holt, R. J., Culmer, P., et al. (2012). Predicting the effect of surface texture on the qualitative form of prehension. *PLoS One* 7:e32770. doi: 10.1371/journal.pone.0032770

Fleming, J., Klatzky, R. L., and Behrmann, M. (2002). Time course of planning for object and action parameters in visually guided manipulation. *Vis. Cogn.* 9, 502–527. doi: 10.1080/13506280143000557

Fleming, R. W. (2014). Visual perception of materials and their properties. *Vis. Res.* 94, 62–75. doi: 10.1016/j.visres.2013.11.004

Fleming, R. W. (2017). Material perception. *Annu. Rev. Vis. Sci.* 3, 365–388. doi: 10.1146/annurev-vision-102016-061429

Fleming, R. W., Wiebel, C., and Gegenfurtner, K. (2013). Perceptual qualities and material classes. *J. Vis.* 13, 1–20. doi: 10.1167/13.8.9

Freud, E., Robinson, A. K., and Behrmann, M. (2018). More than action: the dorsal pathway contributes to the perception of 3-D structure. *J. Cogn. Neurosci.* 30, 1047–1058. doi: 10.1162/jocn_a_01262

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin Harcourt.

Gordon, A. M., Forssberg, H., Johansson, R. S., and Westling, G. (1991a). The integration of haptically acquired size information in the programming of precision grip. *Exp. Brain Res.* 83, 483–488. doi: 10.1007/BF00229825

Gordon, A. M., Forssberg, H., Johansson, R. S., and Westling, G. (1991b). Visual size cues in the programming of manipulative forces during precision grip. *Exp. Brain Res.* 83, 477–482. doi: 10.1007/BF00230004

Grafton, S. T. (2010). The cognitive neuroscience of prehension: recent developments. *Exp. Brain Res.* 204, 475–491. doi: 10.1007/s00221-010-2315-2

Han, G., and Choi, S. (2010). "Extended rate-hardness: a measure for perceived hardness," in *Haptics: Generating and Perceiving Tangible Sensations.* EuroHaptics 2010. Lecture Notes in Computer Science, Vol. 6191, eds A. M. L. Kappers, J. B. F. van Erp, W. M. Bergmann Tiest, and F. C. T. van der Helm, (Berlin: Springer).

Jeannerod, M. (1981). "Intersegmental coordination during reaching at natural visual objects," in *Attention and Performance IX*, eds J. Long, and A. Baddeley, (Hillsdale, NJ: Lawrence Erlbaum Associates), 153–168.

Jeannerod, M. (1984). The timing of natural prehension movements. *J. Mot. Behav.* 16, 235–254. doi: 10.1080/00222895.1984.10735319

Johansson, R. S. (1996). "Sensory control of dexterous manipulation in humans," in *Hand and Brain: The Neurophysiology and Psychology of Hand Movements*, eds A. M. Wing, P. Haggard, and J. R. Flanagan, (San Diego, CA: Academic Press), 381–414. doi: 10.1016/B978-012759440-8/50025-6

Marneweck, M., Lee-Miller, T., Santello, M., and Gordon, A. M. (2016). Digit position and forces covary during anticipatory control of whole-hand manipulation. *Front. Hum. Neurosci.* 10:461. doi: 10.3389/fnhum.2016.00461

Milner, A. D. (2017). How do the two visual streams interact with each other? *Exp. Brain Res.* 235, 1297–1308. doi: 10.1007/s00221-017-4917-4

Mon-Williams, M., and Murray, A. H. (2000). The size of the visual size cue used for programming manipulative forces during precision grip. *Exp. Brain Res.* 135, 405–410. doi: 10.1007/s002210000538

Newman, C., Atkinson, J., and Braddick, O. (2001). The development of reaching and looking preferences in infants to objects of different sizes. *Dev. Psychol.* 37, 561–572. doi: 10.1037/0012-1649.37.4.561

Paulignan, Y., Jeannerod, M., MacKenzie, C., and Marteniuk, R. (1991a). Selective perturbation of visual input during prehension movements. 2. The effects of changing object size. *Exp. Brain Res.* 87, 407–420. doi: 10.1007/BF00231858

Paulignan, Y., MacKenzie, C., Marteniuk, R., and Jeannerod, M. (1991b). Selective perturbation of visual input during prehension movements 1. The effects of changing object position. *Exp. Brain Res.* 83, 502–512. doi: 10.1007/BF00229827

Paulun, V. C., Gegenfurtner, K. R., Goodale, M. A., and Fleming, R. W. (2016). Effects of material properties and object orientation on precision grip kinematics. *Exp. Brain Res.* 234, 2253–2265. doi: 10.1007/s00221-016-4631-7

Paulun, V. C., Kleinholdermann, U., Gegenfurtner, K. R., Smeets, J. B. J., and Brenner, E. (2014). Center or side: biases in selecting grasp points on small bars. *Exp. Brain Res.* 232, 2061–2072. doi: 10.1007/s00221-014-3895-z

Petrenel, L., Sigaud, O., and Babic, J. (2017). Unifying speed-accuracy trade-off and cost-benefit trade-off in human reaching movements. *Front. Hum. Neurosci.* 11:615. doi: 10.3389/fnhum.2017.00615

Pryde, K. M., Roy, E. A., and Campbell, K. (1998). Prehension in children and adults: the effects of object size. *Hum. Mov. Sci.* 17, 743–752. doi: 10.1016/S0167-9457(98)00024-4

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Rocha, N. A. C. F., de Campos, A. C., Silva, F. P. D. S., and Tudella, E. (2013). Adaptive actions of young infants in the task of reaching for objects. *Dev. Psychobiol.* 55, 275–282. doi: 10.1002/dev.21026

Schubotz, R. I, Wurm, M. F., Wittmann, M. K., and von Cramon, D. Y. (2014). Objects tell us what action we can expect: dissociating brain areas for retrieval and exploitation of action knowledge during action observation in fMRI. *Front. Psychol.* 5:636. doi: 10.3389/fpsyg.2014.00636

Sharan, L., Rosenholtz, R., and Adelson, E. (2009). Material perception: what can you see in a brief glance? *J. Vis.* 9:784. doi: 10.1167/9.8.784

Sharan, L., Rosenholtz, R., and Adelson, E. (2014). Accuracy and speed of material categorization in real-world images. *J. Vis.* 14:12. doi: 10.1167/14.9.12

Smeets, J. B., and Brenner, E. (1999). A new view on grasping. *Motor Control* 3, 237–271. doi: 10.1123/mcj.3.3.237

Touvet, F., Roby-Brami, A., Maier, M. A., and Eskiizmirliler, S. (2014). Grasp: combined contribution of object properties and task constraints on hand and finger posture. *Exp. Brain Res.* 232, 3055–3067. doi: 10.1007/s00221-014-3990-1

Turella, L., and Lingnau, A. (2014). Neural correlates of grasping. *Front. Hum. Neurosci.* 8:686. doi: 10.3389/fnhum.2014.00686

Weir, P. L., MacKenzie, C. L., Marteniuk, R. G., Cargoe, S. L., and Frazer, M. B. (1991). The effects of object weight on the kinematics of prehension. *J. Mot. Behav.* 23, 192–204. doi: 10.1080/00222895.1991.10118362

Woods, A. T., and Newell, F. (2004). Visual, haptic and cross-modal recognition of objects and scenes. *J. Physiol. Paris* 98, 147–159. doi: 10.1016/j.jphysparis.2004.03.006

Zatsiorsky, V. M., and Latash, M. L. (2008). Multifinger prehension: an overview. *J. Mot. Behav.* 40, 446–476. doi: 10.3200/JMBR.40.5.446-476

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Paper II

Keywords: feedforward control of grip force, visual material properties, everyday life objects.

Title: Coffee to go! Examination of grip force anticipation when lifting paper cups.

Abstract:
Grip force scaling in object lifting has been extensively studied in controlled experiments using simple objects and the precision grip. To obtain a resemblance to the real-world function, we examined grip force scaling for the whole hand when handling objects that varied in material properties. Here, we examined the anticipatory fingertip force control when lifting paper cups that varied in glossiness, content weight and content visibility. The force data was collected using flexible force sensors attached directly to the fingertips of the participants, which allowed us to obtain force profiles based on the individual digits, as opposed to the single force profile normally obtained using the precision grip. Moreover, it allowed us to use experimental objects and motor movements, that mimic the ones that we use in our daily lives. Overall, we found functional differences in grip force control during the five-digit hold, in which peak grip force was not only guided by the weight of the cups, but also by their glossy surface appearance and the visibility of their content. We conclude that surface gloss has a significant role in grip force scaling where distinct functional strategy is applied for expected- and perceived properties and individual digits

Introduction:

Grip force anticipation in object lifting is often examined using the precision grip, in which the participants are instructed to lift either a grip-device (i.e., grasp-lift manipulandum) with force sensors attached to the device's handle, or the object itself, using only the thumb and the index fingers. This method provides precise measurements and highly controlled experiments, but it also creates grasping force with confined grip dynamics that does not necessarily represent everyday life grasping behavior, which is considerably more versatile in nature (see e.g., Bullock, Zhent, Rosa, Guertler, and Dollar, 2013; Cutkosky, 1989; Napier, 1956; and Sperling and Jacobson-Sollerman, 1977).

An alternative approach to collect force data on adaptable grasping behavior is to allow five-digit grasp and use sensors that are either attached directly to the experimental objects, or to the fingertips of the participants. This opens up the opportunity to examine grip force control and grip force adaptation using a broader range of objects that vary not only in weight but also in texture, friction, shape, and density. In a recent study, Bergmann Tiest and Kappers (2019), examined the importance of visual and haptic information about materials on early grip force application, in which they attached force sensors directly on the experimental stimuli, and found that grip force correlated highly with the real-time sensation of the surface friction (texture). In their experiment they used steel blocks covered with various materials, such as textile, sandpaper, and plastic films with either a wood-like or a metallic appearance, and placed the force sensors between the block and the covering material. Although, this approach allowed for direct interactions with the different types of surface texture while carefully measuring the grip force when lifting the objects, the density remained approximately equal as the insides were the same, creating mismatch between what is seen and what is covertly present.

In the current study, we used flexible force sensors attached directly to the fingertips rather than the object itself to achieve natural grasping behavior for the given experimental objects. We used devices that are readily accessible to measure the grip force of the whole hand when interacting with everyday life objects, namely paper cups. In a previous study (Ingvarsdóttir and Balkenius, 2020), a kinematic analysis of grasp preparation using the same experimental stimuli as used for the current study, revealed that grasp preparation was not only guided by the object mass, but also by the visually inferred hardness and the glossy surface appearance of the cups. More interestingly, the results suggested distinct temporal grasping behavior for the different combinations of expected hardness and perceived glossiness. The participants needed a longer grasp preparation time for lifting the cups they expected to be hard, compared to the cups they expected to be soft. Moreover, adding surface gloss to the expected hard cups increased the grasp preparation time significantly, and continued to increase the shinier the expected hard cups were rated. In contrast, the grasp preparation time decreased substantially when surface gloss was added to the cups expected to be soft, and continued to decrease the shinier the soft cups got rated. The results show that surface gloss has a different role in grasp control for objects with expected hard and expected soft properties. Here, we continue our previous investigation of grasp control and visually inferred material properties, and examine the grip force scaling occurring after object contact is made.

Grip force control in object lifting is largely guided by the mass of the object and the reactive haptic feedback, in which heavy objects require greater grip force to be securely lifted compared to light objects (e.g., Zatsiorsky and Latash, 2008; Marneweck et al., 2016). In addition, studies have shown that the anticipated object weight is involved in the early stages of grasping, that is in the grip force profiles occurring between the object contact and before the true mass and weight becomes known when the object is lifted (e.g., Gordon et al., 1993; Flanagan et al. 1995; Flanagan and Wing 1997; Buckingham et al.2009; Hermsdörfer et al., 2011). Considering that grip force is applied before direct information about the mass has become available to the person, it is too early for haptic information to be involved in the force control. Thus, an early grip force rate, that is, the first few milliseconds after the initial object contact, could also reflect grip force control as based on visual object properties and prior knowledge rather than tactile feedback alone. The anticipation of a required grip force to successfully lift an object, is generally believed to be reflected in the measurement of the increment of the grip force rate that occurs after the initial object contact has been made, and before the object is lifted from the table (Johansson and Westling, 1988, Gordon et al., 1993, Gordon et al., 1994; Li et al. 2009; Nowak et al. 2007b). To measure the role of expected object weight in early grip force scaling, a peak grip force rate (peak GFR) is computed. Peak GFR is the first peak of the grip force rate occurring between the initial object contact and the moment of the object lift. The force increment represented by the GFR represents the expected object weight, in which the rule is the larger the object weight, the quicker the grip force rate (Johansson and Westling, 1988).

What about object properties other than object weight that can be visually inferred, like surface gloss and hardness? Here, we examined grip force and anticipatory grip force control before object lifting when participants lifted paper cups that varied in visual material properties and weight. The following three questions were addressed: First, *does five-digits grip force scaling account for material-based predictions?* We investigated whether having a liquid in the cups, a material with a higher mass than paper, would lead to changes in force control. Adding liquid does not only change the visual composition of the experimental object, but also its weight and density. Filled cups are heavier and denser than empty cups. Moreover, we examined if changes in the visual appearance of the cups themselves lead to changes in force control, by applying a glossy varnish on otherwise matte material. Changing the appearance of the material opens up a different interpretation of the material properties and what they afford, such as hardness and friction. In addition, we also put lids on top of few of the cups, to determine whether the visibility of the mass distribution of the cups had any contrasting effect on grip force scaling for the different types of surface appearances and object weights.

Second, *do predictions based on anticipated vs. perceived material properties result in distinct grip force scaling?* We examined the object lifts in the first block of trials, to see in what way knowledge of material properties based on visual information controls grip force. Then, we compared those trials with the object lifts in the last block of trials, which are force scaling based on recent experience and sensorimotor information. We also examined the grip force scaling based on the subjective ratings of the material properties, and grip force magnitudes at two different timepoints after object contact.

Third, *do individual digits have different grip force profiles due to changes in material properties?* We examined the individual fingertip forces to see if their grip force profiles varied depending on the different types of material properties, that is, if some digits were guided by object weight while others were guided by surface characteristics.

2. Material and Methods

2.1. Participants

Twenty right-handed participants (eight females) with the average age of 29 ± 6 years, were recruited on Lund University campus to participate in the study. Four participants were removed from the analysis due to technical error in the data collection, in which force signals were missing for all five fingers on more than one trial. Instead, we report an analysis for 16 participants (6 females), with the average age of 28 ± 6 years old. A sensitivity power analysis conducted using the power analysis software G∗Power 3 (Faul et al., 2007), for our sample size of 16 participants revealed a minimum detectable effect size of .28 Cohen's F, given a power of 80% and .05 alpha. Allowing us to detect a medium effect size given our sample size. The study was conducted according to the Declaration of Helsinki (1964) and all participants gave their written consent prior to the study.

2.2. Stimuli

The stimuli were 8 drinking cups made of paper with the capacity for 35 cl of liquid, and measured 114 mm in height, 57 mm outer diameter at the base of the cup, and 87 mm at the top. The cups were manipulated in terms of surface appearance (matte vs. shiny), content weight (light vs. heavy), and visibility of the content to the observer (visible vs. hidden). That said, 4 cups were presented in their original matte appearance, whereas the appearance of the remaining 4 cups was altered using an application of high-gloss varnish to create a shiny surface. Within each group of surface appearance, 2 cups were empty and thus light weight (11 g) and the other 2 cups were heavier (252 g) as they contained water. To examine the effect of the visibility of the content when reaching for the cups, half of the cups (N=4) had their content hidden with a lid, in which a thin paper was attached to the top of the cups, whereas the other half had no lid and thus a visible content. The amount of water poured into the cups was easily visible to the observer when the cups had no lid. The following list clarifies the 8 types of stimuli. 1) matte and empty cup without lid, 2) matte and empty cup with lid, 3) matte and filled cup without lid, 4) matte and filled cup with lid, 5) shiny and empty cup without lid, 6) shiny and empty cup with lid, 7) shiny and filled cup without lid, and 8) shiny and filled cup with lid. Examples of stimuli are seen in figure 1.

A Source Four® jr™ ellipsoidal reflector spotlight with 575W high performance lamp and 50° beam angle (Electronic Theater Controls, Inc) was positioned on the left of the participant with a distance of 140 cm from location A, casting a light diagonally to *enhance* the surface appearance of the stimuli.

Figure 1. Examples of the experimental stimuli. (top left) a cup with shiny appearance, (bottom left) a cup with shiny appearance and hidden content, (top right) a cup with matte appearance (bottom right) a cup with matte appearance and hidden content. The filled cups had liquid up to 80% of the height of the cups, which was easily visible to the participants.

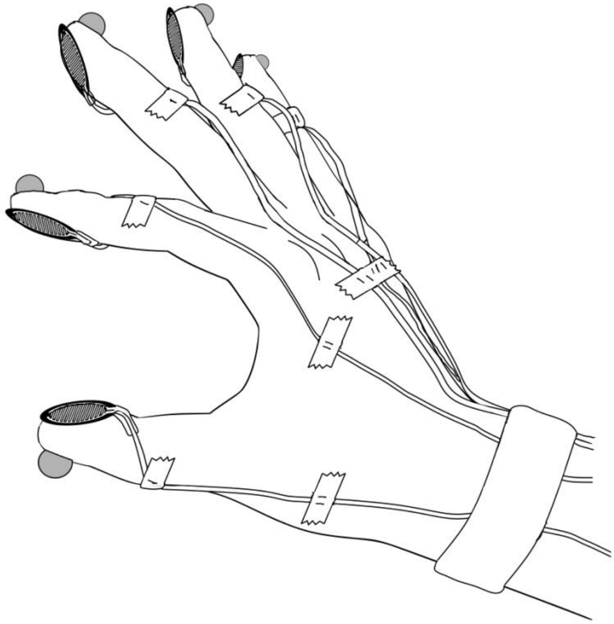

Figure 2: An illustration of a hand with attached force sensors.

2.4. Force Sensors

The force data was collected using five touch sensors. The sensors are 0.5" circular force sensitive resistors (Interlink Electronics) that are thin and flexible. The sensors were calibrated using a digital force gauge. We used 40 measurements to estimate the following calibration curve where x is the measurement from the sensor and y is the force in Newton: $y = 0.0604x - 9.9276$, $R2 = 0.91$. Each finger had a sensor attached to its fingertip using double-sided adhesive tape, and the cables included with the force sensors were loosely attached to the fingers and the arm using surgical tape and Velcro bands to prevent the cables to get in the way during the experimental task. A Phidget InterfaceKit 8/8/8 board attached to the cables included with the force sensors read the pressure applied on the sensors and a script run on a HP Probook 4310s, with Intel ® core TM2 Duo CPU T6670 2.20 GHz recorded the force data.

2.4. Motion Capture System

We used motion capture to collect information on the timepoints of events, where each cup had 3 reflective markers, 7 mm in diameter, attached to its top surface using double-sided adhesive, and one reflective marker (7 mm) was attached to each fingernail individually (N=5) using same technique. The Qualisys motion capture system (Qualisys AB, Gothenburg, Sweden) operated with the tracking software Qualisys track manager (QTM) and 8 high-speed infrared cameras (Oqus 5+ series) recorded the position of the reflective markers at the sampling rate of 100 Hz. The markers positions were then exported into MATLAB 8.1.0.604 and R (R2013a) for further evaluation. Although we recorded the movement of all five fingers, we were mainly interested in the kinematic data for behavioral event segregation for the current study. We therefore, only report motion capture data for the index finger as the data from the remaining fingers are irrelevant for the analysis in the current paper.

Calibration was done for each participant using a stationary L-shaped reference object (200 mm x 350 mm at size) placed on the experimental desk, and a T-shaped wand (600 mm) that was moved through the volume to define the experimental area. A suitable calibration for the study had tracking residuals per camera below 0.4 mm on the average.

The participants were asked to clap their hands three times to create visible pattern in both the motion capture data and the force sensor data for the two measures to be synchronized.

2.5. Rating Scales

A 7-points semantic differential scale was used to collect the subjective experience of the cups. The participants rated the 8 types of cups on two occasions during the study, once before interacting with the cups in the first block of trials, and once after lifting the cups in the last block of trials. Three rating scales were used for rating the cups; heaviness, glossiness, and hardness. The rating scales were the following. *Heaviness: if you were to lift the object how heavy would it feel? Low values represent light and high values represent heavy, ranging from (1) Very light, (2) Light, (3) Somewhat light, (4) Neutral, (5) Somewhat heavy, (6) Heavy, (7) Very heavy. Glossiness: how shiny is the object? Low values represent matte and high values represent shiny, ranging from (1) Very matte, (2) Matte, (3) Somewhat matte, (4) Neutral, (5) Somewhat shiny, (6) Shiny, (7) Very shiny. Hardness: if you were to touch the object, would you be able to change the shape of the object using your hand? Low values represent soft (easily with a small force) and high values represent hard (difficult with a small force), ranging from (1) Very soft, (2) Soft, (3) Somewhat soft, (4) Neutral, (5) Somewhat hard, (6) Hard, (7) Very hard.*

2.6. Design and Procedures

Figure 3 illustrates the experimental set up. The participant sat in front of the experimental table with the left hand resting under the table and the right hand resting on the start position area in front of them on the table. The participant was instructed to rest the right hand on that position before starting each trial and again after its completion. Before the start of each trial, the experimenter took a cup from a collection of cups hidden behind a dividing wall and placed it on location A on the experimental table. During this process the participants were asked to close their eyes to prevent them from inferring the object's properties based on the experimenter's interaction with the object.

Next, the participant was signaled to open the eyes and start the trial. On each trial, the participant visually inspected the cup on location *A* before lifting the cup using a pinch grip. Then, transported the cup at a comfortable pace to location *B*, released the grip, and moved their hand back to the starting position. The distance between location *A* and *B*, and the distance between the start position and location *A* were 40 cm, respectively. Both locations *A* and *B*, and the storage area behind the dividing wall had thin coasters to prevent any sound feedback that could hint the weight of the cup when the cups contacting the experimental table.

Each participant went through 40 trials in 5 blocks in which each type of cup was presented total 5 times, once in each block of trials. The total runtime was approximately 45 minutes. All trials were randomized and counterbalanced across participants. In the first block of trials the participant was instructed to verbally rate the cup in front of them before reaching for it, using the three types of rating scales (i.e., heaviness, hardness, and glossiness) to get the pre-reaching ratings of the cups. In the final block of trials (block nr 5) the participant was instructed to rate the cup again, but after the last transportation movement, in order to achieve post-transportation ratings. No subjective ratings were collected for the three blocks of trials in between. After completion of the experimental task the participant was debriefed and thanked.

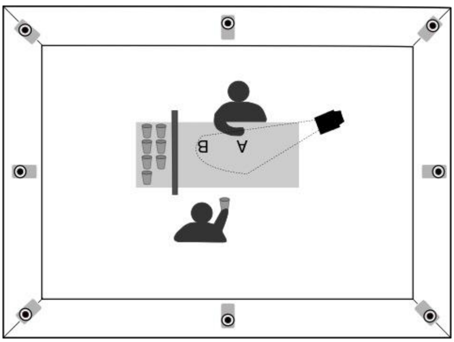

Figure 3. A schematic drawing of the experimental setup. In the motion capture studio, the participant sat in front of a desk with eyes closed while the experimenter placed the cups on location A one at a time. On the start of each trial, the participant opened the eyes, and moved its right hand from the start position to reach for the cup on location A, lifted the cup and transported it to location B before releasing the grip to move the hand back to the start position. The cups were illuminated using a spotlight positioned on the left side of the participant. Cups not used in the trial were hidden from view by a dividing wall.

2.6. Data processing and statistical analysis

The motion capture system recordings were coded for the index finger and the 8 experimental stimuli to segregate the behavioral events, that is, reaching, object contact, and object lifting, using the Qualisys software. The velocity units were obtained in MATLAB 8.1.0.604 using the MoCap Toolbox (Burger and Toiviainen, 2013), in which the three-dimensional motion capture data was transformed into Euclidean norms to get one informative vector instead of three (X, Y, Z directions) to compute the velocity magnitudes.

Event segregation was obtained using kinematic- and grip force thresholds. In addition, a graphical display of the grip force profiles as a function of time within each trial was utilized for confirming the detection of the behavioral events. In terms of motion, the onset of reaching was determined when the participant moved the hand from the starting position at a velocity speed above 30 mm/s, and the onset of object contact when the velocity speed returned under 30 mm/s**,** and finally the onset of object lifting when the velocity speed exceeded above 30 mm/s again. In terms of grip force application, the moment of object contact was determined when the first finger to touch the cup had a force application over 0.1 N and remained above that value.

The results were analyzed in the following manner, using the statistical software R (R Core Team, 2018). For each participant and trial, we first computed the grip force for the pinch grip by averaging the force profiles for the five digits on the right hand. That way, we obtained a single measurement for the overall grip force. Next, we computed the grip force for the individual digits by averaging the force measurements of the thumb together with one individual digit at a time, creating 4 different pairs of grip force profiles: The thumb and the index finger, the thumb and the middle finger, the thumb and the ring finger, and the thumb and the little finger.

The following variables were obtained and analyzed for the two sets of grip force averages, the pinch grip using all five digits and the pairs of individual digit force. First, we found the maximum grip force value occurring after the object contact, *peak grip force (peak GF)*. Next, we found the grip force rate as a function of time relative to the movement onset, and computed the peak values *(peak GFR)* occurring after the object contact and before object lift-off. Other variables of interest were peak GF at two timepoints after object contact to investigate the development of peak GF. First, we examined the peak GF occurring very soon after the object contact had been made, at maximum 10ms after the object contact. That timepoint could indicate what the participants anticipated before they corrected their grip force in order to achieve successful lifting. Given that the time point 10 ms after object contact is considered too early for any haptic information to be registered, but still has GF data based on the cup's properties, it is possible that the grip force magnitudes will reflect what the participants are expecting. (see e.g., Bergmann Tiest and Kappers, 2019). Second, we examined the peak GF occurring at maximum of 50 ms after the object contact for comparison purposes.

Analysis of variance was performed using repeated measures ANOVA to examine the effect of object weight, surface gloss, content feedback on grip force control, and effect sizes were reported as generalized eta squared (η^2_G). The three rating scales were analyzed using the following approach. To reduce individual rating bias, the ratings were normalized for each participant by subtracting the overall mean for each rating scale from each individual rating within the same rating scale, using the *Bbmisc* package in R (Bischl, et al., 2017). Regression analysis were performed using Linear mixed effect models (LMMs) using the lme4 package in

R (Bates et al., 2015) where all models had random intercept for each participant, and deviations from normality examined via Q-Q probability plots.

3. Results

3.1. Does five-digits grip force scaling account for material-based predictions?

The force measurements for the five digits were averaged for each participant and trial to obtain the average force representing the pinch grip. Next, a three-way repeated measures ANOVA was performed to examine the overall effect of object weight (empty vs. filled), surface appearance (matte vs. shiny), and content visibility (visible vs. hidden) on peak grip force (peak GF) and peak grip force rate (peak GFR), respectively.

Figure 4a illustrates the results for peak GF and shows that the participants did not only apply more grip force when lifting filled and consequently heavy cups, but they also applied more grip force when lifting cups with shiny properties or when the content of the cups could not be inferred visually prior to the object lifting. Peak GF was significantly higher when lifting filled cups (24.01 ± 3.82 N) compared to empty cups (22.12 ± 5.07 N), $F(1, 15) = 6.03$, $p < 0.05$, $\eta^2_G = 0.05$, and significantly higher when lifting cups with shiny appearance (23.85 ± 4.61 N) compared to matte appearance (22.28 ± 4.43 N), $F(1, 15) = 5.85$, $p < 0.05$, $\eta^2_G = 0.03$.
Cups with hidden content had also larger peak GF (23.79 ± 4.37 N) compared to cups with visible content (22.33 ± 4.68 N), $F(1, 15) = 5.57$, $p < 0.05$, $\eta^2_G = 0.02$, indicating that the participants applied more grip force when they were unsure about the weight of the cups. No significant two-way or three-way interactions were found, all *p*s = n.s.

Using repeated measures ANOVA analogous to the analysis performed on the overall data, we separated the two visibility conditions, and further examined the peak GF. For cups with only visible content, we found both object weight and surface appearance had significant effect on peak GF, where filled cups had higher peak GF (23.47 ± 4.29 N) compared to empty cups (21.19 ± 4.84 N), $F(1, 15) = 5.97$, $p < 0.05$, $\eta^2_G = 0.06$, and shiny cups had higher peak GF (23.28 ± 5.10 N) compared to matte cups (21.37 ± 4.07 N), $F(1, 15) = 5.47$, $p < 0.05$, $\eta^2_G = 0.05$. This was not the case when the content of the cups was hidden to the participants, as no significant difference in peak GF was found between the two levels of object weight (Empty: 21.0 ± 0.96 N; Filled: 22.2 ± 0.65 N), $F(1, 15) = 2$, *p* = n.s. or the two levels of surface appearance (Matte: 21.4 ± 0.86 N; Shiny: 21.5 ± 0.80 N), $F(1, 15) = 0.09$, *p* = n.s. when the content was hidden, confirming that the participants applied similar grip force for the different types of cups when the content was hidden.

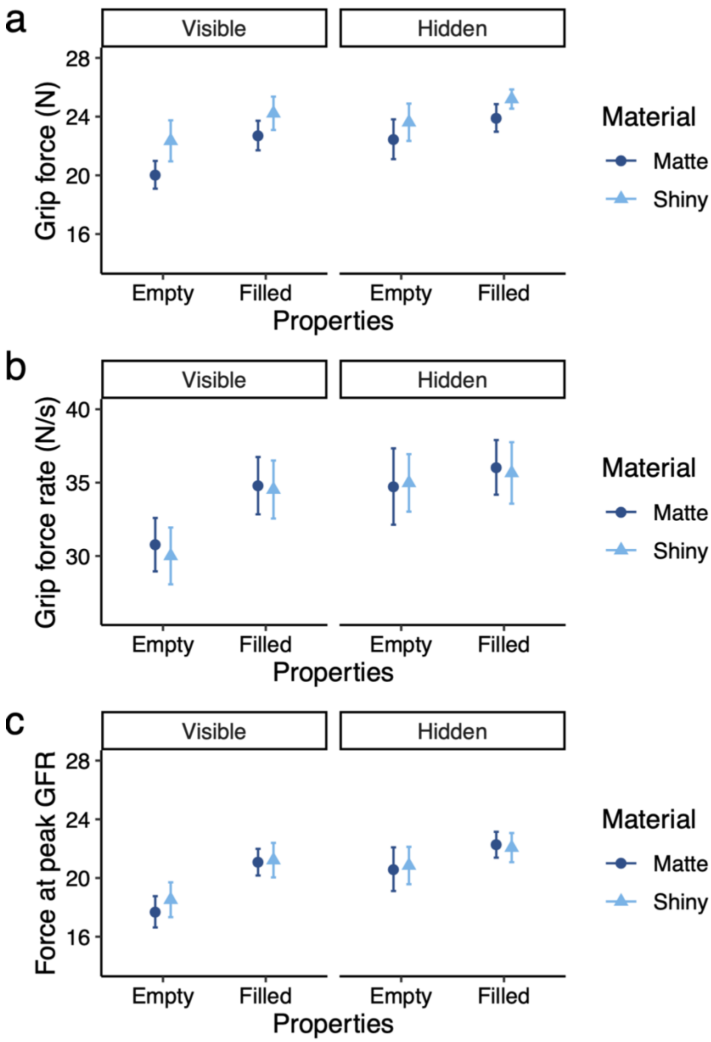

Figure 4. a) Average peak grip force, **b)** average peak grip force rate, and **c)** grip force at average peak grip force rate, for the different types of cups, *surface appearance (matte vs. shiny), object weight (empty vs. filled),* and *content visibility (visible vs. hidden).*

The anticipatory control of grip force is shown in Figure 4b. We found peak GFR varied for the different types of object weight and content visibility, in which we found significant interaction effect on peak GFR between the weight of the cups and the visibility of the content, $F(1, 15) = 6.92$, $p < 0.05$, $\eta^2_G = 0.01$, where filled cups had higher peak GFR compared to empty cups, but that difference in peak GFR was larger when the content was visible (filled: 34.70 ± 7.73; empty: 30.40 ± 7.40) compared to hidden (filled: 35.80 ± 7.80; empty: 34.90 ± 9.06). Furthermore, a separate analysis on peak GFR for cups with either visible or hidden content, revealed significant difference between the two object weights when the participants could see the content of the cups, $F(1, 15) = 6.22$, $p < 0.05$, $\eta^2_G = 0.08$, whereas no significant difference in peak GFR between the two object weights were found when the participants could not see the content of the cups, $F(1, 15) = 0.4$, p = n.s. Overall, peak GFR was significantly higher when lifting cups with hidden content (35.35 ± 8.40 N/s) compared to cups with visible content (32.52 ± 7.81 N/s), $F(1, 15) = 6.42$, $p < 0.05$, $\eta^2_G = 0.03$.

To highlight the difference in grip force magnitudes based on anticipation (i.e., peak GFR) and perceived feedback (i.e., peak GF), we examined the grip force magnitudes applied at the moment of peak GFR. Figure 4c shows both object weight and content visibility to have significant effect on the grip force magnitude at the moment of maximum GFR, in which greater GF was found for filled cups (21.65 ± 3.92) compared to empty cups (19.41 ± 5.11), $F(1, 15) = 7.00$, $p < 0.05$, $\eta^2_G = 0.06$, and cups with hidden content (21.44 ± 4.67) compared to visible content (19.62 ± 4.53) $F(1, 15) = 7.57$, $p < 0.05$, $\eta^2_G = 0.04$. A separate analysis for the two types of content visibility revealed a significant difference in GF at peak GFR between filled and empty cups when the content was visible, $F(1, 15) = 8.10$, $p < 0.05$, $\eta^2_G = 0.11$, whereas the difference was non-significant when the content was hidden, $F(1, 15) = 2.29$, p = n.s.

3.2. Do predictions based on anticipated vs. perceived material properties result in distinct grip force scaling?

Figure 5 shows the grip force profiles obtained from the first and the last (5th) block of trials to demonstrate the development of grip force control over the course of the study. In particular, figure 5a illustrates the difference in peak GF between the two blocks of trials, where repeated measures ANOVA on the first block of trials revealed significant main effect of object weight on peak GF, $F(1, 15) = 14.02$, $p < 0.05$, $\eta^2_G = 0.08$, and an interaction effect between the object weight and the surface appearance of the cups on peak GF, $F(1, 15) = 5.01$, $p < 0.05$, $\eta^2_G = 0.02$. The participants applied larger grip force when lifting the filled cups (19.80 ± 5.06 N) compared to empty cups (16.72 ± 6.13 N), and treated the two types of surface appearances similarly when lifting the filled cups (matte: 19.71 ± 4.71 N; shiny: 19.88 ± 5.47 N), whereas larger difference in peak GF was found for matte and shiny cups when the cups were empty (matte: 15.05 ± 6.31 N; shiny: 18.38 ± 5.56 N). There was also an interaction effect between the content visibility and the surface appearance of the cups on peak GF, $F(1, 15) = 5.44$, $p < 0.05$, $\eta^2_G = 0.04$. When the participants could not visually infer the content of the cups, they applied similar grip force when lifting either matte (19.11 ± 5.39 N) or shiny cups (18.83 ± 4.65 N), whereas shiny cups were lifted with greater peak GF (19.45 ± 6.34 N) compared to matte cups (15.65 ± 6.15 N) when the participants could see the content of the cups. In comparison, only object weight was found to have significant difference between its levels in terms of peak GF for the last block of trials, in which participants applied higher peak GF when lifting filled cups (20.02 ± 4.93 N) compared to light cups (17.49 ± 7.26 N), $F(1, 15) = 14.02$, $p < 0.05$, $\eta^2_G = 0.08$.

We also examined the grip force development for the first milliseconds after the object contact to see how the participants adjusted their grip force according to the tactile feedback.

We therefore examined the changes in grip force magnitudes occurring at two time points within the first block of trials, in particular the peak GF within the first 10 ms after object contact as it demonstrates force control based on visually inferred properties, and peak GF within the first 50 ms after object contact when the tactile sensation of the cups had become available to the participants and could therefore be learned. Repeated measures ANOVA analysis on the peak GF occurring within the first 10 ms after the object contact revealed interaction effect between the object weight and the surface appearance, $F(1, 15) = 4.95$, $p < 0.05$, $\eta^2_G = 0.01$, in which the participants applied more grip force when preparing to lift empty cups with shiny surface appearance (5.42 ± 5.02 N) compared to matte appearance (3.89 ± 4.28 N), but applied similar grip force for the two types of surface appearances when the cups were filled (matte: 4.31 ± 5.29 N; shiny: 4.17 ± 5.46 N). Further examination on the peak GF occurring within the first 50 ms after object contact showed surface appearance to have a dominant effect on the grip force application, as it was the only object property to have significant effect on peak GF. The participants applied more grip force when preparing to lift cups with shiny surface appearance (11.00 ± 4.95 N) compared to matte appearance (9.54 ± 4.75 N), $F(1, 15) = 4.95$, $p < 0.05$, $\eta^2_G = 0.01$, irrespective of their content or visibility of it. In comparison, for the last block of trials we found the participants had developed a strategy in which all cups were grabbed similarly in terms of grip force for the first 10 ms and 50 ms after object contact, as no significant difference in peak GF between the levels of different types of object properties were found for the two time points respectively (all *p*s = ns.).

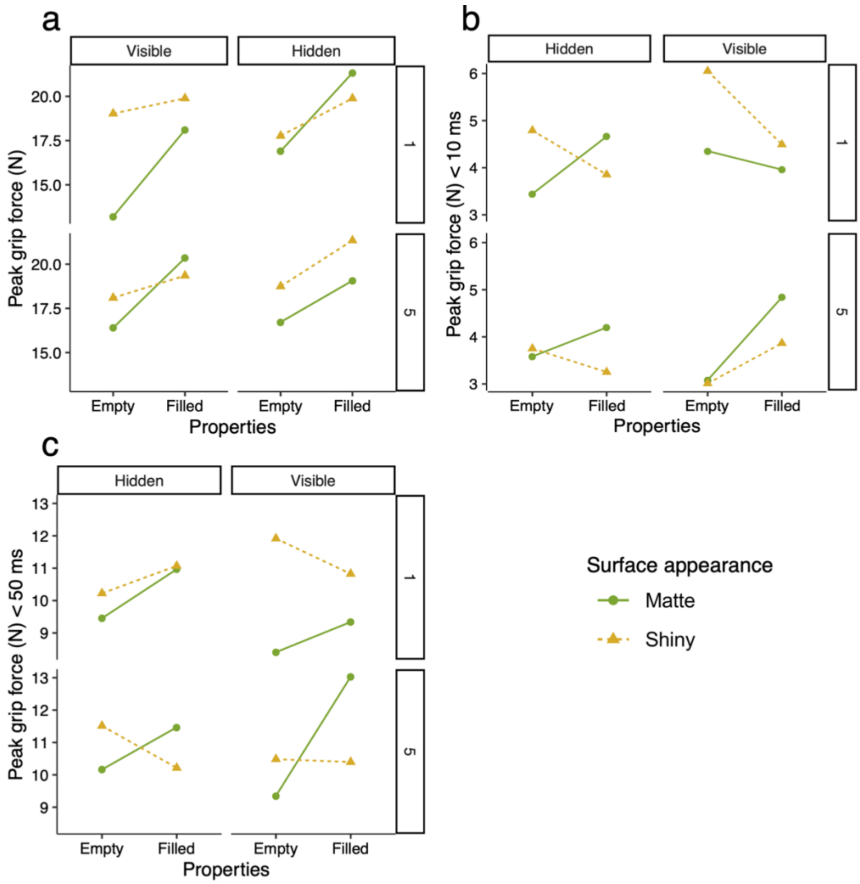

Figure 5. a) The average peak grip force, b) the average peak grip force within the first 10 milliseconds after the object contact, and c) the average peak grip force within the first 50 milliseconds after the object contact, for the three experimental properties (object weight, surface appearance, and content visibility) as obtained from the first (1) and the last (5) block of trials.

The distinct force profiles found for the object lifts obtained in the first and the last block of trials called for further examination, we therefore examined the correspondence between the physical properties of the cups, and how the participant experienced the three targeted object properties, heaviness, glossiness, and hardness, using the pre-reaching ratings from the first block of trials as a measure of expected properties, and the post-lifting ratings from the last block of trials as a measure of perceived properties, respectively. First, A LMM regression analysis based on the pre-reaching ratings revealed a significant correspondence between expected *heaviness* and object weight (b = 3.12, SE = 0.31, $\chi^2(1)$ = 90.19, $p < 0.001$), and content visibility (b = 1.37, SE = 0.31, $\chi^2(1)$ = 5.99, $p < 0.05$), specifically the interaction between the object weight and content visibility, b = -3.56, SE = 0.31, $\chi^2(1)$ = 126.77, $p < 0.001$. In terms of surface appearance, we found expected *glossiness* to correspond with surface gloss, b = 2.87, SE = 0.31, $\chi^2(1)$ = 342.85, $p < 0.001$, which suggests a strong categorical distinction for the two types of surface appearances. Finally, for expected *hardness* we found significant correspondence with the object weight, b = 1.12, SE = 0.25, $\chi^2(1)$ = 9.30, $p < 0.01$, especially as an interaction term with content visibility, b = -1.37, SE = 0.35, $\chi^2(1)$ = 20.94, $p < 0.001$. Second, a LMM analysis based on the post-lifting ratings, demonstrated a correspondence between perceived *heaviness* and object weight, b = 3.69, SE = 0.24, $\chi^2(1)$ = 984.18, $p < 0.001$, as well as correspondence between perceived *glossiness* and surface gloss, b = 2.88, SE = 0.31, $\chi^2(1)$ = 213.15, $p < 0.001$, whereas perceived *hardness* corresponded largely with object weight, b = 1.25, SE = 0.25, $\chi^2(1)$ = 143.51, $p < 0.001$. Figure 6 illustrates the average scores (normalized) of the three rating scales, heaviness, glossiness, and hardness, that were collected at two timepoints in the study, a) pre-reaching ratings collected in the first block of trials before the participants made object contact, and b) post-lifting ratings that were collected in the last block of trials after the participants had repeated experience of lifting the cups.

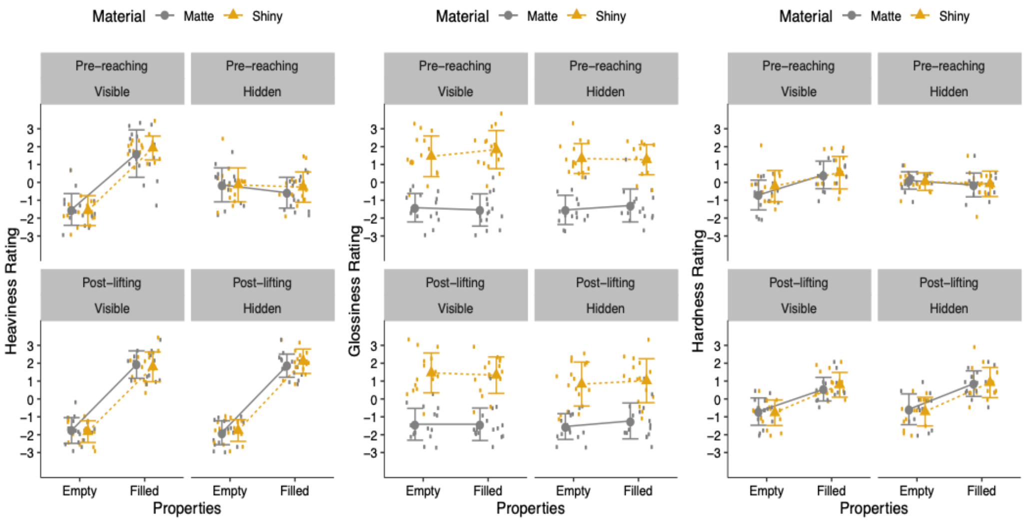

Figure 6. Rated heaviness, glossiness, and hardness as collected in block 1 before interacting with the cups (pre-reaching ratings) and in block 5 after several trials of interacting with the cups (post-lifting ratings). The figure shows normalized and averages and individual observations for each rating scale for the different types of cups. Error bars are standard deviations. Overall, the figure reveals clear categorical distinction for two of the scales, heaviness and glossiness, but less distinction is seen for the hardness scale.

Next, a linear mixed-model regression analysis was conducted separately for each of the three rating scales, heaviness, glossiness, and hardness, to examine the relationship between the rated material properties and the grip force magnitudes for the two types of content visibility, i.e., cups with visible or hidden content. We found significant main effects for all of the three rated properties, that is rated heaviness, $b = 0.76$, SE = 0.24, $\chi^2(1) = 7.18$, $p < 0.01$; rated glossiness, $b = 0.65$, SE = 0.27, $\chi^2(1) = 4.48$, $p < 0.05$; and rated hardness, $b = 2.06$, SE = 0.47, $\chi^2(1) = 13.04$, $p < 0.001$ on peak GF, and an interaction effect between rated hardness and content visibility on peak GF, $b = -1.80$, SE = 0.69, $\chi^2(1) = 6.63$, $p < 0.05$. This opted for further examination, and thus we examined the expected- and perceived properties separately. The analysis on the pre-reaching ratings obtained in the first block of trials revealed significant correspondence between the peak GF and the expected hardness, $b = 1.51$, SE = 0.58, $\chi^2(1) = 6.84$, $p < 0.01$, but not with expected heaviness, $b = 0.57$, SE = 0.30, $\chi^2(1) = 3.61$, p = .06. or expected glossiness, $b = 0.42$, SE = 0.26, $\chi^2(1) = 2.53$, p = n.s. Conversely, regression analysis on the post-lifting ratings obtained in the last block of trials revealed significant correspondence between peak GF and the perceived heaviness, $b = 0.46$, SE = 0.22, $\chi^2(1) = 4.26$, $p < 0.05$, including the perceived hardness, $b = 1.14$, SE = 0.44, $\chi^2(1) = 6.74$, $p < 0.01$ respectively, whereas perceived glossiness had no significant relations with peak GF in the last block of trials, $b = 0.42$, SE = 0.27, $\chi^2(1) = 2.32$, p = n.s. The grip force magnitudes (peak GF) as a function of rated properties are illustrated in figure 7.

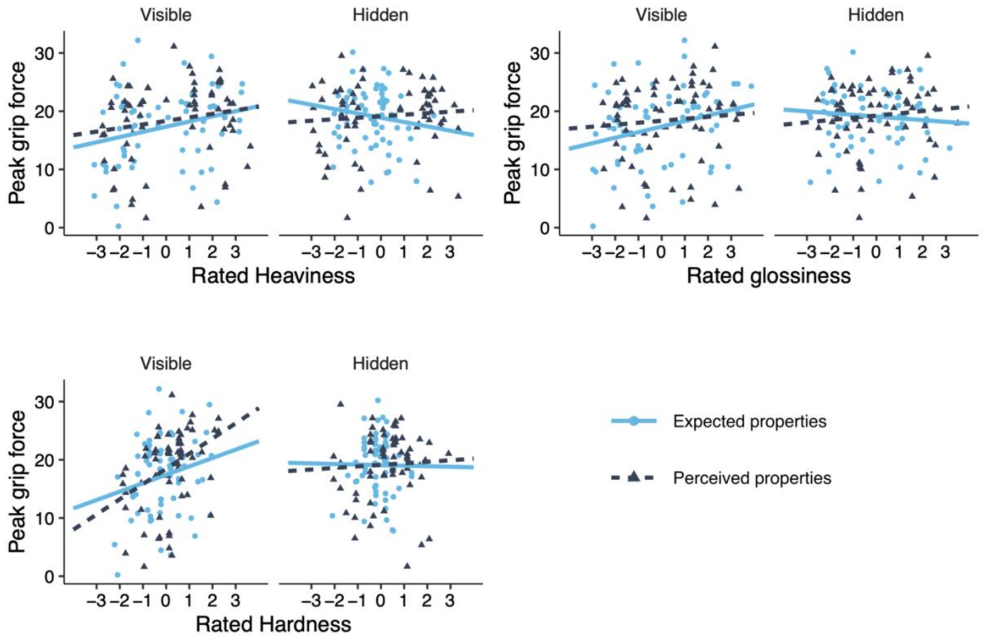

Figure 7. Peak grip force (peak GF) as a function of expected- and perceived properties as based on the three rating scales and grouped by content visibility, i.e., whether the content of the cups was visible or hidden. The figure shows individual ratings per visibility condition and separate regression lines for pre-reaching ratings (expected properties) and post-lifting ratings (perceived properties). The figure shows grip force magnitudes to increase the heavier, glossier, or harder the cups got rated when the participants could see the content of the cups, but to remain similar for the different rating scores when the content was hidden.

3.3 Do individual digits have different grip force profiles due to changes in material properties?

We extended the analysis on the five-digits pinch grip and computed the force profiles for individual pairs of digits to examine how the participants distributed the force across the digits. New grip force variables were computed to obtain accurate force magnitudes for the digits, in which the forces applied by each digit on the right side of the cups were individually averaged with the force applied by the thumb on the left side of the cup. This resulted in four different pairs of grip force variables: the average grip force between the thumb and the index finger (index); the thumb and the middle finger (middle); the thumb and the ring finger (ring); and the thumb and the little finger (little). Finally, for each participant and trial, we computed the peak GF for each pair of digits. Overall, a linear mixed-model regression analysis on the peak GF revealed various grip force magnitudes for the different pairs of digits, $\chi^2(3) = 268$, p < 0.001; index: $b = 0.65$, SE = 0.18, middle: $b = 5.69$, SE = 0.66, ring: $b = 6.42$, SE = 0.65, little: $b = 10.59$, SE = 0.64, which called for further examination. We therefore, analyzed the digits separately for the first and the last block of trials to further examine the involvement of visual- and perceived properties in grip force control.

For the first block of trials, we found larger range in peak GF between the different pairs of digits, especially between the digits closes to the base of the cups (range: 14.36 N – 27.38 N), in comparison to the last block of trials where the grip force magnitudes for the different pairs of digits were more clumped together (range: 16.46 N – 25.35 N). Moreover, for the first block of trials, we found the digits closest to the base of the cups carried most of the load when lifting the cups as those digits had the largest force magnitudes, especially the little finger, which had the largest difference in peak GF when lifting empty (M = 20.8 N) and filled cups (M = 25.0 N, Δ = 4.2 N). In terms of surface appearance, the digit with the largest difference in force magnitudes when lifting cups with either matte (M = 15.5 N) or shiny appearance (M = 20.5 N, Δ = 5 N) was the ring finger, which also had the largest difference in force magnitudes when lifting cups with either visible (M = 15.8 N) or hidden content (M = 20.3 N; Δ = 4.5 N). For the last block of trials, the index finger had a dominant role in grip force application, as it had the largest difference in peak GF for the two object weights (Empty: M = 7.30 N, Filled: M = 12.9 N, Δ = 5.6 N), and for the two types of content visibility (Visible: M = 10.8 N, Hidden: M = 9.35 N, Δ = 1.45 N), whereas the ring finger had the largest difference in peak GF between matte and shiny cups (Matte: M = 16.0 N, Shiny: M = 20.0 N, Δ = 5 N). A Kruskal-Wallis test on peak GF confirmed that the grip force magnitudes were significantly different for the different pairs of digits, both for the peak GF obtained from the first block of trials, $\chi^2(3) = 21.21$, $p < 0.001$, and from the last block of trials, $\chi^2(3) = 23.85$, $p < 0.001$. Overall, the middle digit (f0f2) had the smallest variation in peak GF, or standard deviation of 1.65 N, which made it the least affected digit by the object properties.

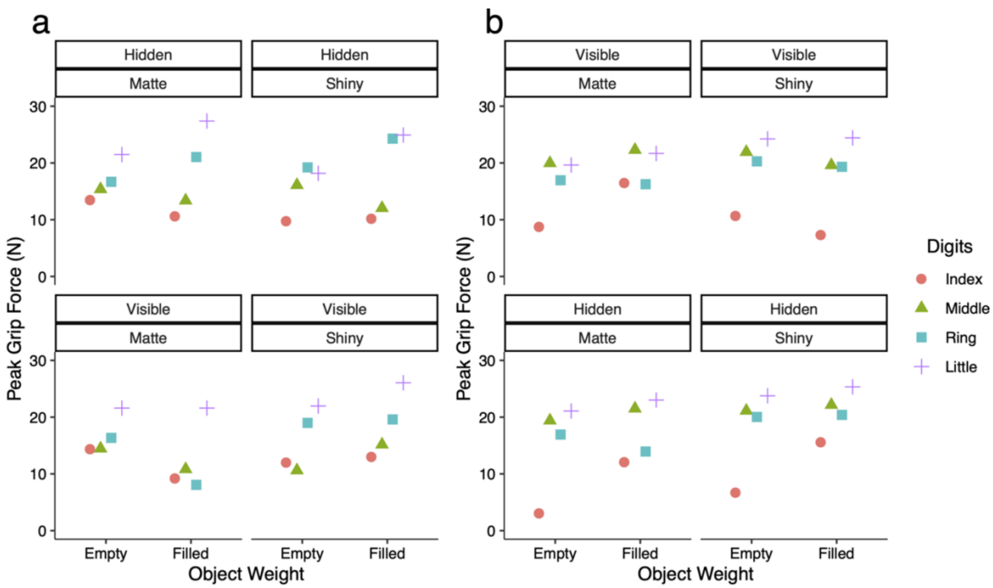

Figure 8. Peak GF for each of the four grasp pairs; thumb and the index finger (index), thumb and the middle finger (middle), thumb and the ring finger (ring), thumb and the little finger (little), both for the first block of trials (a) and the last block of trials (b). The digits positioned closest to the base of the cups had the largest grip force magnitudes and were mostly guided by the weight of the cups and their surface appearance.

4. Discussion

Skilled object manipulation requires accurate grip force scaling, which is guided by predictions that are based on learned associations between e.g., weight and material (Buckingham et al. 2009; Gordon et al. 1993), or weight and size (Gordon et al. 1991). These findings are primarily based on grip force data collected using the precision grip for experimental control. Here, in a grip-lift task, we examined grip force scaling in a real-world scenario resembling a daily activity, in which participants lifted paper cups using a five-digit hold, to further explore the role of material properties in grip force scaling. We asked participants to lift and transport paper cups that varied in surface appearance (matte vs. varnished surface), weight (empty vs. filled cups) and content visibility (visible vs. hidden content) while measuring their five-digit grip force using light and flexible force sensors attached directly to the fingertips. In this study we sought to answer the following three questions. First, does five-digits grip force scaling account for material-based predictions? Second, do predictions based on anticipated vs. perceived material properties result in distinct grip force scaling? Third, do individual digits have different grip force profiles due to changes in material properties?

Overall, we found five-digits grip force scaling when lifting objects using the whole hand to be guided by predictions that are based on material properties. The participants did not only scale their grip force according to the object weight, in which they applied more force (peak GF) when lifting the filled and heavier cups compared to the empty cups, but the participants did also apply greater grip force when lifting cups with glossy varnish as opposed to cups with matte appearance. Thus, the five-digits grip force scaling certainly accounts for object properties other than weight, specifically surface properties that are at first hand perceived via sight. Oddly, we did not find any effect of surface appearance on peak grip force rate (peak

GFR), which normally is considered to be a measurement that reflects the anticipatory control of the grip force scaling (Johansson and Westling, 1988). Instead, we found object weight and content visibility to have considerable effect on the rate of grip force magnitudes, where filled cups had quicker GFR, specifically when the content was visible to the participants, whereas cups with hidden content had similar peak GFR for filled and empty cups.

We further examined the grip force magnitudes occurring within the timespan between the onset of object contact and the point of maximum grip force rate (GF at peak GFR), and found similar results as we did for peak GFR alone. Evidently, grip force scaling was different depending on whether the participants could visually infer the weight of the cups based on their content prior to lifting them or not, i.e., whether the content was visible or hidden. The overall results for peak grip force when lifting cups with hidden content revealed similar grip force magnitudes across the different types of cups. Moreover, the grip force magnitudes were generally larger in comparison to the grip force magnitudes applied when lifting cups with visible content, which indicates a possible "better safe than sorry" strategy, where high grip force is applied to the object when in doubt of its weight to prevent torque (see e.g., Gorniak, Zatsiorsky, and Latash, 2009).

When examining the role of anticipated- and perceived properties in grip force scaling, we found grip force predictions based on anticipated properties resulted in different grip force scaling compared to predictions based on perceived properties. In general, when lifting objects, people rely on internal representations to ensure stable grasp, where previous object manipulation are stored and then retrieved for planning a future manipulation (Johansson and Westling, 1988, Gordon et al., 1993, Gordon et al., 1994; Li et al. 2009; Nowak et al. 2007b). People anticipate the required grip force for lifting the targeted object, in which the grip force scaling is at first based on the expected object weight or expected frictional properties based on visual cues (e.g., size and material appearance) and prior experience, then once the object has been touched and manipulated, the internal representations are updated according to the provided sensorimotor feedback (Johansson and Westling, 1984; van Polanen and Davare, 2015). Our participants scaled their grip force differently depending on whether the anticipated grip force requirements were based on expected or perceived properties, as the comparison between the grip force magnitudes obtained from the first block of trials and the grip force magnitudes obtained from the last block of trials revealed. We found object weight and surface appearance both affected the scaling of the grip force during the first object lifts, whereas the grip force scaling occurring in the last block of trials was dominated by object weight alone. That said, visual information about material properties, specifically surface gloss, has a significant role in early grip force scaling, but is quickly replaced by the learned tactile information when the participants have experienced the true weight of the cups by interacting with them. Similar effect has previously been demonstrated when lifting symmetric objects using only two digits by Crajé, Santello, and Gordon (2013), in which they showed people used visual density cues for plastic and brass, to scale their grip force on initial lifts.

To further examine the anticipatory feed-forward control of grip force scaling and the role of visual material properties before tactile information becomes available, we examined the five-digits grip force magnitudes at two timepoints during the initial object lifts. Namely, the peak grip force occurring within the first 10 milliseconds after the object contact, as well as within the first 50 milliseconds after the object contact. For the first 10 milliseconds after the object contact, we found both object weight and surface appearance had significant effect on peak grip force. Empty cups with shiny surface had larger grip force magnitudes compared to empty cups with matte appearance, whereas the two surface appearances had similar grip force magnitudes when the cups were filled. Interestingly, when examining the peak grip force occurring within the first 50 milliseconds after object contact, we found surface appearance had a dominant effect on scaling the grip force, in which larger grip force magnitudes were found for cups with shiny appearance compared to matte. For comparison purposes we examined the grip force magnitudes occurring within the same timepoints, however, in the last block of trials, and found no significant difference between the different types of cups. Suggesting, the different object properties had equal role in scaling the grip force for the first milliseconds after object contact, at least after the true physical weight of the cups had been learned. We also collected the subjective experience of the cups' object properties and examined those ratings in relations to the grip force measurements.

Based on the properties-ratings, collected before lifting the cups, we found both expected heaviness and expected hardness corresponded well with object weight and content visibility, while expected glossiness corresponded well with surface gloss alone. However, we only found expected hardness to have significant correspondence with the grip force magnitudes in the first block of trials. In comparison, the ratings collected in the last block of trials after the participants had repeatedly manipulated the cups, revealed both perceived heaviness and perceived hardness corresponded well with object weight, while perceived glossiness corresponded well with surface gloss. Interestingly, both perceived heaviness and perceived hardness were found to have significant correspondence with the grip force magnitudes in the last block of trials. It is possible that during the initial object lifts, the participants scaled their grip force according to the visual cue for hardness based on their knowledge of material density, while later object lifts are sensory driven and therefore controlled by the learned object weight and surface hardness. Another possibility is that the participants have difficulty distinguishing between hardness and density, and thus the hardness ratings might actually represent rated density. Further experimentation will be required to test this. However, in our previous study using similar experimental stimuli, we found similar tendency to assign filled cups with harder qualities compared to empty cups, and concluded that the hardness ratings apply correctly, given that filled cups are less deformable as the liquid creates a force against the insides of the cups (Ingvarsdóttir and Balkenius, 2020). Moreover, and similarly to the results in current study, we found hardness ratings did not only apply to the amount of liquid in the cups, because a difference in hardness ratings was found between matte and shiny cups that were empty, in which empty shiny cups were considered to have harder quality compared to empty matte cup, making it less likely that the hardness ratings were primarily based on the liquid's density.

Here we have shown that surface gloss has a role in early grip force scaling before object lifting. Previously, we showed that surface gloss affected the temporal component of reaching during initial lifting of similar cups as used in the current study, while later lifts and spatial grip position on the cups are predominantly controlled by the weight of the cups. In Ingvarsdóttir and Balkenius (2020), we found the adjustment time, i.e., the deceleration time in reaching before object contact, was affected differently by the two surface appearances, depending on whether the cups had previously been rated to have hard or soft qualities. Shiny cups assigned with hard surface properties tended to have longer adjustment time in reaching, compared to shiny cups with assigned soft properties that had quicker adjustment time in reaching, whereas matte cups required similar adjustment time, irrespective of their hardness assignment.

We were also curious to see if the object properties caused distinct force profiles for each individual digit pair. Our analysis showed the distribution of peak grip force across the digits varied depending on whether the cups were filled or empty, had shiny or matte surface, and whether their content was visible or hidden. Overall, we found the digits closest to the base of the cups carried most of the load as they had the largest grip force magnitudes. Interestingly, however, we also found the grip force magnitudes spread differently across the different digits for the first block of trials and for the last block of trials. The range in grip force was larger during the first object lifts, whereas the grip force magnitudes of the digits that had significant roles in the last block of trials were more clumped together. Each digit had a distinct pattern of grip force control that was different for the first- and the last object lifts. For the first block of trials, we found the grip force magnitudes applied by the little finger corresponded best with the weight of the cups, while the ring finger corresponded best with the surface appearance and the content visibility. Comparatively, the grip force magnitudes applied in the last block of trials, revealed a larger role for the index finger, which had a great correspondence with the object weight and the content visibility, as well as a continuing role of the ring finger, which corresponded greatly with the surface appearance of the cups. A plausible explanation for the distinct grip force pattern per digit could be that each digit has a different role in manipulating the cups, especially for the first encounters, in which anticipated grip force requirements are based on expectations. It is likely that the participants apply most of the grip force to the lower part of the cups, using the digits closest to the base of the cups, to secure a space above the vertical grip force axis for repositioning the grip in case the cups are heavier or more slippery than expected and could thus slide down. While in the last block of trials, when the participants have learned the true weight and frictional properties of the cups, the grip force is distributed further up spatially, i.e., the index finger has a larger role in supporting the cups while lifting them. All in all, the digit that was least affected by the object properties, was the middle digit, which indicates its role is mostly to support the hold, irrespective of object properties. Anticipatory control based on visual information is important for skillful manipulations of objects, and in communications, such as successful handovers, i.e., when one person, or even a humanoid robot, hands an object over to another person or robot (see e.g., Controzzi, et al., 2018). Studies on material properties and their roles in early grip force planning when handling objects using the whole hand are therefore logical future step, specifically if the goal is to resemble real-world scenarios. A detailed examination of individual digits and early grip force planning when handling objects made of various materials would therefore be a worthwhile future study.

In sum, early grip force scaling when lifting paper cups using five-digits hold, is not only guided by the object weight but also by the surface appearance of the cups, specifically the surface gloss. We also found our participants scaled their grip force differently depending on whether their anticipated grip force requirements were based on expected or perceived properties, and that individual digits had distinct grip force profiles that depended on the object properties of the cups.

References

Bergmann Tiest, W. M. and Kappers, A. (2019). The influence of visual and haptic material information on early grasping force. *Royal Society Open Science*, *6*(3), [181563]. DOI: 10.1098/rsos.181563

Buckingham, G., Cant, J. S., and Goodale, M. A. (2009). Living in a material world: how visual cues to material properties affect the way that we lift objects and perceive their weight. *Journal of Neurophysiology,102*(6), 3111–3118. DOI: 10.1152/jn.00515.2009

Bullock, I. M., Zheng, J. Z., Rosa, S. D. L., Guertler, C., Dollar, A. M. (2013) Grasp frequency and usage in daily household and machine shop tasks. *IEEE Transactions on Haptics, 6*(3), 296 – 308. DOI: 10.1109/TOH.2013.6

Controzzi, M., Singh, H., Cini, F. et al. (2018) Humans adjust their grip force when passing an object according to the observed speed of the partner's reaching out movement. *Experimental Brain Research, 236*, 3363–3377. https://doi.org/10.1007/s00221-018-5381-5.

Crajé, C., Santello, M., and Gordon, A. M. (2013) Effects of visual cues of object density on perception and anticipatory control of dexterous manipulation. *PLoS ONE, 8*(10), e76855. DOI: 10.1371/journal.pone.0076855

Dafotakis, M., Sparing, R., Eickhoff, S. B., Fink, G. R., and Nowak, D. A. (2008). On the role of the ventral premotor cortex and anterior intraparietal area for predictive and reactive scaling of grip force. *Brain Research*, *1228*, 73-80. DOI:10.1016/j.brainres.2008.06.027

Flanagan, J. R., and Wing, A. M. (1997). The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads. *Journal of Neuroscience, 17*(4), 1519-1528. DOI: 10.1523/JNEUROSCI.17-04-01519.1997

Flanagan, J.R., Bowman, M.C., and Johansson, R.S. (2006). Control strategies in object manipulation tasks. *Current Opinion in Neurobiology 16*(6), 650–659. DOI: 10.1016/j.conb.2006.10.005

Gordon, A. M., Forssberg, H., Johansson, R. S., and Westling, G. (1991). Visual size cues in the programming of manipulative forces during precision grip. *Experimental Brain Research. 83*(3), 447–482. DOI: 10.1007/bf00229824.

Gordon, A., Westling, G., Cole, K., and Johansson, R. (1993). Memory representations underlying motor commands used during manipulation of common and novel objects. *Journal of Neurophysiology, 69*(6), 1789-1796. DOI: 10.1152/jn.1993.69.6.1789

Gordon, A., Forssberg, H., and Iwasaki, N. (1994). Formation and lateralization of internal representations underlying motor commands during precision grip. *Neuropsychologia, 32*(5), 555-568. DOI: 10.1016/0028-3932(94)90144-9.

Gorniak, S.L., Zatsiorsky, V.M. and Latash, M.L. (2009). Hierarchical control of static prehension: I. Biomechanics. *Experimental Brain Research, 193*, 615–631. https://doi.org/10.1007/s00221-008-1662-8.

Hermsdörfer, J., Li, Y., Randerath, J., Goldenberg, G., and Eidenmüller, S. (2011). Anticipatory scaling of grip forces when lifting objects of everyday life. *Experimental Brain Research, 212,* 19-31. DOI: 10.1007/s00221-011-2695-y.

Ingvarsdóttir, K.Ó. and Balkenius, C. (2020) The Visual Perception of Material Properties Affects Motor Planning in Prehension: An Analysis of Temporal and Spatial Components of Lifting Cups. *Frontiers in Psychology, 11,* 215. DOI: 10.3389/fpsyg.2020.00215

Johansson, R.S., and Westling, G. (1984). Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental Brain Research. 56*, 550–564. DOI:10. 1007/bf00237997.

Johansson, R.S., and Westling, G. (1988). Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Experimental Brain Research, 71*(1), 59-71. DOI: doi.org/10.1007/BF00247522

Johansson, R. S. and Edin, B. B. (1993). Predictive feed-forward sensory control during grasping and manipulation in man. *Biomedical Research, 14*(4), 95 - 106. DOI: 10.1523/JNEUROSCI.17-04-01519.1997

Marneweck, M., Lee-Miller, T., Santello, M., and Gordon, A. M. (2016). Digit position and forces covary during anticipatory control of whole-hand manipulation. Front. Hum. Neurosci. 10:461. doi: 10.3389/fnhum.2016.00461

van Polanen, V., and Davare, M. (2015). Sensorimotor memory biases weight perception during object lifting. Frontiers in Human Neuroscience. 9:700. DOI: 10.3389/fnhum.2015.00700

Wolpert, D.M., and Flanagan, J.R. (2001). Motor prediction. *Current Biology 11*(18), R729–R732. DOI: 10.1016/s0960-9822(01)00432-8

Zatsiorsky, V. M., and Latash, M. L. (2008). Multifinger prehension: an overview. J. Mot. Behav. 40, 446–476. doi: 10.3200/JMBR.40.5.446-476

Bischl, B., Lang, M., Bossek, J., Horn, D., Richter, J. and Surmann, D. (2017). BBmisc: Miscellaneous Helper Functions for B. Bischl. R package version 1.11. https://CRAN.R-project.org/package=BBmisc (accessed June 02, 2020).

Cutkosky, M. R. (1989). On Grasp Choice Grasp Models and the Design of Hands for Manufacturing Tasks. *IEEE Trans. Robotics and Automation, 5*(3), 269 – 279.

Napier, J. (1956). The Prehensile Movements of the Human Hand. *Journal of Bone Joint Surgery, 38*(4), 902 – 913.

Sperling, L. and Jacobson-Sollerman, C. (1977). The Grip Pattern of the Healthy Hand during Eating. *Scandinavian J. Rehabilitation Medicine, 9*(3), 115 - 121.

Paper III

Material properties and visual judgment of bounce height

Kristín Ósk Ingvarsdóttir [1*], Christian Balkenius[1]

[1]Lund University Cognitive Science, Lund University, Helgonavägen 3, 221 00 Lund, Sweden.

*** Correspondence:**
Corresponding Author
Kristin.ingvarsdottir@lucs.lu.se

Keywords: material properties1, bouncing object2, roughness3, glossiness4, learned associations5

Abstract

An extensive research has been done on perception of bouncing objects, yet little is known about how material properties are incorporated into the visual assessment of bounce height, in particular the material properties of the surface plane affording the bounces. The bounce height of a bouncing ball is controlled by the physical elements of the ball, but also by the nature of the surface plane it bounces on, such as whether it has hard or soft qualities. The appearance of the ground could therefore potentially affect the judgment of bounce height by either reducing or enhancing the perceived height of the bounce, depending on the learned association between the visual- and the intrinsic properties of the surface plane. To investigate the influence of material properties on visual judgments of bounce height, we asked participants to compare the bounce heights of two successive presentations of animated balls bouncing on either roughly or smoothly textured planes. In a second study we examined the estimation of the same bounce heights, when specular highlights were added to the surface planes, to further examine the role of its visual appearance in perceived bounce height. We found the participants judged the bounce heights differently depending on the visual characteristics of the surface planes, where rough surface planes increased the estimated bounce heights, compared to smooth surface planes. This effect of rough surface planes was also found in Experiment 2 for the shiny surface planes, however the judged bounce height differed less between the two textures, rough and smooth. The current study proposes that the visual assessment of bounce height incorporates more than simply the physical elements of the ball and the bounce, as our results demonstrates that different surface appearances can cause changes in judged intrinsic properties, and consequently perceived bounce height. We propose that perceived bounce height incorporates learned associations between *what materials look like* and *what materials behave like.*

Introduction

One of many challenges in research on material perception is to understand how visual characteristics of surfaces are used to judge intrinsic physical properties that are normally perceived by other means than sight. In a daily activity such as lifting an object, we need to visually infer intrinsic properties that are normally perceived haptically. For instance, prior to lifting an object, we anticipate its weight before receiving the sensorimotor feedback provided by the object contact. This anticipation is based on the object's visual appearance, and our previous experience of interacting with similar objects in similar scenes (Johansson and Westling, 1988). In that way, we can prepare the required grip force for lifting the object without slippage.

Weight anticipation has been extensively studied in object lifting, in particular in relations to object size, where studies on size-weight illusions during object comparison demonstrate how judgment of intrinsic properties are affected by the visual appearance of the object such as its size. The perceptual illusion occurs when people judge a smaller object to be heavier than a larger object. even when the two objects have identical mass. The illusion continues to persist perceptually while the motor movement of the hand (i.e., the grip) is adjusted according to the actual mass of the objects (e.g., Flanagan and Beltzner, 2000). Another example of judgment of intrinsic object properties based on extrinsic properties, are material-weight illusions. The illusions demonstrate that the assessment of weight, is highly influenced by the visual characteristics of the material from which the object is made. An object made from a light material (e.g., Styrofoam) is judged to be heavier than an object made from a heavy material (e.g., stone), despite them being altered to have equal mass and size. Similar to the size-weight illusion, the material-weight illusion is independent of the forces applied when lifting the objects (Buckingham, Cant and Goodale, 2009; Charpentier, 1891; Ellis & Lederman, 1999; Harshfield & DeHardt, 1970; Seashore, 1899; Wolfe, 1898).

Intrinsic properties of materials are often expressed visually. For instance, spilled milk has different visual characteristics than spilled yoghurt due to their differences in viscosity (Adelson, 2001; van Assen, Barla, and Fleming, 2018), and thickness and stiffness of fabrics can be inferred based on how they move and fold (Xiao, Bi, Jia, Wei, and Adelson, 2016). When judging hardness of compliant and hard materials, people normally rely on tactile information such as surface deformation, vibrations, and force/finger-displacement ratio, as hardness is more readily encoded by haptics (Bergmann Tiest & Kappers, 2009; Bergmann Tiest, 2010). However, due to learned associations between material appearances and their intrinsic properties, people are able to judge the hardness based on only the visual features of the material (e.g., Fleming, Wiebel, & Gegenfurtner, 2013). Indicating that perceived hardness is not bound to the tactile sensory modality alone (see Baumgartner, Wiebel and Gegenfurtner, 2013 for a study on correspondence between visual and haptic perceptual spaces of materials, and for studies on associations between auditory information and perceived hardness see Arnott, et al., 2008; Giordano & McAdams, 2006; Vickers, 1981; and Vickers, 1987). On the basis of these studies, the conjecture is that people will incorporate a generic knowledge about materials and their visual properties when estimating the intrinsic properties of an object.

When a ball hits a surface plane and bounces back into the air, its bounce height is affected by the physical elements of the ball, its mass and elasticity, as well as environmental variables such as the friction of the surface plane and the aerodynamic drag. For instance, a ball bouncing on a hardwood floor that has hard qualities affords higher bounce height than a ball bouncing on a carpet floor, because the hard floor absorbs less kinetic energy than the softer floor. To investigate the effect of generic knowledge of material properties on the perception of bounce height, we simulated a spherical object with familiar appearance bouncing on surface planes with various material properties. In two experiments using paired-comparison task, we asked participants to visually judge bounce heights presented together with either smooth or rough surface plane, with and without specular highlights. The goal was to examine if alterations of the visual properties of the surface plane would give rise to changes in perception of bounce height, where we theorized that our participants would judge the bounce heights on visually different surface planes differently due to learned associations.

Although roughness is typically defined as a tactile property, its salient appearance is informative enough to be visually assessed as a distinctive feature for various materials (e.g., Adelson, 2001; Anderson, 2011; and Bergmann Tiest & Kappers, 2007). On the other hand, visual glossiness is a complex material property to assess due to various environmental factors, such as the position of the light source and the observer's viewing position to name few. The light source position controls the amount of light reflected of the surface, and high-gloss surfaces contain images of its surrounding environment in its reflection, that are easily changed when the viewer position is changed. Consequently, visual assessment of glossiness appears to be based on surface highlights rather than involving estimation of specular reflectance parameter of a surface. Specular highlights in form of white patches that are visually qualitatively different (e.g., size, brightness) for various surface characteristics and level of glossiness (Marlow, Kim and Anderson 2012; Beck & Prazdny, 1981; Marlow & Anderson, 2013).

In summary, we argue that hardness of a surface plane can be visually assessed when observing a familiar object bouncing on that surface due to prior experiences with similar scenes and materials. Moreover, we argue that perceived bounce height incorporates expectations of 'bounce affordance' that vary with the visual appearance of the materials. We expected rough surface planes to be assigned high bounce heights, based on the assumption that materials with rough characteristics (e.g., stone and wood) are often perceived as having a hard quality (see e.g., Fleming, et al., 2013 for more material classes with distinctive typical qualities). Moreover, we expected the addition of specular highlights to the rough surface texture would either enhance or diminish the previous found effect (if any), depending on whether the shiny surface planes are expected to have hard or soft quality.

Experiment 1: Roughness

The visual assessment of bounce height was examined using paired comparison. In particular, we investigated the effect of the material properties roughness and smoothness on the perception of bounce height.

Methods (Experiment 1)

Participants

Thirty right-handed participants (19 females and 11 males) recruited on Lund University campus (*Mean* age = 26 years, *SD* = 4 years) took part in the study. Participants gave their consent to participate in the experiment. All participants reported normal or corrected-to-normal vision.

Apparatus

An Apple Mac OS X 10.8.5 (Intel HD Graphics 4000 512 MB) running MATLAB (The MathWorks Inc., 2013) and Psychophysics Toolbox 3 (Brainard, 1997) controlled the experiment. Stimuli were presented on a Samsung Wide Screen LCD Monitor with the aspect ratio 16:10 (1920 x 1200 resolution) and a 75 Hz refresh rate. All participants were tested individually in a quiet darkened room seated at a viewing distance of 70 cm form the screen of the LCD monitor. Participants responded using two response keys, F (left) and K (right), on a standard QUERTY keyboard. Background illumination of the LCD monitor was held constant at 0.1 cd/m^2.

Stimuli

The stimuli were 10 animations of a ball bouncing once on a plane having either a smooth or a rough surface appearance. The stimuli were rendered and animated using two distinct software products; a) the computer graphics software Blender to create the visual appearance of the ball and the surface planes (Blender Foundation, 2013), and b) the game engine Unity 3D to create the geometrical shape of the ball and the plane, including the animation of the five different bounce heights (Unity Technologies, 2013). The ball's appearance was created using a material downloaded from the Blender Open Material Repository (Engel, 2007). We selected a material resembling the appearance of a basketball, an orange rubber-like surface with black ribs, because Swedish students typically have knowledge of balls with such appearance and bounce properties.

The surface planes, on which the ball bounced, were created in Blender as well. A burned orange colored base with matte appearance was assigned to the surface plane. The matte appearance was created using a Lambertian diffuse shader, with the brightness intensity kept at the default level (0.8). Such shader is typically used for simulating materials that have a low level of specular reflection, as the shader imitates a light being isotropically scattered in all directions. The result is a balanced appearance throughout the surface plane, which prevents that reflection spots appear when a light hits the surface plane. Preventing reflection spots to appear on the surface was important for current study, as they would have indicated a surface plane made of an intermediately reflective material, which was not the goal for Experiment 1. We created the surface planes using properties commonly seen in ground surfaces and floor covering with hard qualities, like concrete flooring, laminate, linoleum, and marble, etc. The rough appearance of the surface plane was obtained in Blender, with a Voronoi texture map with irregularly shaped cells with noise size of 0.1. A procedural texture map proved to be suitable for simulating objects with a rough appearance such as hammered metal or rocky mountain terrains (Worley, 2003). The smooth uniform appearance was obtained by not applying any procedural texture map to the surface plane. When the visual appearances of the ball and the surface planes were ready, they were *baked* (rendered) as textures in Blender and then imported into Unity to be assigned to the two geometrical meshes, for the ball and the surface plane. Examples of the two surface planes, with smooth texture and rough texture, and the ball's appearance, are seen in Figure 1.

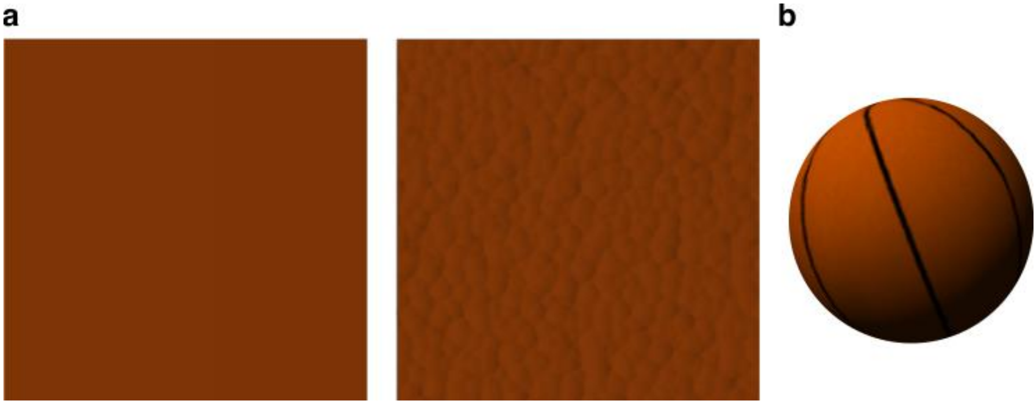

Figure 1. a) The visual appearance of the two rendered surface planes, smooth (left) and rough (right), b) the visual appearance of the bouncing ball.

The movement of the ball bouncing on the surface plane was simulated using Unity 3D's built-in physics engine PhysX® (NVIDIA Corporation, 2013). The ball's bounce height was systematically varied in 5 steps from 31 mm to 43 mm vertical height as measured from the base of the plane (2° 31'-3° 31' visual angle), while the surface appearance of the ball was held constant. Each animation started when the ball was on the surface plane rather than in mid-air. This was done to prevent the participants from making predictions of the coming bounce height before the ball bounced on the surface plane. The bounce animations were created using the Unity's physics engine, by adding a Rigidbody component to the surface planes and assigning a physics material to it. The five bounce heights were achieved by changing the bounciness settings of the physics material in an increment of five steps (0.900, 0.925, 0.950, 0.975, 1). The bounciness was set using Unity's arbitrary physics units where 0 represents no bounce at all and 1 is the maximum possible bounce of that object. All other physics variables such as gravity, mass, angular drag, and the colliding object's friction were held constant across all the animated stimuli. The bounce settings were the same for the rough and smooth surface textures.

A directional light was used in Unity to simulate environmental illumination such as sunlight. The light source was positioned above the plane and slightly to the side to make the bumps on the plane cast shadows. An aim was to ensure that the participants had sufficient visual information about the surface texture to be able to make judgments about its roughness. Both the surface plane and the ball were made to cast and receive shadows, although studies have shown that moving cast shadows can influence the perceived trajectory of an object by inducing its apparent motion (Kersten et al., 1996; Kersten, Mamassian, & Knill, 1997; Mamassian, Knill, & Kersten, 1998). More importantly, cast shadows are known to be important for localizing objects in three-dimensional scenes, and without them, objects appear to float above the ground rather than being in contact with it (Imura et al., 2006; Sugano, Kato & Tachibana, 2003; and Yonas & Granrud, 2006). This being the case, we determined it was important to have cast shadows visible during the bounce impact when the ball touched the surface planes, in order for our participants to perceived the full impact of the ball bouncing on the surface plane.

A custom JavaScript was used to record and export each stimulus animation rendered in Unity, into QuickTime movie format for stimulus presentation. The dimensions of the stimuli were 13.5 cm x 10.2 cm (11° x 8° 20' visual angle) and the duration of each animation was roughly 1 sec.

Design

Each participant completed two tasks; a paired-comparison task followed by a rating task. In the paired comparison task the participants observed 10 animations of a ball bouncing on either a rough or smooth surface plane with 5 different bounce heights each. The animations were organized into a total of 18 different combinations of pairs, using a diamond shaped pairing combination. This design has been proved to be efficient in avoiding problems of collinearity when studying discrimination of paired stimuli (Hellström, 1979, 2003; Patching, Englund & Hellström, 2012). The two types of presentation orders consisted of 9 pairing combinations each. In one arrangement, the first stimulus contained a smooth surface plane while the second stimulus contained a rough surface plane. In another arrangement, the first stimulus contained the rough surface plane and second stimulus contained the smooth surface plane. The participants completed 4 blocks of test trials preceded by one block of practice trials. Each test-block comprised of 18 combinations of stimulus pairs (9 for each presentation order) where each stimulus pair was presented 4 times in each block, resulting in total of 288 trials per participant. All trials, including the practice trials were randomized. The diamond shaped factorial combination for one presentation order is illustrated in Figure 2. The two graphs demonstrate the average bounce height, and the intra-pair difference in bounce height, for each combination of stimulus pairs. The average bounce heights were 34 mm, 37 mm, and 40 mm, and were calculated by adding the bounce height of the first presented stimulus (e.g., 37 mm) to the second presented stimulus (e.g., 31 mm), divided by two. The intra-pair differences in bounce height were ±6 mm, and 0 mm, and were calculated by subtracting the bounce height of the first presented stimulus (e.g., 37 mm) from the bounce height of the second presented stimulus (e.g., 43 mm).

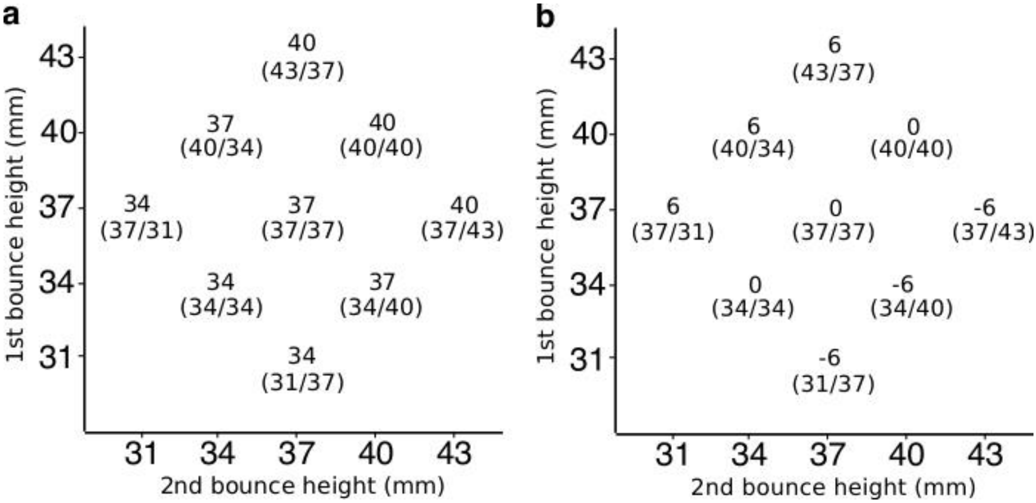

Figure 2. Diamond shaped pairing combinations, demonstrating the 9 different combinations of stimulus pairs for each presentation order; rough surface plane presented first vs. smooth surface plane presented second, and vice versa. (a) A combination demonstrating the average bounce height for each stimulus pair, ranging from lower left to upper right, 34 mm, 37 mm, and 40 mm. (b) A combination demonstrating the intra-pair difference for each stimulus pair, ranging from lower right to upper left, -6 mm, 0 mm, and 6 mm.

The response assignment was counterbalanced over participants. 15 participants were instructed to use the index finger of their left hand to press the left key when they judged the first stimulus of each stimulus pair to have higher bounce height than the second stimulus, and the index finger of their right hand to press the right key when they judged the second stimulus to have higher bounce height than the first stimulus. The other 15 participants were instructed to use a reversed order of key presses – the index finger of their left hand to press the left key if they judged the second stimulus to have higher bounce than the first stimulus, and the index finger of their right hand to press the right key if they judged the first stimulus to have higher bounce than the second stimulus.

The idea underlying the rating task was to examine if participants perceived the intended appearance of the various surface planes. Participants rated the surface property of the planes of the same 10 animations as shown in the paired-comparison task. The presentation order of the animations was randomized, and participants were requested to rate the surface property of the planes shown in each animation using a 5-points semantic-differential scale with the following three bipolar dimensions:

- Roughness: high values represent rough, and low values represent smooth.
- Glossiness: high values represent shiny, and low values represent matte.
- Hardness: high values represent hard, and low values represent soft.

Procedure

Following written instructions, participants first completed the paired-comparison task. Each trial comprised of a pair of animated balls bouncing on a plane, one bounce on a rough plane, and one bounce on a smooth plane. The task was to compare the bounce heights of the two bounces in each stimulus pair, and to indicate which one of them was higher by pressing one of the response keys. All participants were instructed to respond as quickly and accurately as possible.[1] On each trial, a fixation cross was presented for 500 ms, followed by the first stimulus for duration of approximately 1000 ms. The inter-stimulus interval duration was 500 ms. The second stimulus was then presented for a duration of 1000 ms. The inter-trial interval duration was 3000 ms. Participants could respond following the completion of the second stimulus. After completing the paired-comparison task, the participants rated the appearance of the plane of the same 10 animations, one at a time, using the three rating scales provided. On average participants took 60 minutes to complete the experimental session. The experimental layout is seen in figure 3.

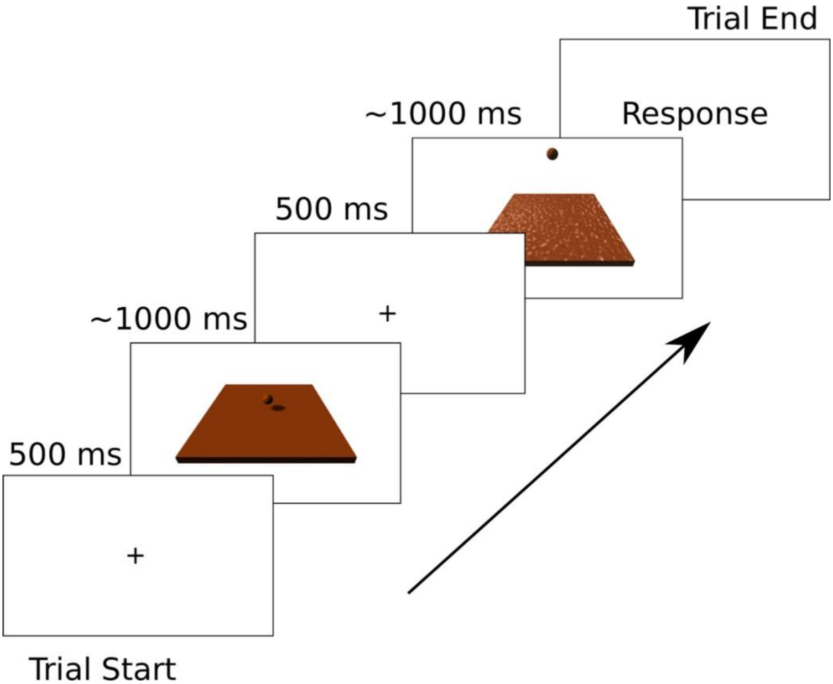

Figure 3. The experimental layout for the paired-comparison task. On each trial, the participants were presented sequentially with pairs of bounce heights presented with either smooth or rough surface plane. The task was to judge, which of the stimulus, the first presented bounce height or the second, contained a higher bounce height. After completing the paired comparison task, the participants judged the material properties, roughness, glossiness, and hardness, of the surface plane from the 10 animations using a 5-points semantic-differential scale.

[1] Response time was measured but since the data showed a similar pattern of results as logit P, and to avoid obfuscating the analysis and obscuring the main findings of the study, these data are not reported in this paper.

Data analysis

The analysis was conducted in three parts. First, the binary responses obtained in the paired-comparison task were analyzed following procedures described by Patching et al. (2012), in which the overall effect of surface appearance on judgments of bounce height was presented in terms of log odds of the proportion, *P*, of choosing the first stimulus to have higher bounce height compared to the second stimulus: logit $P = \log_e [P/(1 - P)]$. The choices were analyzed using General Linear Mixed Effects Model (GLMM) with a binomial distribution and a logistic link, which allowed us to examine judged bounce height as a binary response, while taking into account the repeated measures of each participant. Model evaluation was based on Akaike Information Criterion (AIC), where a full regression model was compared to nested models using likelihood ratio tests. A model with the lowest complexity, deviance, and AIC score, was judged to be the most suitable model to describe the data, and was therefore selected for further regression analysis.

Second, the metric effect of judged bounce height was calculated using the following procedures. For each stimulus presentation order [rough surface plane first and smooth surface plane second, vs. smooth surface plane first and rough surface plane second] we obtained the predicted probabilities for each bounce height by conducting a GLMM analysis on the log odds of choosing the first stimulus to have higher bounce height, using the intra-pair difference between the stimuli (-6 mm, 0 mm, +6 mm) as the main fixed effect, with random intercept and slope for the effect of intra-pair difference. Next, we regressed the predicted probabilities on the five stimuli bounce heights (ranging from 31mm to 43mm) to find the estimated metric effect for their judged bounce height.

Finally, we analyzed how the participants rated the surface planes using the three rating scales, roughness, hardness, and glossiness, by averaging the ratings scores and calculating the standard errors for each stimulus. For the purpose of assessing how much the participants agreed on the appearance of the surface planes, we examined the correlation between the rating scores among the participants, in other words the inter-rater reliability of the subjective ratings. We averaged the subjective ratings in order to have one rating score per surface texture (rough vs. smooth) and participant. Then we calculated the intra-class correlation coefficients for each rated property: roughness, hardness, and glossiness, and ran an additional F-tests. Following Shrout and Fleiss (1979), a poor reliability was considered anything less than 0.5 ICC, a moderate reliability were values between 0.5 and 0.75 ICC, a good reliability were values between 0.75 and 0.9 ICC, and everything above 0.9 ICC stood for excellent reliability (Provoost, et al., 2019).

All analyses were conducted in the R environment (R Core Team, 2013), using the *lme4* package (Bates et al., 2015) for the mixed modeling, and the *psych* package (Revelle, 2018) for the inter-rater reliability.

Results and Discussion (Experiment 1)

The binary choices from the paired comparison task were modeled as the log odds of choosing the first bounce height to be higher than the second bounce height using General Linear Mixed Effects Model (GLMM) with a binomial distribution and a logistic link. Five models with different variations of fixed effects (dummy coded) were evaluated. All models had by-participant intercepts and random slopes for the fixed effects of interest. The results from the model evaluation are shown in Table 1. A model with the lowest AIC score was chosen for further analysis, which was a model with *surface texture* [first bounce height on a smooth surface vs. first bounce height on a rough surface] and *average bounce height* [34 mm, 37 mm, 40 mm] entered as fixed main effects (AIC = 11819), and with by-participant intercept and random slope for surface texture and bounce height, as random effects. As can be seen in Table 1, the inclusion of *response assignment* [left key for the first bounce height and right key for the second bounce height, vs. the reversed order] as an effect in model 1 reduced the model fit and was therefore left out of the final model (model 2). Moreover, having an interaction between surface texture and average bounce height, as seen in model 5, did not improve the model fit either, and was therefore excluded from the final model as well.

Table 1. *The log likelihood, deviance, degrees of freedom, and Akaike Information Criterion scores (AIC) for each evaluated model of judged bounce height.*

Model ID	Model	Log Likelihood	Deviance	DF	AIC	Comparison
1	Keys + Surface + Bounce	-5895.6	11791	15	11821	
2	Surface + Bounce	-5895.6	11791	14	11819	mod1, mod2
3	Surface	-5897.9	11796	12	11820	mod2, mod3
4	Bounce	-5899.2	11798	13	11824	mod2, mod4
5	Surface + Bounce + Surface:Bounce	-5894.5	11789	16	11821	mod2, mod5

Note. Regression model number 2, which included surface texture and average bounce height as fixed main effects, was the best fitted model for our analysis.

A graphical presentation of the effect of surface texture over average bounce height, revealed the magnitude and direction of log odds of judging the first bounce height to be higher than the second, varied with changes in both appearance of the surface planes and bounce height. Figure 4a, shows the log odds (logit P) to go from negative to positive with increased bounce height. Indicating that physically higher bounce heights were judged to be higher than physically lower bounce heights. Moreover, both presentation orders were found to have negative log odds, in which the log odds for the presentation order, *smooth surface first and rough surface second*, are larger than the reversed presentation order, *rough surface first and smooth surface second*, indicating an underestimation of the first presented stimulus and a preference for choosing rough surface planes for higher bounce heights.

Using GLMM analysis on the final model (model 2) revealed both surface texture and average bounce height had significant effects on the judgment of bounce height. The probability of choosing the first bounce height to be higher than the second, were significantly larger when the bounce was presented on a rough surface plane ($b = 0.13$, $SE = 0.05$, $Z = 2.76$, $p < 0.01$), and increased with increasing bounce height, in which the higher average bounce height (40 mm) was found significantly different than the remaining average bounce heights, when keeping the 34 mm bounce height as the reference level ($b = 0.11$, $SE = 0.06$, $Z = 1.99$, $p < 0.05$). The predicted probabilities for choosing the first stimulus to have higher bounce height were larger for bounce heights presented with rough surface planes (34mm = 46%, 37 = 45%, 40 = 47%), compared to bounce heights presented with smooth surface planes (34mm = 40%, 37 = 42%, 40 = 45%). Figure 4b shows the estimated coefficients for model 2.

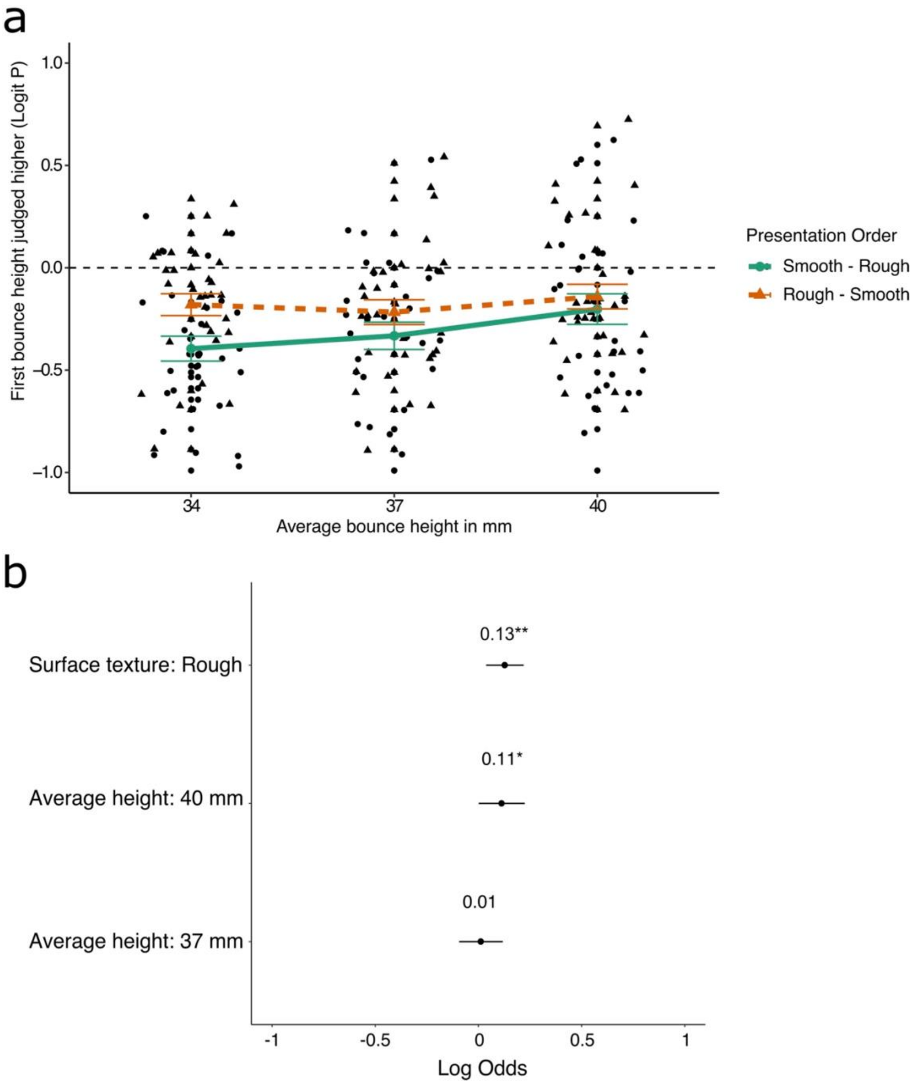

Figure 4. a) The Logit P (log odds) of judging the first presented bounce height to be higher than the second, over average bounce height for the two presentation orders [smooth surface first and rough surface second vs. rough surface first and smooth surface second]. The graph includes individual observations displayed using jittering to prevent overplotting. b) The estimated coefficients for the fixed effects *surface texture* and *average bounce height*, ordered from the largest (top) to the smallest (bottom). The log odds of judging the first bounce height to be higher than the second, significantly increased, both when the bounce was presented on rough surface texture, and when the physical bounce height was higher in mm.

The overall estimated metric effect on judged bounce height based on the predicted probabilities (y) of judgement of physical bounce height of the first presented stimulus (x) for the two surface textures was on average 1 mm higher for rough surface planes ($y = -2.07 + 0.07x$, estimated metric: 31mm, 33mm, 36mm, 39mm, 42mm, M=36.2mm), compared to smooth surface planes ($y = -2.02 + 0.07x$, estimated metric: 30mm, 32mm, 35mm, 38mm, 41mm, *M*=35.2mm). Although the participants estimated bounce heights on rough surface planes to be higher in mm compared to the bounce heights on smooth surface planes, the estimated metric effect for the judged bounce height was lower for the both surface textures compared to the actual bounce heights (31mm, 34mm, 37mm, 40mm, 43mm) by 1mm for rough surface planes and 2mm for smooth surface planes.

For the subjective ratings given by the participants, we found a reasonable correspondence between the intended and the rated surface appearances. On the average, the rough surface plane was rated rougher (M = 3.87, SD = 0.61) than the smooth surface plane (M = 1.63, SD = 0.73), and although both surface planes were rated matte, the rough plane got rated less shiny (M = 2.33, SD = 0.62) than the smooth plane (M = 2.77, SD = 1.08). Moreover, both surface planes were rated to have hard qualities, and the hardness ratings increased with increasing bounce height (see figure 5). Interestingly, smooth surface plane was rated with harder property (M = 3.68, SD = 0.69) than the rough surface plane (M = 3.26, SD = 0.65).

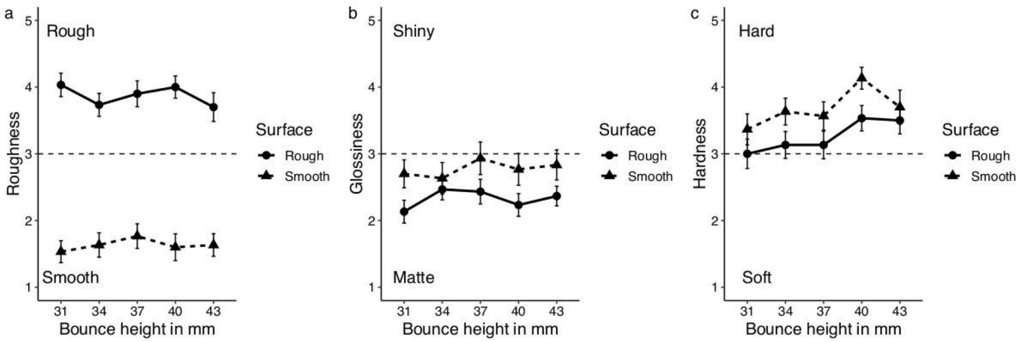

Figure 5. The mean rating scores and standard errors of the rating scales (a) *roughness, (b) glossiness,* and c) *hardness*, for the five bounce heights presented with either rough or smooth surface plane. Graph c) demonstrates an increment in rated hardness that follows the increasing bounce height.

The inter-rater reliability analysis on the subjective ratings for the two types of surface planes, revealed an excellent consistency among the participants for rated roughness, in which the ICC was 0.98 with 95% confidence interval ranging from 0.90 to 1, ($F(1,29) = 61$, $p < 0.001$, representing an excellent reliability for rated roughness. An excellent consistency was also found for rated hardness, with ICC of 0.91, and 95% confidence interval ranging from 0.60 to 1, ($F(1,29) = 18$, $p < 0.001$. Suggesting that the hardness estimation is based on distinctive surface appearance of the planes, rather than then the bounce alone. To test this we examined the average bounce height as a single factor influencing the hardness ratings and found no significant consistency between the participants for rated hardness, ICC = 0.24, with 95% confidence interval ranging from 0.11 to 0.65, ($F(1,29) = 1.39$, p = n.s. Finally, no significant inter-rater reliability was found for rated glossiness, indicating a low consistency among the participants for rated glossiness, ICC = - 0.01 with 95% confidence interval ranging from - 0.09 to 0.07, ($F(1,29) = 0.71$, p = n.s.

Experiment 2: Glossiness

Experiment 1 revealed the material properties roughness and smoothness influenced the perception of bounce height, where participants paired higher bounce heights with rough surfaces. In Experiment 2, we examined the perception of bounce height using surface planes with specular highlights. The aim was to see if the specular properties diminished or enhanced previously found effect for rough surfaces.

Methods (Experiment 2)

Participants

Thirty right-handed participants (17 females and 13 males) recruited on Lund University campus (*Mean* age = 25 years, *SD* = 3 years) took part in the study. Participants gave their consent to participate in the experiment. All participants reported normal or corrected-to-normal vision.

Stimuli, Design, and Procedure

The experimental setup for experiment 2 was identical to that described for Experiment 1 apart from the changes in the visual appearance of the surface planes. Instead of the matte appearance, specular properties were added to the surface planes, on which the ball bounced. The specular highlights were created using the same procedures as for the surface planes as used in Experiment 1, except here the Lambertian diffuse shader was replaced with a Blinn specular shader. This is a physically based specular model that creates bright highlights for shiny appearances to simulate materials such as glass, water and plastics. The Blinn shader was set to have white specular color at the intensity of 0.5 and hardness set to 1. The goal was to not imitate the characteristics of any specific material category, but rather to simulate an intermediate glossiness, which simulates a typical floor appearance (see Figure 6a) and add it to the pre-existing surface textures (smooth and rough) (see Figure 6b). For physical comparison, a spectral analysis was conducted on the four rendered images used for the surface planes (see Figure 6c). The experimental procedures and data analysis were otherwise the same as in Experiment 1.

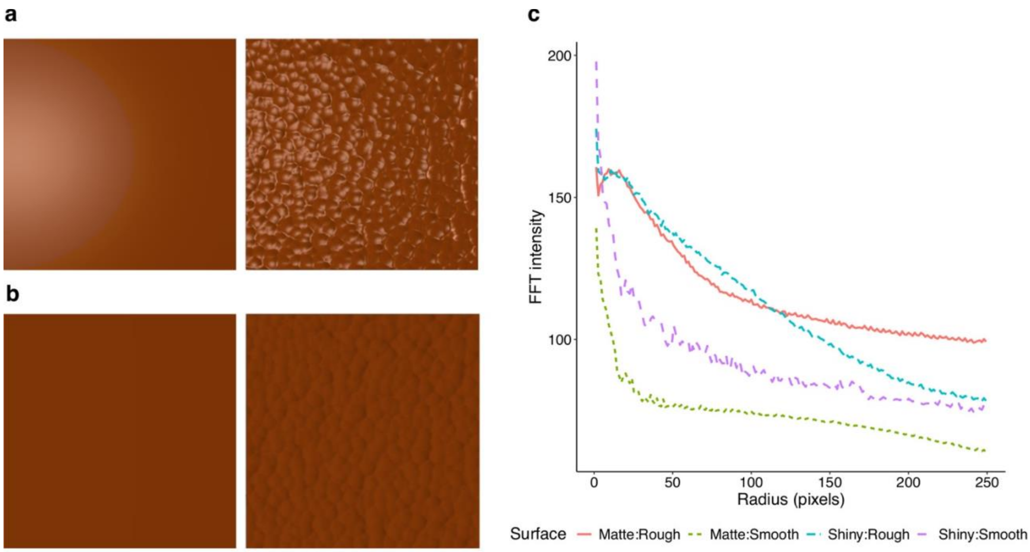

Figure 6. a) The visual appearance of the surface planes with added specular highlights, smooth (left) and rough (right). b) The matte surface planes used in Experiment 1, smooth (left) and rough (right) textures. c) A spectral graph for the four types of surface planes used in experiment 1 and 2 computed using radial transform Fourier analysis. The x-axis shows the radius measured as pixels, and the y-axis shows the normalized integrated fast Fourier transformation (FFT) intensity. The graph demonstrates the intensity difference between the four images, especially between the matte and shiny surface planes.

Results and Discussion (Experiment 2)

Similar to experiment 1 a model comparison based on the AIC weights, revealed that *response assignment* and the interaction term between surface texture and average bounce height, reduced the fit of the models. They were therefore not included in the final model (see table 2). Based on the Akaike Information Criterion score and the Deviance score, the most suitable model for our analysis, was a model with *surface texture* and *average bounce height* entered as dummy coded fixed effects, and with by-participant intercept and random slope for surface texture and bounce height as random effects (model 2 in table 2).

Table 2. *The Akaike Information Criterion scores (AIC) for models on judged bounce height presented with shiny surfaces.*

Model ID	Model	Log Likelihood	Deviance	DF	AIC	Comparison
1	Keys + Surface + Bounce	-5944.7	11889	15	11919	
2	Surface + Bounce	-5921.6	11843	14	11871	mod1, mod2
3	Surface	-5922.8	11846	12	11870	mod2, mod3
4	Bounce	-5953.6	11907	13	11933	mod2, mod4
5	Surface + Bounce + Surface:Bounce	-5920.9	11842	16	11874	mod2, mod5

Note. The table shows the log likelihood, deviance, degrees of freedom, and AIC scores, for each evaluated model of judged bounce height. Regression model number 2, which included surface texture and average bounce height as fixed effects, was the best fitted model for our analysis.

A graphical presentation of the effect of surface texture over average bounce height, revealed the magnitude and direction of log odds of judging the first bounce height to be higher than the second varied with changes in both appearance of the surface planes and bounce height. Figure 7a, shows the magnitudes and the direction of the log odds (logit P) to go from negative to positive with increased bounce height, indicating that the first bounce height is judged to be higher when it was physically higher than the second bounce heights. Moreover, the log odds for the presentation order, *smooth surface first and rough surface second*, have negative values that are larger than the reversed presentation order [*rough surface plane first and smooth surface plane second*], which indicates a preference for rough surface planes when choosing higher bounce heights.

The effect of the surface appearance of the planes on the judgment of bounce height was assessed in terms of log odds of choosing the first stimulus to have higher bounce height than the second stimulus. The binomial GLMM analysis showed the coefficients for the log odds to be larger for bounce heights presented with rough textured surface plane (b = 0.16, SE = 0.04, p < 0.001), compared to smooth surface plane. Interestingly, average bounce heights, using the 34 mm bounce height as the reference level, were found to have non-significant effect on the log odds (for 37 mm, b = 0.00, SE = 0.06, p = n.s.; for 40 mm, b = 0.08, SE = 0.06, p = n.s.). The predicted probabilities for choosing the first stimulus to have higher bounce height were the following (rough first and smooth second: 34mm = 50%, 37 = 48%, 40 = 51%), compared to bounce heights presented with smooth surface planes (smooth first and rough second: 34mm = 45%, 37 = 46%, 40 = 47%). Moreover, in order to create a contrast between the average mid-height (37 mm) and the average highest bounce height (40 mm), we changed the reference level of the average bounce height from 34 mm to 40 mm. Such changes did not result in significant difference between the estimated coefficients, as both contrasts were found non-significant (for 34 mm bounce height, b = -0.08, SE = 0.06, p = n.s.; and for 37 mm bounce height, b = - 0.08, SE = 0.06, p = n.s). The estimated coefficients for the initial model are represented in Figure 7b.

Based on the predicted probabilities (y) of judgement of physical bounce height of the first presented stimulus (x), the overall estimated metric effect on judged bounce height for rough surface planes (y = -2.18 + 0.07x) was 0.6 mm higher (estimated metric: 32mm, 35mm, 38mm, 41mm, 44mm, M= 38mm), compared to the estimated metrics for smooth surface planes (y = -2.16 + 0.07x, estimated metric: 32mm, 34mm, 37mm, 40mm, 44mm, M= 37.4mm). Overall, judged bounce height on rough surface planes was estimated to be 1mm higher than the actual stimulus bounce height (stimulus bounce height: 31mm, 34mm, 37mm, 40mm, 43mm, M=37mm), whereas judged bounce height on smooth surface planes was estimated to be 0.4mm higher than the actual bounce height.

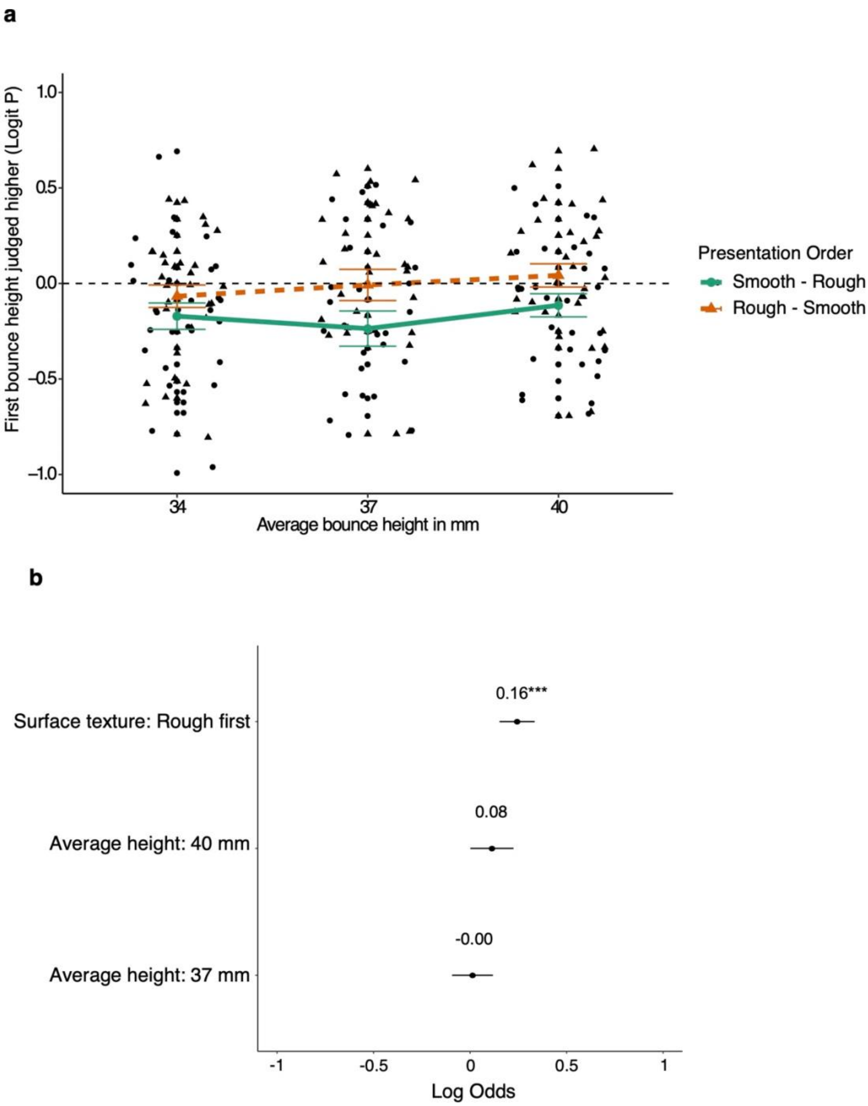

Figure 7. a) The effect of surface texture over average bounce height in terms of log odds (logit P) of judging the first bounce height to be higher than the second bounce height. b) The estimated coefficients for the fixed effects surface texture and average bounce height for specular surface planes, ordered from the largest (top) to the lowest. The log odds of judging the first bounce height to be higher significantly increased when the bounce was presented together with rough surface plane. Unlike the results from Experiment 1, average bounce height was found non-significant.

For the subjective ratings given by the participants, we found a reasonable correspondence between the surface appearances and the ratings. The participants rated the rough surface texture to be rougher (*M* = 3.92, *SD* = 0.85) and less shiny (*M* = 3.14, *SD* = 0.99) than the smooth surface texture (Roughness: *M* = 1.81, *SD* = 0.91; Glossiness: *M* = 3.45, *SD* = 1.08). Both surface textures were, however, rated shiny on the average. For the hardness ratings we found the two surface textures to be rated very similar, where both surface textures were rated to have hard quality (rough surface texture: *M* = 3.26, *SD* = 0.65; smooth surface texture *M* = 3.26, *SD* = 0.65). Moreover, we found the hardness ratings of the surface planes to be influenced by the absolute bounce height, as the hardness ratings followed the increasing bounce height (see figure 8).

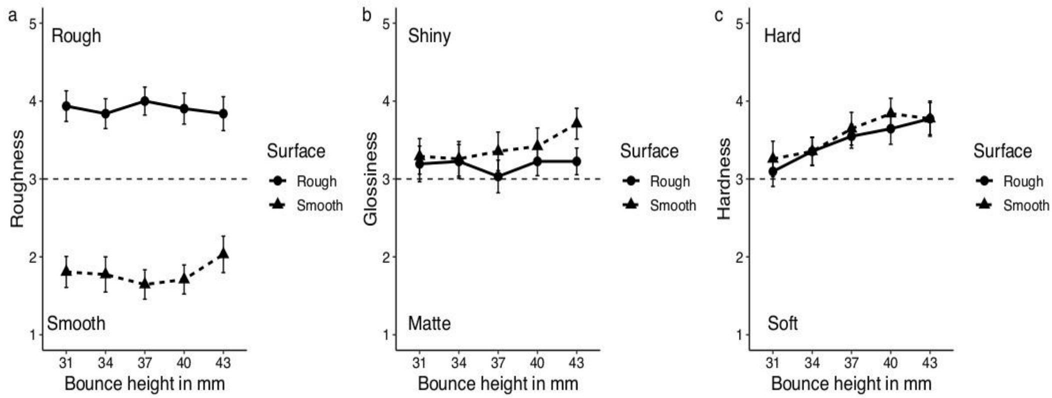

Figure 8. The mean rating scores and standard errors of the rating scales a) *roughness, b) glossiness,* and c) *hardness* for the 10 surface planes with added specular highlights and either rough surface texture or no texture (smooth appearance). A correspondence between the surface texture and rated roughness (a) is seen, including an increment in rated hardness in line with increasing bounce height (c). Less distinction is found between the two types of surface planes and rated glossiness (b) as well as rated hardness (c).

We found excellent inter-rater reliability among the participants for rated roughness with the average ICC of 0.98 with 95% confidence interval ranging from 0.88 to 1 ICC, ($F(1,29) = 43$, $p < 0.001$. A good reliability was found for rated glossiness, 0.88 ICC with 95% confidence interval ranging from 0.37 to 1 ICC, ($F(1,29) = 8.4$, $p < 0.001$, but no significant reliability was found for rated hardness, -0.02 ICC with 95% confidence interval ranging from -0.1 to 0.07 ICC, ($F(1,29) = 0.05$, p = n.s. Keeping the bounce height as the single predictive variable in the analysis did not improve the inter-rater reliability for rated hardness, ICC = -0.01, with 95% confidence interval ranging from -0.07 to 0.04 ICC, ($F(1,29) = 0.49$, p = n.s.

General Discussions

While moving around as active perceivers we learn about the physical properties of the world and how to interact with them. We learn that a ball will bounce higher on a cement ground compared to a carpeted floor, and we create associations between the typical appearances of those grounds and their intrinsic properties such as hardness quality. A knowledge we rely on when estimating the intrinsic properties of new materials. In the current study we examined how changes in the visual appearance of a surface plane resulted in systematic changes in judged bounce height of an object. It was theorized that participants would judge bounce height differently depending on the surface features of the plane on which the ball bounced, due to generic knowledge of materials obtained from earlier encounters with similarly looking surfaces.

In experiment 1 we tested matte surface planes with rough and smooth textures. The best fitted model to describe our data from the paired-comparison task, was a model with *surface texture* and *average bounce height* as fixed effects. A GLMM analysis revealed judgment of bounce height to be significantly affected by the two fixed effects, in which we found increased probability of perceiving a higher bounce when it was presented on rough surface planes compared to bounces presented on smooth surface planes. This suggests that the participants were more likely to perceive a higher bounce height on a rough surface plane. The estimated metrics of judged bounce height support this, although the participants judged bounce heights on both surface textures to be lower than the actual stimulus bounce height.

In terms of the subjective ratings, a correspondence was found between the intended appearance of the surface planes and rated roughness and glossiness. Interestingly, a high consistency between the different participant ratings was found for rated hardness, suggesting that the rated hardness of the visual appearance of the matte surface planes is somewhat objective. Based on the results from the paired-comparison task, it was expected that the participants would rate the rough surface plane to have harder quality than the smooth surface plane. Since, the participants chose higher bounce heights with rough surface planes more frequently than they chose the smooth surface planes. However, in contrast to what was expected, the participants rated the smooth surface planes to have harder quality than the rough surface planes. One possible explanation for such difference, could lie in the context of the two tasks and what was expected of the participants. In the paired comparison task, the participants were expected to focus on the bounce height without any explicit instructions to pay attention to the characteristics of the surface plane, whereas in the rating task, the participants were asked to observe the surface plane and rate it. Regardless, both surface planes were rated with hard qualities on the average, suggesting the planes were perceived to have the intrinsic properties required to afford the ball to bounce on it.

In experiment 2 we tested the same surface planes as in Experiment 1, but with the addition of specular highlights. This was done to examine whether the effect found for surface texture in Experiment 1 was either enhanced or decreased if specular properties were added to the surface planes, creating a qualitatively different material. Using similar method as used in Experiment 1 to analyze the data, we found that the participants estimated the bounce height to be higher on the rough surface texture. Nevertheless, the log odds as a function of average bounce height varied much less in comparison to the results obtained in Experiment 1. Moreover, the GLMM analysis revealed that only surface texture had a significant effect on the judgment of bounce height, where the average bounce height was found to be non-significant. Indicating that each average bounce height level (34mm, 37mm, 40mm) was judged equally likely to have higher bounce height. Furthermore, removing average bounce height as a fixed effect from the final model improved the model fit in terms of AIC score. Although the participants favored rough surface plane when choosing higher bounce height in the paired-comparison task. It appears that the participants disagreed on how soft or hard the surface planes were, as the inter-rater reliability for rated hardness was found non-significant for specular surface planes. This suggests that hardness assessment is a very subjective property when judging surfaces with specular appearance. Regardless, on the average the participants judged the shiny surfaces to have hard qualities. Moreover, the estimated metrics of judged bounce height revealed participants to overestimate bounce heights presented on specular surfaces, especially bounce heights presented with rough surface planes. In both experiments, the participants judged bounce heights on rough surface planes to be higher in mm compared to bounce heights on smooth surface planes. However, the participants underestimated the bounce heights presented with matte surfaces, whereas bounce heights presented with shiny surface planes were overestimated, that is estimated to be higher than the actual bounce heights. Suggesting a relationship between surface glossiness and perceived bounce height, where shiny surfaces ‘afford’ higher bounce heights compared to matte surfaces.

Overall, the results demonstrate an involvement of learned associations between the distinctive appearances of materials and their functional qualities in assessment of bounce height. In both experiments, the participants assigned a hard quality to all of the surface planes, suggesting a relationship between the physical appearance of the surfaces and their intrinsic properties. A similar relationship has been demonstrated before by Vrins, Wit and van Lie (2009), who showed that the generic knowledge of physical properties of materials can influence whether a participant sees an object as a whole or not when partially occluded. Using an amodal completion task, Vrins et al., found that when participants were presented briefly with a round object with the appearance of either soft or hard material (e.g., a cucumber slice or a metal disc), and partially attached to a block with the appearance of either soft or hard material (e.g., a butter or a brick), participants reported a completed circular slice only if the slice was penetrated into a soft block (butter) as opposed to a hard block (brick). Indicating that the perceived completion was shaped by the knowledge that both the metallic disc and the cucumber slice would remain whole if pushed into a soft butter, but not if pushed into a brick. Similar to cognitive stereotypes, where judgments are based on the notion that similarly looking things are likely to yield similar behavior (Buckingham et al., 2009; Macrae, Milne & Bodenhausen, 1994), we anticipate the intrinsic characteristics of a material from its signature visual characteristics, based on a pool of previously experienced similar examples.

In both experiments, the hardness ratings followed the increment of the bounce height in a step-like manner, where surface planes with higher bounce heights got rated harder than the surface planes with physically lower bounce heights. Regardless, based on the performance in the paired comparison task, we would have expected surface planes with rough texture to be rated with harder quality than surface planes with smooth texture. On the contrary, for matte surface planes the average subjective rating for hardness was larger for smooth surfaces compared to rough surface planes, whereas the surface planes with specular appearance were rated similar despite their difference in roughness. Moreover, we found a high consistency between the different participants ratings for the two matte surface textures, rough and smooth, which indicates that hardness is somewhat an objective material property. At least when it comes to matte surfaces, as low consistency was found for the shiny surfaces. A high consistency between participants in rated hardness has been reported before by Fleming, et al. (2013), who also found negative correlations between rated roughness and glossiness, similar to our findings, when judging images of various material classes. Like in our study, their participants rated rough surfaces with less glossiness. Such negative correlation between roughness and glossiness is sensible given that when a light hits a rough surface it reflects in many directions, creating a diffused (matter) reflection.

There are few items and challenges to consider for a follow-up analysis. In the current study, we simulated surfaces using textures and appearances commonly used for simulating floor covering and ground surfaces, in particular the material properties roughness and glossiness. An interesting follow up would be to include other common attributes such as elasticity, slipperiness, and compliance. Especially slipperiness, since surface friction has been reported to influence glossiness perception, in which surfaces with low friction are perceived to be shinier (Adams, Kerrigan, & Graf, 2016). Moreover, it would be interesting to look into well-defined material categories of surfaces (e.g., grass, sand, concrete, wood, carpet, and steel), to see if similar effect on bounce perception is achieved using higher-level categories of materials. In our experiments we kept the bounce physics (e.g., gravity, mass, angular drag, and friction) constant for all of the four surface planes tested, besides the bounciness settings, which was varied in five steps in order to achieve the five bounce heights for our stimuli. Changing the physics settings in future experiments would be an interesting addition to our work, as it could tell us more about perceived kinematics of a bounce, and how such variables are affected by the context the bounce is presented in.

In summary, we found that visually different surface planes induced different perception of bounce height. We have shown that the hardness of a material can be visually assessed when observing a familiar object bouncing on it. Moreover, we have shown that the perceived bounce height is different depending on the distinctive characteristics of the surface plane. Overall, rough surface planes were chosen more frequently compared to smooth surface planes when choosing higher bounce heights. In the second experiment we found that the addition of specular highlights enhanced this effect, but resulted in smaller difference between the two surface textures, rough and smooth, as well as lower consistency between the participants in perceived hardness. We conclude that the visual assessment of bounce height is not only based on the physical properties of the bounce itself, but it also incorporates generic knowledge of what materials look like and what they afford.

Acknowledgements

The authors are thankful to Lund University Humanites Lab, Rasmus Bååth, and Geoffrey Roy Patching for helpful discussion.

References

Adams, W. J., Kerrigan, I. S., & Graf, E. W. (2016). Touch influences perceived gloss. *Scientific reports, 6*, 21866. doi:10.1038/srep21866

Adelson, E. H. (2001). On seeing stuff: the perception of materials by humans and machines. In B. E. Rogowitz, & T. N. Pappas, (Eds.), *Human Vision and Electronic Imaging VI. Proceedings of the SPIE,* 4299 (pp. 1- 22).

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

Arnott S. R. Cant J. S. Dutton G. N. Goodale M. A. (2008). Crinkling and crumpling: An auditory fMRI study of material properties. *NeuroImage*, *43*(2), 368–378.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Baumgartner, E., Wiebel, C. B., & Gegenfurtner, K. R. (2013). Visual and haptic representations of material properties. *Multisensory Research, 26*(5), 429 – 455.

Beck, J., & Prazdny, S. (1981). Highlights and the perception of glossiness. *Attention, Perception, & Psychophysics, 30*(4), 407-410.

Bergmann Tiest, W. M., & Kappers, A. M. L. (2007). Haptic and visual perception of roughness. *Acta Psychologica, 124*(2), 177 – 189.

Bergmann Tiest, M. W., & Kappers, A. M. L. (2009). Cues for haptic perception of compliance. *IEEE Transaction on Haptics, 2*(4), 189 – 199.

Bergmann Tiest, W. M. (2010). Tactual perception of material properties. *Vision Research, 50*(24), 2775 – 2782.

Blender Foundation (2013). 3D modeling and rendering package (Version 2.65) [Computer software]. Amsterdam: Netherland, Blender Institute. Available from: http://www.blender.org/

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 443 - 446.

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,* 3111 – 3118.

Charpentier, A. (1891). Analyse expérimentale: De quelques éléments de la sensation de poids [Experimental study of some aspects of weight perception]. *Archives de Physiologie Normale et Pathologique, 3*, 122 – 135.

Ellis, R. R. & Lederman, S. J. (1999). The material-weight illusion revisited. *Perception & Psychophysics, 61*(8), 1564 – 1576.

Engel, C. (2007). Blender Open Material Repository [Online database]. Frankfurt: Germany. *Basket Ball* [Blender file]. Available from: http://matrep.parastudios.de/index.php?action=view&material=305-basket-ball

Flanagan, J. R., & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size-weight illusion. *Nature Neuroscience, 3*(7), 737-741. Doi: 10.1038/76701

Fleming, R., Wiebel, C. & Gegenfurtner, K. R. (2013). Perceptual qualities and material classes. *Journal of Vision, 13*(8), 1 – 20. Doi:10.1167/13.8.9

Giordano B. L. McAdams S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *Journal of the Acoustical Society of America*, *119*(2), 1171–1181.

Harshfield, S. P., & DeHardt, D. C. (1970). Weight judgment as a function of apparent density of objects. *Psychonomic Science, 20*, 365–366.

Hellström, Å. (1979). Time errors and differential sensation weighting. *Journal of Experimental Psychology: Human Perception & Performance, 5*, 460-477.

Hellström, Å. (2003). Comparison is not just subtraction: effects of time- and space-order on subjective stimulus difference. *Perception & Psychophysics, 65,* 1161-1177. Doi:10.3758/BF03194842

Imura, T., Yamaguchi, M. K., Kanazawa, S., Shirai, N., Otsuka, Y., Tomonaga, M., & Yagi, A. (2006). Perception of motion trajectory of object from the moving cast shadow in infants. *Vision Research, 46*(5), 652 – 657.

Kawachi, Y., Kawabe, T., & Gyoba, J. (2011). Stream/bounce event perception reveals a temporal limit of motion correspondence based on surface feature over space and time. *i-Perception*, *2*(5), 428–439. DOI:10.1068/i0399

Kersten, D., Knill, D. C., Mamassian, P., & Bülthoff, I. (1996). Illusory motion from shadows. *Nature, 379*, 6560 – 6531.

Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadow induce apparent motion in depth. *Perception, 26,* 171 – 192.

Macrae, C. N., Milne, A. B., & Bodenhausen, G.V. (1994). Stereotypes as energy-saving devices: a peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66,* 37-47.

Mamassian, P., Knill, D. C., & Kersten, D. (1998). The perception of cast shadows. *Trends in Cognitive Science (Regul. Ed.), 2,* 288 – 295.

Marlow, P. J., Kim, J., & Anderson, B. L. (2012). The perception and misperception of specular reflectance. *Current Biology, 22,* 1909 – 1913.

Marlow, P. J., & Anderson, B. L. (2013). Generative constraints on image cues for perceived gloss. *Journal of Vision, 13*(14), 2. doi: 10.1167/13.14.2.

NVIDIA Corporation (2013). PhysX (Version 3.3) [Computer software]. Santa Clara, CA: USA. Available from: http://www.nvidia.com/

Patching, G.R., Englund, M.P., & Hellström, Å. (2012). Time- and space-norder effects in timed discrimination of brightness and size of paired visual stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 915-940.

Provoost, S., Ruwaard, J., van Breda, W., Riper, H., and Bosse, T. (2019). Automated Sentiment Analysis of Online Cognitive Behavioral Therapy Patient Texts: An Exploratory Study. *Frontiers in Psychology, 10*, 1065. DOI: 10.3389/fpsyg.2019.01065

R Core Team (2013). R: foundation for statistical computing (Version 3.1.0) [Computer software]. Vienna: Austria. Available from: http://www. r-project.org/

Revelle, W. (2018) psych: Procedures for Personality and Psychological Research (Version 1.8.12) [Computer software]. Northwestern University, Evanston, Illinois:USA. Available from: https://CRAN.R-project.org/package=psych.

Seashore, C. E. (1899). Some psychological statistics: II. The material weight illusion. University of Iowa Studies in Psychology, 2, 36 – 46.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428. DOI: 10.1037/0033-2909.86.2.420

Sugano, N., Kato, H. & Tachibana, K. (2003). The effects of shadow representation of virtual objects in augmented reality. *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality, 76- 83.*

The Mathworks Inc. (2013). MATLAB (Version 8.1.0.604) [Computer software]. Natick, Massachusetts: USA. Available from: http://se.mathworks.com/products/matlab/

Unity Technologies (2013). Game Creation System (Version 4.2.2f1) [Computer software]. San Francisco, CA: USA. Available from: http://www.blender3d.com/

van Assen, J. J. R., Barla, P., & Fleming, R. W. (2018). Visual features in the perception of liquids. *Current Biology, 28,*(3), 452–458.

Vickers, Z. M. (1981). Relationships of chewing sounds to judgments of crispness, crunchiness and hardness. *Journal of Food Science, 47,* 121 – 124.

Vickers Z. M. (1987). Crispness and crunchiness—Textural attributes with auditory components. In H. R. Moskowitz (ed.), *Food texture: Instrumental and sensory measurement* (45 – 66). New York: Marcel Dekker.

Vrins, S., de Wit, T. C., & van Lier, R. (2009). Bricks, butter, and slices of cucumber: Investigating semantic influences in amodal completion. *Perception, 38*, 17 – 29.

Wolfe, H. K. (1898). Some effects of size on judgments of weight. *Psychological Review, 5*(1), 25 – 54.

Worley, S. (2003). Cellular texturing. In D. S. Ebert, F. K. Musgrave, D. Peachy, K. Perlin, and S. Worley (eds.), *Texturing and Modeling: a Procedural Approach* (3rd edition)(135-157). USA: Elsevier Science.

Xiao, B., Bi, W., Jia, X., Wei, H., & Adelson, E. H. (2016). Can you see what you feel? Color and folding properties affect visual-tactile material discrimination of fabrics. *Journal of Vision, 16*(3), 1-15. doi: 10.1167/16.3.34

Yonas, A., & Granrud, C. E. (2006). Infants' perception of depth from cast shadows. *Perception & Psychophysics, 68*, 154 – 160.

Paper IV

Learning Material Properties from Visual and Proprioceptic Features in a Humanoid Robot

Kristín Ósk Ingvarsdóttir, Birger Johansson, Christian Balkenius

Lund University Cognitive Science

Abstract

We show how a humanoid robot can learn to estimate the weight and hardness of objects during manipulation and to use these features to categorize the objects. The robot uses 'proprioceptive' feedback from servos controlling the harm to measure both hardness and weight without using any sensors within the actual hand. The different objects are categorized in a self-organizing feature map using the combination of weight and hardness.

In addition to the proprioceptive sense, the robot uses a visual system to identify the the visual characteristics of objects as well as their identity. The visual processing of the objects consists of foreground segmentation, followed by color and texture analysis to produce a visual feature vector for each object. The proprioceptive and visual processing combine to produce multi-modal representations of the objects. This can be used to associate between the modalities and may also be used to reduce the amount of energy used when lifting lighter objects.

Our experiments with the robot show that the method can learn object properties from visual and manual exploration. The proposed method is one component in a system that learns object affordances.

1 Introduction

Already from a young age we learn about the physical properties of the world we live in and to incorporate that knowledge into our actions to perform daily tasks, such as reaching for a soft toy or a hard building block to play with. These motor actions appear effortless, but are nevertheless remarkable, considering the great diversity of physical properties that needs to be recognized in order to pick up the toys and handle them without mistakes. Moreover, toys and other objects are made from all kinds of materials (e.g. plastic, metals, fibers etc.) all of which have their own typical characteristics, both visual and physical, that influence how we interact with them.

Many physical properties that typically describe the characteristics of materials, such as roughness, weight and hardness, are normally perceived directly using touch. We actively explore the surface of objects to know how rough or smooth the objects are, and we weigh the objects using our hands to estimate their weight, and we test how compliant or hard the objects are by applying force to their surface (Bergmann Tiest, 2010; Lederman, 1974; Lederman and Klatzky, 1987; Okamoto et al., 2013). Usually before objects are touched to be interacted with, we visually explore the objects to locate them in space. We also use the provided visual information to estimate the size and the shape of the objects to prepare the grip size of our hands to hold the objects accurately (Johansson and Cole, 1992), and we estimate the weight of the objects based on their size and type of materials Saccone et al., 2019, and prepare the grip force that is required to manipulate the objects without slippage (Zoeller et al., 2019). Perceived hardness/softness is based on the compliance of the material the object is made from and is measured as how much an object deforms when force is applied to it (mm/N) (Lederman and Klatzky, 1987). For a study on perceived compliance see e.g. Zoeller et al. (2019). Hardness differs from other modalities in that it requires an active manipulation of the object to be perceivable.

Other material characteristics, such as surface gloss, are perceived using sight but may nevertheless convey tactile information (e.g. friction), see e.g.Chadwick and Kentridge (2015); Paulun et al. (2016), which is interesting considering that a number of studies have demonstrated similarities in material representations between the visual and the haptic domains, see e.g. (Baumgartner et al., 2013; Vardar et al., 2019). For instance, a high correspondence has been demonstrated between perceived softness based on tactile and visual information of compliance (Cavdan et al., 2021). Furthermore, a number of studies have demonstrated the involvement of implicit prior knowledge of the physical properties when planning interactions with object, in which the planning of the haptic exploration is guided by learned associations between the typical visual features of the objects and previous sensorimotor experience of similarly looking objects (Buckingham et al., 2009, 2011; Flanagan and Wing, 1997; Flanagan and Beltzner, 2000; Wolpert and Flanagan, 2001).

The ability to recognize physical properties of materials based on visual features is present at an early age. Paulus and Hauf (Paulus and Hauf, 2011) demonstrated a developmental progression of material knowledge in 9-, 11-, and 13- month old infants using a preferential reaching task, in which 11 month old infants were able to visually recognize previously preferred light object over a heavy object based on their visual features alone. More interestingly, the 13 month old infants were able to generalize the previously learned association between the material properties and visual characteristics, to novel objects made of same materials as the previously explored objects. That said, material knowledge is a complicated process that involves information from all of our senses as well as our memories of previous interactions with similar materials.

J. J. Gibson (1966; 1979) described the link between perception and action by defining the concept affordance, and argued that we as active agents perceive action possibilities (affordances) of objects directly through vision (Şahin et al., 2007). According to Gibson, the design of a chair affords the action *to sit* and the handle of a coffee cup affords the action *to hold*. Although Gibson's idea of the term affordance usually applies to the geometrical shape of an object, one could borrow the concept to describe perceived action possibilities based on other object properties than shape. After all, objects do not only come in various geometrical forms, but they also come in various types of materials that differ in physical properties, in which each property requires a specific type of handling. For instance, a fluffy cube with round edges made of synthetic fibres has soft physical properties and would afford the action *to squash*, whereas a similarly shaped wooden cube with sharp edges and hard physical properties would require a greater force to be deformed. Despite their similarities in geometrical shape, the perceived action possibilities are different due to their different material properties.

Several studies have investigated how robots can learn affordances by interacting with objects (Fitzpatrick et al., 2003; Gonçalves et al., 2014; Montesano et al., 2008; Stoytchev, 2005), how a robot could learn about tools use (Mar et al., 2015) and the dynamic properties of objects (Tikhanoff et al., 2013). Methododlogies includes Bayesian networks(Montesano et al., 2008), convolutional networks (Nguyen et al., 2016), and more recently, metric learning (Hjelm et al., 2019).

Although the shape of an object is usually the primary focus in studies of affordances, a number of techniques have been developed to allow humanoid robots to sense the hardness of manipulated objects. Matsuoka (1995) used competitive neural networks to learn the hard-

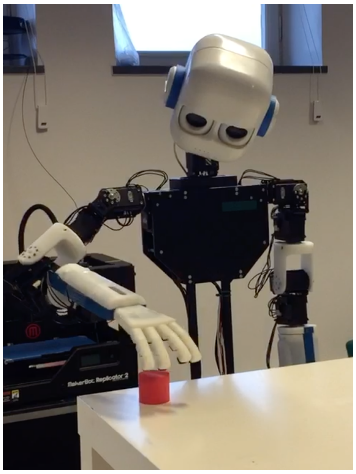

Figure 1: The robot Epi the picking up an object.

ness of objects from measurements made using force-sensitive resistors in combination with a potentiometer-based angle sensor in each finger. A similar idea was used to recognize both hardness and texture using self-organizing maps (Johnsson and Balkenius, 2008b,a, 2009). Regoli at al. (Regoli et al., 2017) also used this approach for the iCub that repeatedly squeezed an object to determine its hardness. This can be seen as exploratory movements that serve to obtain information about the object (Hoelscher et al., 2015). An alternative method is to use optical techniques to record the deformation of the finger as it touches an object (Yussof et al., 2008). Common to all these methods is that they look at the sensory signal over time as an object is touched and the change to the sensory signal over time reflects the hardness of the object.

The weight of an object is in principle is easier to determine. It can be measured by using piezoelectric sensors directly (Choi et al., 2006) or by using deflection sensors in the limbs (Sugaiwa et al., 2010).

Below, we show that both hardness and weight can be estimated without any specific sensors by using feedback from the servos involved in the manipulation of an object. Furthermore, we show how a robot can learn an associative mapping between visual features and proprioceptice features (hardness and weight). The proposed mechanism could constitute a part of an affordance learning system.

We approach the problem in two experiments. In the first, we test if the robot is able to measure weight and hardness using a small number of objects as a proof of concept. Based on the results of the first experiment we subsequently developed a system based on different learning mechanisms that is able to learn both weight and hardness dimensions and we tested the system in a systematic way with a larger number of objects. Finally, we designed a model of how the robot can learn a map between the different modalities.

The humanoid robot Epi

For the experiments, we used the humanoid robot Epi that has been developed at Lund University Cognitive Science (Johansson et al., 2020) (Fig. 1). It has two arms with five degrees of freedom each, three in the shoulder, one in the elbow, and one in the wrist. Each joint is controlled by a Dynamixal MX-106 servo that allows position control and produces a large number of feedback signals that can be read through a serial interface. These include, the

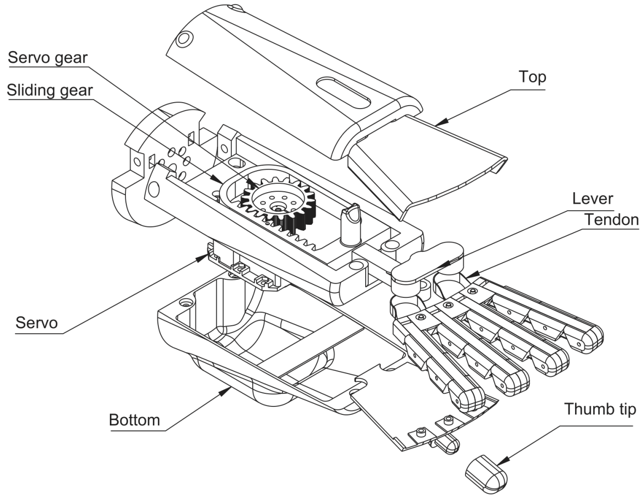

Figure 2: The design of the robot hand used in the experiments. The four movable fingers are controlled by a single servo through a gear rack that drags rubber tendon using a whippletree mechanism within the palm of the hand to distribute the force from the servo to the different fingers.

current position, the current used, temperature etc, and specifically the a value called 'load' that roughly corresponds to the current effort of the servo. Load has relatively low resolution, but by integrating over time, a better sensitivity can be obtained.

Epi has two hands with a single servo that controls all fingers except for the stationary thumb (Fig. 2). Each of the movable digits are controlled by an individual tendon made from 3D-printed polyurethan plastic. The force onto each of the tendons is controlled by a whippletree mechanism (Fukaya et al., 2000) within the palm of the hand that distributes the forces between the fingers and makes the fingers automatically grasp around objects.

The rubber design can withstand very strong forces and it is not possible to manually break the tendons by dragging or tearing. The joints continue into the tendon that then seamlessly connect to tendon of the next finger (Johansson et al., 2020).

The fingers have no dedicated touch sensors. Instead, the robot uses feedback from the servo controller to determine if the fingers are touching during a grasp.

The head contains two laterally moving eyes with HD-cameras. The head itself can turn sideways and tilt up and down.

The robot is controlled by the Ikaros[1] system that is an infrastructure for system-level brain modeling that also implements a number of analysis and control modules that were used in this study (Balkenius et al., 2010).

2 Experiment 1

In the first experiment feedback data from the hand and arm of the robot was analyzed while it manipulated different objects. We also recorded images of the objects from the robot eyes. The data was used to analyse the weight and hardness of the objects as well as their visual appearance.

[1]www.ikaros-project.org

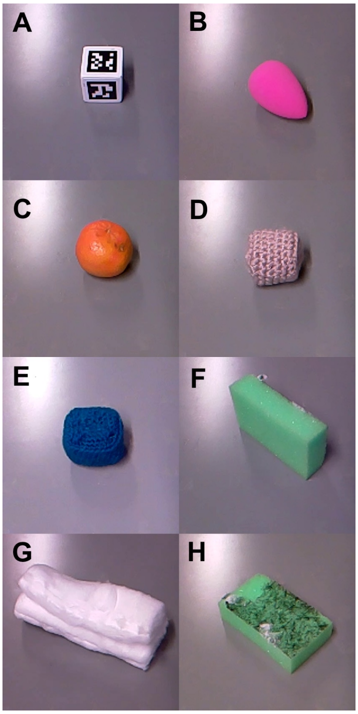

Figure 3: The objects used for the experiment as viewed by the robot. They differ in hardness, weight as well as color and texture.

Objects

Eight objects were used in the experiment (Fig. 3). They were selected based on their different visual appearance and different hardness. The objects from hardest to softest were (A) a wooden cube with a black and white texture, (B), a pink foam ‘egg’, (C) a clementine, (D) a crocheted light pink cube, (E) a crocheted blue cube, (F) a light green sponge, (G) a piece of white cotton, and (H) the light green sponge in another orientation.

Procedure

The robot performed the same motion for each object. The behavior was a pick-and-place operation where the robot first grasps an object and then lifts it to transport it to another location. The movement of the robot is controlled by the motion sequencer in Ikaros that produces a fixed behavior pattern. This behavior is controlled using position control with a resolution of 40 Hz to allow the arm to trace the desired trajectory. The objects that were manipulated were always located at the same location so no eye-hand coordination was necessary.

When the motion sequencer begins a new action, it triggers the modules that do hardness and weight estimation, to let them start their measurements (see below). It also triggers the visual system to make it store images of the objects from the left and right eyes of the robot before manipulation. The motion sequencer communicates with the robot through a servo control module that sends commands to the servos in the robot and also collects the feedback signals from those servos. Specifically, the load signals of the hand and the shoulder servos are sent onward to the hardness and weight estimators.

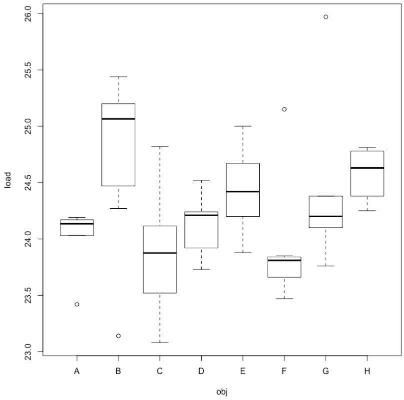

Figure 4: Weight measurements for the eight objects (in arbitrary units) based on servo load in the shoulder joint. There is a relatively small difference in the weight of the objects.

Weight analysis

To estimate the weight of an object we used measurements of load from the servo at the first shoulder joint which would be expected to result in the largest effect due to the larger leverage at this point. We here used the average of the load measurements over the whole action to reflect that the servo is influenced by the weight of the object throughout the action.

We tested whether the robot could measure the weights of the different object during the pick-and-place behavior. Fig. 4 shows the results. There is not a large difference in the measurements for the different objects, reflecting their relatively similar weights.

To test a larger range of weights, we let the robot lift a box containing weights several time. We used the weights 50 g, 150 g and 200 g and measured the load on the first shoulder servo throughout the movement. Fig. 5 shows the result. There is an almost linear relationship between weight and the load measurement.

Hardness analysis

To analyze the hardness of a grasped object, the robot uses the time series consisting of the load feedback from the servo controllers during the closing of the grip. We have earlier shown that the time series generated at a pressure sensor during a grasp contains information about the hardness of an object (Johnsson and Balkenius, 2008b,a, 2009). Here, we instead use feedback signals from the servo. For a soft object, there is nearly no resistance during the closing of the hand. For a harder object, resistance is larger. We analyzed the load of the servo over time and concluded that the average of the load during the closing of the hand was a good predictor of hardness. Surprisingly, the lead measurement appears to be the inverse of what we expected. Unfortunately, the computation of this signal in the Dynamixel servo is undocumented (Mensink, 2008), so we can not determine why this is the case without further experimentation with the servos. Nevertheless, it is a useful signal that can substitute for pressure sensors in the fingers and can be mapped to a hardness dimension (See Fig. 6).

The results show that the variability is relatively low, but there are a few outliers. The sponge laying flat on the table, object H, also has a larger variability than the other objects. The reason for this is that this object in this orientation is very sensitive to the exact location of the grasp since it contains both softer and harder material. The outliers are a result of the same problem. When the robot grasps slightly differently, it may measure the hardness differently.

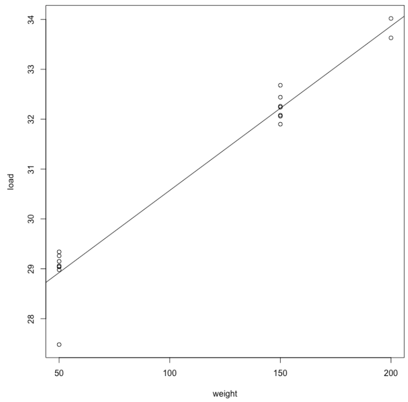

Figure 5: The load at the first shoulder servo for three objects with the weights 50 g, 150 g and 200 g recorded during the whole pick-and-place operation. There is an almost linear relationship between the integrated load and weight. Each circle represents one measurement.

This can easily be overcome by allowing the robot to squeeze the same object several times and average the measurements, similar to how a human would measure hardness by repeatedly squeezing.

Image analysis

The image of each object from the left eye was analyzed to produce a visual description of the object. Before visual analysis of each object, foreground pixels where detected using using the Stauffer-Grimson background subtraction method (Stauffer and Grimson, 1999). This method learns a gaussian mixture model of the background pixels that is later used to detect pixels with a color that differs from the learned background. Since this method depends on a stationary camera, it can only be used while the robot is looking at the same location on the table in front of it, but that is sufficient for the current study. When the foreground pixels have been detected, morphological operators are used to fill out the pixels to a simply connected region.

Next, the foreground pixels are analyzed in two ways. The first is based on a color histogram (Swain and Ballard, 1991; Van De Sande et al., 2010). The output of this processing stage is a vector with 64 elements, each of which codes for the relative amount of color within a certain color band. For simplicity, this is done in the RGB-color space since that is sufficient here, but in general, a transformation to a more adequate color space could be useful.

The second component uses a HOG-description of the foreground pixels (Dalal and Triggs, 2005). Since we are not interested in object detection for current study, but rather in how visual surface properties can be associated with material properties, a single HOG-descriptor is applied to the whole foreground area. The output is a histogram with eight elements that represent the distribution of gradients in the foreground patch of the image.

Mapping visual features to material properties

The model used to let Epi associate visual properties to material properties is shown in figure. 7. There are essentially two pathways from the sensory systems to a central associative mechanism. One has its origin in visual processing, starting at the top of the figure, while the other starts in the motor control system and is drawn at the bottom of the figure.

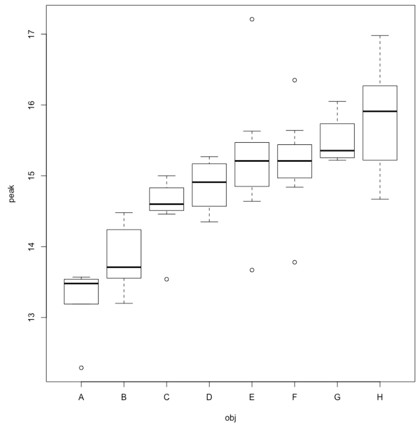

Figure 6: The estimated hardness of each object ranging from hard (A) to soft (H) based on the measured servo load during the closing of the hand (peak). Each box represents six grasps of an object and shows the variability of the measurement.

The results of the weight and hardness analysis are sent to a pair of associator modules that learn associative mappings from the visual descriptors of an object to the hardness and weight of that object. The expected material properties are calculated as a linear mapping from the visual feature vector.

The associator uses the LinearAssociator module in Ikaros that can learn the mapping on-line, either using the delta rule or directly by estimating the optimal mapping using the least square method. The latter method was used here and the association was calculated by left division of a matrix containing the visual feature vectors for each objects in the columns with the column vector containing the estimated hardness or weight for each object:

There are two comparators in the model that compares the expected hardness and weight to the estimated sensory properties and produce a surprise signal when expectations are not met. These modules compute the difference between the expected properties computed from visual features and the currently estimated value from the interaction with the object.

The robot was allowed to test each of the objects once to learn an associate mapping from visual features to hardness and weight. Given the low number of objects, the robot was able to learn a perfect mapping. The error in the expected hardness and weight was less than 10^{-14} for all the objects. This is what would be expected if the visual description vectors were orthogonal to each other. To test this we computed the confusion matrix for the different descriptors using cosine similarity between every pair of objects (Fig. 8). As can be seen, the description vectors are not orthogonal, but there is still a good separation between the objects.

As can be expected, the two green sponges (object F and H) are most similar with a similarity of 0.91. The most dissimilar objects are the clementine (C) and the piece of cotton (G) with a similarity of 0.21. Surprisingly, the cotton also matches the green sponges. However, this is an effect of the RGB-color space used for the color histogram that does not handle white very well.

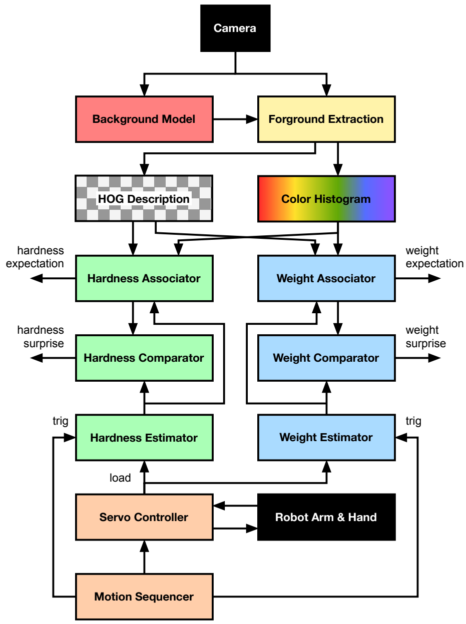

Figure 7: Overview of the implemented model. See text for explanation.

	A	B	C	D	E	F	G	H
A	1.00	0.54	0.30	0.74	0.56	0.68	0.81	0.77
B	0.54	1.00	0.26	0.55	0.43	0.63	0.63	0.62
C	0.30	0.26	1.00	0.33	0.26	0.26	0.21	0.28
D	0.74	0.55	0.33	1.00	0.53	0.67	0.72	0.79
E	0.56	0.43	0.26	0.53	1.00	0.52	0.54	0.57
F	0.68	0.63	0.26	0.67	0.52	1.00	0.84	0.91
G	0.81	0.63	0.21	0.72	0.54	0.84	1.00	0.88
H	0.77	0.62	0.28	0.79	0.57	0.91	0.88	1.00

Figure 8: Confusion matrix for the visual description vectors for the eight objects. The visual coding is sufficiently different for the different objects to allow perfect mapping to the material properties.

Discussion

Through repeated experiences of weighing or squeezing an object and associating that weight and hardness with the visual characteristics of that object humans develop internal models of material knowledge and incorporate that knowledge into our future actions. We implemented this mechanism in a humanoid robot and tested how well it was able to estimate the material properties during manipulation of objects.

Furthermore, we designed a computational model that can learn to associate visual features with material properties to form expectations of future interaction with the object. Similarly to infants learning physical properties of the world through repeated encounters of seeing, reaching for and grasping various objects, the humanoid robot learns to pair heaviness and hardness with provided visual information.

A limitation of the visual system is that the texture analysis using the HOG-operator is only made at a single scale. It would also be useful to look at other visual texture features to make the analysis of the surface properties more complete.

Another possible extension would be to explore behaviors that more actively tries to measure the hardness and weight of objects, for example, to press the object repeatedly or to move the hand up and down while weighing the object in the way that a human would. Such behaviors would allow the robot to integrate the signals over a longer time and will potentially result in more stable measurements. It is possible that such strategies would result in better measurements for objects with a low weight, as those shown in Fig 4. There also appear to be some interaction between shape, hardness and weight that could be further explored. We address some of these issues in a second experiment.

3 Experiment 2

The promising results of the first experiment encouraged us to further investigate multimodal representations of objects with a larger data set. In the second experiment we used 25 objects and looked at the ability of the robot to estimate the weight and hardness of the objects.

Objects

For the second experiment we used 25 different object (Fig. 9). These were selected to represent different weights and hardness as well as different visual features including size, color and texture. We also used a set of 50 g weights to train the robot on weight perception.

Procedure

To collect training data for weight learning, the robot lifted its arm while holding different metal weights ranging from 50 to 500 g in steps of 50 g. Each metal weight was lifted five times and the time series from the load signal from the servos was recorded together with the position feedback from the servos and a trigger signal that was generated when each measurement movement started. This signal was used to divide the complete time series into separate sequences for teach weight.

The robot used two types of movements to explore the 25 objects. The first consisted in holding the object in the hand while moving it up and down five times. In the second, the robot would squeeze the object five times. These procedures were repeated a number of times for each object.

Load and position feedback from the hand and elbow servos were recorded together with a trig signal that indicated the start of each trial.

Figure 9: The 25 objects used in experiment 2.

Supervised learning of weight

In this experiment we first trained the robot to estimate the weights of different metal weights with specified weights using supervised learning. Figure 10 shows five examples of the time series recorded for the 100 g weight. There is very little variation in the time series between the trails. A neural network was trained to map from the time series for each measurement to the corresponding weight. We used Keras to run networks with different number of layers and different activation functions. However, a linear mapping from the input layer to a single output node proved to be sufficient. The first layer consisted of 230 nodes, the length of the time series, while the output used a single node. Both layers used a linear activation function. The training targets were scaled to the range 0 to 1 to fit the output range of the network during training and were scaled back for the testing of the network. The network was trained on all the data using a mean square error loss function for function approximation. We did not use a separate test or validation set since we intended to test the network on the actual objects in the second phase. The best run of the training phase gave a maximum error of 0.12 g. This shows that the actual weight of the objects is a linear function of load time series during the weighing operation so a simple linear regression would have been sufficient. Fig. 11 shows a box plot of the learned mapping from the time series to weight. As can be seen, there is almost no visible variation in the estimation.

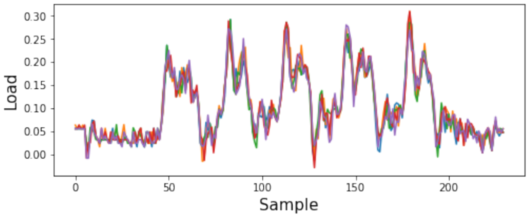

Figure 10: Five superimposed time series for the 100 g weight. The robot shows good repeatability.

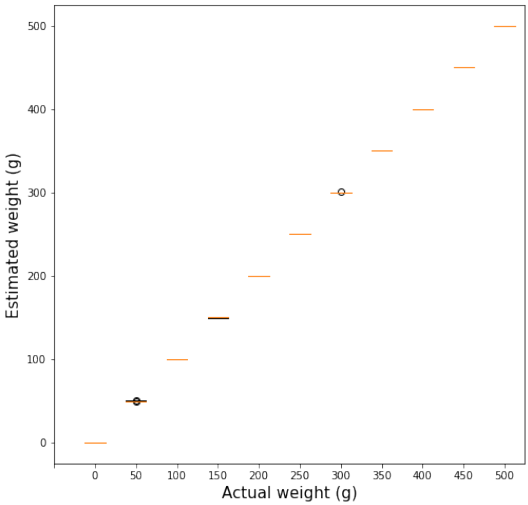

Figure 11: The estimated weight for each of the metal weights. Each box represents five separate weighing operations. There is very little variation and even the outliers, marked with circles, are very close to the median.

Estimating the weight of the different objects

We tested the ability of the robot to weigh the 25 different object after training with the metal weights. The result are shown in figure 12. The average error was 10.3 g. The robot did reasonably well on the heavier objects while the error was much larger for the light objects. For the lighter objects, the weight of the arm and hand has the relatively larger effect on the load feedback. In contrast to the training that used identical weights hanging in the exact same way from the hand, there is also a larger variation in the grasp of the different objects which is likely contributing to the larger variation in this case.

Unsupervised discovery of the weight dimension

Both Experiment 1 and the weight learning here indicate that that weight is a linear function of the load feedback from the servos, it thus seems reasonable that the weight dimension could be learned directly from the data without the actual weight as target. We investigated this using principal component analysis (PCA) on the times series from the servos. We did not use any manual tuning of the process. Instead, the whole time series was used for each object.

Figure 13 shows the amount of variation in the data explained by each of the first principal

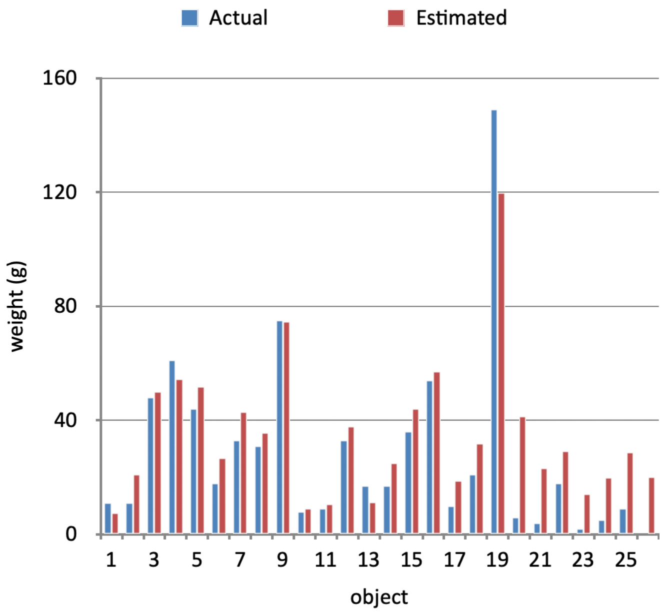

Figure 12: The estimated weight for each of the objects. Each bar represents the average of five independent weighing operations of the same object. The last red bar to the right is the estimation for an empty hand.

components. It is clear that nearly all variation is explained by a single dimension. The variance ratio for the first principal component is 0.82. In Figure 14 we have plotted the measurements for the different objects mapped onto the space spanned by the first two principal components. This suggests that the firsts principal component is indeed weight. It is clear that the principal component analysis has discovered the weight dimension in the time series.

As a last step, we repeated the estimation of the weights of the 25 object using the PCA-based method. To allow a comparison between the PCA-based estimation and the actual weights of the objects, we scaled the estimated weight for each object to obtain the same mean for the estimations as for the actual weights. The results are shown in figure 16 together with the estimates using the supervised learning as described above. It is clear that the PCA-estimates are much less noisy that those for the neural network based method. The RMS error for the neural network based method was 22.4 g while the PCA-based method had a RMS error of 17.9 g.

Unsupervised discovery of the hardness dimension

Given the results in experiment 1 and the success of the PCA-based method for weight estimation, we tried this method also for hardness estimation. An example of the load feedback from the servos during the squeeze operation is shown in figure 17. Like for weight, we performed a principal component analysis of the feedback signals from the servos. In controst to the weight estimations, there is substantial variance in more than one dimension (Fig. 18). The first three principal components have a variance ration of 0.26, 0.26 and 0.07 respectively.

Figure 19 shows the location of the measured objects in the space spanned by the first two principal component of time series recorded during the squeezing operation. The first principal component roughly corresponds to the hardness of the objects but the second principal

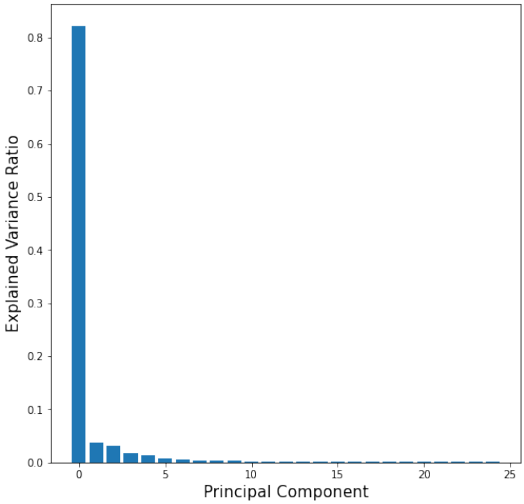

Figure 13: The variance explained by the 25 first principal components.

component also appears to differ for the different objects. Subsequent measurement of the same objects are nicely clustered together indicating both that the robot hand shows good repeatability and that different objects can be distinguished based on the load feedback. The separation of the objects is not perfect but that would not be expected given that some of the objects are rather similar. For example, the two cubes differ only in visual appearance and not in weight, shape or hardness.

A haptic self-organizing map

To see if the objects could be distinguished purely based on feedback from the servos we trained a self-organizing map (Kohonen, 1990) to represent the space of the haptic feedback. We used one the values produced by the PCA-analysis for weight and hardness, one for weight and two for hardness There were thus three measurements for each object as input to the self-organizing map. Before training we normalized the input variables to each have zero mean and a standard deviation of 1. The reason for this is that we did not want one of the variables to dominate the training of the self-organizing map. The

We used the SimpSOM library (Comitani, 1019) for for the simulations and set up a network with 50×50 nodes and trained it for 10000 epochs with an initial learning rate of 0.01.

Figure 20 shows the maps learned as a result of the training for the three input variables. The different patterns for the three dimensions indicate that nodes have learned to react to different combinations of the input values.

Figure 21 shows the weight difference map which is an indication of soft borders between clusters. The image also shows the location of each of the objects in the map as white dots. The self-organiznning map does a reasonably good job at distinguishing between the objects based on the haptic feedback.

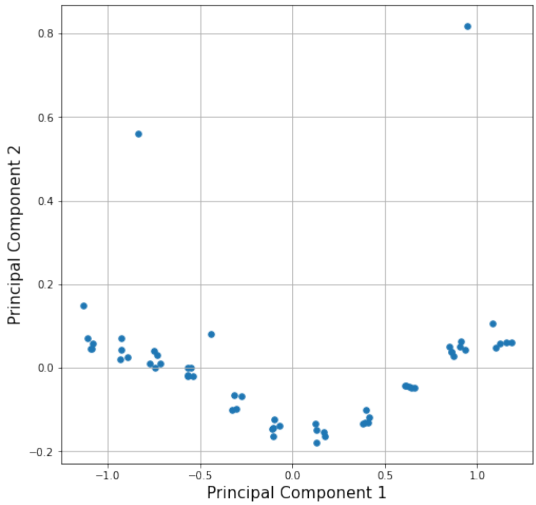

Figure 14: The weight dimension is discovered as the first principal component of the data.

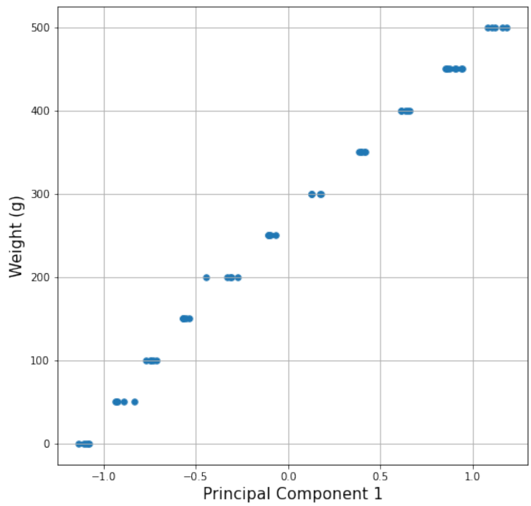

Figure 15: The weight as a function of the first principal component of the measurements. It is clear that the first principal component represents weight.

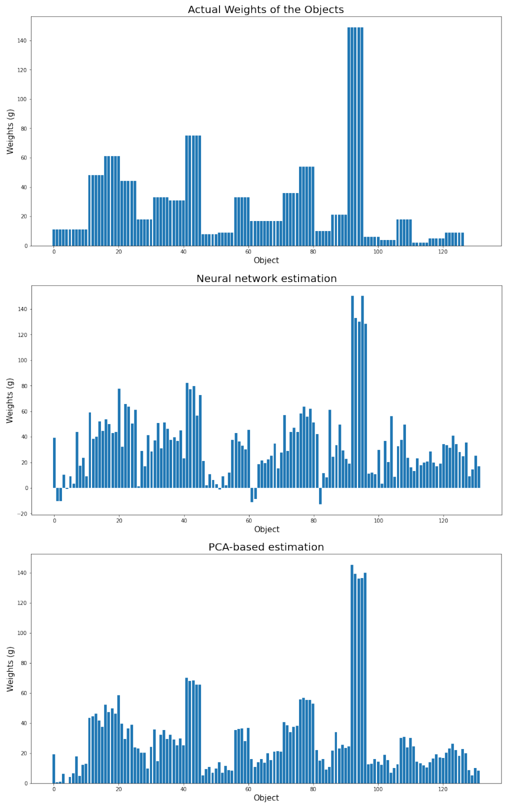

Figure 16: TOP. The actual weights of the objects. MIDDLE. The weights of the objects estimated using the neural network. BOTTOM. The weights of the objects estimated using the first principal component (scaled to the mean of the actual data for comparison). The PCA-based method is clearly less noisy.

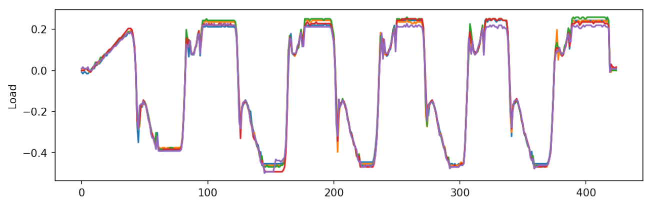

Figure 17: Five consecutive time series for object A. The variability of the measurements is slightly lager than for the weight measurements.

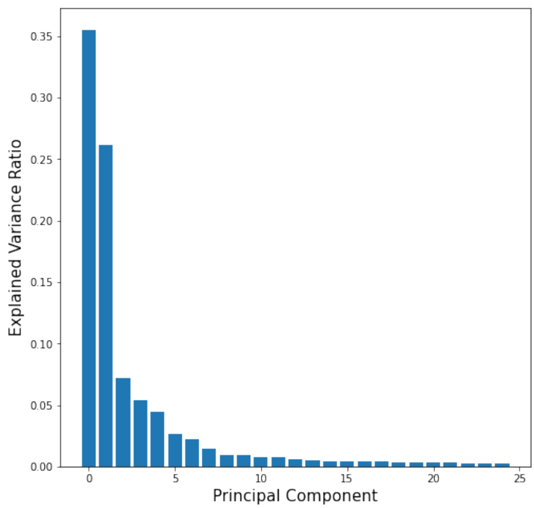

Figure 18: The variance ration for the first 25 principal components of the times series during the squeeze operation.

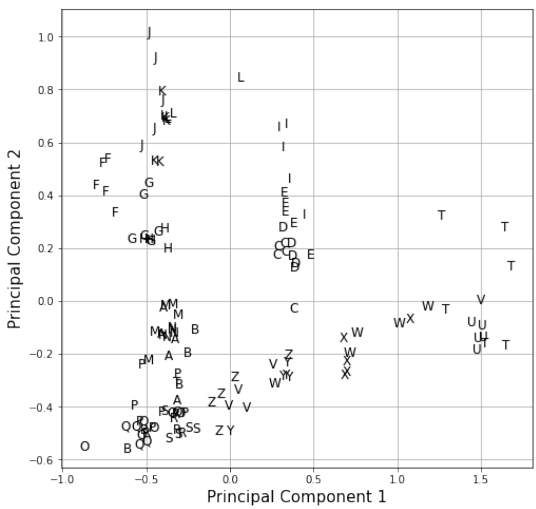

Figure 19: The first two principal components of the feedback from the hand servo during a sequence of squeezes of objects A-Z. Objects with similar properties such as A and B or G and H are represented very close to each other in the space.

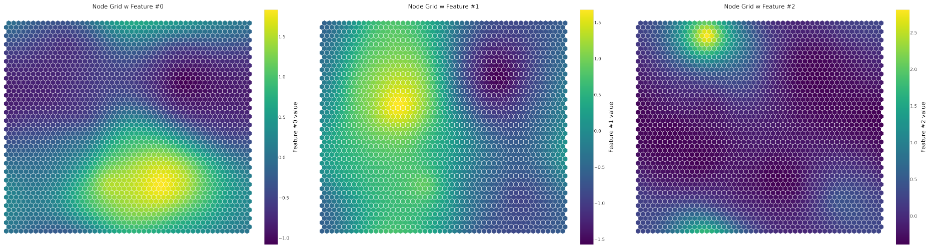

Figure 20: The maps for each of the three features. The different feature dimensions are mapped preferentially to different locations in the map.

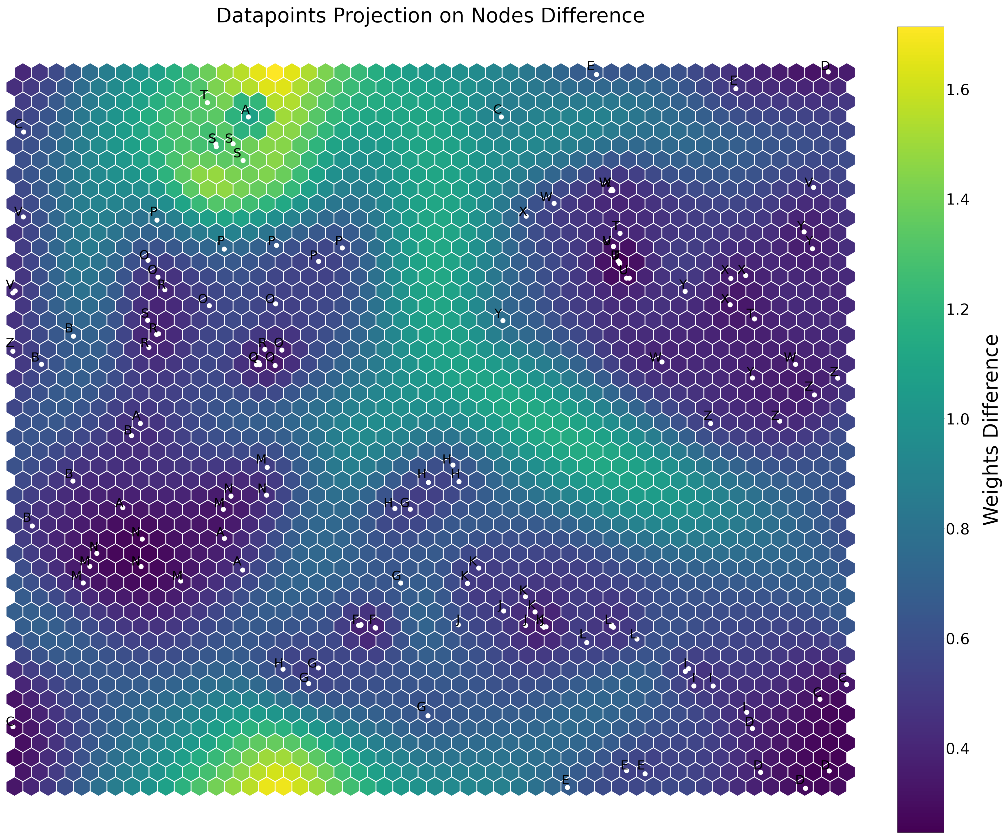

Figure 21: The location of each of the objects on the self-organizing map plotted on a graph of the node's weight difference.

Discussion

In the second experiment we further investigated the information in the haptic feedback from the servos during manipulation of 25 objects. We first showed that a simple neural network could be trained to very accurately estimate the weight of each of the object. Furthermore, we shoved that weight and hardness dimensions could be recovered from the data using principal component analysis. Finally, trained a self-organizing map to code the different object on a two-dimensional map. It is interesting to note that the shape of the object influences the perception of hardness. Since the robot hand has five fingers that are connected to a single servo, it will matter whether all fingers meet resistance from the object or not. This will make a larger object where all fingers are stopped in their movement feel harder than a smaller object like the cubes where one or two of the fingers can still bend almost completely before reaching the object.

Although the noise in the PCA-based estimation was lower than that of the neural network, it is possible that the neural network method could in principle give similar results if a more thorough training method had been used to avoid overfitting. After all, both methods learn a simple linear mapping of the data. The robustness and simplicity of the PCA-based method still gives it an advantage. There are no parameters to set and the result is deterministic.

The self-organizing map used to learn the different object from haptic feedback shows that there is sufficient information in this feedback to distinguish between object with different haptic properties. Although the separation between the different objects is far from perfect, there is a clear mapping of similarity on the map surface. Similar objects are mapped to similar locations. For example, the three larger soft objects C, D and E are all represented in the lower right corner of the map. Similarily object G and H that are virtually identical are represented in the middle of the map, while F and K who both have a cylindrical shape and similar weight and hardness are represented further to the right.

4 General Discussion

We trained a humanoid robot to recognize objects based on proprioceptive signals from the arm and hand as well as through visual features. A novel feature of the robot used is that it can estimate hardness, and to some extent shape, without having any sensors in the fingers. This makes the hand very robust since there are no fragile parts in the actual fingers. Furthermore, we investigated how the robot could learn to perceive the weight of different objects and showed that very accurate weight measurements can be learned by both supervised and unsupervised methods. Finally, we tested how a self-organizing map could learn to represent a haptic space where the different object are mapped according to similarity.

There are several limitations of the current approach. For example, the weight and hardness perception of the robot hand depends on doing the exact same motion every time. It would be interesting to also include the position feedback from the servos in the analysis to potentially allow the robot to sense both weight and hardness also for different speeds of the hand and arm motions. More advanced haptic exploration behaviors could also be explored.

We have earlier investigated other robot hands with a large set of sensors (Johnsson and Balkenius, 2008b,a, 2009), but the current study shows that in some cases a simpler design may be sufficient. This may be of practical use in cheaper robots where the number of sensors have to be limited. Another use of the hardness sensing described here, is that it can also be used to detect if the robot has dropped an object. This is something we will explore further in the future. However, we also want to combine the 'proprioceptive' sensing used here with traditional contact and bend sensors on the fingers to allow the hands to sense shape as well as hardness. Contact senors are very important for manipulation since they tell the robot that the shape of the hand contains information about the object.

The visual analysis can also be extended in several ways. A better color analysis will be necessary to cope with changing illumination — a problem that we did not address here. It would also be possible to use more advanced texture recognition algorithms. Other properties such as the shininess of an object could also be measured since it will typically influence the properties of the object, for instance are brighter looking objects judged to weigh less than dark looking objects (Walker and Walker, 2010), and glossier objects are judged to be more slippery than matte objects (Adams and Graf, 2016).

We also want to test the generalization ability on a larger set of objects. Presumably, visual properties such as as size and texture could be used to guess the weight and hardness of an object even if there will obviously be a much greater error than for previously encountered objects. The main next step will be to allow the robot to use a more advanced visual feature analysis to estimate the weight and hardness also for previously unseen objects. Although there is obviously no correct answer in these cases, it would nevertheless be interesting to see if the robot could use the size and surface texture of objects to make a reasonable guess.

It is likely that a deep network could be trained to do map visual input to a feature vector describing the glossiness and texture of an object or perhaps directly to expected weight and hardness. This would probably require an even larger set of objects to include the possibility for generalization between object with similar properties.

It is interesting to note that the mechanisms used here for tactile categorization could shed some light on the corresponding mechanisms in the brain. Like our model, humans do not usually have access to an objective measurement of the weight and hardness of an object. Instead, we need to rely on our our own senses to learn the relevant dimensions. Our model shows that dimensionality reduction of proprioceptive time series may play a role in this process. Here we used PCA for this purpose, but dimensionality reduction is a property of many neural network models and could presumably also take place in the brain using similar mechanisms (Beyeler et al., 2019). The self-organizing map used to categorize the objects also has at least a superficial similarity to how cortex represents perceptual information.

On use of the system in a robot would be to limit the energy consumption during manipulation of objects. This is especially important for robots that runs off battery. Typically, the servo motors are allowed to use an excessive amount of energy to move along the programmed paths. However, by reducing the maximum current available, the energy consumption can be significantly reduced. We currently experiment with this feature of the servos we use. By identifying an object using vision before it is lifted, the robot can set the available current to the minimum necessary for the operation based on its expectation of the weight of the object.

We currently use position control to move the robot joints, but if we instead used force control, this mechanism would be even more useful. In that case, the force would need to be adjusted so that the robot does not lift light objects with too much force. An error in the predicted weight would in this case be similar to when we try to lift, for example, a milk package that we believe to to be full but in fact is empty.

In summary, we have shown how a humanoid robot can detect weight and hardness using feedback signals from its servos without the need for dedicated tactile sensors. Furthermore, we have shown how a visual analysis of surface properties such as color and texture can be used to predict material properties through an associative learning mechanism.

References

Adams, W. Kerrigan, I. and Graf, E. (2016). Touch influences perceived gloss. *Sci Rep*, 6:21866.

Balkenius, C., Morén, J., Johansson, B., and Johnsson, M. (2010). Ikaros: Building cognitive models for robots. *Advanced Engineering Informatics*, 24(1):40–48.

Baumgartner, E., Wiebel, C., and Gegenfurtner, K. (2013). Visual and haptic representations of material properties. *Multisensory Research*, 26(5):429–455.

Bergmann Tiest, W. M. (2010). Tactual perception of material properties. *Vision Research*, 50(24):2775–2782.

Beyeler, M., Rounds, E. L., Carlson, K. D., Dutt, N., and Krichmar, J. L. (2019). Neural correlates of sparse coding and dimensionality reduction. *PLoS Computational Biology*, 15(6):e1006908.

Buckingham, G., Cant, J. S., and Goodale, M. A. (2009). Living in a material world: How visual cues to material properties affect the way that we lift objects and perceive their weight. *Journal of Neurophysiology*, 102(6):3111–3118.

Buckingham, G., Ranger, N. S., and Goodale, M. A. (2011). The material–weight illusion induced by expectations alone. *Attention, Perception, & Psychophysics*, 73(1):36–41.

Cavdan, M., Drewing, K., and Doerschner, K. (2021). The look and feel of soft are similar across different softness dimensions. *Journal of Vision*, 21(10):20.

Chadwick, A. C. and Kentridge, R. W. (2015). The perception of gloss: A review. *Vision Research*, 109:221–235.

Choi, B., Lee, S., Choi, H. R., and Kang, S. (2006). Development of anthropomorphic robot hand with tactile sensor: Skku hand ii. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3779–3784. IEEE.

Comitani, F. (1019). SimpSOM (version 1.3.4), doi:10.5281/zenodo.2621560.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893. IEEE.

Fitzpatrick, P., Metta, G., Natale, L., Rao, S., and Sandini, G. (2003). Learning about objects through action-initial steps towards artificial cognition. In *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*, volume 3, pages 3140–3145. IEEE.

Flanagan, J. R. and Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size–weight illusion. *Nature Neuroscience*, 3(7):737–741.

Flanagan, J. R. and Wing, A. M. (1997). The role of internal models in motion planning and control: Evidence from grip force adjustments during movements of hand-held loads. *Journal of Neuroscience*, 17(4):1519–1528.

Fukaya, N., Toyama, S., Asfour, T., and Dillmann, R. (2000). Design of the tuat/karlsruhe humanoid hand. In *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)(Cat. No. 00CH37113)*, volume 3, pages 1754–1759. IEEE.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Allen and Unwin.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Houghton Mifflin.

Gonçalves, A., Saponaro, G., Jamone, L., and Bernardino, A. (2014). Learning visual affordances of objects and tools through autonomous robot exploration. In *2014 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 128–133. IEEE.

Hjelm, M., Ek, C. H., Detry, R., and Kragic, D. (2019). Invariant feature mappings for generalizing affordance understanding using regularized metric learning. *arXiv preprint arXiv:1901.10673*.

Hoelscher, J., Peters, J., and Hermans, T. (2015). Evaluation of tactile feature extraction for interactive object recognition. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pages 310–317. IEEE.

Johansson, B., Tjøstheim, T. A., and Balkenius, C. (2020). Epi: An open humanoid platform for developmental robotics. *International Journal of Advanced Robotic Systems*, 17(2):1729881420911498.

Johansson, R. S. and Cole, K. J. (1992). Sensory-motor coordination during grasping and manipulative actions. *Current Opinion in Neurobiology*, 2(6):815–823.

Johnsson, M. and Balkenius, C. (2008a). Associating SOM representations of haptic submodalities. *Towards Autonomous Robotic Systems*, 2008:124–129.

Johnsson, M. and Balkenius, C. (2008b). Recognizing texture and hardness by touch. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 482–487. IEEE.

Johnsson, M. and Balkenius, C. (2009). Experiments with self-organizing systems for texture and hardness perception. *Robotics and Autonomous Systems*, 4:53–62.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480.

Lederman, S. J. (1974). Tactile roughness of grooved surfaces: The touching process and effects of macro- and microsurface structure. *Perception & Psychophysics*, 16:385–395.

Lederman, S. J. and Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19(3):342–368.

Mar, T., Tikhanoff, V., Metta, G., and Natale, L. (2015). Self-supervised learning of grasp dependent tool affordances on the icub humanoid robot. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3200–3206. IEEE.

Matsuoka, Y. (1995). Embodiment and manipulation learning process for a humanoid hand. Technical Report 1546, MIT Artificial Intelligence Laboratory.

Mensink, A. (2008). Characterization and modeling of a dynamixel servo. *Trabajo Individual de Investigación en el Electrical Engineering Control Engineeringde la University of Twente.*

Montesano, L., Lopes, M., Bernardino, A., and Santos-Victor, J. (2008). Learning object affordances: from sensory–motor coordination to imitation. *IEEE Transactions on Robotics*, 24(1):15–26.

Nguyen, A., Kanoulas, D., Caldwell, D. G., and Tsagarakis, N. G. (2016). Detecting object affordances with convolutional neural networks. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2765–2770. IEEE.

Okamoto, S., Nagano, H., and Yamada, Y. (2013). Psychophysical dimensions of tactile perception of textures. *IEEE Transactions on Haptics*, 6(1):81–93.

Paulun, V. C., Gegenfurtner, K. R., Goodale, M. A., and Fleming, R. W. (2016). Effects of material properties and object orientation on precision grip kinematics. *Experimental Brain Research*, 234(8):2253–2265.

Paulus, M. and Hauf, P. (2011). Infants' use of material properties to guide their actions with differently weighted objects. *Infant and Child Development*, 20(4):423–436.

Regoli, M., Jamali, N., Metta, G., and Natale, L. (2017). Controlled tactile exploration and haptic object recognition. In *Advanced Robotics (ICAR), 2017 18th International Conference on*, pages 47–54. IEEE.

Saccone, E., Landry, O., and Chouinard, P. (2019). A meta-analysis of the size-weight and material-weight illusions. *Psychon Bull Rev*, 26(4):1195–1212.

Şahin, E., Çakmak, M., Doğar, M. R., Uğur, E., and Üçoluk, G. (2007). To afford or not to afford: A new formalization of affordances toward affordance-based robot control. *Adaptive Behavior*, 15(4):447–472.

Stauffer, C. and Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. In *cvpr*, page 2246. IEEE.

Stoytchev, A. (2005). Behavior-grounded representation of tool affordances. In *Proceedings of the 2005 ieee international conference on robotics and automation*, pages 3060–3065. IEEE.

Sugaiwa, T., Fujii, G., Iwata, H., and Sugano, S. (2010). A methodology for setting grasping force for picking up an object with unknown weight, friction, and stiffness. In *2010 10th IEEE-RAS International Conference on Humanoid Robots*, pages 288–293. IEEE.

Swain, M. J. and Ballard, D. H. (1991). Color indexing. *International journal of computer vision*, 7(1):11–32.

Tikhanoff, V., Pattacini, U., Natale, L., and Metta, G. (2013). Exploring affordances and tool use on the icub. In *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 130–137. IEEE.

Van De Sande, K., Gevers, T., and Snoek, C. (2010). Evaluating color descriptors for object and scene recognition. *IEEE transactions on pattern analysis and machine intelligence*, 32(9):1582–1596.

Vardar, Y., Wallraven, C., and Kuchenbecker, K. J. (2019). Fingertip interaction metrics correlate with visual and haptic perception of real surfaces. In *Proceedings of the IEEE World Haptics Conference (WHC)*, pages 395–400, Tokyo, Japan.

Walker, P. Francis, B. and Walker, L. (2010). The brightness-weight illusion. *Exp Psychol.*, 57(6):462—469.

Wolpert, D. M. and Flanagan, J. R. (2001). Motor prediction. *Current Biology*, 11(18):R729–R732.

Yussof, H., Ohka, M., Takata, J., Nasu, Y., and Yamano, M. (2008). Low force control scheme for object hardness distinction in robot manipulation based on tactile sensing. In *2008 IEEE International Conference on Robotics and Automation*, pages 3443–3448. IEEE.

Zoeller, A., Lezkan, A., Paulun, V., Fleming, R., and Drewing, K. (2019). Integration of prior knowledge during haptic exploration depends on information type. *Journal of Vision*, 19(4):20.