Expectations affect the perception of material properties

Lorilei M. Alley* Justus Liebig University, Giessen, Germany ✉

Alexandra C. Schmid* Justus Liebig University, Giessen, Germany ✉

Katja Doerschner Justus Liebig University, Giessen, Germany
Bilkent University, Ankara, Turkey ✉

Many objects that we encounter have typical material qualities: spoons are hard, pillows are soft, and Jell-O dessert is wobbly. Over a lifetime of experiences, strong associations between an object and its typical material properties may be formed, and these associations not only include how glossy, rough, or pink an object is, but also how it behaves under force: we expect knocked over vases to shatter, popped bike tires to deflate, and gooey grilled cheese to hang between two slices of bread when pulled apart. Here we ask how such rich visual priors affect the visual perception of material qualities and present a particularly striking example of expectation violation. In a cue conflict design, we pair computer-rendered familiar objects with surprising material behaviors (a linen curtain shattering, a porcelain teacup wrinkling, etc.) and find that material qualities are not solely estimated from the object's kinematics (i.e., its physical [atypical] motion while shattering, wrinkling, wobbling etc.); rather, material appearance is sometimes "pulled" toward the "native" motion, shape, and optical properties that are associated with this object. Our results, in addition to patterns we find in response time data, suggest that visual priors about materials can set up high-level expectations about complex future states of an object and show how these priors modulate material appearance.

Introduction

The material an object is made of endows it with a certain utility or function: chairs are usually rigid, to afford sitting; spoons are hard to afford eating; and towels are soft and absorbent to afford drying. The material choices for many types of man-made objects (keys, cups, cushions) tend to be restricted, and thus many objects that we encounter have a "typical" material: bookshelves are made from wood, pillows from cloth and down, and keys from metal. Over a lifetime of interacting with objects, we learn not

only what they look like, but also how their material "behaves" under different types of forces. For example, we know that porcelain vases shatter when knocked over, that soft Jell-O wobbles when poked, and that rubber balls bounce when thrown at the wall. Strong associations are formed between the object's shape and its optical and mechanical material properties.

Observers rely on these associations when estimating material qualities, which can lead to paradoxical findings. For example, physically, the softness of an object (compliance) is independent of its optical properties (color, gloss, or translucency), and with sophisticated production techniques, it is possible to manufacture soft objects with any set of optical properties. Yet, when judging the softness of an object in static images, research has shown that that optical characteristics of the material affected observers' judgements, for example, a velvety cube was perceived as softer than a chrome cube of the same shape and size (Paulun et al. 2017; Schmidt et al., 2017). To account for these results, Paulun et al. (2017) propose an indirect route to perception, where image cues activate the memory of a particular material along with its learned associations, including its typical mechanical properties (e.g., "This looks like honey—so it's probably quite runny"). They suggest that this route operates in parallel with a more direct one, where image cues directly convey something about the properties of the material (e.g., Paulun et al., 2017; Paulun et al., 2015; van Assen & Fleming, 2016; van Assen et al., 2018; Marlow & Anderson, 2016; see Figure 1 for an illustration).

Learned associations about material properties are not only evoked by the optical characteristics of an object, but may also be elicited by its shape: Experience with soft materials or liquids seems to create strong associations between shape deformations and perceived material qualities (e.g., Bi et al., 2019; Kawabe, 2018; Kawabe et al., 2015; Mao et al., 2019; Paulun et al., 2015; Paulun et al., 2017; Schmid & Doerschner, 2018; Schmidt et al., 2017; van Assen & Fleming, 2016; van Assen et al., 2018). Similarly, if a shape

Figure 1. Contribution of prior associations and image cues on perceived material qualities. The role of predictions or associative mechanisms in material perception is not well understood. (A) The perception of material qualities (such as gelatinousness) can be influenced by prior associations between dynamic optics, shape, and motion properties. (A) Watching the green (*left*) object deform may evoke an association with green Jell-O (B), and may therefore be perceived as wobblier and more gelatinous than the matte object, despite both objects wobbling in identical ways (as shown in Supplementary Movies S1 and S2). (C) Alternatively, the green object may be perceived as wobblier owing to larger image differences between frames, and potentially higher motion energy, as illustrated on the right (Doerschner et al., 2011, Doerschner, Kersten, & Schrater, 2011). A combination of associative and modulatory mechanisms is also possible. The difference in motion energy in images of the translucent object in (C) is about seven (6.8) times larger than that of the matte one, purely owing to the difference in optical properties between these two objects.

reminds the observer of a specific object, its typical mechanical material properties might also become activated. Consequently, recognizing an object should cause observers to activate strong predictions about the object's material properties, and how it should behave under physical forces. Here we ask whether these predictions in turn affect how we perceive material properties.

The human brain uses prior knowledge to continuously generate predictions about visual input to make quick decisions and guide our actions (for a review, see Kveraga et al., 2007). Predictions about material properties should be no exception to this: To avoid small daily disasters, we need to be able to predict how bendable and heavy a cup is before picking it up. How prior knowledge is combined with visual input has been investigated in several areas of vision. For example, researchers have studied how color memory for objects influences perceived color (e.g., Hansen et al., 2006; Olkkonen & Allred, 2014), or how memory modulates the perception of motion direction or dynamics of binocular rivalry (Chang & Pearson, 2018; Scocchia et al., 2013; Scocchia et al., 2014). Within the framework of Bayesian models (e.g., Ernst & Banks, 2002; Hillis et al., 2002; Kersten et al., 2004; Landy et al., 1995) the influence of prior knowledge on the integration of different types of sensory input has been looked at more formally. In particular, cue conflict scenarios have proven extremely useful to generate insights about the complex interplay of prior selection and the weighting of sensory input in the perception of object properties (e.g., Knill, 2007a, b).We will use an experimental paradigm analogous to cue conflict by juxtaposing indirect (prior) and direct (sensory) information in the perception of material properties to formally test whether expectations about an object's mechanical properties are generated.
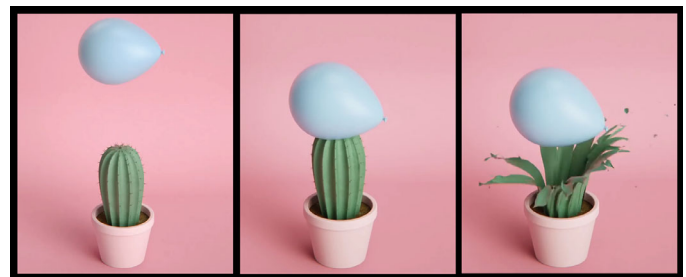

Figure 2. Three frames from "Preposterous" by Florent Porta. Artists have played with our expectation of how objects and their materials should behave. In this study, we compare material perception for falling objects that deform in surprising and unsurprising (i.e., expected) ways. Retrieved from https://vimeo.com/191444383.

Violating the expected mechanical properties of materials necessarily involves image motion. Figure 2 illustrates this phenomenon. Shown below are three frames from an animation by Florent Porta. The first panel sets up the viewer's expectations about the objects material properties (i.e., the balloon will pop when it comes into contact with the spines of the cactus). As the movie proceeds, our expectations about the event to unfold are violated, and the viewer is quite surprised when the cactus pops like the balloon would. In the present study, we directly test how such expectations affect the perception of material properties by comparing material perception for falling objects that deform in surprising and expected ways.

We anticipate our results using the Bayesian framework for analogy. Suppose the task of an observer watching the movie in Figure 2 was to rate the cactus' rigidness. If the observer recognizes the object as a cactus, it is probably safe to assume that they have a strong prior belief about how rigid cacti are, but it is

also possible that other, weaker prior beliefs exist. When confronted with a popping cactus, the visual system may either veto all cues that previously suggested that the cactus was rigid (e.g., Landy et al., 1995) and judge the cactus as very soft, or the visual system may down-weigh cues to rigidness (Knill, 2007a). If cue vetoing occurs, we would expect to find no difference in the perception of mechanical material properties (e.g., perceived rigidness in our example) between objects that deform in surprising (popping cactus) and expected (popping balloon) ways, because in either case the strongest cue is used (i.e., the visual sensory input that shows the cactus is popping and thus really not that rigid). If, however, a down-weighing of the visual input occurs, the visual prior (on the typical rigidness of cacti) would exert an influence when observers rate the rigidness of the popping cactus, and we would find they are rated as somewhat more rigid than popping balloons. The Bayesian framework predicts another result: when expectations are violated—as in the exploding cactus scenario—the visual system needs to update its internal model of the world (the generative model; e.g., see Kersten et al., 2004) to minimize the prediction error for future tasks. This updating (or prediction error correction) is thought to be a reiterative process, which may take time to complete. In fact, a recent study by Urgen and Boyaci (2019) modelled an individuation task and showed that, for surprising trials, the model predicted longer duration thresholds. By analogy, we expect to find that it takes observers longer to perform perceptual tasks when judging material attributes of surprisingly deforming objects (e.g., popping cacti) than ones that deform as expected (e.g., popping balloons).

In this study, we use a violation of expectation paradigm to investigate how predictions about object deformations based on object knowledge influence how we perceive the material of an object. To manipulate object familiarity, we rendered two types of objects: familiar objects, for which there exist strong predictions about their mechanical material properties; and novel, unfamiliar objects, for which no strong predictions should exist. For each of these object types, we rendered two motion sequences that showed how objects deformed when being dropped on the floor: an expected sequence, in which objects' deformations were consistent with the observers' prior beliefs about the mechanical properties of the material; and a surprising sequence, in which they were not. The novel objects inherited their optical and motion properties from a corresponding familiar object. Participants rated the objects in these movies on various material attributes. Although we expected to find differences in ratings and response times between expected and surprising events for familiar objects (discussed elsewhere in this article), we also expected to find differences—although attenuated—for novel shapes,

owing to optical cues and/or the fact that they were bounded three-dimensional (3D) objects potentially eliciting prior expectations (e.g., a rigidity prior) about these objects' mechanical material properties. We included two additional experimental conditions in which observers rated the same attributes on static images that showed each familiar and novel object during the first and last frame of the motion sequence to explicitly assess existing priors on material properties, in addition to the influence of shape recognizability (i.e., after the object had deformed) on ratings, respectively.

Methods

Stimuli

Objects

We used two types of objects: familiar and novel. Figure 3 gives an overview of all objects used in the experiment. To create a stimulus set with a broad range of typical material classes, we choose 15 familiar objects belonging to one of five material mechanics: nondeforming (wooden chair, metal key, metal spoon), wobbling (red and green Jell-O, custard), shattering (wineglass, terracotta pot, porcelain teacup), wrinkling (linen, velvet, and silk curtain), and splashing (milk, honey, and water drops). All objects were rendered (Blender 2.77a, Stichting Blender Foundation, Amsterdam, the Netherlands) with their typical optical material properties, for example, a metallic-looking key, green transparent Jell-O, or a silky appearing curtain). Objects were located in a room that had brown walls and a hardwood, polished floor, and were illuminated by the Campus environment map (Debevec, 1998). The 3D meshes of the familiar objects were obtained or adapted from TF3DM (www.tf3dm.com) and TurboSquid (www.turbosquid.com) or created by hand. Unfamiliar shapes (Glavens; Phillips, 2004) inherited their optical and mechanical properties from the corresponding familiar object.

The objects were rendered at approximately the same size as each other so that, for example, the key was the same "physical" (simulated) size as the chair, even though in real life chairs are larger than keys. We chose to do this so that the objects would hit the ground at the same time, and behave in a similar way under gravity. This factor is important for some of our analyses (see Analysis—Response time).

Deformations

For each object, we rendered short movies that showed an object falling from a height and interacting with the ground. To manipulate surprise in our

Figure 3. Stimuli expected motion. Shown are all 15 familiar (*top*) and corresponding novel objects (*bottom*) used in the experiments. Objects are organized according to their typical material kinematics, that is, how they deform under force. Note that individual scenes are scaled to maximize the view of the object (*first frame*, *left columns*), or to give a good impression of the material kinematics (*last frame*, *right columns*).

	Non deforming	Wobbling	Shattering	Wrinkling	Splashing
Expected	Painted chair	Green Jell-O	Clay pot	Linen curtain	Honey droplets
	Brass key	Red Jell-O	Wine glass	Silk curtain	Milk droplets
	Metal spoon	Custard	Teacup	Velvet curtain	Water droplets
Surprising	Milk drops	Wine glass	Custard	Green Jell-O	Painted chair
	Red Jell-O	Honey droplets	Linen curtain	Metal spoon	Clay pot
	Velvet curtain	Brass key	Water droplets	Teacup	Silk curtain

Table 1. Overview of Objects and Conditions in the Experiment. Columns show the five categories of typical material kinematics (material class) in the experiment; rows show which objects occurred in the expected and surprising conditions. Every object in the expected condition would also appear exactly once in the surprising condition, where it would be rendered with very atypical mechanical properties, for example, the silk curtain would splash (see Figure 3 and Supplementary Figure 1 for corresponding renderings). Novel objects had no familiar shape but had the same optical and mechanical material properties as the familiar objects in this table.

experiment, an object could either behave as expected—for example, a glass would shatter—or it could inherit the mechanical material behavior from another object—for example, milk drops would stay rigid upon impact (Figure 4B shows two more examples of surprising material behaviors; also see Supplementary Figure 1). We created corresponding expected and surprising movies for novel objects (Figure 3, Figure 4C, and Supplementary Figure 1). Table 1 shows which objects inherited which material mechanics in the

surprising condition. All experimental movies can be downloaded at https://doi.org/10.5281/zenodo.2542577.

Our stimulus set was balanced in the sense that for every type of deformation (splashing, wrinkling, etc.), we found three familiar objects that "naturally" have these material kinematics. In the surprising condition the deformations were paired with other familiar objects such that a given deformation might be perceived as unusual for this type of object (e.g., a chair splashing).
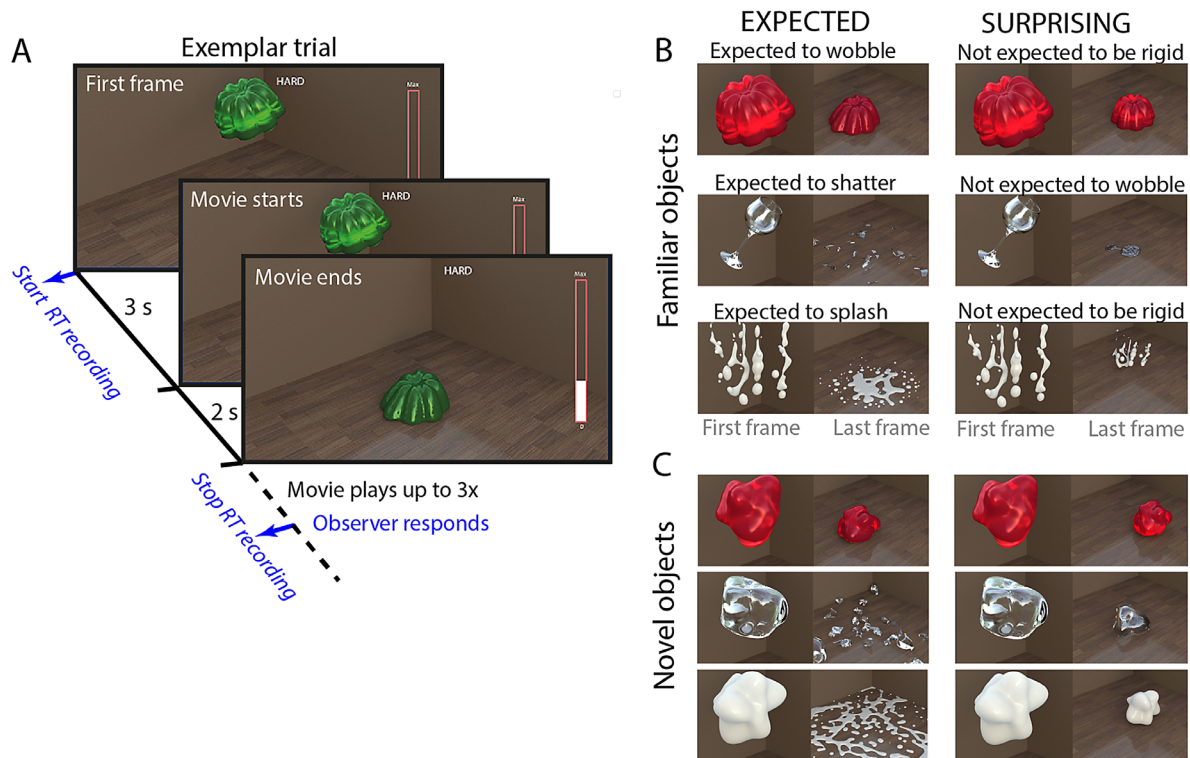
Figure 4. Trial and stimuli. (A) An exemplar trial. (B, C) A subset of familiar objects (B) and corresponding novel objects (C) used in the experiments. Familiar objects could either behave as expected, or in a surprising manner. Note that this distinction (expected vs surprising) is only meaningful for familiar objects. Note that individual scenes are scaled to maximize the view of the object (*first frame*), or to give a good impression of the material kinematics (*last frame*). Figure 3 (expected condition) and Supplementary Figure 1 (surprising condition) show corresponding views for the entire stimulus set. The objects were rendered at approximately the same size as each other, so that for example the key was the same "physical" (simulated) size as the chair, even though in real life chairs are larger than keys. We chose to do this so that the objects would hit the ground at the same time, and behave in a similar way under gravity. This is important for some of our analyses (see Analysis—Response time).

Animations

Each movie consisted of 48 frames, depicting an object suspended in air, which then fell to the ground. Impact occurred exactly at the 11th frame for all objects. The largest extent of the objects in the first frame varied between 6.91 (clay pot) and 12.6 (spoon) degrees visual angle. The largest extent of the objects in the last frame depended on the deformation, but varied between 48.85° visual angle for shattering or splashing items and 4.29° visual angle for rigidly falling items. Object deformations were simulated using the Rigid Body, Cloth, and Particles System physics engines in Blender. For technical specifications about the Particle System simulations, we refer the reader to the parameters listed in (Schmid & Doerschner, 2018).

Apparatus

The experiment was coded in MATLAB 2015a (MathWorks, Natick, MA) using the Psychophysics Toolbox extension (version 3.8.5, Brainard, 1997; Kleiner et al., 2007; Pelli, 1997), and presented on a 24-5/8" PVM-2541 Sony (Sony Corporation, Minato, Tokyo, Japan) 10-bit OLED monitor, with a resolution of 1024×768 and a refresh rate of 60 Hz. Videos were played at a rate of 24 frames per second. The participants were seated approximately 60 cm from the screen.

Task and procedure

Main (motion) experiment

Observers were asked to watch a short video clip to the end and then to rate the object they saw on one of four attributes (hardness, gelatinousness, heaviness, and liquidity) as quickly (but as accurately) as they could. We choose the attributes such that they would capture some aspect of the mechanical material qualities of the objects. For example, a splashing object is likely to be rated as very *liquid*, and a nondeforming object not; a

wiggling object is likely to be rated as very *wobbly* but a shattering object not. To familiarize observers with the rating task, the use of the slider bar and the key presses, they completed four practice trials with two objects that did not occur in the actual experiment and with two rating adjectives that also did not occur in the experiment (e.g., rate how shiny this object is).

The experiment was organized into four blocks, with one block per attribute. Before the block started, the observer was familiarized with the rating question and then proceeded with a button press to start the trials. On every trial, a reminder of the question of this block remained at the top of the screen, for example, *hard* (for "How hard is the object?"), together with the first frame of a movie, which was held static for three seconds before the movie was played to the end. After this the movie clip repeated two more times (without the hold at the beginning), if necessary. Participants were asked to first watch the video until it finished (i.e., the first play through) and then to rate the object as quickly as they could (while still maintaining accuracy). They indicated their rating by using the mouse to adjust the height of a slider bar placed on the right side of the screen ([Figure 4](#)A). A zero setting indicated the absence of an attribute, for example, not *gelatinous* at all, whereas a maximum setting would correspond to the subjective maximum value of an attribute. The trial was completed when the observer pressed the space key on the keyboard, after which time the next trial would immediately begin. Response time was measured from the beginning to the end of a trial (between spacebar presses). The slider position of the previous trial was carried over the new trial to give the experimental interface a more natural feel to it.

Participants completed 240 trials in total (2 surprise conditions [expected, surprising] × 2 object types [familiar, novel] × 4 attributes × 15 objects). Surprise condition and object type were the two relevant manipulations in the experiment. While the order of blocks was the same for all observers (*hard*, *gelatinous*, *heavy*, *liquid*), the trial order in each block was randomized.

First frame/last frame experiment

The tasks, setup, and procedures were identical to that of the motion experiment. In contrast with the motion experiment, in these experiments, the first/last frame of each of the videos was held on-screen for the same duration as a single presentation of the video.

Participants

Motion experiment

Twenty-five participants, mean age of 24.8 years, 18 female, participated in the experiment; 23 were right handed and all had self-reported normal or corrected-to-normal vision.

First frame experiment

Fifteen participants, mean age of 26.40 years, 13 female, participated in this experiment; 13 participants were right handed and all had self-reported normal or corrected-to-normal vision.

Last frame experiment

Fourteen participants, mean age of 26.85 years, 10 female, participated in the static last frame condition of experiment 1; 12 were right handed. All participants had self-reported normal or corrected-to-normal vision.

All participants were native German speakers, and the experiment was given entirely in German. The experiment followed the guidelines set forth by the Declaration of Helsinki, and participants were naïve as to the purpose of the experiment. All participants provided written informed consent and were reimbursed at a rate of €8 per hour.

Analysis

Rating differences: Expected versus surprising

To measure the influence of object knowledge on perceived material properties, we computed rating differences between expected and surprising conditions. To do this, we first computed average ratings (across all observers), for each object type (familiar, novel), each attribute (how hard, how gelatinous, how heavy, how liquid), each material class (nondeforming, wobbling, shattering, wrinkling, splashing), and each outcome (expected, surprising). Note, that in expected and surprising conditions, how an object deformed was the same, but which objects would deform was different. For example, honey, milk, and water would splash in the expected splashing condition, and the chair, pot, and curtain would splash in the surprising conditions.

After obtaining average ratings we computed rating differences between expected and surprising outcomes for each object type, each attribute and each material class. This resulted in 20 differences scores (5 material classes × 4 attributes), in the familiar object condition and 20 in the novel object condition (see Supplementary Figure 2A). If object knowledge influences perceived material properties, then ratings should differ between familiar objects that behave as expected, and familiar objects that behave in a surprising way.

We expected this difference to be overall larger for familiar objects than for novel objects. Thus, to obtain an overall measure of whether expectations affect the perception of object properties at all we took the

absolute value of each of these 40 difference score as an index (effect of expectation index, $\in$; also see Supplementary Figure 2B):

$$\in = \mid \text{av. ratings Expected} - \text{av. ratings Surprising} \mid \quad (1)$$

and assessed whether $\in_{\text{Familiar Object}} > \in_{\text{Novel Object}}$, with a paired t-test.

Prior pull

The four attributes that observers rated in our experiments are related to the kinematic properties of objects. Such properties can be best estimated when we observe how an object interacts with another one; for example, if two objects collide and one deforms more than the other, we can tell that one object is softer. Conversely, it should be difficult to judge such properties (hardness, gelatinousness, etc.) in static frames, unless we rely on our previous experiences observing how objects interact. Therefore, one could use the first frame ratings of familiar objects as a measure of prior knowledge about the material qualities from familiar shape and optics associations. Ratings of moving novel objects, in contrast, could be used as a measure of how much the image motion, generated by the kinematics of the material ('sensory' route), influences the rating (no influence of familiar shape, equating for the effect of familiar optical properties). These two conditions make similar predictions (i.e., yield similar ratings) when the material behavior is expected but make different predictions (i.e., yield different ratings) when the material behavior is surprising. We operationally define and measure the "prior pull" as the distance between familiar and novel object motion ratings in the direction of first frame ratings. This property could be stated as a conditional statement: a prior pull occurs if:

$$FF_{\text{Familiar Objects}} > M_{\text{Familiar Objects}} > M_{\text{Novel Objects}}$$

or if

$$FF_{\text{Familiar Objects}} < M_{\text{Familiar Objects}} < M_{\text{Novel Objects}}$$

If one of these two conditions is true, then the magnitude of the prior pull can be computed as:

$$\mid M_{\text{Familiar Objects}} - M_{\text{Novel Objects}} \mid$$

We assess whether prior pull occurs more frequently in the surprising motion condition compared with the expected motion condition.

Response times

The time taken to make each judgment (response time) was measured. We reasoned that rating the material properties of materials that behave surprisingly might involve the reiterative correction of a prediction error by the visual system, and this error correction might be associated with an increase in response time when rating objects that behave in a surprising way. Before computing the difference in response time between expected and surprising conditions, we preprocessed response time data as follows: we subtracted the time to impact (3 seconds static first frame + 0.45 seconds to impact) from the raw response times so that a response time of zero would now indicate time of impact. For each participant, we averaged their response times over objects and attributes. A repeated measures analysis of variance was performed on this averaged response time data to determine whether response times differed between expected and surprise conditions, or familiar and novel object conditions.

Exclusions

For the first 10 participants in the motion experiment, data points for the novel object matched to the surprising key object were excluded owing to a stimulus presentation error that showed the rigid version of the object instead of the wobbling one (10 subjects × 4 attributes = 40 data points excluded). For the remaining 15, subjects this error was fixed.

Data points that were faster than 0.75 seconds after impact (fastest possible button press) were excluded, because a shorter response time than this would indicate that observers started pressing the space bar (for the next trial) before watching the impact frame of the movie (which would be a violation of the instructions). Response latencies that were longer than 2 standard deviations above the mean were also excluded. By opting for a 2*SD cutoff as opposed to the traditional cutoff of 3*SD (Magnussen et al., 1998) we excluded more of the longer RTs, which occurred more for familiar objects that behaved surprisingly (see Supplementary Figure 3). Therefore, this cutoff criterion was more conservative.

After these exclusions, approximately 6% of the data were excluded for response times that were too fast or too slow according to this criterion (approximately 1.3% too fast and approximately 4.7% too slow; see Supplementary Figure 3 for a breakdown of exclusions per experimental condition).

Results

Ratings and the effect of expectation ($\in$)

Figure 5 provides an overview of average observer ratings in in all experiments and conditions. Each of the five polar plots in each row shows average ratings for three (symbolized by icons) of the 15 familiar
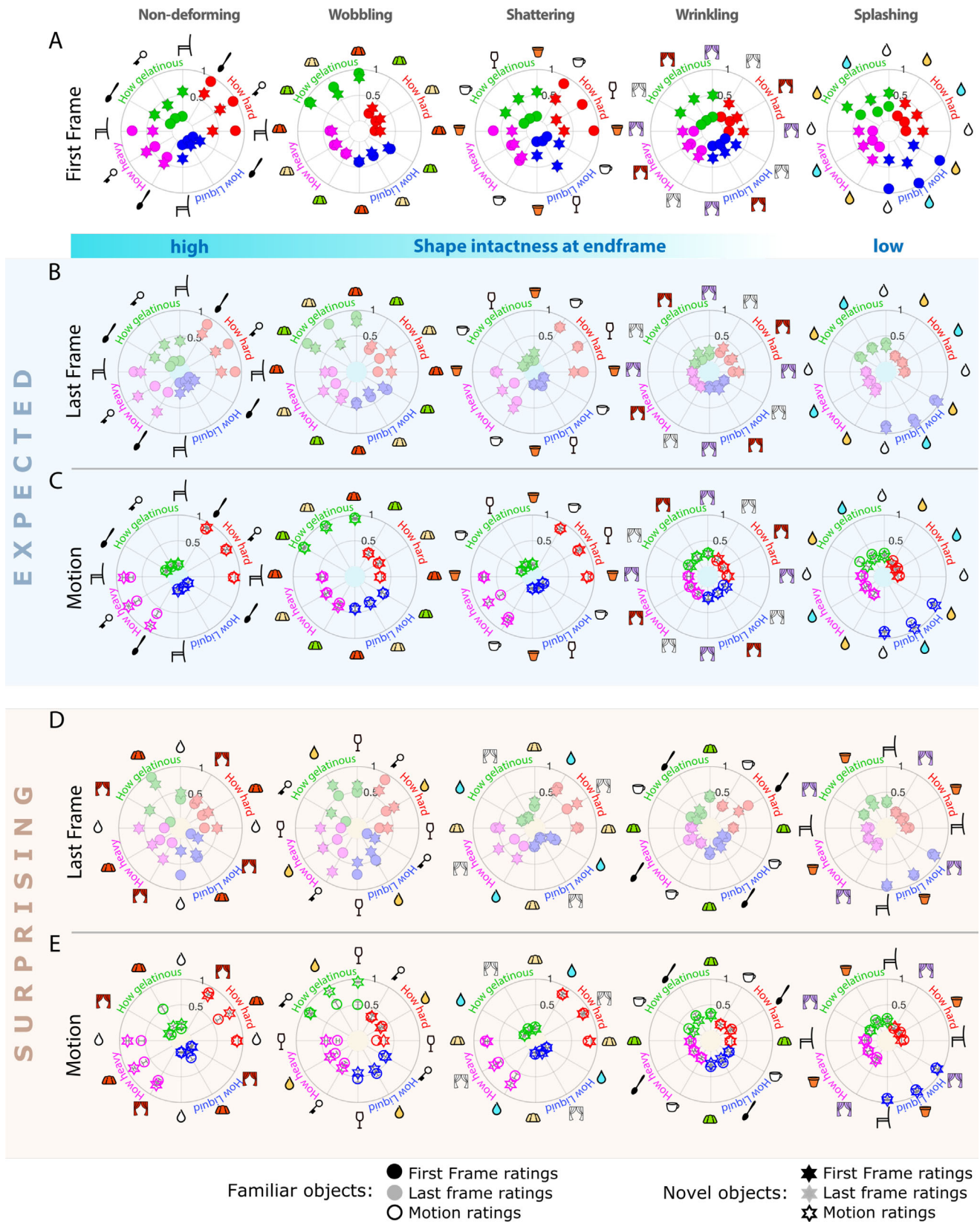
Figure 5. Ratings results from all experiments. Shown are average observer ratings of all experiments for the four different questions about material qualities, for example, how hard, liquid, heavy, or gelatinous an object appears. Each column shows the data across experiments for one particular type of material kinematics (wobble, splash, etc.). Icons symbolize individual familiar objects (chair, key, cup, pot, glass, spoon; *blue droplet*, water; *yellow droplet*, honey; *white droplet*, milk; *violet curtain*, silk; *red curtain*, velvet; *white*

→

←

curtain, linen; *yellow custard*, red and green Jell-O). Each rating question and corresponding data are coded in the same color (*red*, How hard?; *blue*, How liquid?; *purple*, How heavy?; and *green*, How gelatinous?). Ratings could vary between 0 (lowest) and 1 (highest). The *circle* and *star* symbols correspond with ratings of familiar and novel objects, respectively. The symbol style—filled, desaturated, and open—correspond with the different experiments—that is, first frame, last frame, and motion, respectively. Standard deviations denote 1 SE of the mean. Overall, ratings of familiar and novel objects tended to overlap much more in the motion condition (C, E), in particular the expected motion condition (C), than compared with the first frame (A) and last frame (B, D) experiments. In the surprising conditions, the object identity and how the object deforms after falling onto the floor mismatch (e.g., shattering water, splashing chairs, etc.). Data from the first frame (A) and motion conditions (C, E) are replotted in Figure 7 to illustrate the influence of prior expectation on material quality judgements.

objects (*circles*) and corresponding novel objects (*stars*) in response to each of the four questions (how hard [*red*], gelatinous [*green*], heavy [*magenta*], or liquid [*blue*]). The grouping of the objects into one polar plot was done according to how these objects (familiar and corresponding novel) would deform (i.e., nondeforming, wobbling, shattering, wrinkling, or splashing). Note that, in the surprising conditions the object identity and how the object deforms after falling onto the floor mismatch (e.g., shattering water, splashing chairs, etc.) Overall, observers' responses made sense. For example, in Figure 5A static rigid objects were all rated as hard (*red circles*), Jell-O was rated as very gelatinous (*green circle*), and liquids were rated as very liquid (*blue circles*). Ratings for novel objects (*stars*) that inherited material properties from corresponding familiar objects tended to overlap with that of the familiar objects; however, this overlap tended to vary between the different experiments (Figure 5A–5E): for example, overlaps in the expected motion condition (Figure 5C) were much larger than in the first frame (Figure 5A) or last frame conditions (Figure 5B, 5D), likely owing to the different types of information available in each case (e.g., image motion or not, surprising motion or expected motion, etc.). In the regression analysis provided elsewhere in this article, we aim to explain the differences in ratings between familiar and novel objects, allowing for different high and low-level factors to account for the data.

To determine whether object knowledge influences perceived material properties, we computed the effect of expectation Index ($\in$) for familiar and novel objects. We postulated that this effect would be on average different (greater) for familiar objects because expectations about material properties should not be triggered as strongly when the object shape is unfamiliar. This is what we found, $t(19) = 3.41$, $P < 0.003$ (also see Figure 6). This result supports our hypothesis that judgments of material qualities are not based purely on the observed material mechanics, but are also affected by prior knowledge about the typical material mechanics of a (familiar) object. Note that the effect of expectation index for novel objects was not zero, which suggests that observers might generate some predictions about the mechanical
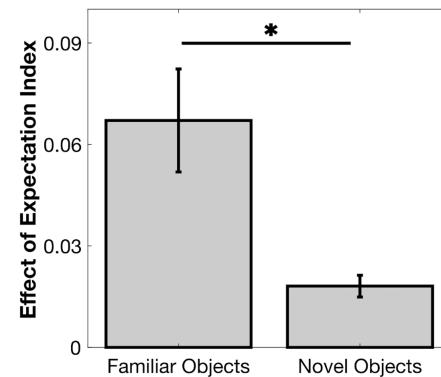

Figure 6. Effect of the expectation index ($\in$). Shown are the averages across participants, material types and rating attributes for familiar and novel objects. $\in$ was calculated as the absolute value of the average rating differences between expected and surprising conditions (see analysis section, Equation 1). Error bars are one standard error of the mean.

properties based on the estimated optical properties of these objects, and/or based on associations they have with the general material properties of a bounded convex shape, for example, a rigidity prior (Grzywacz & Hildreth, 1987; Ullman, 1976; but also see Jain & Zaidi, 2011). We break down these specific influences on material judgements and based on this develop a linear regression model that tries to account for the observed rating differences in expected and surprising conditions.

Prior pull

In some cases, the directionality of the rating differences (expected vs surprising) is directly interpretable. For example, the spoon that wrinkled surprisingly (Figure 5E, fourth plot) was rated as harder (*red circle*, 0.25) than the linen, silk, or velvet curtains that wrinkled expectedly (Figure 5C; 0.18). Prior knowledge about spoons being hard seems to have led to increased ratings of *hardness* compared with their soft curtain counterparts, despite all of these objects wrinkling. Thus, our results suggest that prior object

knowledge about hardness pulled ratings of *hardness* toward this expectation. We visualize this prior pull in the data in Figure 7.

Prior pull can only occur if ratings on a given quality in first frame (*filled dots*) and surprising motion (*open stars*) conditions differ substantially, as indicated by the length of the dark gray lines between these two types of data points in Figure 7. We would expect to find this occurring more frequently in the surprising motion condition, and our results confirm this: first frame ratings differed significantly from novel object motion ratings for 48 of 60 or 80% of conditions in the surprising motion condition, and only 24 of 60 or 40% of conditions in the expected motion condition (also see Table 1). A meaningful prior pull occurs only when the ratings of atypically moving (surprising) familiar (*unfilled circles*) and novel objects (*unfilled stars*) do not overlap and when ratings of atypically moving (surprising) familiar objects are pulled in the direction of first frame ratings (*filled circles*; see Figure 7C, ii). Significant cases of the prior pull effect are highlighted in yellow in Figure 7 (see Supplementary Table 1 for significance tests). If, instead, ratings of moving familiar (*open circles*) and novel objects (*open stars*) overlap completely, this shows that the object shape prior did not exert any significant influence on the rating (see Figure 7C, ii). We found that the more the familiar object remained intact and thus recognizable in the surprising motion condition (Figure 7B, *left plots*), the more likely the prior exerted an influence over the material appearance (also see Supplementary Figure 4, which assesses object intactness).

That only a subset of our objects exhibits a pull toward the prior is consistent with literature that shows that, when sensory input is unambiguous, it will dominate the percept, and prior knowledge does not have an effect (Summerfield & De Lange, 2014). In our case, particular combinations of shape and optics (often involving "shape recognizability") at the end of the animation may lead observers to essentially "ignore" sensory information. For example, the rigid Jell-O is perceived as gelatinous despite not wobbling. Interestingly, prior associations with specific combinations of motion, shape, and optics at the end of the movie may "enhance" differences between familiar and novel object ratings (e.g., in the case of the wobbling glass, the particular combination of translucent optics forming what looks like a puddle might evoke strong representations of liquidity, so despite the wobble, it is not considered gelatinous).

One might argue that the reason why we see relatively few instances of significant prior pull is that novel objects might also elicit expectations (after all, they are bounded shapes with particular optical properties). The correlations in Figure 8 between first frame and motion ratings, however, suggests that, for the most part, this is not the case; although familiar object

first frame ratings predicted expected motion ratings extremely well (Figure 8A; $R^2 = 0.9$), novel object first frame ratings predict motion ratings very poorly (Figure 8B; $R^2 = 0.07$). These results are also supported by the average interobserver correlations in Figure 8C. Interobserver correlation is much lower for novel objects in the first frame experiment than for familiar objects suggesting that static images of novel objects do not elicit consistent (typical) expectations about kinematic material properties across observers. Thus, for the most part, familiar object first frame ratings are a good measure of observers' prior expectations about the material qualities of these objects. In contrast, novel object motion ratings are an ideal measure of sensory input because the motion from the material behavior includes information about both the mechanical deformations, as well as any additional effect of image motion from optics (Doerschner et al., 2011a, b).

Linear regression models

Given that there are a few cases where novel objects do seem to generate correct predictions about the material outcome (those at the bottom left and top right of Figure 8B), and because some of the magnitude of the prior pull may be explained by shape recognizability at the end of the movie, we tested a linear regression model that predicted the direction and magnitude of the difference between ratings of moving familiar and novel objects from optics and shape prior pull, and shape recognizability at the end of the animation (Figure 9). From our data we modelled 3 different association routes that could potentially influence the rating: optics associations (H1), shape associations at the beginning of the movie (H2), and shape associations at the end of the movie (H3).

Each predictor in the high-level model was multiplied by a weight G that took into account the predictability of the "behavior" of a given stimulus (computed as the difference between first frame and motion ratings in the expected condition) and the extremeness of the rating (computed as the twice the difference between first frame ratings in the expected condition and 0.5). Thus, predictors in the high-level model were: H1 = H1 × G, where G = predictability × extremeness: The optics and bounded shape prior (computed as the difference between first frame ratings of novel objects in the expected condition and moving stimuli ratings of novel objects in the expected condition). The rationale is that this predictor captures to what extent the reflectance properties and the fact that observers see a bounded shape will lead to prior expectations about the material kinematics. H2 = H2 × (1 − G): The shape familiarity (computed as the overall differences between first frame ratings of familiar objects and corresponding ratings of moving novel objects (across both expected
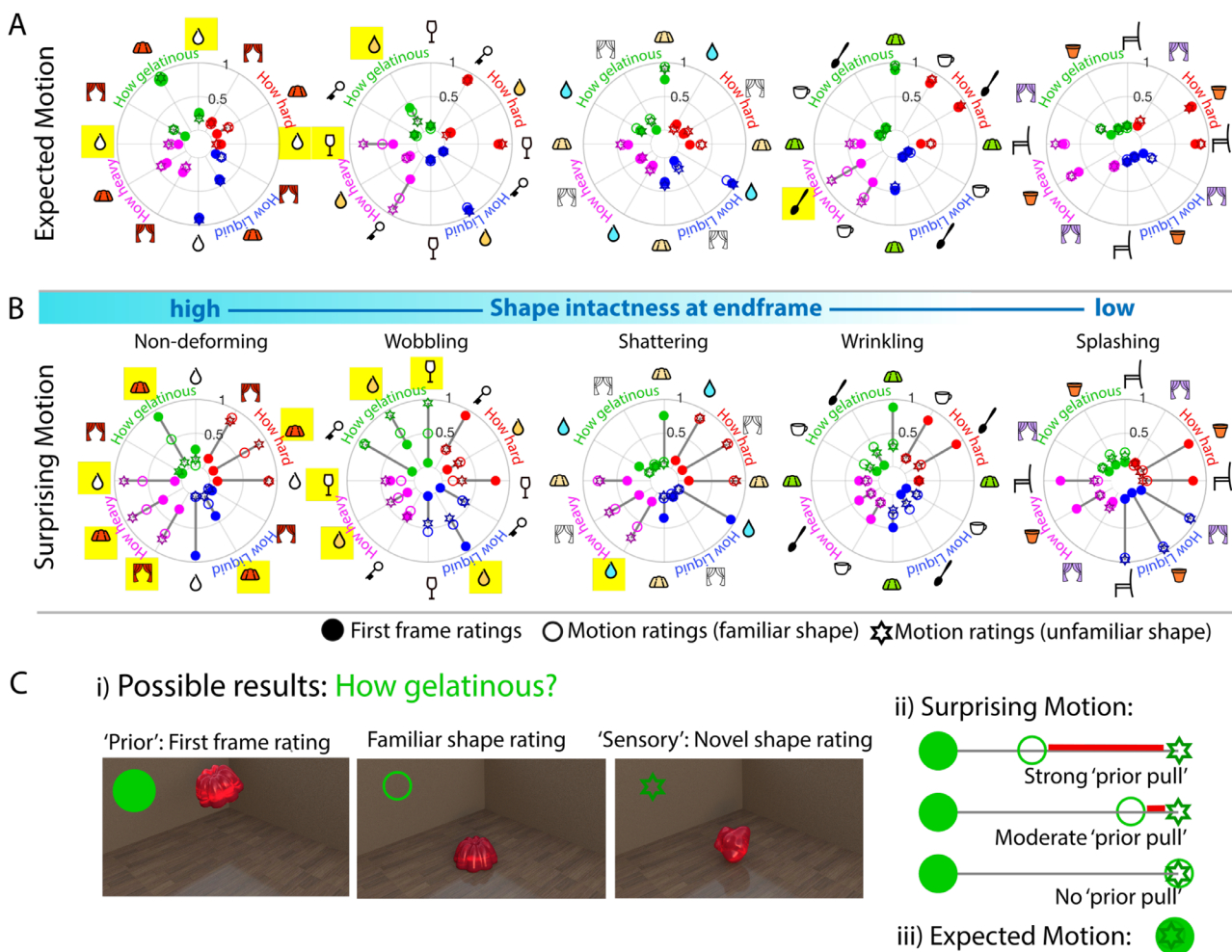
Figure 7. Material quality ratings and prior attraction. Results of the first frame and motion conditions from Figure 5 are replotted, keeping the same symbols and notation. (A) Average observer ratings from three conditions (i.e., first frame familiar objects [*filled circles*], typically behaving familiar objects [*unfilled circles*] and corresponding [*moving*] novel objects [*unfilled stars*]) tended to overlap. The difference between first frame ratings for familiar objects and ratings of moving novel objects is indicated by a *dark grey line*. The organization of objects follows that in (B). (B) Same as (A), but here, ratings of atypically behaving familiar objects are plotted as *unfilled circles* (organized by type of motion), and ratings of corresponding novel objects—that is, unfamiliar shapes—which inherit their optical and kinematic qualities from a familiar object—as *unfilled stars*. The motions are arranged according to how much the object remains intact and recognizable after impact on the floor (also see Supplementary Fig. 4). *Yellow highlighted symbols* show a statistically significant prior pull. See main text for more detail. The yellow highlighted cases show that prior pull occurred more in conditions where the object was still intact and recognizable at the end of the movie (objects that behaved rigidly or wobbled). Supplementary Table 1 lists corresponding statistics and *P* values. (C, i) How we measure how much the rating of an atypically moving familiar object (*middle*) overlaps with the rating of a material-matched moving novel object (*right*), or conversely, how much it is pulled toward ratings of a static view of the familiar object (*left*). (C, ii) Possible results. For example, seeing an image of red Jell-O in its classical shape, observers tend to expect that it is quite gelatinous. When they see an object with the same optical properties that falls and does not wobble when it hits the floor, they rate it as very nongelatinous, that is, we have a large rating difference (*gray line*). When a classically shaped red Jell-O falls on the floor and does not wobble, observers could either rate it similar to the novel object—after all it does not wobble at all (no prior pull)—or it could be rated as somewhat more gelatinous, despite the sensory input, possibly because prior experience influences the appearance, making observers perceive wobble when there is not (prior pull, *red line*). (C, iii) When the familiar object moves exactly as expected, and when there is no strong influence of shape familiarity on material judgements, all three ratings will overlap.
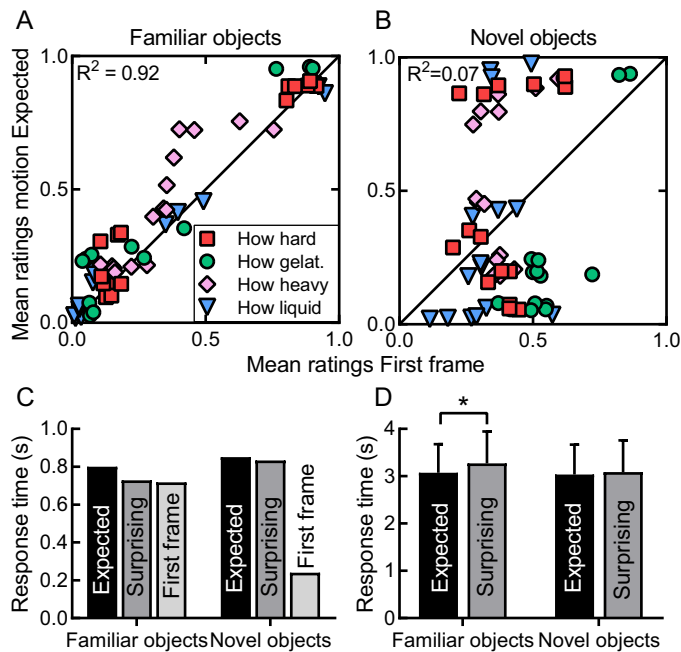
Figure 8. Prediction strength response time differences and interobserver correlations. (A) Correlation between mean first frame ratings and mean expected motion ratings for familiar and novel objects. (B) A high correlation indicates that the first frames (still images) of objects are highly predictive of the objects' kinematic properties, and thus are in good agreement with ratings in the expected motion condition, where objects fall and deform according to their typical material kinematics. This is clearly not the case for novel objects, suggesting that these objects do not elicit strong prior expectations about how an object will deform. (C) Average interobserver correlation for expected and surprising motion trials, as well as the first frame experiment. Note that only for novel objects, this latter correlation was quite low, suggesting that still images of unfamiliar objects do not elicit a strong prior in observers about the material qualities measured in this experiment. (D) Response time data averaged across all observers for expected (*black*) and surprising trials (*medium gray*). Stars indicate significant differences, *P* < .001. Error bars are 1 standard error and show variability between subjects.

and surprising conditions). The rationale is that this predictor captures to what extent expectations generated by the familiar shape of the object can explain the rating differences, not controlling for the effects of optics and bounded shape (which are captured by H1). $H3 = H3 \times (1 - G)$: The last frame shape recognizability (computed as the overall differences between last frame ratings of familiar and novel objects (across both, expected and surprising conditions). The reason for including this predictor is our hypothesis that the conflict between prior and surprising motion should be larger if the object would be still recognizable

after falling to the ground (as in the wobbling or rigid motion conditions; see Table 1).

We can model the data moderately well with this high-level model (see analysis, R^2 between 0.266 and 0.59, depending on the question), but not perfectly, potentially owing to specific motion–shape–optics interactions (Schmid & Doerschner, 2018).

Although retinal size was approximately equated between novel and familiar objects, their area and 3D volume were not. Moreover, shape and volumes within familiar object stimuli varied quite a bit. A voluminous sphere-like object will behave quite differently than a spoon (imagine how much each of them would splash or wobble upon impact if made from liquid or jelly). Therefore, such stimulus variations might to some degree account for rating differences between the familiar and corresponding novel objects and might thus provide alternative explanations to our effect of expectation hypothesis. To test the degree to which such low-level effects account for the rating data we developed a low-level regression model, which contained three predictors that aimed to capture the most salient low-level differences between stimuli.

L1: Motion energy difference (this was computed in two steps): (1) Take the sum of the absolute value of all consecutive image differences, starting with the impact frame, for example, sum(abs(f11–f12), abs(f12–f13), …, abs(f39–f40)) (Tschacher et al., 2014), for all experimental conditions/stimuli (60 in total) and normalize these values by object size (number of pixels corresponding to the object on frame 11 MEn), (2) For each object in the expected and surprising conditions, compute the differences of MEFamiliar and MENovel conditions. Familiar and novel objects were often different in 3D volume and shape; thus, differences in ratings might be attributable to resulting differences in motion energy when these two classes of objects deform.

L2: First frame object size (computed as number of pixels in the first frame corresponding to the object). Differences in the area taken up by familiar and novel objects in the first frame might account for rating differences, for example, if the familiar object was smaller, it might have generated predictions to be less heavy than the corresponding novel object (irrespective of the familiar shape of the object).

L3: Last frame object size (computed as numbers of pixels in the last frame corresponding to the object). This predictor is related to H3 in the high-level model (discussed elsewhere in this article). Differences in the area taken up by familiar and novel objects in the last frame might account for rating differences, for example, if the novel object 'splashed' more (thus took up more area in the last frame) it might have been rated more liquid simply owing to this fact.

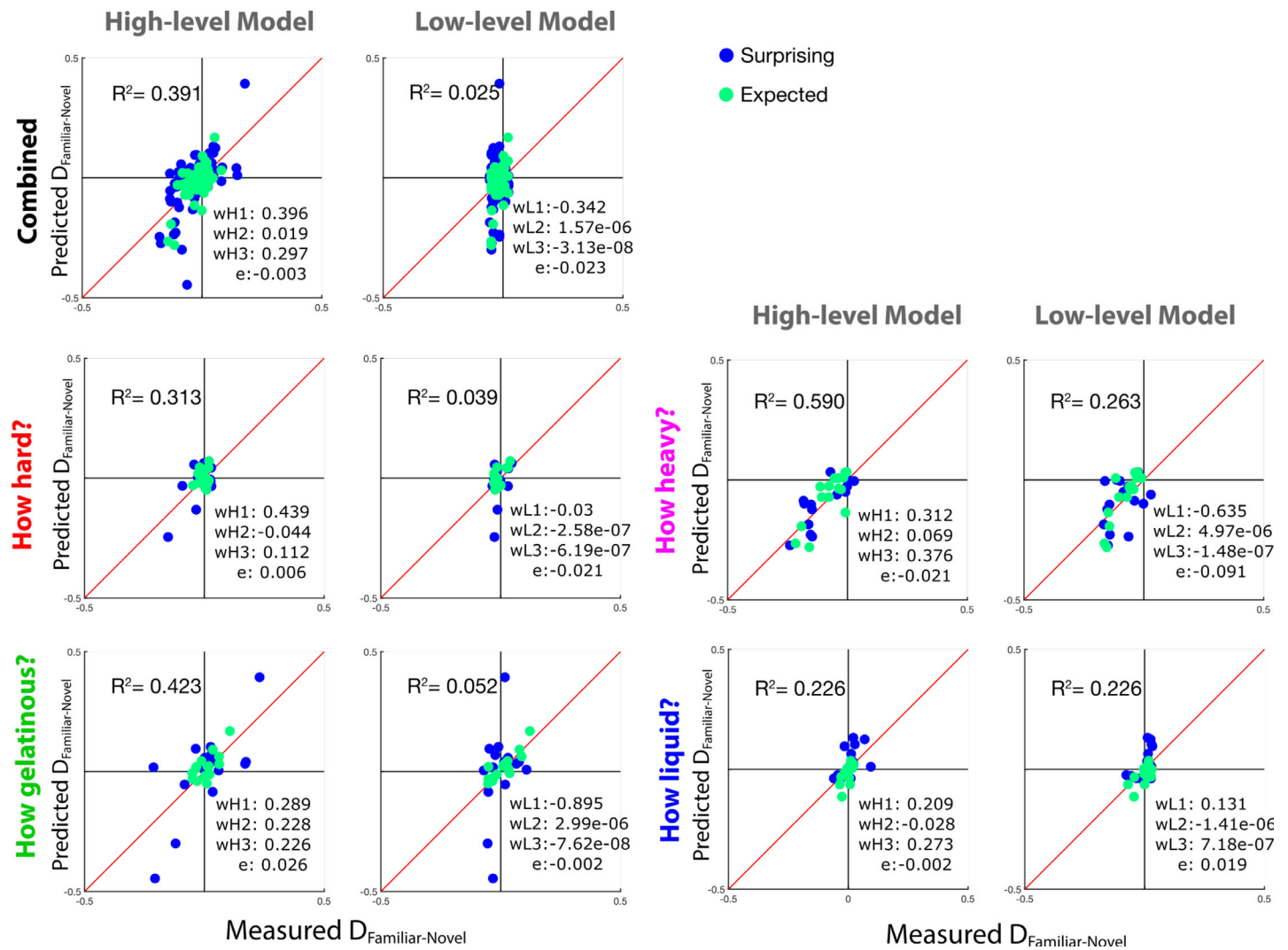Such a model performs extremely poorly, $R^2 = 0.025$, $P > 0.05$ (Figure 9). This finding suggests that rating

Figure 9. Low and high level linear regression models predicting the difference between ratings of moving familiar and control objects $D_{familiar-novel}$. We developed two models with the aim to account for the differences we observed in ratings of familiar and novel moving objects. The computation of the individual predictors is described in the main text (Analyses section). In the lower right inset of each plot we show the weights (w) of each predictor (H1: bounded shape and optics prior; H2, shape familiarity; H3, last frame shape recognizability; L1, motion energy difference [also see Supplementary Figure 6]; L2, object size differences first frame; L3, object size differences last frame). Overall, the high-level model was more successful in predicting this difference than the low-level model (top two panels: combined). However, this pattern varied as a function of rating questions: the high-level model performed best for ratings of gelatinousness and hardness, whereas the low-level model performed as good as the high-level model for ratings of liquidness. The latter is likely due to the fact that ratings of liquidness might be strongly modulated by how much a substance physically spreads in the image.

differences are not due to differences in object size or image motion.

Response times

We find a small but significant increase in response time in the familiar object surprising condition—which is the condition that most strongly juxtaposes prior expectation with sensory evidence (Figure 8D). There was no significant response time difference between expected and surprising novel objects conditions. This was assessed with a two-way repeated measures analysis of variance, which revealed a main effect of object familiarity (participants took longer to respond to familiar versus novel objects), $F(1, 24) = 14.72$, $P = 0.0008$, and a main effect of surprise (participants took longer to respond to surprising versus expected

events), $F(1, 24) = 13.24$, $P = 0.013$. However, these main effects must be interpreted in light of the significant interaction between object type and surprisingness, $F(1, 24) = 6.395$, $P = 0.0184$. Follow-up *t*-tests (Bonferroni corrected) show that participants took longer to respond to surprising versus expected stimuli when objects were familiar, $t(24) = 4.911$, $P < 0.0001$. However, this difference was not significant for novel objects, $t(24) = 1.334$, $P = 0.1946$.

Discussion

Visual perception is not a one-way (bottom-up) road; how we process visual input is influenced by expectations about the sensory environment, which develop from our previous experience and learning about existing regularities in the world, that is, associating things or events that co-occur. Expectations have been shown to facilitate visual processing in the case of priming, to modulate the frequency of a particular percept in b-stable stimuli, and to change our interpretation of ambiguous stimuli (see Kveraga et al. [2007] or Panichello et al. [2013], for a review). However, the stimuli used in these experiments have been fairly simple (static images of objects), and it has been shown that learning associations can also include fairly complex phenomena. For example, recently, Bates et al. (2015; also see Battaglia et al., 2013; Kubricht et al., 2016) showed that humans can learn to predict how different liquids flow around solid obstacles (also see other examples for predicting of motion trajectories of rigid objects; e.g., Flombaum et al., 2009; Gao & Scholl, 2010; Soechting et al., 2009). Although the authors attributed human performance to an ability to "reason" about fluid dynamics, here we explicitly test whether existing perceptual expectations about material properties can set up rather complex predictions about future states, and whether—and to what extent—these expectations influence material appearance. We show that the qualities of surprising materials (Itti & Baldi, 2009) are perceived different to expected ones that behave the same (Figure 7), and that surprise leads to increases in processing time of the stimuli (Figure 8D, in line with Baldi & Itti, 2010; Itti & Baldi, 2009; or recently Urgen & Boyaci, 2019). Our method provides a general technique to differentiate the extent to which material qualities are directly estimated from material kinematics versus being modulated by prior associations from familiar shape and optical properties, and contributes thus an essential piece to the puzzle of how the human visual system accomplishes material perception (Adelson, 2001; Anderson, 2011; Fleming, 2014; Fleming et al., 2015; Komatsu & Goda, 2018; Maloney & Brainard, 2010).

Knowledge affects (material) perception

There are countless demonstrations showing that knowledge affects how we perceive the world: from detecting the dalmatian amidst black and white blotches, identifying an animal in the scene (Thorpe et al., 1996), deciding on the identity of a blob (Oliva & Torralba, 2007), or being a greeble or bird expert (Gauthier et al., 2003), knowledge directly influences our ability to perform in these instances. Object knowledge does not just facilitate categorical judgments, it also affects the estimation of visual properties such as color or motion (Hansen et al., 2006; Olkkonen & Allred, 2014; Scocchia et al., 2013). How exactly knowledge alters and facilitates neural processes in visual perception is a topic of ongoing research (e.g., Gauthier et al., 2000; Kveraga et al., 2007; Rahman & Sommer, 2008).

In contrast, the role of predictions or associative mechanisms in material perception is not well understood (Paulun et al., 2017; Schmidt et al., 2017; van Assen et al., 2018). Knowledge about materials entails several dimensions and can include taxonomic relations: gold is a metal, metals are elements with physical properties, metals are usually malleable and ductile. These classes of metal also have their own perceptual regularities: gold looks yellowish, often has a very shiny, polished and smooth appearance, feels cool to the touch, and so on. Our experimental results suggest that identifying a material (i.e., knowing what it is) not only coactivates its typical optical qualities, but also elicits strong predictions about the typical kinematic properties and resulting material behaviors. For example, liquids are not only translucent or transparent, they also tend to run down, splash, or ooze. Importantly, we seem to have quite specific ideas of what running down, splashing, or oozing should look like (e.g., Dövencioglu et al., 2018 probed such ideas explicitly), supposedly because we have ample visual (but also haptic) experiences with liquids, and thus opportunities to learn the regularities (statistical or other) associated with a specific material category. Our results show that these specific ideas, or priors, about material behaviors interfere with the bottom-up processing of visual information, leading to predictable differences in ratings of material properties between expected and surprising conditions.

Interestingly, Sharan et al. (2014) measured reaction times in a material categorization task while participants judged objects made from real (e.g., a cupcake made from dough) and fake materials (e.g., a knitted cupcake figure). They found a substantial decrease in reaction times at very short presentation times in the real condition. Given our results that show that the response time increases in the surprise condition for familiar objects, it would be interesting to know if their decrease is primarily driven by an increase in reaction time

for the surprising stimuli in their experiment, that is, fake ones—because for those objects, the object shape/identity (a natural/edible thing) was in conflict with the material of the object (plastic/inedible).

Prior pull

The yellow highlighted cases in Figure 7 show that prior pull occurred twice as much in the surprising than the expected condition, and more in conditions where the object was still intact and recognizable at the end of the movie (objects that behaved rigidly or wobbled). Prior pull in the expected condition also occurred where estimation (sensory input) and associative (prior knowledge) accounts made different predictions (gelatinousness of the honey, heaviness of the wine glass, Figure 7A). In these cases, the expected events were not so expected—and this phenomenon may be related to shape properties or the size of the splashing of the liquids. Although we controlled for the effects of image motion from optics (e.g., specular highlights), perhaps other low-level image differences exist between familiar and novel objects that could be driving differences in ratings. However, we were able to rule this out, because a linear regression model that incorporates these low-level effects as predictors performs extremely poorly. This result suggests that rating differences were not due to differences in object size or image motion.

We did not aim to test an exhaustive list of material attributes, but to determine whether effects of prior associations on visual input might depend on the type of material attribute judged, and on how the object behaves under external forces. We found that mechanical qualities like hardness and liquidity seem to be more directly estimated from material kinematics in "shape-destroying" conditions (splashing, shattering, wrinkling), but prior associations play a modulatory role—to the extent where material kinematics can even be ignored (e.g., red and green Jell-O)—when shape remains somewhat intact. These latter conditions seem to create more of a cue conflict and are more ambiguous. On the other hand, qualities like gelatinousness and heaviness (which are much more difficult to estimate directly from mechanical deformations) were more affected by familiar shape and optics associations.

Alternative explanations for rating differences

One might argue that the prior pull we demonstrated here is not perceptual, but in fact is due to a particular cognitive strategy of some observers (i.e., explicitly ignoring the motion information and thus rating material qualities of atypically moving familiar objects

as they rated objects on the first frame, while other observers' ratings were 100% identical to novel object ratings). This would have resulted in bimodal rating distributions and/or low interobserver correlation in the object motion condition, neither of which we found (Supplementary Figure 5 and Figure 8C, respectively).

Another argument against the cognitive strategy approach is supported by response time patterns in our experiments. A small but significant increase in response time in the familiar object surprising condition—which is the condition that most strongly juxtaposes prior expectation with sensory evidence—would be consistent with the idea of recurrent prediction error correction (Urgen & Boyaci, 2019; Figure 8D). Importantly, we do not find evidence for a response time advantage in expected familiar objects condition compared with novel, also consistent with Urgen and Boyaci (2019), which suggests that this increase in response time cannot simply be due to the fact that observers positioned the slider in advance to the wrong position, that is, to a position that would be more consistent with a rating based on the first frame information only. If observers adopted such a strategy, we should have also seen faster response times for expectedly moving familiar objects.

Although we do not believe that cognitive strategies were driving our results per se, we do acknowledge that it is unlikely that ratings tasks directly measure the appearance of materials, because it is not known what the relevant perceptual dimensions of visual experience are. Finding the relevant perceptual dimensions of materials is an active area of research (e.g., Schmid & Doerschner, 2018; Toscani et al., 2019). Many studies have investigated the visual perception of properties like softness or gloss (see Fleming, 2017, for a review), and we believe our participants' ratings in our study results reflect perception to the same degree as these other studies.

It is quite striking that the same material deformations were rated differently in the expected and surprising conditions. Toscani et al. (2013) showed that depending on the task, observers pointed their gaze at specific points on the stimulus, for example, near the brightest regions on an object for lightness judgements. One possibility could thus be that the task, object knowledge, and expectations about the material behavior guided eye movements of observers in our experiment to specific locations on the stimulus. Although in the expected conditions fixation patterns might have been optimal with respect to the task, for example, observers correctly anticipated how the object would deform (or shatter, splash, wobble, etc.), it is possible that in the surprising conditions the wrong expectation guided eye movements to the wrong locations on the object, which in turn lead to a different sampling of information and ultimately influenced

their judgements of material qualities. We are currently investigating this possibility directly.

A Bayesian account of rating and reaction time differences

We believe that the results of this study fit well within the Bayesian framework, which offers an account of how prior knowledge is integrated with sensory input. Our experiment constitutes a situation not unlike classical cue-conflict experiments (e.g., Ernst & Banks, 2002; Knill, 2007b), where the sensory cues may conflict with one another and/or the prior belief. Although we do not aim to model our results formally, we still believe that this analogy is useful in interpreting our findings. We first focus on the rating differences that we found in the expected and surprising conditions. In the latter, in many instances ratings were pulled toward the expected material property, not the signaled one. For example, a wrinkling spoon was rated harder than any of the wrinkling control objects. In this latter example, we have two cues to object hardness: object shape (a spoon) and the object behavior (it is wrinkling). Here, the cues to hardness are in conflict: the (familiar) shape of the object (in the first frame, which observers saw for three seconds) suggest a very hard object, whereas the subsequent motion information when the spoon impacts suggests a very soft material. In this situation the visual system has two possibilities. The first is that, in light of the strong sensory (i.e., image motion) information, it could simply completely reject the idea that this object has any degree of hardness and veto the familiar shape cue completely, which classic cue combination would suggest (e.g., outlier rejection; Landy et al., 1995). As a result, we would see no difference between wrinkling spoon, wrinkling curtains or wrinkling novel objects. The second possibility is that the visual system entertains multiple priors (strong and weaker ones) about the state of the world, and that, depending on the sensory input, it adjusts the weights of these priors (Knill, 2007b). In our concrete example, this would mean that the visual system may entertain multiple priors on spoon hardness—say from softish to very hard, and that instead of vetoing the idea that the spoon was ever hard, the hardness cue from a familiar shape is only weighed down and integrated with other available cues (e.g., image motion information). This would yield a hardness rating different from that of the wrinkling control objects, or curtains, and this is what we see in our data.

Thus, the robust cue integration model seems to offer a good explanation for the differences in rating data between novel and familiar objects in the Surprise condition. This model was originally developed by Knill (2007b) to account for observer behavior in situations with large cue conflicts, in which our experiment clearly is. However, it seems to only account well for the data when the shape of the object is still recognizable toward the end of the animation (Figure 7; e.g., nondeforming, or wobbling in the surprise condition). When instead the shape is unrecognizable upon impact (e.g., when splashing), we see more of a situation analogous to cue-vetoing, that is, there is no difference in perceived material quality between familiar and novel object ratings. However, a destruction of the shape has made the (familiar) shape cue completely unavailable which might have changed the integration situation entirely, which makes an interpretation in these cases difficult.

When faced with violated expectations, as in our surprising condition, the visual system needs to update the generative model to minimize the prediction error (i.e., the error between the expected state and measured state of the world) to perform future tasks (Urgen & Boyaci, 2019). Because this updating is a reiterative process, we reasoned that it would take observers longer to perform perceptual tasks when judging material attributes of surprisingly deforming objects. This is exactly what we found. One might criticize that response times for ratings were much longer than times measured in classic reaction times studies (e.g., see a review by Eckstein, 2011, on visual search). However, it is not all that uncommon to consider response times of 2 seconds and longer, as in categorical color perception (Boynton & Olson, 1990; Okazawa et al., 2011). Note that we treated response time data as conservatively as classic reaction time studies, for example, by removing data points that were two standard deviations above the mean.

Our study bears resemblance to the work by Fujisaki et al. (2014), who investigated how different kinds of information sources, namely visual and auditory, are combined in material categorization and material property rating tasks. Some of their stimulus conditions were not unlike our cue conflict scenarios (e.g., combining a visual glass stimulus with a bell pepper sound). They found that for the material rating task, the integration of the two types of information follows a weighted average rule, where the weights depend on the reliability of the respective signals. This reliability was in part related to the task: for example, participants gave higher weights to visual cues for judgements of color and gave higher weights to auditory cues for judgments of pitch or hardness. Also, in our results in Figure 9 we see systematic changes in the regression weights of both, high- and low-level models as a function of rating task: for example, in the low-level model, the visual cue motion energy (wL1) receives more weight explaining rating differences between familiar and novel objects than for ratings of, for example, heaviness, or similarly in the high-level model: the predictor shape familiarity (wH2) played a much larger role for explaining rating

differences between familiar and novel objects for ratings of gelatinousness that for ratings of liquidness. To determine whether this shift in weights is indeed associated with the reliability of the respective cues requires further experimentation, where cue reliability would be manipulated directly.

Estimation versus association

Our work shows that previously acquired object–material associations play a central role in material perception and are much more sophisticated than previously appreciated. Previously there have been conflicting findings in the literature about the relative influence of optics, shape, and motion cues to the perception of material properties (e.g., Aliaga et al., 2015; Schmid & Doerschner, 2018; van Assen & Fleming, 2016). For example, although some work proposes that perceived material qualities like softness are strongly influenced by motion and shape cues, which completely dominate optical cues (Paulun et al., 2017), other work showed that both optical and mechanical cues affect estimates of softness (Schmidt et al., 2017). Schmidt et al. (2017) suggested that shape recognizability after deformation (e.g., recognizing that an object used to be a cube in Paulun et al.'s study) affects how reliable shape cues versus surface optical cues are when judging material properties, thus leading to shape cues dominating over optical cues in Paulun et al.'s (2017) study. Our results back up the idea that familiar shape affects observers' ratings of material properties. The interactions between familiar shape, optical and motion properties is something that future material perception studies should consider and investigate further.

This study not only shows that existing perceptual expectations about material properties can set up rather complex predictions about future states of materials, it also extends a growing theme in the material perception literature that studying the perception of kinematic material qualities can serve as a tool to guide investigations of the neural mechanisms about material properties, because it provides insight into components (high and low level) that make up material perception as a whole (Schmid & Doerschner, 2019).

Conclusions

This work shows that the visual system can predict the future states of rigid and nonrigid materials. Such predictions can be activated by the shape of an object and—to a lesser extent—also by the optical qualities of a surface. Understanding how high-level expectations are integrated with incoming sensory evidence is an essential step toward understanding how the human visual system accomplishes material perception.

Keywords: material perception, kinematics, prediction, expectation violation, cue conflict

References

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

Aliaga, C., O'Sullivan, C., Gutierrez, D., & Tamstorf, R. (2015). Sackcloth or silk? The impact of appearance vs dynamics on the perception of animated cloth. In L. Trutoiu, M. Guess, S. Kuhl, B. Sanders, & R. Mantiuk (Eds.), *ACM SIG- GRAPH Symposium on Applied Perception* (pp. 41–46). New York: ACM, doi:10.1145/2804408.2804412.

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

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks, 23*(5), 649–666, https://doi.org/10.1016/j.neunet.2009.12.007.

Bates, C. J., Yildirim, I., Tenenbaum, J. B., & Battaglia, P. W. (2015). Humans predict liquid

dynamics using probabilistic simulation. In *Cognitive Science Society* (pp. 172–177), Retrieved from https://pdfs.semanticscholar.org/1867/112a8deb73d72d5e08553e7d065346b7308f.pdf?_ga=2.153744048.96345254.1562848276-1034217185.1562848276.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences of the United States of America, 110*(45), 18327–18332, https://doi.org/10.1073/pnas.1306572110.

Bi, W., Jin, P., Nienborg, H., & Xiao, B. (2019). Manipulating patterns of dynamic deformation elicits the impression of cloth with varying stiffness. *Journal of Vision, 19*(5), 18, https://doi.org/10.1167/19.5.18.

Boynton, R. M., & Olson, C. X. (1990). Salience of chromatic basic color terms confirmed by three measures. *Vision Research, 30*(9), 1311–1317, https://doi.org/10.1016/0042-6989(90)90005-6.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436, Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9176952.

Chang, S., & Pearson, J. (2018). The functional effects of prior motion imagery and motion perception. *Cortex, 105*, 83–96, https://doi.org/10.1016/j.cortex.2017.08.036.

Debevec, P. (1998). Rendering synthetic objects into real scenes. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques - SIGGRAPH '98* (pp. 189–198). New York, NY: ACM Press, https://doi.org/10.1145/280814.280864.

Doerschner, K., Kersten, D., & Schrater, P. R. (2011a). Rapid classification of specular and diffuse reflection from image velocities. In *Pattern Recognition* (Vol. *44*, pp. 1874–1884). Oxford, UK: Pergamon, https://doi.org/10.1016/j.patcog.2010.09.007.

Doerschner, K., Fleming, R. W., Yilmaz, O., Schrater, P. R., Hartung, B., & Kersten, D. (2011b). Visual motion and the perception of surface material. *Current Biology, 21*(23), 2010–2016, https://doi.org/10.1016/j.cub.2011.10.036.

Dövencioglu, D. N., van Doorn, A., Koenderink, J., & Doerschner, K. (2018). Seeing through transparent layers. *Journal of Vision, 18*(9), 25, https://doi.org/10.1167/18.9.25.

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision, 11*(5), 14–14, https://doi.org/10.1167/11.5.14.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429–433, https://doi.org/10.1038/415429a.

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

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

Fleming, R. W., Nishida, S., & Gegenfurtner, K. R. (2015). Perception of material properties. *Vision Research, 115*, 157–162, https://doi.org/10.1016/J.VISRES.2015.08.006.

Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In *The Origins of Object Knowledge* (pp. 135–164). Oxford, UK: Oxford University Press, https://doi.org/10.1093/acprof:oso/9780199216895.003.0006.

Fujisaki, W., Goda, N., Motoyoshi, I., Komatsu, H., & Nishida, N. (2014). Audiovisual integration in the human perception of materials. *Journal of Vision, 14*(4), 12, 1–20, https://doi.org/10.1167/14.4.12.

Gao, T., & Scholl, B. J. (2010). Are objects required for object-files? Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition, 18*(1), 82–109, https://doi.org/10.1080/13506280802614966.

Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology, 20*(3–6), 507–523, https://doi.org/10.1080/02643290244000275.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*, 191–197, https://doi.org/10.1038/72140.

Grzywacz, N. M., & Hildreth, E. C. (1987). Incremental rigidity scheme for recovering structure from motion: Position-based versus velocity-based formulations. *Journal of the Optical Society of America. A, Optics and Image Science, 4*(3), 503–518, https://doi.org/10.1364/josaa.4.000503.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience, 9*(11), 1367–1368, https://doi.org/10.1038/nn1794.

Hillis, J. H., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science, 298*(5598), 1627–1630, https://doi.org/10.1126/science.1075396.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research,*

49 (10), 1295–1306, https://doi.org/https://doi.org/10.1016/j.visres.2008.09.007.

Jain, A., & Zaidi, Q. (2011). Discerning nonrigid 3D shapes from motion cues. *Proceedings of the National Academy of Sciences of the United States of America, 108*(4), 1663LP–1668LP, https://doi.org/10.1073/pnas.1016211108.

Kawabe, T. (2018). Linear motion coverage as a determinant of transparent liquid perception. *i-Perception, 9*(6), 204166951881337, https://doi.org/10.1177/2041669518813375.

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

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*(1), 271–304, https://doi.org/10.1146/annurev.psych.55.090902.142005.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception, 36*(14), 1–16.

Knill, D. C. (2007a). Learning Bayesian priors for depth perception. *Journal of Vision, 7*(8), 13, https://doi.org/10.1167/7.8.13.

Knill, D. C. (2007b). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision, 7*(7), 5, https://doi.org/10.1167/7.7.5.

Komatsu, H., & Goda, N. (2018). Neural mechanisms of material perception: Quest on Shitsukan. *Neuroscience*, *392*, 329–347, https://doi.org/10.1016/j.neuroscience.2018.09.001.

Kubricht, J., Jiang, C., Zhu, Y., Zhu, S.-C., Terzopoulos, D., & Lu, H. (2016). Probabilistic simulation predicts human performance on viscous fluid-pouring problem. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 1805–1810). Retrieved from https://www.seas.upenn.edu/∼cffjiang/research/cogsci16/nips16.pdf.

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition, 65*(2): 145–168, https://doi.org/10.1016/j.bandc.2007.06.007.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35*(3), 389–412, https://doi.org/10.1016/0042-6989(94)00176-M.

Magnussen, S., Idås, E., & Myhre, S. H. (1998). Representation of orientation and spatial frequency in perception and memory: A choice reaction-time analysis. *Journal of Experimental Psychology.*

Human Perception and Performance, 24(3), 707–718, https://doi.org/10.1037//0096-1523.24.3.707.

Maloney, L. T., & Brainard, D. H. (2010). Color and material perception: Achievements and challenges. *Journal of Vision, 10*(9), 19–19, https://doi.org/10.1167/10.9.19.

Mao, R., Lagunas, M., Masia, B., & Gutierrez, D. (2019). The effect of motion on the perception of material appearance. In *ACM Symposium on Applied Perception 2019* (pp. 1–9). New York, NY: ACM, https://doi.org/10.1145/3343036.3343122.

Marlow, P. J., & Anderson, B. L. (2016). Motion and texture shape cues modulate perceived material properties. *Journal of Vision, 16*(1), 1–14, https://doi.org/10.1167/16.1.5.

Okazawa, G., Koida, K., & Komatsu, H. (2011). Categorical properties of the color term "GOLD". *Journal of Vision, 11*(8):4, 1–19, https://doi.org/10.1167/11.8.4.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*(12), 520–527, https://doi.org/https://doi.org/10.1016/j.tics.2007.09.009.

Olkkonen, M., & Allred, S. R. (2014). Short-term memory affects color perception in context. *PLoS One, 9*(1), e86488, https://doi.org/10.1371/journal.pone.0086488.

Panichello, M. F., Cheung, O. S., & Bar, M. (2013). Predictive feedback and conscious visual experience. *Frontiers in Psychology, 3*(620), 1–8, https://doi.org/10.3389/fpsyg.2012.00620.

Paulun, V. C., Kawabe, T., Nishida, S., & Fleming, R. W. (2015). Seeing liquids from static snapshots. *Vision Research*, *115*, 163–174, https://doi.org/10.1016/J.VISRES.2015.01.023.

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442, Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9176953.

Phillips, F. (2004). Creating noisy stimuli. *Perception, 33*(7), 837–854, https://doi.org/10.1068/p5141.

Rahman, A. R., & Sommer, W. (2008). Seeing what we know and understand: How knowledge shapes perception. *Psychonomic Bulletin & Review, 15*(6), 1055–1063, https://doi.org/10.3758/PBR.15.6.1055.

Sharan, L., Rosenholtz, R., & Adelson, E. H. (2014). Accuracy and speed of material categorization in real-world images. *Journal of Vision, 14*(9):12, 1–24, doi:10.1167/14.9.12.

Schmid, A. C., & 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, https://doi.org/10.1167/18.1.14.

Schmid, A. C., & Doerschner, K. (2019). Representing stuff in the human brain. *Current Opinion in Behavioral Sciences*, *30*, 178–185, https://doi.org/10.1016/J.COBEHA.2019.10.007.

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

Scocchia, L., Valsecchi, M., Gegenfurtner, K. R., & Triesch, J. (2013). Visual working memory contents bias ambiguous structure from motion perception. *PLoS One, 8*(3), e59217, https://doi.org/10.1371/journal.pone.0059217.

Scocchia, L., Valsecchi, M., & Triesch, J. (2014). Top-down influences on ambiguous perception: The role of stable and transient states of the observer. *Frontiers in Human Neuroscience*, *8*, 979, https://doi.org/10.3389/fnhum.2014.00979.

Soechting, J. F., Juveli, J. Z., & Rao, H. M. (2009). Models for the extrapolation of target motion for manual interception. *Journal of Neurophysiology, 102*(3), 1491–1502, https://doi.org/10.1152/jn.00398.2009.

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745–756, https://doi.org/10.1038/nrn3838.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520–522, https://doi.org/10.1038/381520a0.

Toscani, M., Valsecchi, M., & Gegenfurtner, K. R. (2013). Optimal sampling of visual information for lightness judgments. *Proceedings of the National Academy of Sciences of the United States of America, 110*(27), 11163–11168, https://doi.org/10.1073/pnas.1216954110.

Toscani, M., Yücel, E. I., & Doerschner, K. (2019). Gloss and speed judgments yield different fine tuning of saccadic sampling in dynamic scenes. *i-Perception, 10*(6), 2041669519889070, https://doi.org/10.1177/2041669519889070.

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology, 5*, 1323, https://doi.org/10.3389/fpsyg.2014.01323.

Ullman, S. (1976). The interpretation of structure from motion. *Proceedings of the Royal Society of London: B. Biological Sciences*, *203*, 405–426.

Urgen, B. M., & Boyaci, H. (2019). When expectations are not met: Unraveling the computational mechanisms underlying the effect of expectation on perceptual thresholds. *BioRxiv, 545244*, https://doi.org/10.1101/545244.

van Assen, J. J. R., Barla, P., & Fleming, R. W. (2018). Visual features in the perception of liquids. *Current Biology, 28*(3), 452–458.e4, https://doi.org/10.1016/j.cub.2017.12.037.

van Assen, J. J. R., & Fleming, R. W. (2016). Influence of optical material properties on the perception of liquids. *Journal of Vision, 16*(15), 12, https://doi.org/10.1167/16.15.12.