Seeing liquids from static snapshots ☆

Vivian C. Paulun [a,*], Takahiro Kawabe [b], Shin'ya Nishida [b], Roland W. Fleming [a]

[a] *Department of Experimental Psychology, University of Gießen, Germany*
[b] *NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, Japan*

ABSTRACT

Perceiving material properties can be crucial for many tasks—such as determining food edibility, or avoiding getting splashed—yet the visual perception of materials remains poorly understood. Most previous research has focussed on optical characteristics (e.g., gloss, translucency). Here, however, we show that shape also provides powerful visual cues to material properties. When liquids pour, splash or ooze, they organize themselves into characteristic shapes, which are highly diagnostic of the material's properties. Subjects viewed snapshots of simulated liquids of different viscosities, and rated their similarity. Using maximum likelihood difference scaling (Maloney & Yang, 2003), we reconstructed perceptual scales for perceived viscosity as a function of the physical viscosity of the simulated fluids. The resulting psychometric function revealed a distinct sigmoidal shape, distinguishing runny liquids that flow easily from viscous gels that clump up into piles. A parameter-free model based on 20 simple shape statistics predicted the subjects' data surprisingly well. This suggests that when subjects are asked to compare the viscosity of static snapshots of liquids that differ only in terms of viscosity, they rely primarily on relatively simple measures of shape similarity.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

We encounter different fluids, such as water, honey and shampoo, almost everywhere in our daily lives (Fig. 1). Many such liquids have distinctive visual appearances, allowing us to tell them apart visually and to judge their properties, such as whether they would be slimy, wet or gluey to the touch. Our ability to recognize fluids is partly due to differences in optical characteristics like color or transparency, but we can also distinguish fluids to some extent based on their viscosity, which is one of the key properties determining how they respond to external forces. Less viscous fluids are thin and runny; tend to flow and splash easily; and settle rapidly to fill containers. More viscous fluids like honey are thick, stickier and do not flow easily, while very viscous liquids and gels even tend to pile up into clumps, and change shape very slowly over time, almost like solids.

The ability to perceive viscosity is not only useful in its own right—for example, when judging whether milk has gone off, or whether eggs are sufficiently beaten—but also presumably reflects more general perceptual abilities to recognize objects, textures and materials that have highly mutable appearance. In contrast to many objects, which have stable and well-defined shapes (e.g., shoes, chairs or bananas), a fluid has highly variable structure depending on the particular forces and actions to which it is subjected. This makes liquids a particularly interesting class of material for understanding how the brain identifies features that are common to different samples. Despite this, very little is known about how the visual system recognizes entities like liquids and gels, which do not have a clearly defined structure, and almost no previous research has investigated the perception of viscosity.

Given the extreme physical complexity of fluid flow processes, it seems unlikely that the visual system could accurately estimate the intrinsic physical parameters of liquids through 'inverse optics' (e.g. Barrow & Tenenbaum, 1978; Boyaci, Doerschner, & Maloney, 2004; D'Zmura and Iverson, 1993; Maloney & Wandell, 1986)—that is, by inverting the dynamical equations describing fluid behavior. It seems more likely that the brain abstracts information about fluid properties from various image features that correlate with the intrinsic physical parameters in natural environments. One important source of information is optical motion flow, and we recently found that image speed is a critical cue to viscosity perception (Kawabe et al., in press), in agreement with the natural expectation that viscous liquids flow slowly. However, as Fig. 1 clearly demonstrates, even static snapshots of liquids can yield vivid subjective impressions of liquidity, and it is easy to tell which samples are more or less viscous. This suggests that somehow the visual system is able to abstract information about material properties only from

---

☆ This research is part of an unpublished master thesis by V.C. Paulun.
* Corresponding author at: Department of Experimental Psychology, University of Gießen, Otto-Behaghel-Str. 10F, 35394 Gießen, Germany.
E-mail address: Vivian.C.Paulun@psychol.uni-giessen.de (V.C. Paulun).

static shape cues even though fluids can adopt practically any shape.

Here we isolate static cues to viscosity, in order to investigate how the visual system infers material properties from the way they self-organize into characteristic shapes. Specifically, we sought to measure how the subjective impression of viscosity varies as a function of physical viscosity (i.e., the *psychometric function* for viscosity), and then to identify which shape cues the visual system uses to arrive at an impression of the material properties of pouring liquids. To do this, we presented participants with arrays of static snapshots from computer simulations of flowing fluids, and asked them to judge how similar the different samples appeared to be (Fig. 2). Although computer simulations are imperfect approximations of real physical behavior, they nevertheless provide compelling impressions of liquids with different viscosities, and have the important advantage over real materials that it is possible to precisely control not only the viscosity, but also all other aspects of the scene, such as the lighting, viewpoint, optical properties of the material and the velocity of the source through which liquid enters the scene. To infer the psychometric function relating changes in physical viscosity to changes in subjective appearance, we used maximum likelihood difference scaling (MLDS; Knoblauch & Maloney, 2008; Maloney & Yang, 2003), a technique specifically designed for estimating supra-threshold appearance differences. We measured how subjective viscosity varies at different time points throughout the simulation, and then tested the extent to which simple shape statistics—derived from the images that were presented to participants—predict the subjective variations in viscosity.

2. Experiment

2.1. Materials and methods

2.1.1. Stimuli

We rendered seven 10-s animations of fluids with different viscosities, using Blender 2.61 (Stichting Blender Foundation, Amsterdam, NL), an open-source 3D computer graphics application. As shown in Fig. 2, the scene consisted of a fluid source, a fixed solid sphere, which served as an obstacle for the fluid, and an invisible reservoir ($0.75 \times 0.56 \times 0.39$ m), which filled up over time as the liquid poured into the scene with a constant velocity of 1.8 m/s in the x-dimension. Two unidirectional lamps with constant intensity illuminated the scene from the right side and diagonally from behind the objects. All fluids were greenish and semitransparent (alpha = 0.5). The seven liquids differed only in kinematic viscosity, $v = 10 \times 10^{-y}$ m^2/s with the exponents $y = \{0, 1, 2, 3, 4, 5, 6\}$. For comparison, at 20 °C, water has $v = 1.002 \times 10^{-6}$ m^2/s, oil has $v = 5 \times 10^{-5}$ m^2/s, and honey has $v = 2 \times 10^{-3}$ m^2/s. The simulation resolution was 250 (i.e., the volume was divided into $250 \times 250 \times 250$ cells). Each animation lasted 10 s, resulting in 300 individual frames per liquid, which were saved as 1280×720 pixel PNG files. Our synthesized fluid animations contained rich information enough for the participants to judge the simulated viscosity, since in a preliminary experiment, in which we asked the participants to make a numerical rating of the apparent viscosity for each animation, we found significant correlations between the viscosity ratings and the simulated physical viscosity. In the main experiment, to infer the psychometric function relating changes in physical viscosity to changes in subjective appearance for stationary images using MLDS, we selected the following 15 frames from the sequences: $t = \{3, 24, 45, 67, 88, 109, 130, 152, 173, 194, 215, 236, 258, 279, 300\}$. Images were down-sampled to 569×320 pixels so that four images could be presented on the screen simultaneously. The complete set of stimuli is shown in the Supplemental material (Fig. S1). Participants viewed the stimuli on a laptop (Lenovo ideapad Z570; screen resolution: 1366×768 pixels; refresh rate: 60 Hz), with a glossy LCD display at a freely chosen, comfortable viewing distance (approximately 60 cm).

2.1.2. Participants

Thirteen observers participated in the experiment (10 female, 3 male; mean age = 28.3 years, SD = 9.1 years). Observers reported having normal or corrected-to-normal visual acuity. All participants were unaware of the aims of the study and gave their informed consent prior to participation. The experiment was conducted in accordance with the Declaration of Helsinki, and the procedure was approved by the local ethics committee LEK FB06 at Giessen University (proposal number 2009-0008). Observers took part in the experiment on a voluntary basis and were not paid for their participation.

2.1.3. Procedure

On each trial, subjects were shown two image pairs (a *stimulus quadruple*) that differed from one another in viscosity (Fig. 2). Their task was to report with a key-press which of the two pairs (left or right) contained a larger within-pair difference, in a two-alternative forced choice (2AFC) paradigm. The dissimilarity judgments were used to estimate a perceptual difference scale within a maximum likelihood framework (for details see Knoblauch & Maloney, 2008 and Maloney & Yang, 2003). We used this method to estimate separate perceptual viscosity scales at 15 different time points (i.e., frames) from the simulations. On any given trial the four stimuli in the quadruple had different viscosities, but showed the same time frame.

The experiment consisted of 15 blocks, one for each time frame. Blocks were presented in random order and participants took short breaks between blocks. Each block consisted of 35 trials, one for each of the unique quadruples of the 7 stimuli $\left(\binom{n}{k} \right.$, where $n = 7, k = 4 \Big)$, presented in random order. The positions of the stimuli within a quadruple were randomized. Participants had unlimited time to perform the task.

2.1.4. Analysis

Perceptual scales were computed for each block (i.e. time frame), and each participant separately, using the MLDS package for R (R Development Core Team, 2011) from Knoblauch and Maloney (2008). Accordingly, perceptual scale values were estimated within the framework of a generalized linear model (GLM). The GLM method has the advantage that scale values of the least viscous fluid are set to a fixed value of zero (i.e. $\psi_1 = 0$), whereas all other scale values ($\Psi = \psi_2, \ldots, \psi_7$) are unconstrained. This makes it possible to compare perceptual scales derived from separate MLDS experiments, as is required for a comparison of different time frames.

2.2. Results

Fig. 3a shows the mean perceptual scales of viscosity derived by MLDS, averaged across participants as a function of both physical viscosity and time frame. Across different time frames, most of the scales show a similar sigmoidal shape, with a slow increase of scale values for the less viscous stimuli, followed by steeper changes with intermediate viscosity and finally very small differences or a plateau for the two most viscous fluids (see Fig. 3b and c).

Besides the striking overall similarity of the perceptual scales for different time frames, there are also some notable differences. Dissimilarities between viscosities seem more pronounced for

Fig. 1. Examples of everyday liquids varying in viscosity: water, honey, shampoo, laundry detergent (from left to right). Even such static snapshots provide a clear impression of the fluids' viscosity.
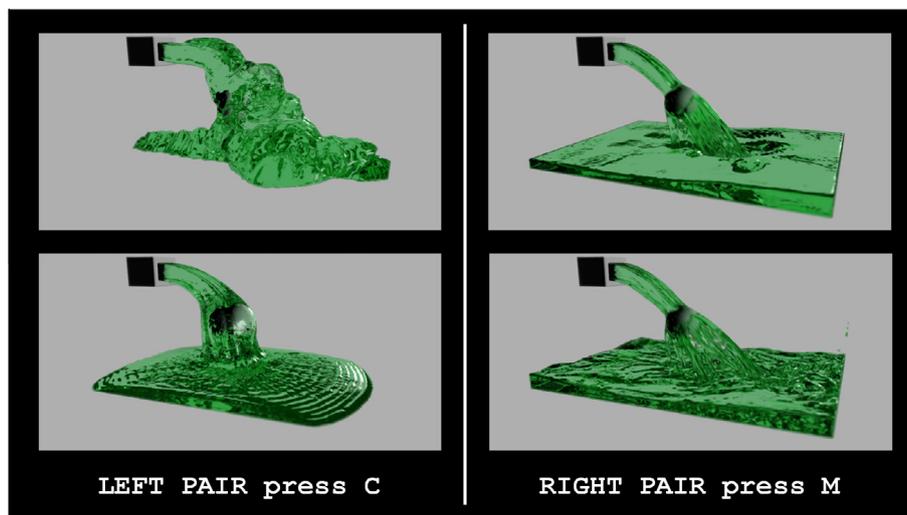

Fig. 2. A stimulus quadruple from the MLDS experiment consisting of four images showing different viscosities at the same time point (here, frame 130). Observers indicated in which pair (LEFT or RIGHT) the within-pair difference between the fluids appeared larger.
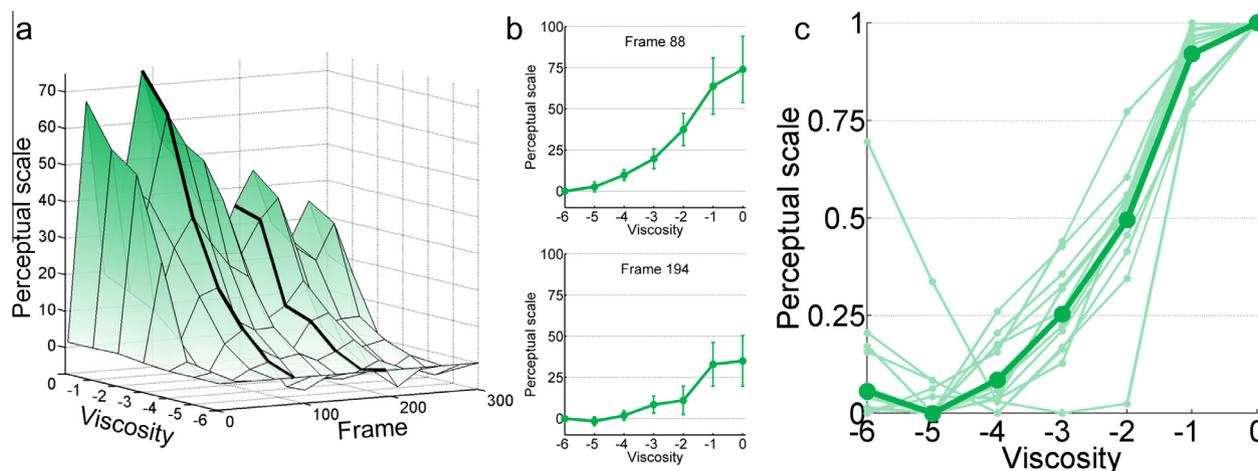

Fig. 3. Results of the MLDS experiment. (a) Mean perceptual scales of fluid viscosity averaged across participants as a function of viscosity (values on the *x*-axis reflect the exponents of the term: $v = 10 \times 10^y$ m^2/s) and time frames (non-normalized values). Note that for non-normalized data, the unit of perceptual scales is arbitrary across time frames and subjects, but that within a given time frame and subject, larger differences correspond to stronger perceived differences between fluids. (b) Example results from two time slices (frames 88 and 194; bold lines in a). These plots show mean data for the indicated time frame averaged across participants (non-normalized data). Error bars indicate standard error. Note that because the range of the data is not normalized, but the GLM fitting method is anchored to zero, the variance accumulates towards the higher end of the physical scale. Thus the error bars do not reflect variance between estimates, but rather differences in estimated standard error of the GLM error term, which determines the overall range of values estimated by the GLM method. Mean perceptual scales for all frames individually are also shown in Fig. 5. (c) Mean perceptual scale of viscosity (normalized data) averaged across participants and time frames (bold line) as well as the individual observers' scales averaged across time frames (thin light green lines). Note that the mean as well as the individual data were normalized to the range {0, 1}.

some frames than for others, suggesting that viscosity perception is not perfectly constant across time, but depends to some extent on the physical similarity in shape between samples. It is important to note that the scales inferred by MLDS from different time frames are computed independently and thus should not be interpreted as falling on a common scale. With this caveat it mind, nevertheless it is interesting to observe that the earliest time frame we tested, (frame 3) showed almost no physical difference in shape between the different viscosities and thus, unsurprisingly, yielded an almost flat perceptual scale. However, note that for this specific

time frame observers are probably not able to order the images reliably with respect to the true physical scale, which is a necessary requirement for difference scaling (Maloney & Yang, 2003). In other frames, e.g. 24, 88 and 109, the reported differences between viscosities were more pronounced, presumably due to richer visual cues caused by interactions between the fluid and the rest of the scene. Example perceptual scales for two time frames, along with their standard errors, are shown in Fig. 3b (a complete set of scales for individual time frames can be found in Fig. 5). Individual differences also become apparent in Fig. 3c, which shows the mean perceptual scale for each participant averaged across time frames (faint lines) as well as the grand mean (bold line). Psychometric functions from most individuals showed broadly similar sigmoidal shapes (although with different scales, reflecting differences in the consistency of their responses). This is also obvious from Fig. 3c in which the psychometric functions are scaled to the same range. An ANOVA with factors for viscosity and subject revealed a significant main effect of viscosity ($F(6, 1274) = 401.87$; $p < .001$), but no main effect for subjects ($F(12, 1274) = 1.22$; $p = .26$). However, there was a significant interaction between the two factors ($F(72, 1274) = 1.81$; $p < .001$), indicating that not all subjects produced the same psychometric curves for viscosity. These individual differences could indicate that participants rely on several different cues, or at least a different number or weighting of such cues in their perception of fluid viscosity.

3. Image-based prediction of the perceptual scale of viscosity

To test the hypothesis that the subjects' percepts of viscosity were derived from simple heuristics capturing the statistical characteristics of the liquid's shape, we sought to model the average perceptual viscosity scale by combining the predictions of a number of simple 2D shape features.

3.1. Materials and methods

3.1.1. Image statistics
To investigate the viscosity-related differences in shape and appearance and identify possible cues for the visual perception of fluid viscosity, twenty 2D shape statistics were calculated for the images shown to observers. These included statistics of curvature, orientation, shape, area and perimeter of objects, and the distribution of pixels in the image. Examples of these statistics are shown in Fig. 4a (verbal and pictorial descriptions of the complete set of statistics can be found in Supplemental material: Table S1, Fig. S2). It is important to note that most of the statistics we used were not independent but strongly related with one another because they capture the same underlying shape characteristics in somewhat different ways. Image statistics were calculated for the alpha maps of our stimuli, i.e. binary images in which every pixel that belongs to the background is black and every pixel that belongs to an object (here the fluid, its rectangular source and the sphere) is white. In case of multiple objects in the image, e.g. the source, the sphere or drops, only the largest object was chosen for subsequent analysis of shape related statistics. All analyses were calculated using Matlab R2007b (The MathWorks Inc., Natick, MA, 2007). Results of these 20 statistics for each viscosity as a function of time are shown in Supplemental material (Fig. S2).

3.1.2. Prediction
We calculated the Euclidian distance between the images of the least viscous fluid and all other fluids with increasing viscosity within the space of each of the individual image statistics. Resulting distances were normalized to the range {0, 1}, and averaged across time frames. This created predicted perceptual scales for

each individual image statistic (see Fig. S2). This procedure was conducted for all 20 image statistics. Although there was no fitting involved, there was one free parameter implied by the normalization to a common range; however, in principle some other normalization may provide a better fit for each curve. To evaluate our predictions we calculated correlations between each prediction and the mean perceptual scale from the MLDS experiment. It is important to emphasize that we did not seek to identify the *specific* image measurements made by the human visual system, but rather to demonstrate that many different measurements predict similar performance, and thus such heuristics provide a plausible approach for human vision to exploit. To reduce the number of predictors, we first embedded all 105 images (15 time frames and seven viscosities) in a common 20 dimensional feature space (defined by the different image statistics), and performed principal component analysis (PCA). To account for the different scales and units for the different image statistics, results of the image analyses were first standardized in *z*-scores before calculating the PCA. We then calculated the Euclidian distance between each viscosity level and the least viscous one within the multidimensional PCA space, equivalent to the distances that were calculated for each image statistic individually. This was done separately for each frame. Resulting distances of different frames were then aggregated and normalized to lie between 0 and 1. The prediction was again evaluated in terms of its correlation with the mean perceptual scale. Note that no explicit parameter fitting was performed in the combination of features (i.e., we did not estimate an optimal weighting of the features to match the subjects' data). This means that individual statistics (or some other combination) could actually out-perform the parameter-free combination in terms of predicting subject performance.

3.2. Results

Even when considered on their own, most of the individual shape statistics can predict the data very well, as indicated by the high correlations between the predictions and the normalized mean perceptual scale. The average correlation across the twenty different image statistics was $r = .95$ ($SD = .04$). Examples of predictions together with mean data can be found in Fig. 4; the complete set of individual predictions is shown in Supplemental material (Fig. S2). The best predictions were derived by statistics of the vertical distribution of pixels (i.e. kurtosis, skewness and the standard deviation), the vertical position of the centroid, and eccentricity. This is in good agreement with the observation that there are substantial differences in the way that differently viscous fluids spread, i.e. highly viscous ones pile up, whereas less viscous fluids settle and spread more evenly on the ground (see above). Thus, these measures of the vertical distribution of the liquid seem to be useful cues that observers could use when comparing the viscosity of the fluids in our scene. Other cues (e.g. the prediction derived from circularity) seem less adequate. However, we do not propose that observers base their judgments on individual shape statistics. Instead, they probably use some combination of features that capture various aspects of the behavior of the fluid, as different shape features are likely to be more or less diagnostic of viscosity in different settings. Thus, we derived a combined prediction from all twenty statistics, which is plotted in Fig. 4, along with the normalized mean perceptual scale across frames and subjects. As is clear from comparing the predicted and observed perceptual scales, the combination of 20 different shape properties leads to a very good prediction of the mean perceptual scale derived by MLDS. The correlation coefficient between the two was very high ($r = .99$) and highly significant ($p < .001$). Thus, the combined prediction was better than most of the single predictions or their mean. It is remarkable that such a high correlation could be
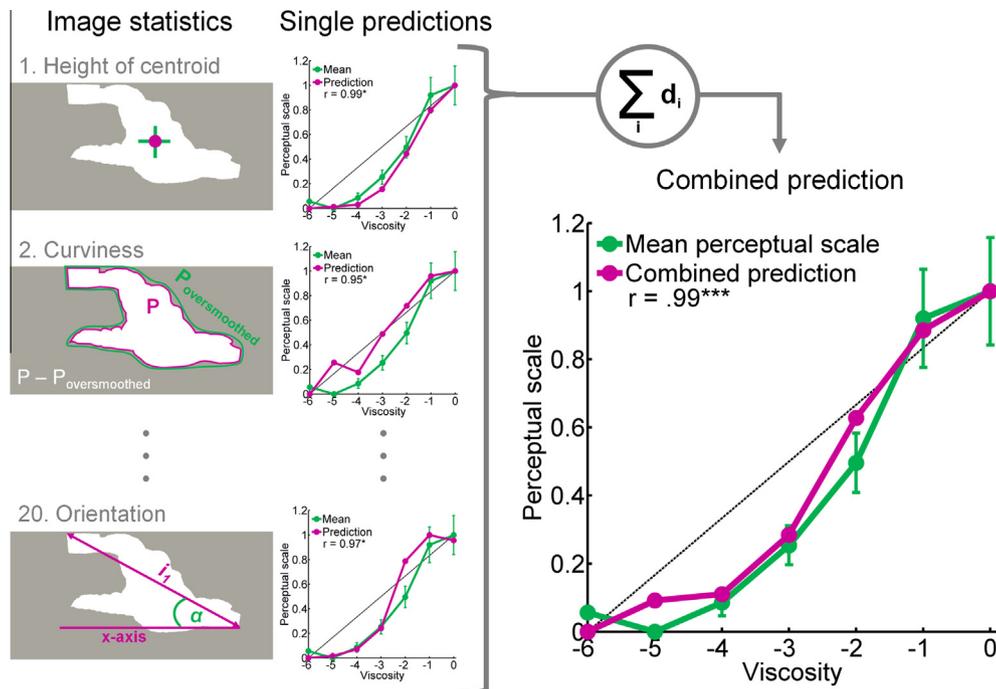
Fig. 4. Image based prediction of the perceptual scale of fluid viscosity. The left panel shows three examples of the 20 shape statistics that were calculated in order to predict the mean perceptual scale derived in our MLDS experiment. Next to this are the corresponding predictions (magenta lines) of individual image statistics together with the normalized mean perceptual scale (green lines: error bars show ± 1SEM), i.e. these plots show how the perceptual scale is predicted if only the indicated image statistic is taken into account. These plots also show the correlation coefficients between the data and prediction. Values on the *x*-axis reflect the exponents of the term: $v = 10 \times 10^y$ m²/s. Results and individual predictions of all 20 image statistics are shown in Supplemental material (Fig. S2). We calculated a PCA across the results of all image statistics for the 105 images used in the experiment. To derive a combined prediction of all statistics we calculated the Euclidian distances (d) between images of all viscosity levels and the first (least viscous) level of each frame (i) in the PCA space and then aggregated the resulting values across time frames. This prediction was then normalized to lay between 0 and 1. The combined prediction (magenta) together with the normalized mean perceptual scale (green) is shown in the rightmost plot, which also indicates the correlation coefficients between the two. Values on the *x*-axis reflect the exponents of the term: $v = 10 \times 10^y$ m²/s.

achieved without any explicit fitting of the model to the data, except by normalizing all predictions to the same range. We believe this reflects the fact subjects' judgments of viscosity are based on statistical measures of the similarity between the shapes, and that many different shape statistics can capture the underlying similarities between the different fluids at each point in time.

This general pattern was also reflected in most individual subjects. To characterize how well the model matched individual subject data, we computed correlation coefficients between each individual subject's mean response curve (across time frames) and the model. The resulting correlation coefficients ranged from 0.51 to 0.99, with a mean of 0.94 and standard deviation of 0.12. Of course, a model that simply computes values from the image without any fitting to the data cannot even in principle predict inter-subject variations. In the future it would be interesting to model the variations between subjects as weighted combinations of the different image measurements in the model.

The different image statistics we used here are highly inter-correlated and measure the same underlying visual correlates of fluid viscosity in different manners. In fact, PCA across all images and statistics showed that only five principal components were necessary to explain more than 95% of the variance; the first two principal components could already explain 77.99% of the variance. This reinforces the idea that many different image measurements yield similar estimates of viscosity similarity.

The MLDS experiment showed that the shape and steepness of the perceptual scales evolved over time. It is interesting to ask whether this temporal evolution can also be predicted by the combination of image statistics. Predictions and mean data are shown in separate plots for each time frame in Fig. 5, along with the corresponding correlation coefficients. Correlations were again very

high (*M* = .96, *SD* = .02). Although not all of the prediction curves seem to capture the data very well, the overall trend appears to be in good agreement for the majority of time frames. Both perceptual and predicted scales are flat for the first time frame (frame 3) and become much steeper in the subsequent frames; for the second half of the frames variations between the frames are much smaller. However, the combination of image statistics predicts even lower variations especially for high viscosities, than can be found in the data. Additionally, the peaks of the two distributions do not coincide (frame 88 for the perceptual scales, but frame 24 for the predictions; this is, however, the second largest peak in the scales derived from the MLDS data).

4. Control experiments

Two control experiments were conducted to (1) explore if observers can reliably judge viscosity when there is no other cue present in the images than the 2D outlines on which our model is based, and (2) to test the extent to which the shape statistic-based predictions generalize to simulations of fluids in a wider range of scenes and contexts. The rating task used for this also helps us to overcome the potential drawback of the 2AFC method, which may encourage participants to base their similarity judgments on relatively low-level properties of the images they compared.

4.1. Materials and methods

4.1.1. Stimuli

Alpha maps of the images used in the MLDS experiment served as stimuli in the first control experiment, see Fig. 6a. The 105 black
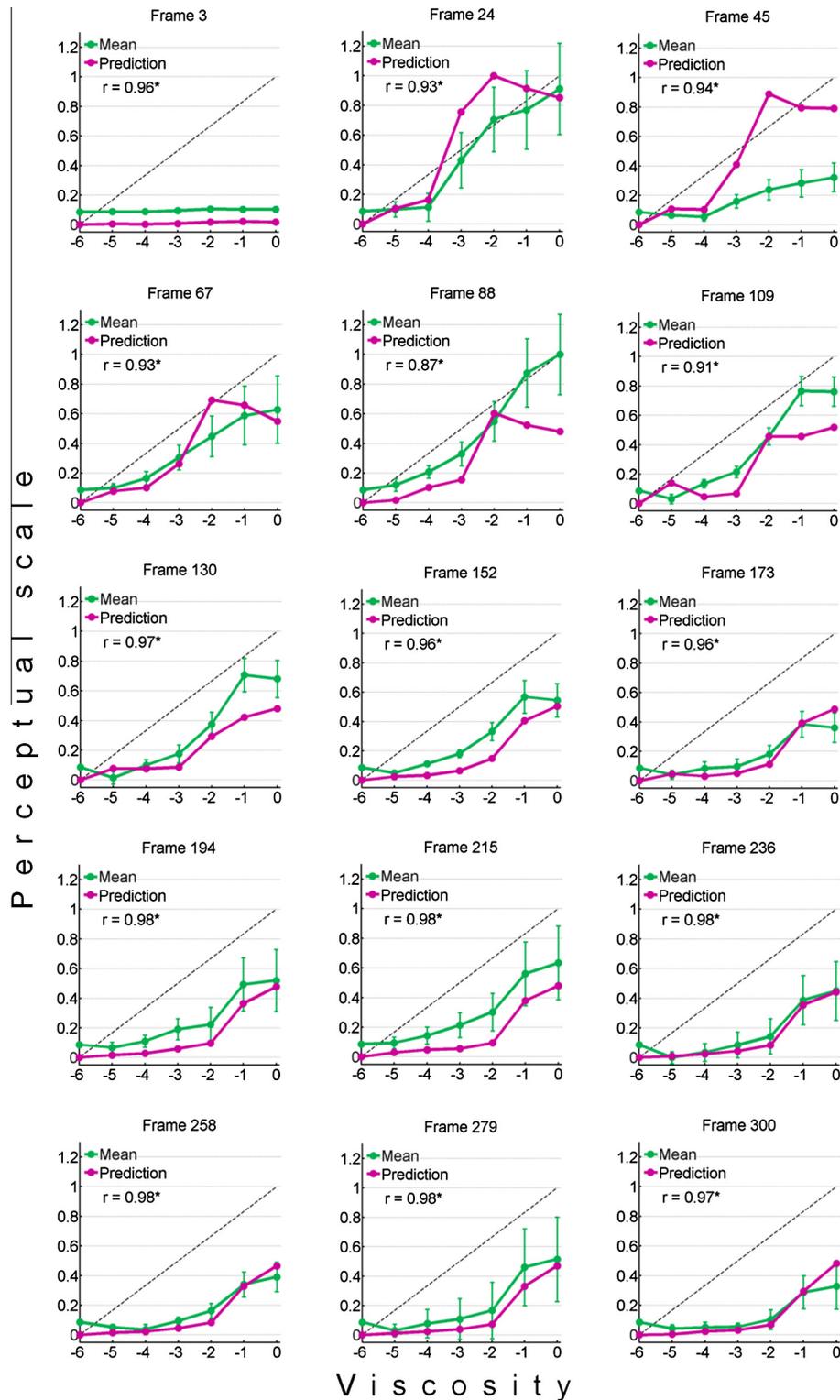
Fig. 5. Perceptual scales and predictions (non-normalized data). Each plot shows the mean perceptual scale of viscosity derived in the MLDS experiment (bold green lines) as a function of physical viscosity. Error bars indicate standard error (note that as in Fig. 3, the error bars reflect differences in the range calculated by the GLM method due to the use of non-normalized data). Data have been scaled to lie on a common scale between 0 and 1 to enable comparison between model and data. Magenta lines show the combined prediction of the 20 image statistics for the indicated time frame. Correlation coefficients given in the individual plots show the corresponding values for the normalized mean data and prediction for each frame individually. Values on the x-axis reflect the exponents of the term: $v = 10 \times 10^y$ m^2/s.

and white images (i.e. 7 viscosities × 15 time frames) were down-sampled to 1024 × 500 pixels. In the second experiment we used renderings of ten different scenes of fluids that have recently been used in another study (Kawabe et al., in press), see Fig. 6b. Each scene was rendered with five different viscosities and we chose

four of the 60 frames {15, 30, 45, 60} for our study. The images had a size of 384 × 384 pixels and the contrast was slightly enhanced to increase the visibility. All stimuli were presented on the same laptop as in the main experiment at a freely chosen, comfortable viewing distance.
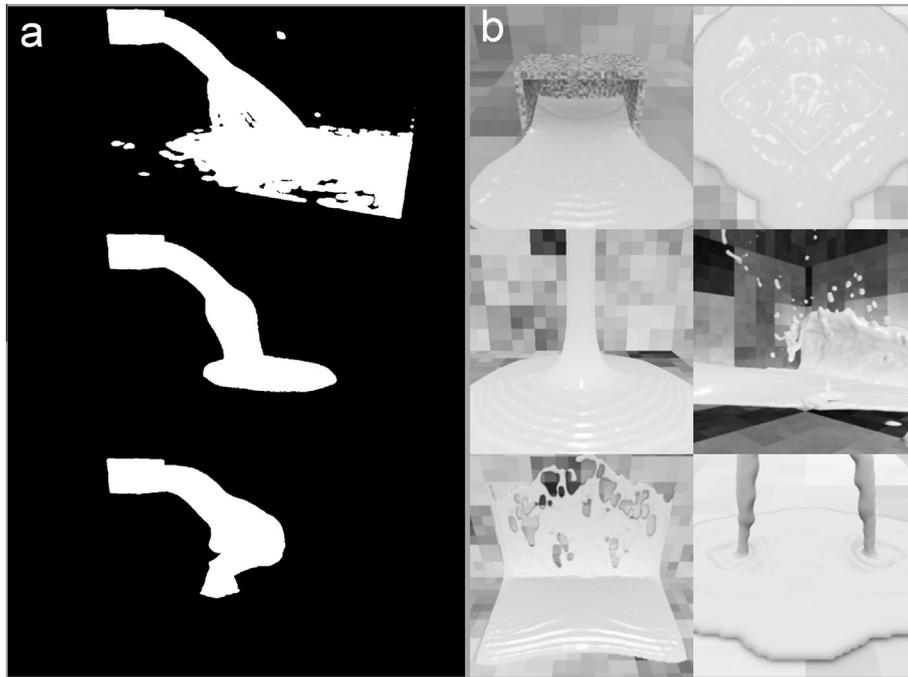
Fig. 6. Examples of stimuli used in the control experiments. (a) Alpha maps of three different viscosities at time frame 24. (b) Six of the ten scenes used in the second rating task. All images here show different viscosities at time frame 30.

4.1.2. Participants

Ten observers with normal or corrected-to-normal visual acuity participated in both control experiments (7 female, 3 male; mean age = 22.8 years, SD = 2.2 years). The order of the control experiments was counterbalanced between participants. Observers were naïve with regard to the aims of the study and gave written informed consent prior to their participation. The experimental procedures were approved by the local ethics committee LEK FB06 at Giessen University (proposal number 2009-0008) and in accordance with the declaration of Helsinki. Observers were paid 8€ per hour of participation.

4.1.3. Procedure

Both experiments consisted of a perceptual rating task. On each trial, one image was presented and remained on the black background until the participant entered a number between 1 (very liquid) and 7 (very viscous) to indicate their subjective rating of the fluid's viscosity in the image. Every image was rated 4 times, resulting in 420 trials with the alpha images (7 viscosities × 15 time frames × 4 repetitions) and 800 trials with different scenes (10 scenes × 5 viscosities × 4 time frames × 4 repetitions). Trial order in each experiment was random. Twenty practice trials were conducted before both tasks, with a pseudo-random order in which each viscosity had to appear at least once in the alpha map rating and each scene and viscosity had to appear at least once in the rating of different scenes. This was done to give the observers an idea of the possible range of viscosities and scenes, aiding them to adjust the range of responses. Participants were informed that all images contained a liquid, whose viscosity they had to rate. No further information was given about the different scenes or the black and white alpha maps, i.e. they were not informed about the presence of the obstacle or source. None of the participants had ever seen the original full renderings from which we derived the alpha maps before completing the task. Thus, if they are nevertheless able to interpret the white "splotches" as liquids flowing in 3D with specific viscosities, some perception of viscosity must emerge even from these impoverished images. Otherwise their judgments

should be random. Participants had unlimited time to perform the task and were debriefed afterwards.

4.1.4. Analysis

The data from each experiment were averaged across repetitions, participants, and time frames. To compare perceptual rating scales with a prediction derived by our model these values were rescaled to a range between {0 and 1}. To predict the perceptual rating of the alpha maps we used the same prediction as for the MLDS task, since this was based on the alpha images we used here. The prediction of the rating scales for the different scenes was based on the exact same image analysis described in the previous section. Here, we used alpha maps of the 200 images used in this experiment as input.

4.2. Results

4.2.1. Rating of alpha maps

Fig. 7a shows the mean perceptual rating scales of viscosity in the alpha images averaged across participants as function of time frames and physical viscosity. Most importantly, participants were able to rate the viscosity of the fluids when they were only presented with the alpha maps of our original stimuli. If this were not the case, we should not find such a clear mapping between physical viscosity and the observers' responses. The overall shape of the rating scale appears similar for many time frames, although varying in the steepness of the curve, see Fig. 7b. The first time frame stands out in the sense that observers did not report differences between different viscosities. However, it should be noted that the black and white images of fluids with different viscosities were almost identical for this time frame and consisted almost exclusively of a white circle and a white rectangle on a black background, i.e. the obstacle and the fluid's source, which were not mentioned to the participants. Thus, they may have assumed one or the other to be a large blob of very viscous fluid and thus always gave the highest rating. The fact that they did not report differences between the fluids is however, consistent with the model prediction, only at a
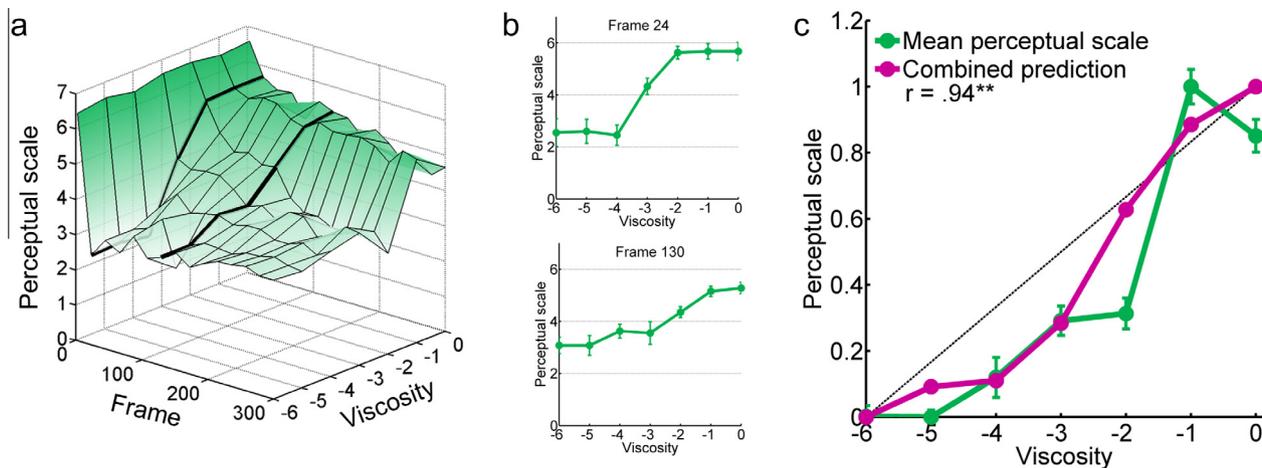
Fig. 7. Results of the alpha map rating experiment. (a) Mean rating scales of fluid viscosity averaged across participants as a function of viscosity (values on the x-axis reflect the exponents of the term: $v = 10 \times 10^y$ m²/s) and time frames. (b) Example results from two time slices (frame 24 and 130; bold lines in a). These plots show the average rating scale ± 1 SEM. Mean perceptual scales for all frames individually are shown in Supplemental material (Fig. S3). (c) Combined prediction derived by our model (magenta) together with mean perceptual scale averaged across time frames and participants (green); error bars show ± 1 SEM. Data shown here was rescaled to the range {0, 1}.
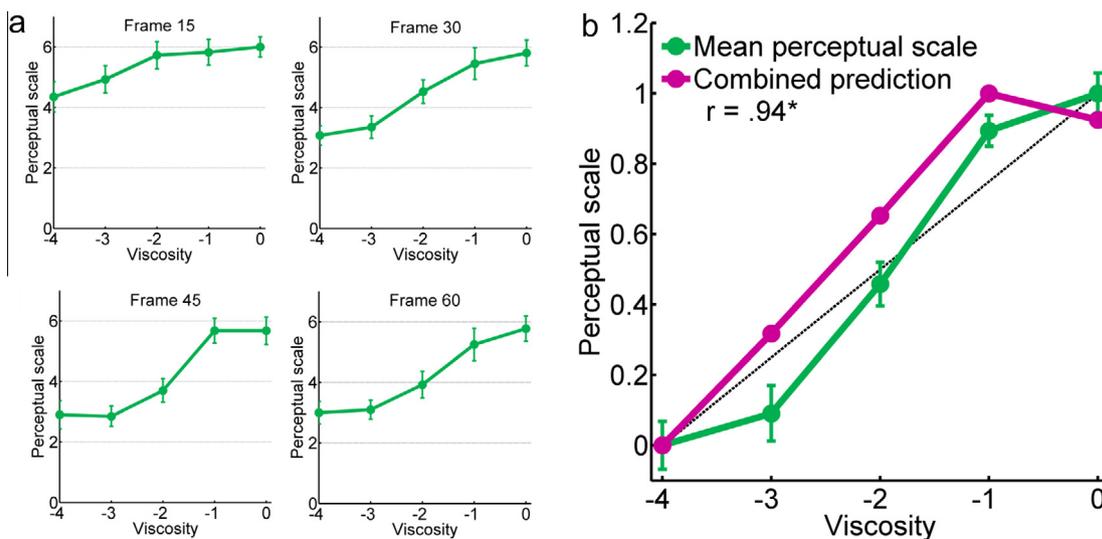

Fig. 8. Results of the viscosity rating in the scene "*Twist*". (a) Mean perceptual scales for the four time frames tested in this experiment averaged across participants. Error bars show ± 1SEM. (b) Combined prediction of the perceptual scale derived by the 20 image statistics together with mean perceptual scale and the correlation coefficient between the two. Data was averaged across the four time frames and 10 observers and then rescaled to the range {0, 1}. Error bars show ± 1 SEM. Values on the x-axes reflect the exponents of the term: $v = 10 \times 10^y$ m²/s. A complete data set together with predictions and corresponding correlations for the remaining nine scenes can be found in Supplemental material (Fig. S4).

different level. Another observation that can be made from Fig. 7a is that observers on average did not use the entire scale to judge the viscosity, but rather only a limited range, as is often the case in rating experiments. For this reason and to be able to compare the rating scale with the prediction we derived from the model, Fig. 7c shows the rating scale averaged across observers and time frames (thick green line) and then rescaled to the range {0, 1} together with the combined prediction (magenta line) and the observers individual scales averaged across time frames and normalized (thin green lines). The correlation between the mean rating scale and the prediction was very high ($r = .94$, $p < .01$).

4.2.2. Rating of different scenes

Fig. 8 shows the results of the rating task for one example scene ("*Twist*"). As in the other rating task, participants did not use the entire range of the 7-point rating scale, avoiding extreme values

on the scale, see for example Fig. 8a. The main objective for the current task was to test how well the model based on simple shape statistics generalized to a wider range of scenes and tasks. Fig. 8b shows the normalized mean rating scale for one scene together with the scale predicted by our model. The high correlation ($r = .94$, $p < .01$) between the two curves in this example confirms the apparent similarity between data and prediction. We found a significant correlation between model and perceptual rating for seven of the ten scenes; their mean correlation was $M = .94$ ($SD = .06$). The correlation for the remaining three scenes was also moderately high ($M = .65$, $SD = .21$). Mean data and predictions together with the correlation coefficients between the two for all other scenes is shown in Supplemental material (Fig. S4). Overall the rating data is captured quite well by the model. As can be expected the model is closer to the data for some scenes than for others since different scenes provide a different quantity and

quality of cues to viscosity. Most importantly, our model *can* predict the perceptual experience in most of the scenes, irrespective of the direction of the fluid flow, its source or the presence of obstacles.

In addition to the rating task described above. We also tested if our model can account for the data that was originally collected with the complete set of stimuli. In Kawabe et al. (in press) participants had to rate the viscosity of each of the 50 animations in a 5-point scale. The results showed that the rated viscosity monotonically increased with the simulated viscosity (see their Fig. 1B). While the stimuli contained both dynamic (motion) and stationary (shape) information about viscosity, viscosity could be judged with certain accuracy even for single stationary images. To test whether the model could predict the rating data, we calculated the shape statistics for all 60 frames of the 50 movies and derived a combined prediction for each movie, as described above. For each of the ten scenes the same model predicted the rating data very well, as indicated by the high correlations between the average rating scale of each movie and the combined prediction (M = .91, SD = .11). The combined prediction for the grand mean averaged across all movies was even higher (r = .98, p < .01). Taken together, this indicates that the combination of shape statistics used in the model captures general visual features for fluid viscosity perception in a variety of different scenes and across different types of subjective task.

5. Discussion

Different liquids adopt distinctive shapes in response to their internal properties, and interactions with gravity and other objects. The purpose of the current study was to measure and predict how subjects use static shape cues to visually distinguish liquids with different viscosities. Using MLDS, we derived perceptual scales for different time points from fluid simulations. We found that the shape of these scales was quite similar across different time frames—although more or less strongly pronounced, depending on the distinctiveness of the fluid shapes at that point in time. For example, at the start of the simulation—when the liquid had barely emerged from the source—all liquids were almost exactly the same shape, leading to a flatter perceptual scale. In contrast, later in the simulation, the shapes of the different liquids were more distinct, leading to a steeper perceptual scale. The perceptual scales were sigmoidal, with smaller perceptual differences for the lower and upper end of the physical scale, and larger perceptual differences in viscosity for the intermediate stimuli. Put intuitively, this suggests a subjective separation of fluids into two broad classes: runny liquids that splash and settle vs. thick fluids that ooze and pile up into clumps. We suggest that this distinction is perceptual in origin, based on the statistical similarity in shape between liquids with different viscosities.

Specifically, we showed that a number of simple 2D shape statistics vary systematically with physical viscosity, in a way that predicts a non-linear perceptual scale similar to the one observed in the MLDS data. These cues capture statistical variations in shape between fluids with different viscosities. Thus, the visual system could use such cues to infer fluid viscosity as a proxy for more complex inverse physics computations. Without any fitting to the data, the combination of shape statistics predicted the mean perceptual scale strikingly well (r = 0.99). The same model also predicts perceptual rating scales of viscosity using another set of rendered fluids of varying viscosities interacting with ten different scenes from another study as well as the rating of the 50 movies of fluids that was reported in the original paper (Kawabe et al., in press). This shows that viscosity-related variations in shape are not specific to our stimuli, but probably reflect more general regularities. For our stimulus regime, the most promising of the cues we tested included those that captured the vertical (and to a lesser extent the horizontal) distribution of the fluid, as well as measures of its elongation (eccentricity) and curviness. This is consistent with the observation that low viscosity fluids tend to settle and spread out on the ground, distributing broadly in the horizontal direction leading to less curvy, more elongated shapes with lower centroids. By contrast, more viscous fluids pile up into curvier mounds and bumps before spreading out more slowly, leading to a higher centroid. Thus, those cues that best predict subjects' percepts, intuitively capture the statistical behavior of the fluids with different viscosities in our scenes, which may account for why they predicted subjective judgments of liquid viscosity. Our work is therefore in line with some similar studies on other material properties. For example, Marlow, Kim, and Anderson (2012) follow a similar approach, investigating the properties of highlights as a cue to the perception of glossiness. They show how interactions between illumination and surface relief can lead to substantial variations in the properties of specular highlights, which have concomitant effects on the perception of glossiness. They asked one set of subjects to judge the properties of the highlights (e.g., contrast, extent), and found that a linear combination of the ratings of the low-level image properties predicted the glossiness ratings made by a different group of subjects. In spirit, this is similar to our approach, except that we derived our predictors directly from the images (rather than ratings from other observers).

Given the high correspondence between shape statistics and the perceptual judgments one might ask if participants in our experiments were really judging viscosity rather than shape itself. This reflects a fairly general question of validity in many perceptual tasks involving 'appearance', which is difficult (maybe even impossible) to overcome completely. Here, we used different types of tasks, in which shape similarity is more (MLDS) or less (rating task) relevant, and gain similar results. Furthermore, the phenomenological impression of liquids with different viscosities is quite strong. Participants expressed no uncertainty about the task; one look at our stimuli yielded a clear and compelling impression of viscosity, so that when we asked them to perform viscosity comparisons or ratings they found this a natural and intuitive judgment to make. Besides this phenomenological argument, it is ultimately questionable why using shape to judge viscosity should be less valid than any other kind of visual viscosity judgment. Viscosity is not an optical feature of a fluid like, for instance, its color. Imagine a bowl containing a still liquid. Without further visual information it is impossible to estimate the viscosity of the liquid; it might not even be obvious to the observer that the material inside the bowl is a *liquid* (although experience with other liquids might lead them to assume that it was). Only after some external force is applied will the viscosity become apparent. At that stage, many features would provide potentially useful visual information about viscosity. Here, we argue that shape is one of them.

It is important to point out that we have measured and modeled perceptual viscosity functions only in the simplest possible case, namely when viscosity was the only parameter that differed between the samples being compared. Under these conditions, shape varies in a highly systematic way between samples, and it is likely that practically any measure of shape similarity would rank order the different fluids correctly. Nevertheless, the fact that even simple 2D statistics can predict the specific non-linear form of the perceptual scale suggests that when subjects are asked to judge viscosity in this context, they rely heavily on some simple measures of shape similarity. We believe this shows that when other factors are held constant, comparisons between liquids are based on intuitive impressions of shape similarity rather than some more sophisticated estimate of the physical parameters of the liquid, derived through inverse physics.

A more challenging case, which we have not explicitly tested so far, is 'viscosity constancy', i.e., the ability of subjects to perceive a given fluid as having the same viscosity in different settings. In the control experiment with additional scenes, subjects had to find a consistent scale to rate viscosity across the different scenes and time points, and we found that they could do so to some extent. However, we did not ask subjects to directly compare apparent viscosity across variations in speed, scene layout, viewpoint, time or any other factors that also contribute to shape. There are almost certainly conditions in which simple image statistics will fail to predict perceptual constancy. For example, early timeframes of different fluids are more similar to one another than an early timeframe of a particular fluid is to a later timeframe of the same fluid. If subjects are able to correctly match fluids at different points in time (i.e., if they exhibit viscosity constancy), this would be challenging for simple image statistics to predict. Only a process that could somehow capture or predict how shape evolves over time could explain constancy across points in time. Additional work should be conducted to identify when simple measures of shape similarity break down and more sophisticated internal models that capture how fluids behave play a role in viscosity perception.

It is also important to note that we are not proposing the visual system uses these *specific* image measurement to represent shape for viscosity perception. There are many other plausible shape statistics that could also play a role, especially, more sophisticated ones that take into account 3D shape, rather than simple 2D outlines. If we were for instance to crop our images so that there is no viscosity information from the outline, we would still be able tell the differences between liquids based on other shape cues inside the object. Indeed it would be foolish to interpret our findings as a process model of how the visual system computes viscosity; that is not the intended purpose of the model. Instead, our goal is to emphasize the more general point that many different shape statistics exist which correlate with changes in viscosity. The transformations in shape that different liquids undergo as they flow under gravity are rich and highly systematic, causing many statistical shape regularities to emerge. If we consider each snapshot as a point in a high-dimensional shape space, then liquids with different viscosities occur at different locations within the space. As long as other factors are held constant, projecting the different points onto almost any shape statistic would make it possible to distinguish the different fluids. This means that many different shape cues could support viscosity perception across different contexts. Here, we intuitively selected twenty easy-to-compute cues, more or less arbitrarily. The fact that so many different cues correlate with viscosity makes demonstrating the causal role of any given image measurement quite challenging. However, we believe that much more important than the role of any specific cue, is the general observation that the visual system could draw on a wide constellation of different measurements that are readily computed by low- and mid-level visual mechanisms. In one of our control experiments we have shown that observers can indeed reliably use the two-dimensional shape to judge viscosity when no other cue is available.

Using weighted combinations of multiple shape properties could give the visual system the flexibility to identify relevant measures of shape similarity in a wide range of settings. For example, while measures of how much the liquid has settled under gravity may be relevant for distinguishing different liquids from one viewpoint (e.g., side view), from other viewpoints of the same scene (e.g. top view), this cue may provide no useful information. At the same time, other cues, such as the smoothness of the contour or the projected area of the liquid may become more diagnostic. We suggest that the visual system flexibly re-weights different measures of shape similarity depending on the context, describing the differences between liquids in different ways depending on the scene.[1] Thus, even though the *physical* viscosity of water remains constant across scenes, in perceptual terms, the *perceived* viscosity of the water that sprinkles out of a watering-can may be in some fundamental way incommensurable with the perceived viscosity of the water that lies in a rippling puddle. Because the shape properties that distinguish water from other liquids in these two contexts are radically different, the visual system may use completely different measurements to determine water's viscosity in these two contexts. Testing this idea and the limits of perceptual constancy are important lines for future investigation.

This line of argument suggests that the visual system may not need to accurately model the physics of the environment in order to identify liquids or work out their properties. Instead it could represent different fluids by exploiting statistical regularities in the *typical appearance* of fluids in the image. Previous work has debated whether the perception of various material perception relies on heuristics derived from image statistics or more sophisticated inverse optics computations (Fleming, 2014; Thompson et al., 2011), e.g., in the perception of surface gloss (Anderson & Kim, 2009; Fleming, Dror, & Adelson, 2003; Kim & Anderson, 2010; Motoyoshi et al., 2007), roughness (Ho, Landy, & Maloney, 2006), transparency (Fleming, Jäkel, & Maloney, 2011) and translucency (Fleming & Bülthoff, 2005). Similarly, the estimation of other physical properties has often been argued to be heuristic or *naive* (Gilden & Proffitt, 1989; McCloskey, Caramazza, & Green, 1980; Nusseck et al., 2007; although see: Hecht, 1996), whereas recent work has suggested explicit mental 'simulation' of physics may what explains our ability to predict complex processes like the tumbling of a tower of bricks (Battaglia, Hamrick, & Tenenbaum, 2013). However, here, we wish to speculatively propose a third alternative that lies between crude heuristics and full inverse optics computations.

We have suggested that rather than using a fixed mapping between image measurements and some internal scale of viscosity, the visual system may flexibly weight different shape features to represent liquids in different contexts. This approach poses a key challenge that requires inference machinery that goes beyond the straightforward application of heuristics. Specifically, having seen only a single fluid in a given context, how does the visual system know *which* aspects of the shape are the relevant ones to use to determine viscosity? How can the visual system predict which shape properties would vary (and by how much) were it to be presented with a *different* fluid in the same context? In other words, how does the brain work out the weights when it does not have multiple liquids to compare?

We speculatively suggest that to weight different shape cues, the visual system may hypothesize—on the fly—which aspects of shape are most relevant, based on perceptual organization processes and previous experience. Observable properties of a shape can provide cues to the (unobservable) underlying generative processes that have yielded that shape (e.g., Feldman, 1992, 1995; Feldman & Singh, 2006; Hoffman & Richards, 1984; Koffka, 1935/1963; Leyton, 1989, 1999; Spröte & Fleming, 2013). In particular, regularities in the shape (e.g., symmetries, periodic elements or other non-generic relationships between features) provide evidence of a systematic (i.e., non-random) generative process. Thus, it makes sense to infer that new sample liquids—which differ in other respects—may preserve such shape features. By contrast, salient features that distinguish a given shape from other previously seen shapes may be key areas where the shape is likely to vary from sample to sample. We suggest that based on such perceptual

---

[1] This appears in our model as a weighting of the different statistics in proportion to the total range of values for different viscosities in a given scene, see Methods.

organization principles, the visual system may be able to cast hypotheses about how the shape is likely to vary from sample to sample. For example, if a particular snapshot of a liquid contains a pronounced puddle, the visual system may hypothesize that other liquids may vary in the size or shape of that puddle, and therefore that measures of the extent of the puddle feature should be weighted strongly in determining viscosity. This active recalibration of the representation based on hypotheses about how appearance might vary is quite different from a passive application of fixed heuristics. Indeed, it is a form of 'internal model' of liquid behavior derived from perceptual organization and previous experience. However, unlike a *physical model* it does not involve explicitly estimating the liquid's physical parameters, but rather on characterizing its distinctive shape features and how they might change under different circumstances. That is, it is a model of the behavior of the *appearance* of the liquid (as it manifests in the image) rather than a model of its *physical behavior*. We call such representations 'statistical appearance models' (Fleming, 2014), because they involve probabilistic inference about likely variations in appearance.

Clearly these speculations go far beyond what can be inferred from the data presented here. Extensive theoretical work is required to turn this speculation into a plausible model and to investigate how symmetries and other regularities can guide the inference of generative processes. Moreover, additional experimental work would then be required to test specific predictions of this approach.

If observers rely on a variety of statistical cues, the resulting estimates will often be imperfect. Depending on the specific cues they use, their estimate may vary. Thus, variability in the quantity or quality of the viscosity cues might be responsible for the differences in perceptual scales between frames in our experiments. Likewise, different observers might have weighted the cues differently, leading to the small observed inter-individual differences.

Here, we focussed on static 2D shape-related measurements, although there are clearly many other cues the brain could use to identify fluids. These include, for example, 3D shape properties, optical cues (e.g., color and translucency), and obviously motion and speed related cues, which we have shown to be important in a separate study (Kawabe et al., in press). In that study, we isolated motion cues and eliminated shape information using arrays of noise patches whose motion statistics matched fluid simulations. Here, by contrast, we have isolated shape cues and eliminated motion, by studying viscosity perception in static snapshots. Other researchers have shown that static snapshots can lead to a vivid impression of objects in motion, and that this can be sufficient to evoke activity in cortical area MT (e.g. Kourtzi, 2004; Kourtzi & Kanwisher, 2000; Senior et al., 2000). Here, we show that specific physical properties that are normally associated with motion, like viscosity, can also be inferred from static snapshots. Under natural viewing conditions, when motion, shape and optical cues are present simultaneously, it seems likely that multiple cues are combined for viscosity perception. Depending on the specific scene characteristics it might be that only some subset of the cues are available or that cue combination or weighting is adapted to current constraints.

The fact that subjects can reliably identify liquids based on their shape also has important consequences for theories of object recognition. Unlike most objects, which have relatively stable shapes, fluids are highly mutable. The fact that a given liquid can take on an enormous variety of different shapes, suggests that the representations of shape cannot be limited to fixed configurations of specific features or parts. Instead, high-level shape representations must also include *statistical* aspects of shape, to capture those characteristics that are common across widely diverging conditions. As mentioned above, Fleming (2014) has recently suggested that the visual system may represent material appearance using statistical generative models, by seeking to predict which image features are most likely to vary systematically across different exemplars. It is interesting to speculate whether the mechanisms that enable the brain to identify liquids across such diverse appearances may also play a role in predicting object appearance across different viewing conditions, which is central to object recognition.

Author contributions

R.W. Fleming and V.C. Paulun contributed to the study design. Stimulus generation, data collection and analysis were performed by V.C. Paulun. R.W. Fleming and V.C. Paulun interpreted the results. Stimuli and preliminary data of a rating task were provided by T. Kawabe, and S. Nishida. V.C. Paulun drafted the manuscript and R.W. Fleming, T. Kawabe, and S. Nishida provided critical revisions. All authors approved the final version of the manuscript for submission.

Declaration of conflicting interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.visres.2015.01.023.

References

Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision, 9*(11), 1–17. http://dx.doi.org/10.1167/9.11.10.

Barrow, H. G., & Tenenbaum, J. M. (1978). Recovering intrinsic scene characteristics from images. In A. Hanson & R. Riseman (Eds.), *Computer vision systems* (pp. 3–26). New York: Academic Press.

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, 110*(45), 18327–18332.

Blender (2.61) [Computer software]. Amsterdam, NL: Stichting Blender Foundation.

Boyaci, H., Doerschner, K., & Maloney, L. T. (2004). Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity. *Journal of Vision, 4*, 664–679. http://dx.doi.org/10.1167/4.9.1, <http://journalofvision.org/4/9/1/>.

D'Zmura, M., & Iverson, G. (1993). Color constancy I: Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America, 10*(10), 2148–2165.

Feldman, J. (1992). Constructing perceptual categories. In *Proceedings of the 1992 I.E.E.E. Conference on Computer Vision and Pattern Recognition* (pp. 244–250). Los Alamitos, CA: I.E.E.E. Computer Society Press.

Feldman, J. (1995). Formal constraints on cognitive interpretations of causal structure. In *Proceedings of the I.E.E.E. Workshop on Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems.* Monterey, CA.

Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences, 103*(47), 18014–18019.

Fleming, R. W. (2014). Visual perception of materials and their properties. *Vision Research, 94,* 62–75.

Fleming, R. W., & Bülthoff, H. H. (2005). Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception, 2*(3), 346–382.

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

Fleming, R. W., Jäkel, F., & Maloney, L. T. (2011). Visual perception of thick transparent materials. *Psychological Science, 22*(6), 812–820.

Gilden, D. L., & Proffitt, D. R. (1989). Understanding collision dynamics. *Journal of Experimental Psychology, 15*(2), 372–383.

Hecht, H. (1996). Heuristics and invariants in dynamic event perception: Immunized concepts or nonstatements? *Psychonomic Bulletin & Review, 3*(1), 61–70.

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2006). How direction of illumination affects visually perceived surface roughness. *Journal of Vision, 6,* 634–648. http://dx.doi.org/10.1167/6.5.8, <http://journalofvision.org/6/5/8/>.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18,* 65–96.

Kawabe, T., Maruya, K., Fleming R. W., & Nishida S. (in press). Seeing liquids from visual motion. *Vision Research*. http://dx.doi.org/10.1016/j.visres.2014.07.003.

Kim, J., & Anderson, B. L. (2010). Image statistics and the perception of surface gloss and lightness. *Journal of Vision, 10*(9), 1–17. http://dx.doi.org/10.1167/10.9.3, <http://journalofvision.org/content/10/9/3/>.

Knoblauch, K., & Maloney, L. T. (2008). MLDS: Maximum likelihood difference scaling in R. *Journal of Statistical Software, 25,* 1–26.

Koffka, K. (1935/1963). *Principles of Gestalt psychology*. New York: Harcourt, Brace & World.

Kourtzi, Z. (2004). But still, it moves. *TRENDS in Cognitive Sciences, 8*(2), 47–49.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience, 12*(1), 48–55.

Leyton, M. (1989). Inferring causal history from shape. *Cognitive Science, 13,* 357–387.

Leyton, M. (1999). *Symmetry, Causality, Mind*. Cambridge, MA: MIT Press.

Maloney, L. T., & Wandell, B. A. (1986). Color constancy: A method for recovering surface spectral reflectance. *Journal of the Optical Society of America, 3*(1), 29–33.

Maloney, L. T., & Yang, J. N. (2003). Maximum likelihood difference scaling. *Journal of Vision, 3,* 573–585. http://dx.doi.org/10.1167/3.8.5.

Marlow, P. J., Kim, J., & Anderson, B. L. (2012). The perception and misperception of specular surface reflectance. *Current Biology, 22*(20), 1909–1913.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science, 210*(4474), 1139–1141.

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. A. (2007). Image statistics and the perception of surface qualities. *Nature, 447,* 206–209.

Nusseck, M., Fleming, R., Langarde, J., Bardy, B., & Bülthoff, H. H. (2007). Perception and prediction of simple object interactions. APGV '07. In *Proceedings of the 4th symposium on applied perception in graphics and visualization* (pp. 27–34). New York, USA: ACM.

R Development Core Team (2011). *R: A language and environment for statistical computing*. 3-900051-07-0. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Senior, C., Barnes, J., Giampietro, V. S., Bullmore, E. T., Brammer, M., & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or 'representational momentum'. *Current Biology, 10,* 16–22.

Spröte, P., & Fleming, R. W. (2013). Concavities, negative parts and the perception that shapes are "complete". *Journal of Vision, 13(14)*(3), 1–23. http://dx.doi.org/10.1167/13.14.3.

Thompson, W. B., Fleming, R. W., Creem-Regehr, S. H., & Stefanucci, J. K. (2011). *Visual perception from a computer graphics perspective*. Wellesley, MA, USA: CRC Press.