Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/5/8/eaaw8845/DC1

Supplementary Materials for

Role of indentation depth and contact area on human perception of softness for haptic interfaces

Charles Dhong*, Rachel Miller, Nicholas B. Root, Sumit Gupta, Laure V. Kayser, Cody W. Carpenter, Kenneth J. Loh, Vilayanur S. Ramachandran, Darren J. Lipomi*

*Corresponding author. Email: cdhong@udel.edu (C.D.); dlipomi@eng.ucsd.edu (D.J.L.)

Published 30 August 2019, *Sci. Adv.* **5**, eaaw8845 (2019)
DOI: 10.1126/sciadv.aaw8845

This PDF file includes:

Comparing the elastic energy of a single feature supported on a substrate

A pillar supported on an elastic substrate deforms and stores elastic energy in three ways: (1) bending of the pillar, (2) stretching of the pillar, or (3) deformation of the underlying substrate.[30] During compression, it is unlikely that the finger will press at an angle greater than 30° from the perpendicular. This is a safe assumption based on our observation of the ways in which the participants interacted with the slabs.

The energy of bending a pillar is given as

$$U_b = \frac{2F_x^2 h^3}{3\pi E r^4} \tag{S1}$$

Where F is the applied force (the x-axis is parallel to the surface of the slab, whereas the z-axis is normal to the surface), r is the radius of the micropatterned pillar, h is the height of the pillar, and E is the Young's modulus.

The energy of compressing a pillar is given as

$$U_p = \frac{F_z^2 h}{2\pi E r^2} \tag{S2}$$

The energy of a pillar deforming the underlying substrate is given as

$$U_D = \frac{8\pi F^2}{27 E r} \tag{S3}$$

Note that all three equations have the same dependence on E and F (to a constant, geometric factor). As a consequence, the relative contribution from bending, compression of a pillar, and deforming the underlying substrate are the same at all moduli and forces. Comparing relative contributions

$$\frac{U_{Total}}{F^2 E r} = \frac{2(\sin(\theta))^2 h^3}{3\pi r^3} + \frac{2(\cos(\theta))^2 h}{2\pi r} + \frac{8\pi}{27} \tag{S4}$$

We obtain dimensionless values of 0.0066, 0.0597, and 0.9308 for the bending, stretching and substrate deformation components, respectively at 30°. The bending component is less than 1% of the stretching and substrate components in all cases. Slabs with a micropatterning of 30% were made of pillars and 50% were made from the negative of pillars (wells). Wells are less susceptible to bending because wells form a contiguous network, whereas pillars are stand-alone features.

Hertzian contact model of a deformable finger

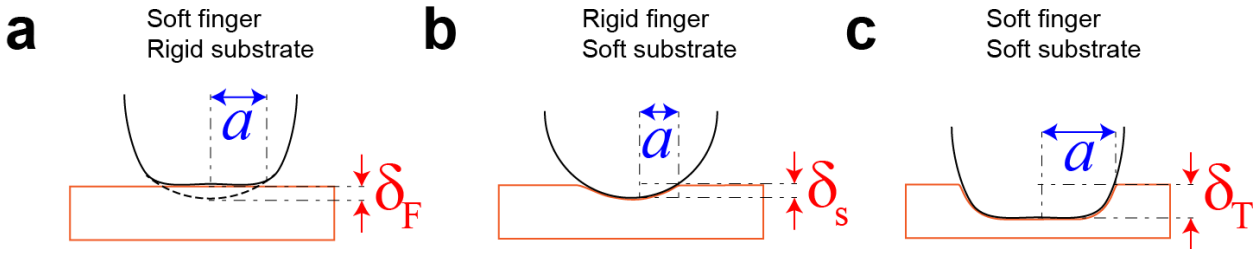

Fig. S1. Schematic of indentation depth and contact area. a) A soft, elastic sphere is pushed into a rigid substrate, which has a contact radius of a and an indentation depth of δ_F. **b**) A rigid sphere indents into a soft, elastic substrate, with a contact radius a and indentation depth into the substrate of δ_s. **c**) A soft, elastic sphere and a soft, elastic substrate indent a total depth δ_T with a contact radius a.

We used the confinement effect of thin films to control the ratio between indentation depth and contact area (or equivalently, the radius), which is otherwise fixed for semi-infinite elastic bodies. In addition to the sample, the finger is also deformable, so we first consider two simpler cases where one object is deformable and the other is rigid. **fig. S1a** considers the case of a deformable finger and shows the contact area and indentation caused by a rigid substrate onto the finger, where the finger is considered as an elastic sphere with a finite thickness. In this case, the finger cannot penetrate into the substrate and deforms inward with a contact radius of a and the interface is displaced by δ_F. In the other case, we consider a deformable planar substrate that is being indented by a rigid finger (**fig. S1b**). In this case, the finger keeps its spherical shape at all applied forces and indents into the substrate by δ_S and a contact radius of a. Finally, we combine a deformable finger pushing into a planar deformable sample with finite thickness (see **fig. S1c**). Qualitatively, the finger acts as a deformable sphere and controls the contact radius. because the contact radius increases with increasing applied force. The contact radius of the finger serves as the radius of a cylindrical flat punch indenting into the substrate. Ordinarily, a cylindrical flat punch has a fixed radius that is insensitive to changes in force but the unique scenario arises shown in **Fig. 1c** arises from two deformable objects with different geometries (planar and spherical).

Model parameters

The model parameters used to calculate all values are given in **Table S1**.

Table S1. Model parameters for the finger and substrate.

Variable	Value	Notes
E_F	750 kPa	Young's modulus of the finger
E_S	0.2, 0.8 or 3 MPa	Young's modulus of the substrate
ν_S, ν_F	0.45	Poisson ratio
R	5 mm	Radius of the finger
h	0.1 to 4.2 mm	Substrate thickness

Two-alternative forced-choice test counterexample

We demonstrate a hypothetical scenario to illustrate the fact that aggregate percentages are not necessarily consistent with head-to-head data. Imagine a scenario where multiple participants were asked to rank several different songs based on which ones were "better" (with no other explanation). Based on participant responses, we could simply average all participant responses and assemble a ranking of which songs were ranked better. Indeed, some songs may even be consistently rated higher than others, for a

variety of cultural reasons. This aggregated percentage, however, is unlikely to be consistent to head-to-head comparisons. It is unlikely that the third-best song was always found to better than the fifth-best song in head-to-head comparisons. It is further unlikely that, across participants, the fourth-best song was better than the fifth-best song in head-to-head comparisons, as well as found to be worse than the third-best song.

Analyzing participant responses

We relate participant responses to the slabs. We calculate an indentation depth and contact area at a range of forces (at a range similar to human touch) for each slab by using the Hertzian contact model. These indentation depths and contact area are then used as the independent variable in the Bradley-Terry model to predict participant responses. Different Bradley-Terry models are constructed by comparing different methods of combining the indentation depth and contact area, as well as comparing, for example, whether the indentation depth follows as the square root, or is linear, with participant responses. These different Bradley-Terry models are then compared to one another through by calculating the Akaike information criterion (AIC) of each model. To compare whether or not a better predictor of participant responses is obtained if we consider the finger is rigid, we recalculate an indentation depth and contact area for each slab, assuming contact with a rigid finger. We then repeat the analysis (calculating different Bradley-Terry models, calculate AIC). Similarly, we verify whether micropatterning had an effect on participant responses by recalculating the contact area and reducing the contact area by the amount predicted from micropatterning. We then repeat the same analysis (calculating different Bradley-Terry models, calculate AIC). A schematic of this analysis shown in **fig. S2**.

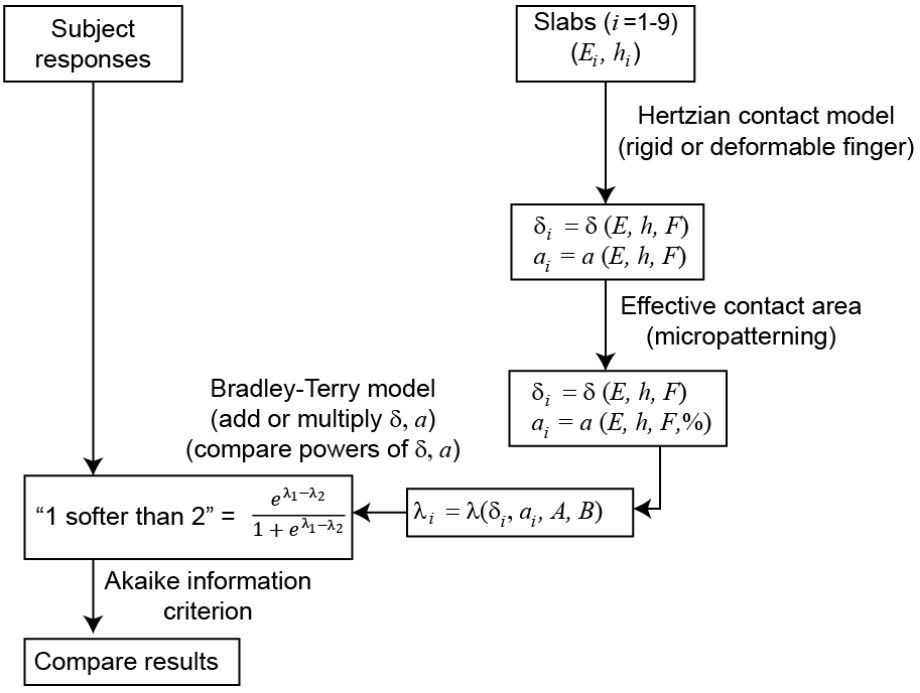

Fig. S2. Flowchart for analyzing participant responses. A Hertzian contact model is used to calculate indentation depth, δ, and contact area (radius, a) at a given force, F. Then the contact area is reduced for some slabs, depending on whether or not the slab was micropatterned. These values are then used to construct λ, which is used to calculate the probability of each participant response to each sample through the Bradley-Terry model. The Bradley-Terry model calculates parameters A and B by minimizing the error between participant responses to all paired comparisons (36 comparisons from 9 samples). The Akaike information criterion of each Bradley-Terry model is calculated, which is used to compare different scenarios (Hertzian models, forces, combinations of δ and a).

AIC for all modeling scenarios

The difference in AIC between the best-fit model and a variety of models is shown in **fig. S3**. These include results from both psychophysical tasks.

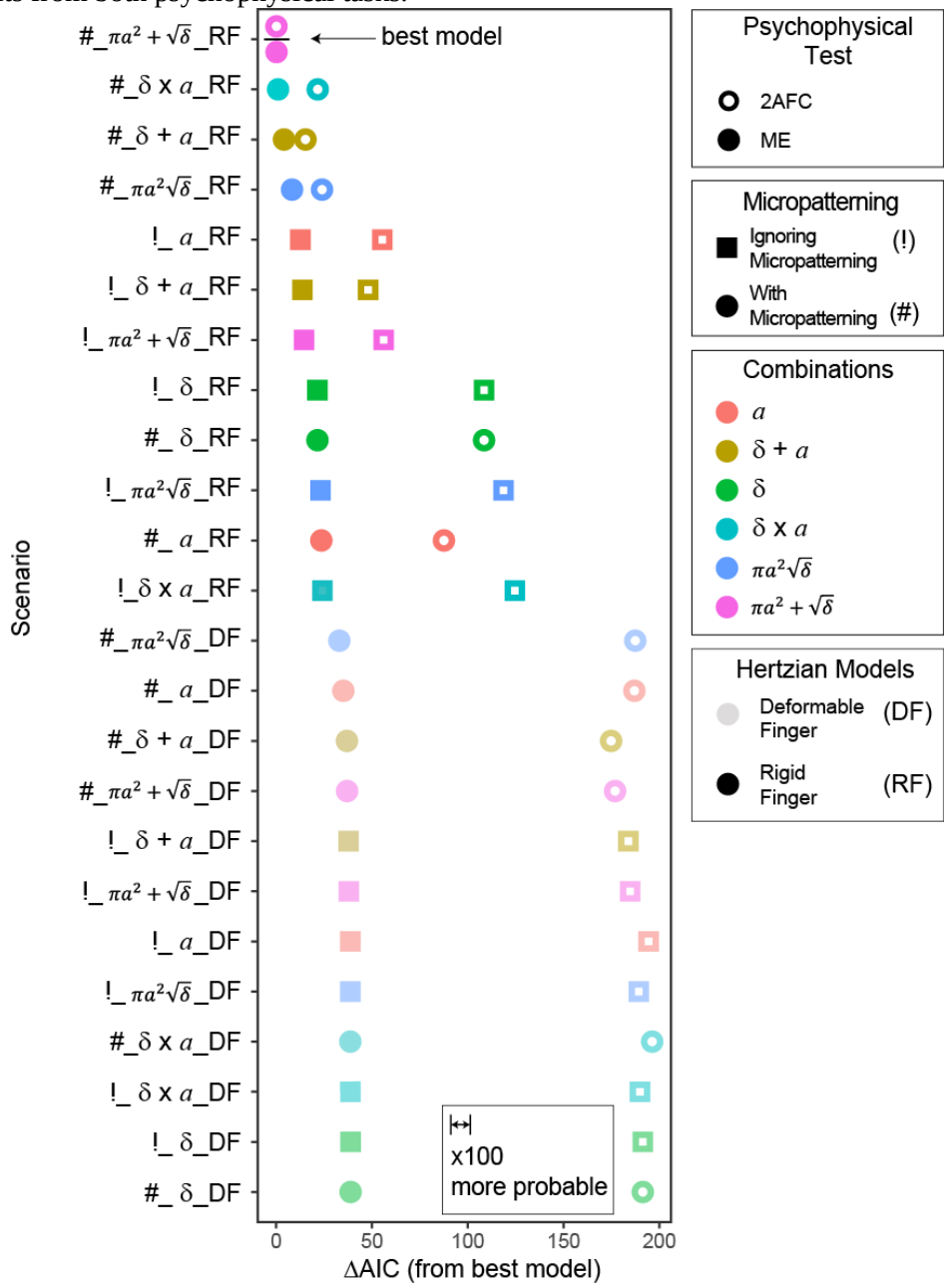

Fig. S3. AIC of all scenarios for both psychophysical tests. Difference in AIC for a variety of scenarios and hypothesis are shown. Open markers represent results from the two-alternative forced choice test (2AFC). Closed markers represent results from the magnitude estimation test (ME). Calculating the contact area including the reduced contact area from micropatterning shown in circular markers, whereas ignoring the reduced contact area shown with circular marker. Title on *y*-axis is redundant with legends, shown for convenience. Naming convention follows accounting for the reduced area from micropatterning (#-includes micropatterning, !-ignores micropatterning), the combination of indentation depth and contact area, and modeling the finger as deformable (DF) or rigid (RF). Note that the scale bar can be translated left or right, but exists on an exponential scale. Therefore, doubling the distance increases the likelihood that one scenario is more probable by 1000-fold.

Finite element model of strain fields

The effect of micropatterned interfaces on the strain fields was simulated and shown in **fig. S4**. A total applied force of 1 N was spread across an interfacial area of 6.4×10^{-5} m^2.

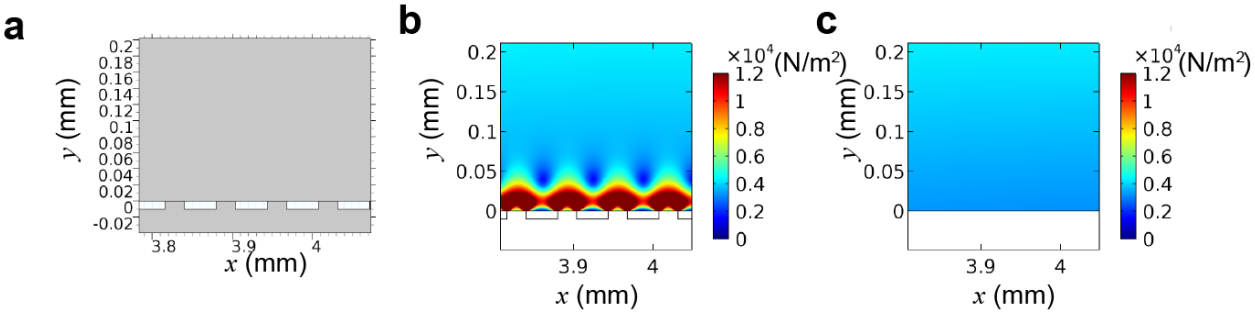

Fig. S4. Finite element modeling of stress between micropatterned surface and finger. In (**a**), model setup of a flat object, representing the soft, outer layer of the finger contacting a micropatterned surface. (**b**) Von Mises stress across a micropatterned surface. (**c**) Same as (b), but for a flat surface. Color scale bar identical for the outputs of both simulations.

The model geometry is shown in **fig. S4a.** Within **fig. S4b**, a heterogenous stress distribution extends between 0.05-0.06 mm into the skin. This heterogenous stress distribution has distinct oscillations in force compared to the flat interface in **fig S4c**.

Testing generalization accuracy using leave-one-out cross-validation

One potential concern is that our equation is overfitted to the participants in our experiment, and that the equation is not generalizable. To address this, we assessed model generalization using leave-one-out cross-validation (LOO CV). For each of the five participants, either Equation 5 (for the two-alternative forced choice test) or Equation 7 (for magnitude estimation) was re-derived using only the other four participants' data, yielding five different equations (one for each participant). These rederived equations were then used to predict the responses for the remaining participant. This LOO CV procedure enables us to estimate the ability of Equation 5 and 7 to predict the responses of new participants whose data did not contribute to the derivation of the equation.

To quantify the accuracy of the predictions in the two-alternative forced choice test, we used a LOO CV model to fit a logistic mixed effects regression with random effect of participant. The intercept term of the model (i.e., model accuracy) was significantly larger than chance ($chance\ \beta_0 = 0, observed\ \beta_0 = 1.30, Wald\ z = 6.19, p < 0.0001$). Furthermore, we compared the prediction accuracy of the LOO CV model to the prediction accuracy of the full (i.e., potentially-overfit) model (**fig. S5**). The LOO CV model was not significantly less accurate than the full model ($Wald\ z = 0.31, p = 0.76$); indeed, the LOO CV model was only 0.55% less accurate than the full model.

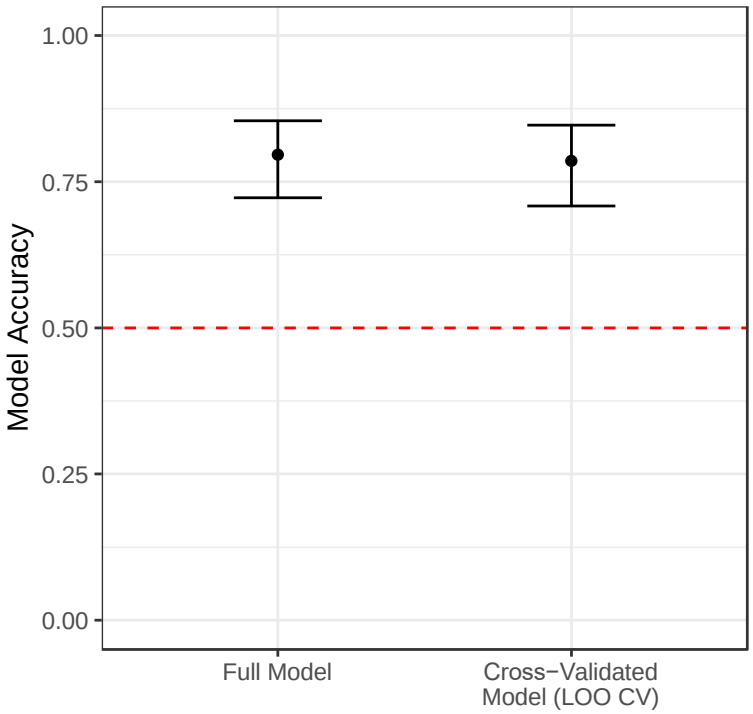

Fig. S5. Leave-one-out cross-validation of participant responses of the two-alternative forced-choice test. Performance of Equation 5 when fit to the full and cross-validated models. Model accuracy is back-transformed from the fixed-term intercept using the inverse logit function. Error bars are 95% confidence intervals. The dotted red line depicts chance performance.

A similar accuracy between the full and cross-validated models both yield suggests that the residual error (~20%) is not due to between-participant differences, but rather due to additional factors (e.g. differences in applied force between participants during the moment of judgement). Furthermore, the similar accuracy suggests that equations 5 is not overfit, and is likely to yield similar accuracy when applied to new participants.

For Equation 7, in magnitude testing, "accuracy" is not as well-defined because participant responses are not binary. Instead, we quantified the average absolute deviation between the prediction of Equation 7 and participant responses. For example, if our equation predicts a softness of 8, and the participant indicates a softness of 6, the absolute deviation for this trial would be 2. As before, we can test generalization by comparing the average absolute deviation for the predictions of the full model to the average absolute deviation for the predictions of a LOO CV model. On average, the model predictions for the LOO CV model were within 1.59 units (on the 1-10 scale) of the participant's true responses, whereas the model predictions for the full model were within 1.54 units (on the 1-10 scale). This is not a statistically-significant difference in average absolute deviation ($t(87.96) = -0.18, p = 0.86$); indeed, the average absolute deviation for the LOO CV model was only 0.047 units (on the 1-10 scale) larger than in the full model.

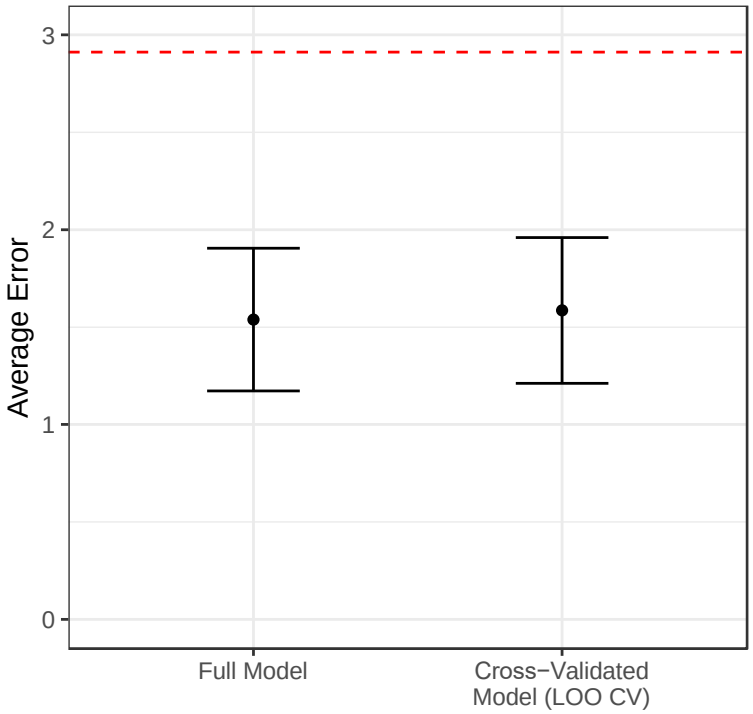

Fig. S6. Leave-one-out cross-validation of participant responses of the magnitude estimation test. Performance of Equation 7 when fit to the full and cross-validated models. Absolute deviation are in units of the 1-10 softness scale. Error bars are 95% confidence intervals. The dotted red line depicts the average deviation of a model for each participant.

As with the two-alternative forced choice test, the insignificant difference between the average absolute deviation of the full model and the leave-one-out model shows that Equation 7 is not overfit. Therefore, additional participants are likely to yield similar results to Equation 7.

Validation of findings on a second set of slabs

We validated our findings on a larger number of participants by repeating the two-alternative forced choice test and magnitude estimation with 10 additional participants on a related, but different set of slabs. These slabs had similar thicknesses and micropatterning, but had lower moduli (63 kPa, 592 kPa, and 810 kPa, instead of 0.1 MPa, 0.8 MPa, and 3 MPa, respectively—see **Table S1**). This set of slabs, as compared to the slabs in the main text, are more similar to one another. Therefore, the psychophysical tasks were somewhat more difficult for the participants.

Table S2. Slab parameters.

Slab		**Young's Modulus** *E* (MPa)	**Thickness** *h* (mm)	**Effective surface area** (% of original surface area)	**Marker**
Yellow	(**Y**)	0.80	4.20	30	■
Blue	(**B**)	0.80	1.40	30	■
Neon green	(**Ng**)	0.80	0.60	50	■
Red	(**R**)	0.60	0.58	100	▲
Green	(**G**)	0.60	0.50	100	▲
Brown	(**Wn**)	0.60	0.30	100	▲
Orange	(**O**)	0.06	0.40	30	●
Purple	(**P**)	0.06	0.20	50	●
Pink	(**K**)	0.06	0.13	100	●

The results from human testing is shown below in **fig. S7**.

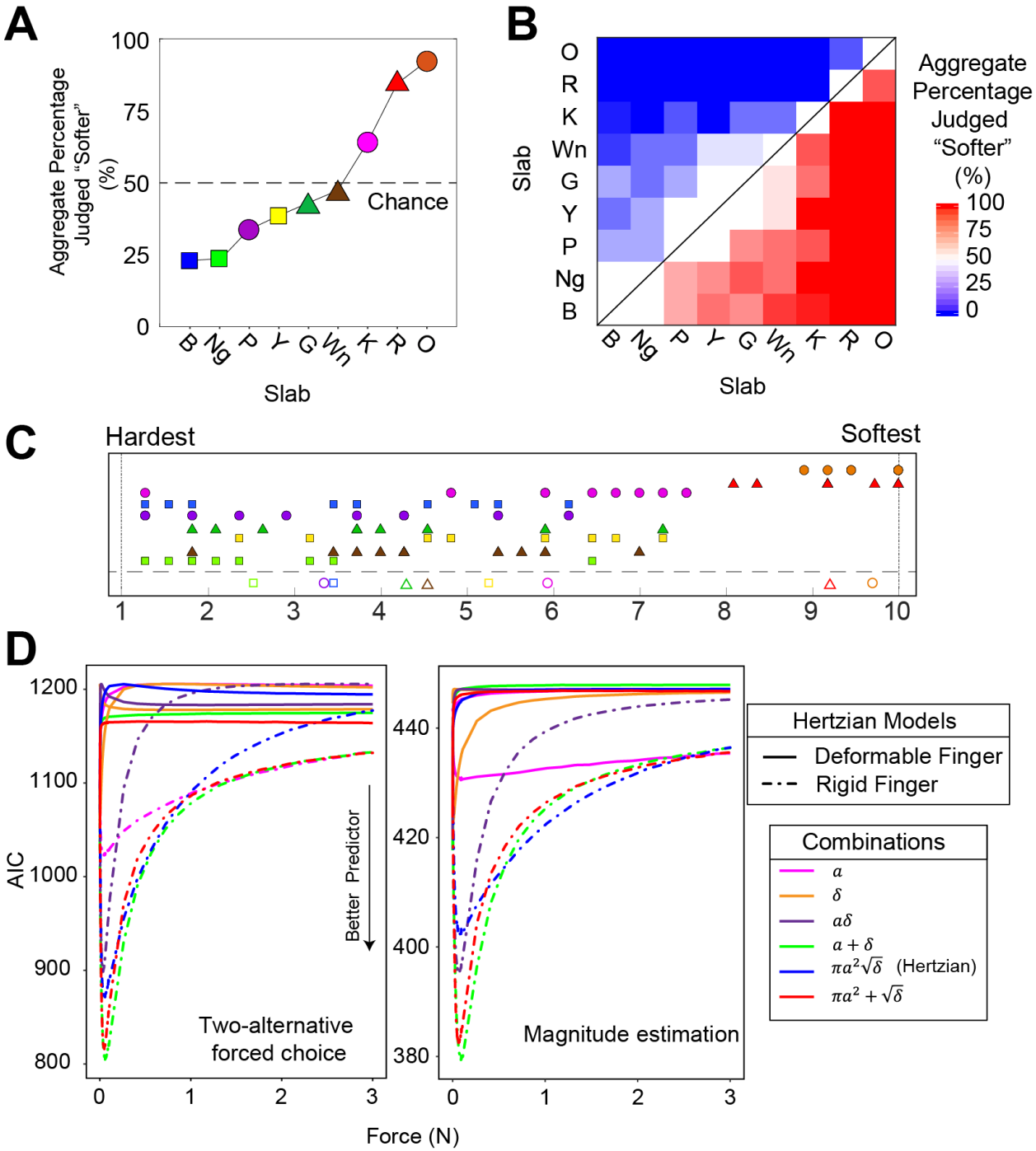

Fig. S7. Validating results with psychophysical testing on a second set of slabs. (**A**) Aggregate percentage of times a slab was judged as softer than another in the two-alternative forced choice test. (**B**) Head-to-head comparisons from the two alternative-forced choice tests. (**C**) Results from the magnitude estimation test. (**D**) Predictive power by AIC of different Hertzian contact models and combinations of indentation depth and contact area. n = 10 in both psychophysical tests.

The results from the two-alternative forced choice test are shown in **fig. S7A-B**. In **fig. S7A**, we see examples where slabs with a lower modulus (**P**) are not often chosen to be soft. Conversely, the second-

softest sample (**R**) was fabricated with an intermediate modulus. We confirm that participants are evaluating the perceived softness of an object on univariate scale as the majority of responses are transitive, as shown in **fig. S7B**. There are three comparisons which are a toss-up (50%), which participants found difficult to distinguish between. Overall, the 33 other comparisons upheld transitivity. The results here support the finding that softness is a univariate and basic tactile sensation preserved between participants.

Participants were asked to evaluate the relative softness of the samples in a magnitude estimation test. These responses are shown in **fig. S7C**. Participants routinely found certain samples softer than others. The order of softness broadly follows the order established by the two-alternative forced-choice test in **fig. S7A**.

A comparison of the different way to model participant responses as a function of the indentation depth and contact is shown in **fig. S7D and S7E** for the two-alternative forced choice test and magnitude estimation, respectively. In both tests, considering the finger as a rigid (dashed lines) object, instead of deformable object was more predictive, confirming the findings in the main text. Furthermore, the contact area and indentation depth were again found to be additive. Finally, the best-fit model in **fig. S7D-E** are when the indentation depth is added with the contact radius (green line). The second best-fit model was the square root of the indentation depth added with the contact area (red line). In the main text, this order was reversed. Instead of a discrepancy, this adds support that both the indentation depth and contact area are important ($p < 2\times10^{-16}$ and $p < 1\times10^{-12}$, for the two-alternative forced choice and magnitude estimation tasks, respectively) and are both independent tactile cues. We note that both the green line and red line represented the top two options from all modeling scenarios and are, in fact, relatively similar in AIC in **fig. S7D-E**. That is, the difference between the green line and red line is not statistically significant. Based on the results here, the probability of judging slab 1 as softer than slab 2 is given as

$$Probability\ of\ Slab\ 1\ softer\ than\ Slab\ 2 = \frac{e^{\lambda_1-\lambda_2}}{1+e^{\lambda_1-\lambda_2}} \tag{S5}$$

$$\lambda_i = 871\ [\mathrm{m}^{-\frac{1}{2}}]\sqrt{\delta_i} + 4.14\times10^6\ [\mathrm{m}^{-2}]\pi a_i^2$$

The relative softness of a slab is given as

$$Softness = -9.78 + 1601[\mathrm{m}^{-\frac{1}{2}}]\sqrt{\delta} + 6.12\times10^6[\mathrm{m}^{-2}]\pi a^2 \tag{S6}$$

Here, note that the magnitude estimation was performed on slabs with a different modulus than the main text. However, relative relationships are still valid. Even with different sets of slabs, the coefficients are relatively similar, ranging from a factor of 1.3 to 4, which confirms the consistency in the human perception of softness. Finally, both models for the two-alternative force choice and magnitude estimation tasks models showed large effect sizes (per Cohen's rule-of-thumb guidelines of >0.25) of with McFadden's pseudo $r^2 = 0.34$ and $r^2 = 0.54$, respectively.

Power analysis

A power analysis was performed on the first set of slabs in the main text. In the two-alternative forced choice task, based on the pseudo r^2 value, a replication using just one participant obtains a power of 0.995 and with ten participants, power > 0.9999. In the magnitude estimation task, replication with two participants has a power of 0.86 and ten participants have a power > 0.9999.