Classification:

Biological Sciences: Neuroscience

Title:

# *How humans grasp three-dimensional objects*

Authors:

Lina K. Klein [a,†] , Guido Maiello [a,†,*], Vivian C. Paulun [a], Roland W. Fleming [a]

[a] Department of Experimental Psychology, Justus-Liebig University Giessen, Giessen 35394, Germany

* Corresponding Author:

Guido Maiello

Department of Experimental Psychology, Justus-Liebig University Giessen, Otto-Behaghel-Str.10F, Giessen 35394, Germany

Email: guido_maiello@yahoo.it

[†] joint first authors; these authors contributed equally to this work

Keywords:

Grasping | Visual Grasp Selection | Precision Grip | Shape | Material

Author Contributions:

GM, LKK, VCP and RWF conceived and designed the study. LKK collected the data. LKK and GM analyzed the data. GM developed the computational model of grasp selection. All authors wrote the manuscript.

## Abstract

We rarely experience difficulty picking up objects, yet of all potential grasp points on an object's surface, only a small proportion yield stable, comfortable grasps. Here, we present extensive behavioral data alongside a computational model that correctly predicts human precision grasping of unfamiliar 3D objects. We tracked participants' forefinger and thumb as they picked up objects of 10 wood and brass cubes configured to tease apart effects of shape, weight, orientation, and mass distribution. Grasps were highly systematic and consistent across repetitions and participants. The model combines five cost functions related to force closure, torque, natural grasp axis, grasp aperture, and visibility. Even without free parameters, we find that the model predicts human grasps with striking fidelity: indeed, it predicts individual grasps almost as well as different individuals predict one another's. Adding fittable weights to the model reveals the relative importance of the different constraints: the combination of force closure, hand posture, and grasp size explains most of human grasping behavior, while our participants cared surprisingly little about minimizing torque and optimizing object visibility. Together, these findings provide a unified account of how we derive effective grasps from objects' 3D shape and material properties to interact with them successfully.

## Significance Statement

Working out how we pick up and interact with objects effectively is one of the most important challenges in behavioral science. Of all the potential contact points on an object's surface, only a small proportion yield effective grasps. Despite this, we rarely experience any difficulty choosing where and how to pick objects up. Here, we present a computational model that unifies the varied and fragmented literature on human grasp selection. We find that the model correctly predicts human grasps across a wide variety of conditions, taking into account the object's 3D shape, material properties and orientation.

## Introduction

In everyday life, we effortlessly grasp and pick up objects without much thought. However, this ease belies its computational complexity. Even state of the art robotic AIs fail to grip objects nearly 20% of the time(1). To pick something up, our brains must work out which locations on the object will lead to stable, comfortable grasps, so we can perform desired actions (Figure 1a). Most potential grasps would actually be unsuccessful, e.g., requiring thumb and forefinger to cross, or failing to exert useful forces (Figure 1b). Even many possible grasps would be unstable, e.g., grasping an object too far from its center, so that it rotates once we try to lift it (Figure 1c). Somehow, the brain must infer which, of all potential grasps, would actually succeed. Despite this, we rarely drop objects or find ourselves unable to complete actions because we are holding them inappropriately. How does the brain select stable, comfortable grasps onto arbitrary 3D objects, particularly objects we have never seen before?

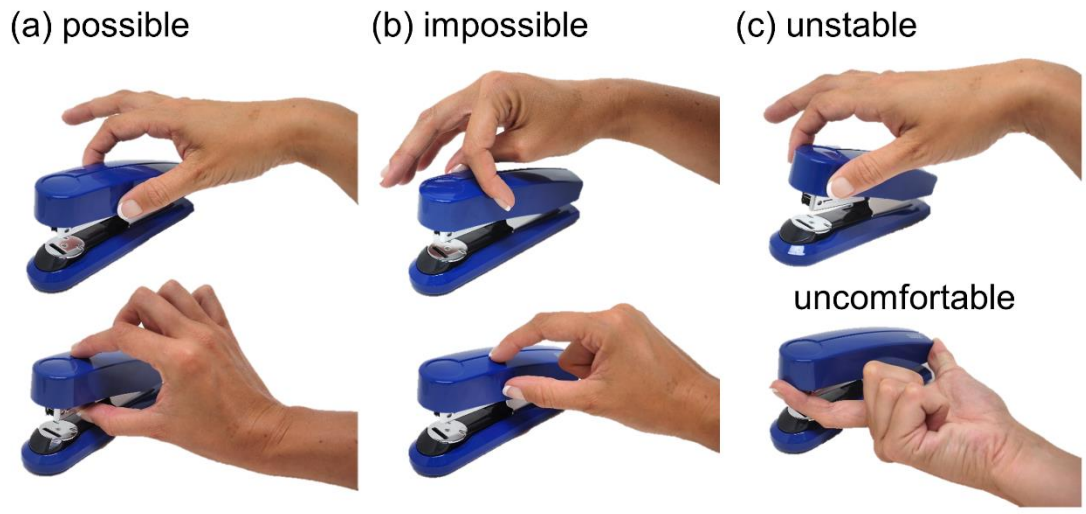



***Figure 1.*** *The computational complexity of grasp selection.* ***(a)*** *Possible* ***(b)*** *Impossible* ***(c)*** *Possible but uncomfortable or unstable grasps.*

Despite the extensive literature describing human grasping patterns and movement kinematics(2–11), little is understood about the computational basis of human grasp selection.

Few authors have attempted to study and model how humans select grasps (e.g. (12, 13)), and even then, only for 2D shapes. This is because, even for two-digit precision grip, many factors influence how we grasp objects. Object shape must be considered, since the surface normals at thumb and index finger contact locations must be approximately aligned (a concept known as force closure(14)), otherwise the object will slip through our fingertips (Figure 1b, bottom). The object's mass and mass distribution must be evaluated, since for grips with high torques (i.e. far from the object's center of mass, or CoM(15–19)) the object will tend to rotate under gravity and potentially slip out of our grasp (Figure 1c, top). The orientation(16, 20–22) and size(23) of each grasp must be considered, since our arm and hand can move and apply forces only in specific ways, and grasps that do not conform to the natural configuration of our hand in 3D space might be impossible (Figure 1b, top), or uncomfortable (Figure 1c, bottom). The hand's positioning may also determine an object's visibility(9, 24–27).

Most previous research on visually guided grasping did not assess the relative importance of these factors, nor how they interact. Here we sought to unify these varied and fragmented findings into a single theoretical and computational framework. We therefore constructed a rich dataset in which we could tease apart how an object's 3D shape, mass, mass distribution, and orientation influence grasp selection. We devised a set of objects made of wood and brass cubes in various 3D configurations (Figure 2), and asked participants to pick them up with a precision grip, move them a short distance and place them at a target location, while we tracked their thumb and forefinger. By varying the spatial configurations of the cubes and orientation of the objects in Experiment 1 we could (1) determine how consistent participants are with themselves and other people, and (2) measure the interactions between allocentric 3D shape and egocentric perspective on those shapes. If actors take the properties of their own effectors into account (e.g., hand orientation, grasp size), we should expect the same 3D shape to be grasped at different locations depending on its orientation relative to the observer(16). In Experiment 2, we varied the mass and mass distribution of the objects (Figure

2c) to test the relative role of 3D shape and mass properties on grasp point selection. If participants take torques into account, identical shapes with different mass distributions should yield systematically different grasps(15, 17–19).

Next, we employed this rich dataset to develop a computational model to predict human grasp patterns. We reasoned that grasps are selected to minimize costs associated with instability and discomfort. Accordingly, we implemented a model that combines five factors computed from the object's shape, mass distribution, and orientation: (i) force closure(14), (ii) torque(15–19) (iii) natural grasp axis(16, 20–22), (iv) natural grasp aperture for precision grip(23) and (v) visibility(24, 25). We find that the model predicts human grasp patterns strikingly well.

## Results:

### *Experiment 1: 3D shape and orientation*

**Human grasps are tightly clustered and represent a highly constrained sample from the space of potential grasps.** We asked 12 participants to grasp four objects made of beech wood presented at two orientations (*Figure 2;* see **Methods**). Figure 3a shows how grasp patterns tend to be highly clustered. In each condition, different grasps have similar size (finger-to-thumb distance) and orientation, and also cover the same portions of the objects. Fitting multivariate Gaussian mixture models to the responses reveals that grasps cluster around only 1, 2, or 3 modes. In Figure 3b we can observe these distinct modes for object U at orientation 2 in a 2D representation of grasp space, where we can also note that human grasps cover only a minute portion of the space of potential grasps. Figure 3c also shows how, for one representative condition, different grasps from the same subjects are more clustered than grasps from different subjects, since individuals predominantly selected only one (70%) or two (27%) modes, and only rarely (3%) grasped objects in three separate locations.

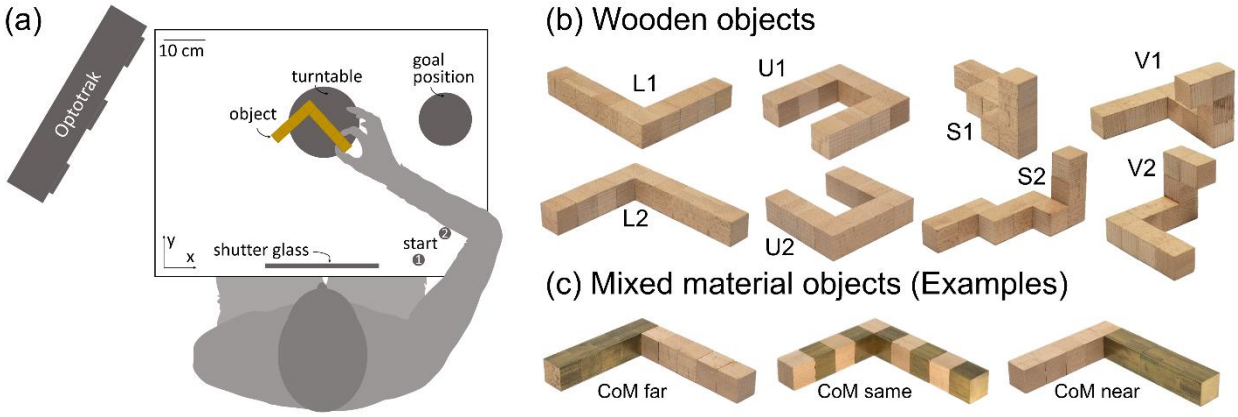
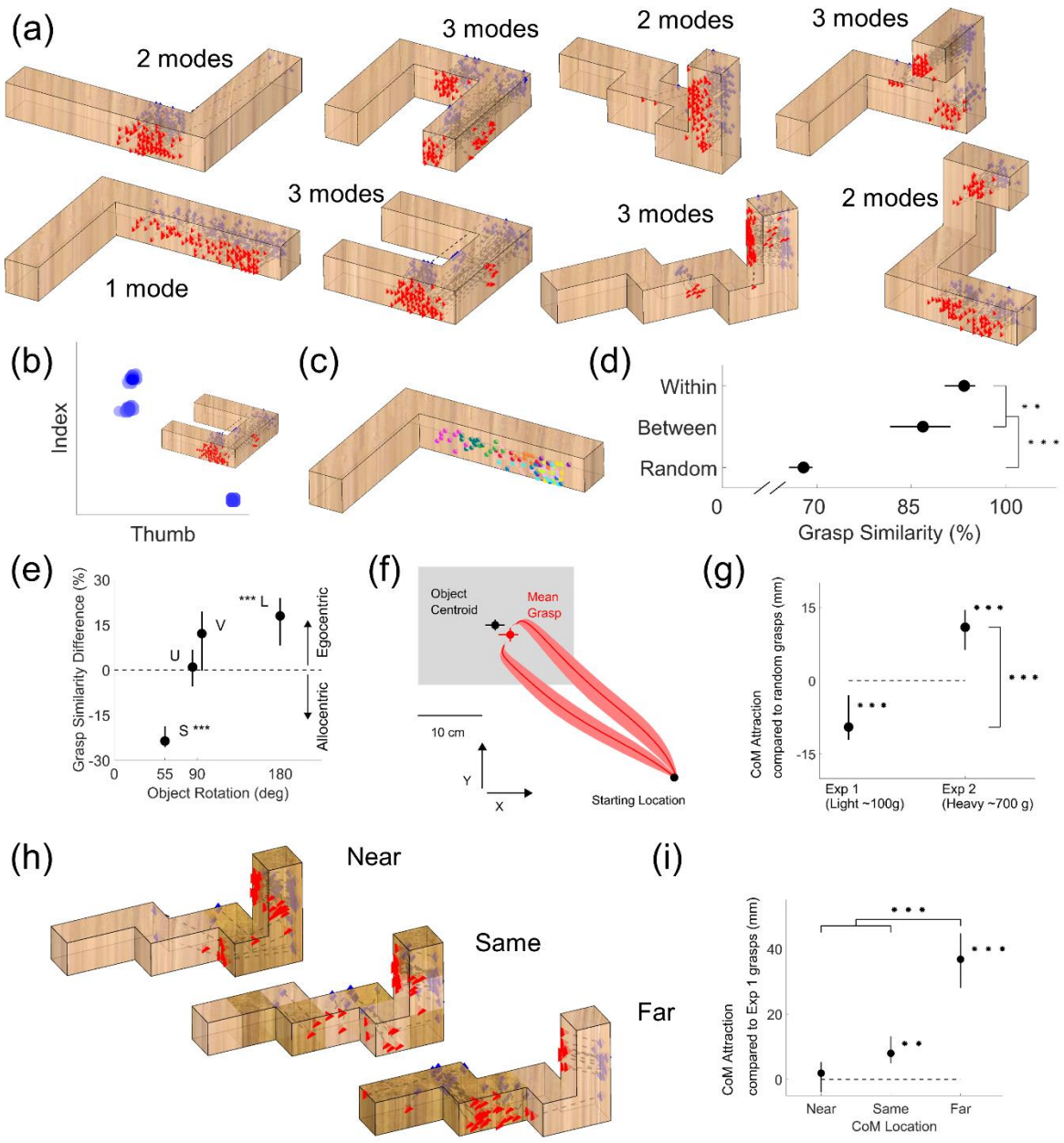



***Figure 2.*** *Setup and stimuli.* ***(a)*** *Experimental setup. Seated participants performed grasping movements with their right hand. Following an auditory signal (coinciding with the shutter window turning transparent) they moved from one of the starting positions to the object and grasped it with a precision grip. They transported and released the object at the goal position and returned to the start position.* ***(b)*** *In Experiment 1 we employed four objects made of wooden cubes. Each object had a unique shape (that here we name L, U, S, V) and was presented at one of two different orientations with respect to the participant.* ***(c)*** *In Experiment 2 the objects had the same shapes as in Experiment 1, but now were made of wood and brass cubes. The brass and wood cubes were organized either in an alternate pattern (middle), so that the CoM of the object would remain approximately the same as for the wooden object, or grouped so that the CoM would be shifted either closer to (right) or away from (left) the participant's hand starting location.*




111

112 ***Figure 3***. *Empirical Results.* ***(a)*** *Human grasps from Experiment 1. Grasps are represented as*

113 *thumb (red triangles) and index finger (blue diamonds) contact positions, connected by dotted*

114 *black lines.* ***(b)*** *Human grasps (blue blobs) for object U, orientation 2, when projected in a 2D*

115 *representation of the space of potential grasps, cluster around three distinct modes.* ***(c)***

116 *Distribution of thumb contact points on object L, orientation 2. Different colors represent grasps*

*from different participants.* ***(d)*** *The level (%) of grasp similarity expected for grasps randomly distributed on the object surface and the observed level of between- and within-participant grasp similarity* ***(e)*** *Difference in grasp similarity across orientations when grasps were encoded in object-centered (allocentric) vs human-centered (egocentric) coordinates, as a function of magnitude of rotation across the two orientation conditions.* ***(f)*** *Average grasp trajectories viewed in the x-y plane (red curves) from start location towards the objects (always contained within the gray shaded region). The average human grasp (red dot) across conditions is biased toward shorter reaching movements compared to the object centroids (black dot).* ***(g)*** *Attraction towards the object CoM for grasps executed onto light (Experiment 1) and heavy (Experiment 2) objects compared to grasps uniformly distributed on the object surfaces (zero reference).* ***(h)*** *Human grasps from Experiment 2 onto object S presented at orientation 2.* ***(i)*** *Attraction towards the object CoM compared to Experiment 1 grasps (zero reference), for Experiment 2 grasps onto heavy objects whose CoM is closer, the same distance as, or farther than the light wooden objects from Experiment 1. In all panels, error bars/regions represent 95% bootstrapped confidence intervals. ** p<0.01, *** p<0.001*

To further quantify how clustered these grasping patterns are we designed a simple metric of similarity between grasps (see **Methods**). Figure 3d shows how both between- and within-subject grasp similarity are significantly higher than the similarity between random grasps due to object geometry ($t(7)=9.76$, $p=2.5*10^{-5}$ and $t(7)=25.11$, $p=4.1*10^{-8}$ respectively). Additionally, within-subject grasp similarity is significantly higher than between subjects ($t(7)=3.89$, $p=0.0060$). Nevertheless, the high level of similarity between grasps from different participants demonstrates that different humans tend to grasp objects in similar ways. The even higher level of within-subject grasp similarity further demonstrates that grasp patterns from individual participants are idiosyncratic, which may reflect differences in the strategies employed by individual participants.

**Findings reproduce several known effects in grasp selection.** First, previous research suggests haptic space is encoded in both egocentric and allocentric coordinates(28), and that grasps are at least partly encoded in egocentric coordinates to account for the biomechanical constraints of our arm and hand(16). Our findings reproduce and extend these observations. For each object we computed grasp similarity across the two orientations in both egocentric and allocentric coordinates. Figure 3e shows that, as the extent of the object rotation increases, grasp encoding shifts from allocentric to egocentric coordinates. Across small rotations (object S, 55 deg rotation), grasps are more similar if encoded in allocentric coordinates ($t(11)=13.90$, $p=2.5*10^{-8}$), whereas for large rotations (object L, 180 degrees) grasps are more similar if encoded in egocentric coordinates ($t(11)=4.59$, $p= 7.8*10^{-4}$). Therefore, both 3D shape as well as movement constraints influence grasps.

Second, Figure 3f shows that participants selected grasps locations that were on average 26 mm closer to the starting location than the object centroid ($t(11)=9.74$, $p=9.6*10^{-7}$), reproducing known spatial biases in human grasp selection (12, 25, 27, 29, 30).

Third, consistent with Kleinholdermann et al (12) but contrary to previous claims(15–19), our findings suggest humans care little about torque when grasping light objects. If actors sought to minimize torque, the selected grasps should be as close as possible to the CoM. In contrast, Figure 3g shows that for the light weight objects in Experiment 1, grasps were on average 9 mm farther from the CoM than the average distance to the object's CoM of grasps uniformly sampled onto the surface of the objects ($t(11)=4.53$, $p=8.6*10^{-4}$).

### *Experiment 2: Mass and Mass Distribution*

**Humans grasp objects close their center of mass when high grip torques are possible.** Due to the low density of beech wood, even the grasps farthest from the CoM in Experiment 1 would produce relatively low torques. Therefore, in Experiment 2 we tested whether participants grasp objects closer to the CoM when higher torques are possible. We did this by using objects

of greater mass and asymmetric mass distributions. Specifically, for each of the shapes in Experiment 1, we made three new objects, each made of five brass and five wooden cubes: two 'bipartite' objects, with brass clustered on one or the other half of the object, and one 'alternating' object, with brass and wood alternating along the object's length. These objects had the same 3D shapes as in Experiment 1, but were nearly tenfold heavier (Figure 2c, see **Methods**).

Figure 3g shows how human grasps are indeed significantly attracted towards the CoM of heavy objects, presumably to counteract the larger torques associated with higher mass. In Experiment 2, grasps were on average 11 mm closer to the object CoM than grasps sampled uniformly from the objects' surfaces ($t(13)=4.94$, $p= 2.7*10^{-4}$), and on average 20 mm closer than the grasps from Experiment 1 ($t(24)=6.63$, $p= 7.4*10^{-7}$). Importantly, participants shifted their grasps towards the CoM—not the geometrical centroid—of the objects (observe how the grasp patterns shift in Figure 3h). Figure 3i shows that when the object CoM was shifted towards the hand starting location, participants did not significantly adjust their grasping strategy compared to Experiment 1 ($t(13)=0.81$, $p=0.43$). Conversely, when the object CoM was in the same position as in Experiment 1, participants shifted their grasps on average by 8 mm towards the CoM ($t(13)=3.92$, $p=0.0017$). When the object CoM was shifted away from the hand starting position, participant grasps were on average 37 mm closer to the object CoM compared to Experiment 1 grasps ($t(13)=8.49$, $p=1.2*10^{-6}$), a significantly greater shift than both the near and same CoM conditions ($t(13)=8.66$, $p=9.2*10^{-7}$ and $t(13)=7.58$, $p=4.0*10^{-6}$). These differential shifts indicate that participants explicitly estimated each object's CoM from visual material cues.

Even with the heavier objects, participants still systematically selected grasp locations that were closer to the starting location than the object centroid ($t(13)=4.03$, $p=0.0014$). However, now participants exhibited only a 9 mm bias, which was significantly smaller than the 26 mm bias observed for the light wooden objects in Experiment 1 ($t(24)=4.67$, $p= 9.6*10^{-5}$).

Together these findings suggest that participants combine multiple constraints to select grasp locations, taking into consideration the shape, weight, orientation, and mass distribution of objects, as well as properties of their own body to decide where to grasp objects. We next sought to develop a unifying model that could predict these diverse effects based on a few simple underlying principles.

***Computational model of human grasp selection.***

Based on the insights gained from our empirical findings, we developed a computational model to predict human grasp locations. The model takes as input 3D descriptions of the objects' shape, mass distribution, orientation, and position relative to the participant, and computes as output a *grasp cost function*, describing the costs associated with every possible combination of finger and thumb position on accessible surface locations (i.e., those not in contact with table). We reasoned that humans would tend to grasp objects at or close to the minima of this cost function, as these would yield the most stable, comfortable grasps. Low cost grasps can then be projected back onto the object to compare against human grasps. It is important to note that this is not intended as a process model describing internal visual or motor representations (i.e., we do not suggest that the human brain explicitly evaluates grasp cost for all possible surface locations). Rather, the model is a way of combining a subset of the factors which are known to influence human grasp selection into a single, unifying framework (12).

For each object, we create a triangulated mesh model in a 3D coordinate frame, from which we can sample (Figure 4a-b). For precision grip, we assume one contact point each for thumb and index finger. Thus, all possible precision grip grasps can be ordered on a 2D plane, with all possible thumb contact points along the x-axis, and on the y-axis, all possible index contacts in the same ordering as for the thumb.

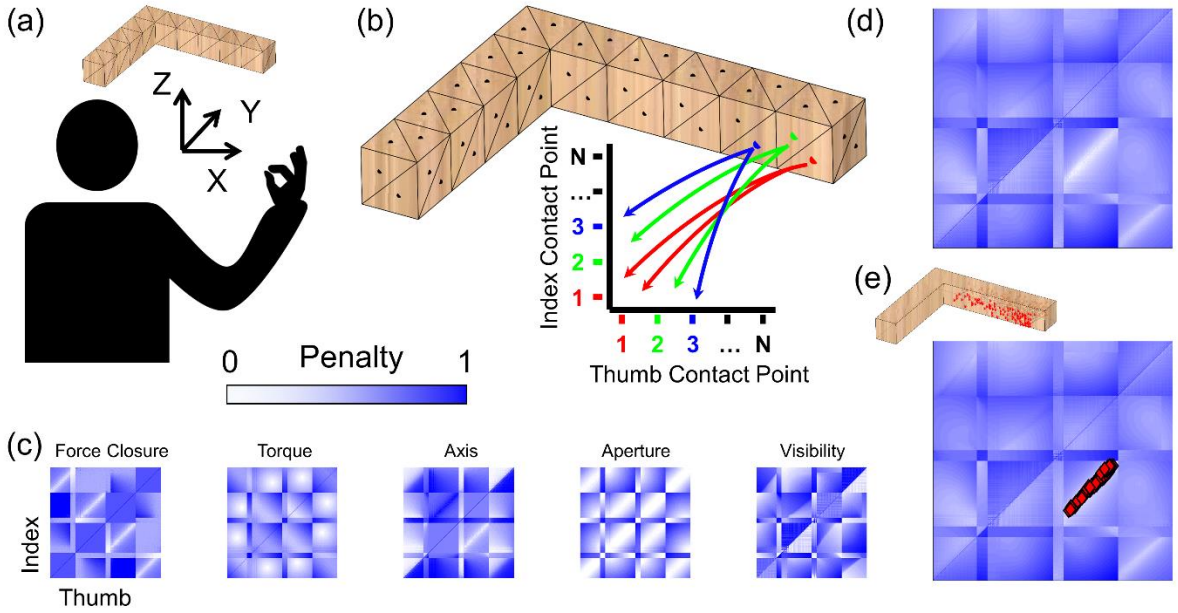



***Figure 4****. A framework that unifies distinct aspects of grasp selection.* ***(a)*** *Mesh model of object in same 3D reference frame as participant poised to execute grasp.* ***(b)*** *Discrete sampling of the surface defines a 2D space containing all potential combinations of index and thumb contact points on the object.* ***(c)*** *Color-coded maps showing penalty values for each potential grasp for each penalty function.* ***(d)*** *Overall penalty function computed as the linear combination of maps in (c).* ***(e)*** *Human grasps projected into 2D penalty-function space neatly align with minimum of combined penalty map.*

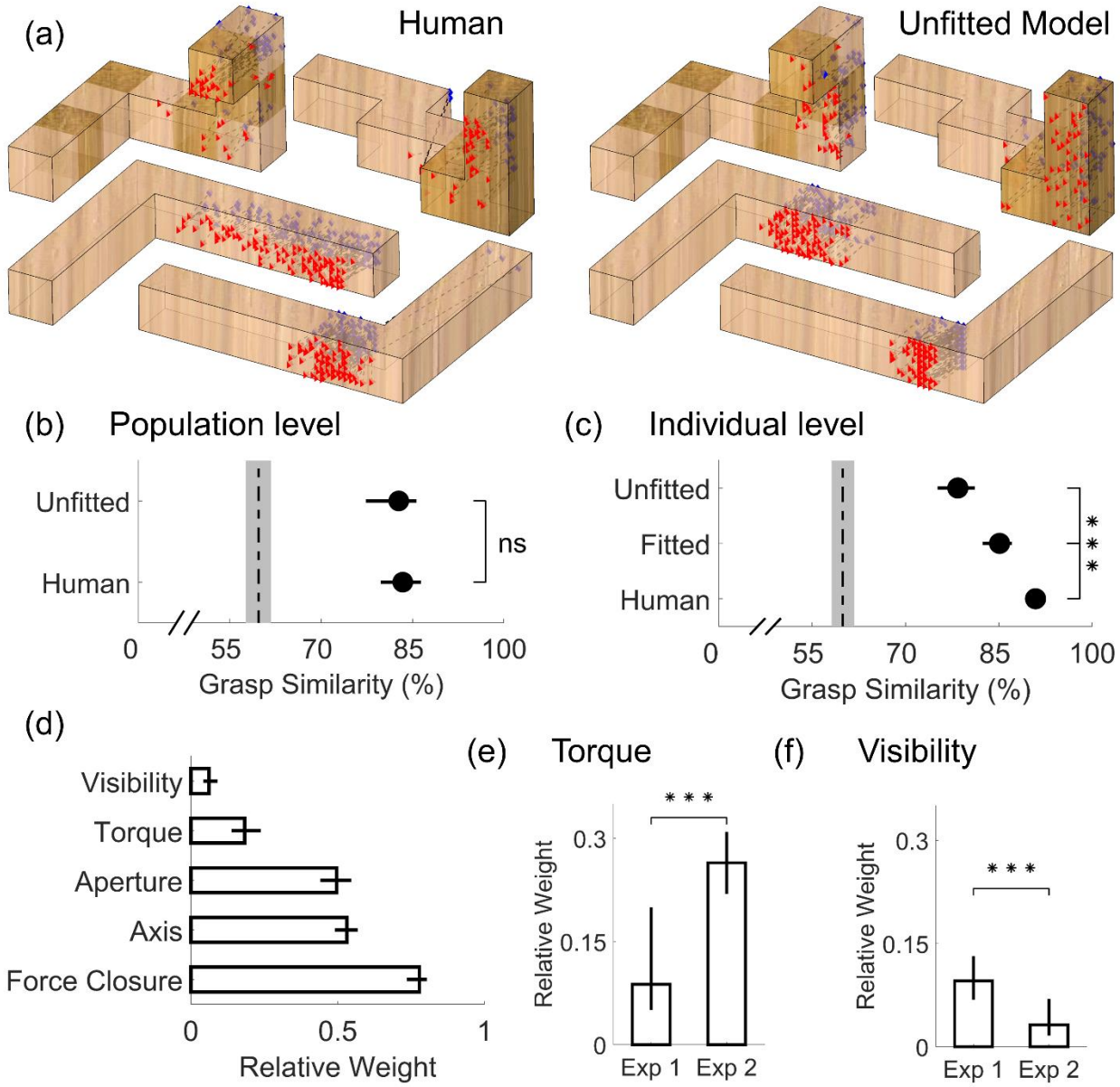
To estimate the cost associated with each grasp, we take a (weighted) linear combination of five *penalty functions*, determined by the physical properties of the graspable object (surface shape, orientation, mass, mass distribution) as well by the physical constraints of the human actuator (i.e. the human arm/hand). Specifically, we consider optimality criteria based on: (i) optimum force closure(14), (ii) minimum torque(15–19), (iii) alignment with the natural grasp axis(16, 20–22), (iv) optimal grasp aperture(23), and (v) optimal visibility(24, 25, 27). (see **Methods** for mathematical definitions). Figure 5(c) shows maps for each penalty function: white indicates low

235 penalty, dark blue high penalty. To compare and combine penalty, values are always
236 normalized to [0,1].
237 **Force closure**: force closure is fulfilled when the two contact-point surface normals, along
238 which gripping forces are applied, are directed towards each other(14). Thus, we penalize
239 lateral offsets between the grasp point normals.
240 **Minimum torque**: grasping an object far from its CoM results in high torque, which causes the
241 object to rotate when picked up(15–19). Large gripping forces would be required to prevent the
242 object from rotating. We therefore penalize torque magnitude.
243 **Natural grasp axis**: when executing precision grip grasps, humans exhibit a preferred hand
244 posture known as the *natural grasp axis*(16, 20–22). Grasps that are rotated away from this axis
245 result in uncomfortable or restrictive hand/arm configurations. We therefore penalize angular
246 misalignment between each candidate grasp and the natural grasp axis (taken from (21)).
247 Unlike force closure and torque, this penalty map is asymmetric about the diagonal.
248 **Optimal grasp aperture**: humans prefer the distance between finger and thumb at contact
249 ('grasp aperture') to be below 2.5 cm(23). We therefore penalize grasp apertures above 2.5 cm.
250 **Optimal visibility**: our behavioral data, and previous studies, suggest humans exhibit spatial
251 biases when grasping. It has been proposed that these may arise from an attempt to minimize
252 energy expenditures through shorter reach movements(24). However, Paulun et al (25) have
253 shown that these biases may in fact arise from participants attempting to optimize object
254 visibility. While our current dataset was not designed to untangle these competing hypotheses,
255 re-analyzing published data (19, 27) confirms that object visibility—not reach length—is most
256 likely responsible for the biases. We therefore penalized grasps that hindered object visibility.
257 We also designed a penalty function for reach length and verified that, since reach length and
258 object visibility are correlated in our dataset, employing one or the other penalty function yields
259 very similar results.
260

We assume that participants select grasps with low overall costs across all penalty functions. Thus, to create the overall grasp penalty function, we take a (weighted) linear sum of the individual penalty maps. The minima of this full penalty map represent grasps that best satisfy all criteria simultaneously. The map in Figure 5d exhibits a clear minimum: the white region in its lower right quadrant.

To assess the agreement between human and optimal grasps, we may visualize human grasps in the 2D representation of the grasp manifold. The red markers in Figure 5(e) are the human grasps from object L at orientation 2, projected in 2D and overlain onto the full penalty map. Human grasps neatly align with the minima of the penalty map.

**Model Fitting.** The simple, equally-weighted combination of constraints considered thus far agrees with human grasping behavior surprisingly well. However, it is unlikely that actors treat all optimality criteria as equally important. Different persons likely weight the constraints differently (e.g., due to strength or hand size). Therefore, we developed a method for fitting full penalty maps to participants' responses. We assigned variable weights to each optimality criterion, and fit these weights to the grasping data from each participant, to obtain a set of full penalty maps whose minima best align with each participant's grasps (see **Methods**).




277

278 ***Figure 5.*** *Computational Results.* ***(a)*** *Grasping patterns predicted through the computational*

279 *framework (right) closely resemble human grasps onto real objects varying in shape,*

280 *orientation, and material (left). Simulated grasp patterns are generated with no knowledge of our*

281 *human data (i.e. model not fit to human grasps).* ***(b)*** *Population level grasp similarity, i.e.*

282 *similarity of human and unfitted model grasps to median human grasp across all participants.*

283 ***(c)*** *Individual level grasp similarity, i.e. similarity of human, unfitted, and fitted model grasps to*

284 *the median grasp of each participant. In panels (b,c), dashed line is estimated chance level of*

*grasp similarity due to object geometry, bounded by 95% bootstrapped confidence intervals.* ***(d)*** *Pattern of fitted weights across Experiments 1 and 2. (e) Relative weight of the minimum torque constraint in Experiments 1 and 2. (f) Relative weight of the visibility constraint in Experiments 1 and 2. Data are means; error bars, 95% bootstrapped confidence intervals. ***p<0.001*

**Model grasps are nearly indistinguishable from measured human grasps.** To compare human and optimal grasps directly, we can sample predicted optimal grasps from around the minimum of the full penalty map (see **Methods**) and project back onto the objects. Figure 5a shows human grasps (left) and unfitted model predictions (right) on a few representative objects (see Supplementary Figure S1 for complete set). Human and predicted grasps have similar size and orientation, and also cover similar portions of the objects.

Figure 5b depicts grasp similarity at the population level, i.e., across participants and between human and unfitted model grasps. Grasp similarity between participants was computed (for each object and condition), as the similarity between the median grasp of each participant and the median grasp across all others. Grasp similarity between human and model grasps was computed as the similarity between the median unfitted model grasp and the median grasp across all participants.

Unfitted model grasps were significantly more similar to human grasps than chance ($t(31)=10.79$, $p=5.0*10^{-12}$), and effectively indistinguishable from human-level grasps similarity ($t(31)=0.31$, $p=0.76$). Note that this does not mean our current approach perfectly describes human grasping patterns; it suggests instead that our framework is able to predict the median human grasping patterns nearly as well as the grasps of a random human on average approximate the median human grasp.

**Fitting the model can account for individual grasp patterns.** In both Experiments, participants repeatedly grasped the same objects in randomized order. Figure 5c depicts how similar human and model grasps are to the median grasp of each individual participant in each

311 experimental condition. Individual subjects are highly consistent when grasping the same object
312 on separate trials. Grasps predicted through our framework with no knowledge of the empirical
313 data were significantly less similar to the median grasps of individual humans ($t(31)=9.33$,
314 $p=1.6*10^{-10}$). This is unsurprising, since the unfitted model predicts the average pattern across
315 observers, but there is no mechanism for it to capture idiosyncrasies of individual humans.
316 Fitting the model to the human data (see **Methods**) significantly improved grasp similarity
317 ($t(31)=5.00$, $p=2.1*10^{-5}$). Note however that model grasp patterns fit to a single participant are
318 still distinguishable from random real grasps by the same individual ($t(31)=4.85$, $p=3.3*10^{-5}$).
319 **Force closure, hand posture, and grasp size explain most of human grasp point**
320 **selection.** The pattern of fitted weights across both experiments (Figure 5d) reveals the relative
321 importance of the different constraints. Specifically, we find that force closure is the most
322 important constraint on human grasping, which makes sense because force closure is a
323 physical requirement for a stable grasp. Next in importance are natural grasp axis and optimal
324 grasp aperture, both constraints given by the posture and size of our actuator (our hand). In
325 comparison, participants appear to care only marginally about minimizing torque, and almost
326 negligibly about object visibility.
327 **Analyzing the patterns of fitted weights confirms our empirical findings.** The model also
328 replicates our main empirical findings in a single step. Figure 5e shows that the relative
329 importance of torque was much greater for the heavy objects tested in Experiment 2 compared
330 to the light objects from Experiment 1 ($t(24)=4.40$, $p=1.9*10^{-4}$). Conversely, Figure 5f shows that
331 the relative importance of object visibility instead decreased significantly from Experiment 1 to
332 Experiment 2 ($t(24)=3.07$, $p=0.0053$). Additionally, by simulating grasps from the fitted model,
333 we are able to recreate the qualitative patterns of all behavioral results presented in Figure 3
334 (see Supplementary Figure S2).
335
336

## Discussion:

We investigated how an object's 3D shape, orientation, mass, and mass distribution jointly influence how humans select grasps. Our empirical analyses showed that grasping patterns are highly systematic, both within and across participants, suggesting that a common set of rules governs human grasp selection of complex, novel 3D objects. Our findings reproduce, unify, and generalize many effects observed previously: (1) both 3D shape and orientation determine which portion of the object people grasp (8, 12, 15, 16, 31–34).; (2) humans exhibit spatial biases even with complex 3D objects varying in shape and mass(12, 25, 27, 29, 30); (3) object weight modulates how much humans take torque into account when selecting where to grasp objects(15–19). We then combined this diverse set of observations into a unified theoretical framework that predicts human grasping patterns strikingly well, even with no free parameters. By fitting the computational model to human behavioral data, we showed that force closure, hand posture, and grasp size are the primary determinants of human grasp selection, whereas torque and visibility modulate grasping behavior to a much lesser extent.

**3D Shape** Behavioral research on the influence of shape on grasping is surprisingly scarce, primarily employs 2D or simple geometric 3D stimuli of uniform materials, and rarely investigates grasp selection (8, 15, 16, 31–34). For example, by using 3D stimuli that only varied in shape by a few centimeters, Schettino et al(33) concluded that object shape influences hand configuration only during later phases of a reaching movement during which subjects use visual feedback to optimize their grasp. Here, we show that distinct 3D shapes are grasped in systematically distinct object locations, and our behavioral and model analyses can predict these locations directly from the object 3D shape.

**Orientation** When grasping spheres or simple geometrical shapes, humans exhibit a preferred grasp orientation (the NGA) (16, 20–22), and most previous work on how object orientation influences grasping has primarily focused on hand kinematics(15, 19, 32, 35). Conversely, with more complex 3D shapes we show that the same portion of an object is selected within a range

of orientations relative to the observer, whereas for more extreme rotations the grasp selection strategy shifts significantly. Therefore, object shape and orientation together determine which portion of an object will be grasped, and thus the final hand configuration.

**Spatial Biases** The spatial biases we observe are consistent with participants attempting to increase object visibility(25, 27), and our data also replicate the finding that these biases are reduced when object weight increases(19, 25).

**Material/Weight/Torque** Goodale at al(15) were among the first to show that participants tend to grasp objects through their CoM, presumably to minimize torque. Lederman and Wing(16) found similar results, yet in both studies low-torque grasps also correlated with grasps that satisfied force closure and aligned with the natural grasp axis. Kleinholdermann et al(12) found torque to be nearly irrelevant in grasp selection, yet Paulun et al(19) observed that grasp distance to CoM was modulated by object weight and material. Our findings resolve these conflicting findings. By using stimuli that decorrelate different aspects of grasp planning, we find that shape and hand configuration are considerably more important than torque for light weight objects, and that the importance of minimizing torque scales with mass. Additionally, shifting an object's mass distribution significantly attracted grasp locations towards the object's shifted CoM, demonstrating that participants could reliably combine global object shape and material composition to successfully infer the object's CoM.

**Computational Modelling** Previous models of grasping have mainly focused on hand kinematics and trajectory synthesis(2–6) whereas we attempt to predict which object locations will be selected during grasping. Our modelling approach takes inspiration from Kleinholdermann et al(12), which to the best of our knowledge is the only previous model of human two-digit contact point selection, but only for 2D shape silhouettes. In addition to dealing with 3D objects varying in mass, mass distribution, orientation, and position, our modeling addresses several limitations of previous approaches. The fitting procedure quantifies the relative importance of different constraints, and can be applied to any set of novel objects to test

how experimental manipulations affect this relative weighting. The modular nature of the model allows additional constraints to be included, excluded or given variable importance. For example, we know that end-state comfort of the hand plays a role in grip selection(36, 37), yet the tradeoff between initial and final comfort is unclear(38). By varying the participants' task to include object rotations, and by including a penalty function penalizing final hand rotations away from the natural grasp axis, it would be possible to assess the relative importance of initial, final (or indeed intermediate) hand configurations on grasp planning. The modelling could also be extended to multi-digit grasping, by adding to each penalty function three dimensions for each additional finger considered (the x,y,z coordinates of the contact point). This approach is consistent with (and complementary to) the approach by Smeets and Brenner(2, 5), who posit that grasping is a combination of multiple pointing movements. Future models should also generalize from contact points to contact patches of nonzero area, as real human grasp locations are not only points but larger areas of contact between digit and object. To facilitate such developments, we provide all data and code (doi:xx.xxxx/zenodo.xxxxxxx upon publication).

**Neuroscience of Grasping** While our model is not meant as a model of brain processes, there are several parallels with known neural circuitry underlying visual grasp selection (for reviews see(39–41)). Of particular relevance is the circuit formed between the Ventral Premotor Cortex (Area F5), Dorsal Premotor Cortex (Area F2), and the Anterior Intraparietal Sulcus (AIP). Area F5 exhibits 3D-shape-selectivity during grasping tasks and is thought to encode grip configuration given object shape(42–44), whereas area F2 encodes the grip-wrist orientation required to grasp objects under visual guidance(45). Both regions exhibit strong connections with AIP, which has been shown to represent the shape, size, and orientation of 3D objects, as well as the shape of the handgrip, grip size, and hand-orientation(46). Additionally, visual material properties, including object weight, are thought to be encoded in the ventral visual cortex(47–51), and it has been suggested that AIP might play a unique role in linking

components of the ventral visual stream involved in object recognition to hand motor system(52). Therefore, the neural circuit formed between F5, F2, and particularly AIP is a strong candidate for combining the multifaceted components of visually guided grasping identified in this work(53–57). Combining targeted investigations of brain activity with the behavioral and modelling framework presented here holds the potential to develop a unified theory of visually guided grasp selection.

## Materials and Methods:

### *Participants*

Twelve naïve participants (5 males and 7 females between the ages of 20 – 31, mean age: 25.2 years) participated in Experiment 1. Fourteen naïve participants (9 males and 5 females between the ages of 21 and 30, mean age: 24.4 years) participated in Experiment 2. Participants were students at the Justus-Liebig-University Giessen, Germany and received monetary compensation for participating. All participants reported having normal or corrected to normal vision and being right handed. All procedures were approved by the local ethics board and adhered to the declaration of Helsinki. All participants provided written informed consent prior to participating.

### *Apparatus*

Experiments were programmed in Matlab version R2007a using the Optotrak Toolbox by V. H. Franz(58). Participants were seated at a table with their head positioned in a chinrest (Figure 2a), in front of an electronically controlled pane of liquid crystal shutter glass(59), though which only part of the table was visible and which became transparent only for the duration of a trial. Objects were placed at a target location, 34 cm from the chinrest in the participant's sagittal plane. Small plastic knobs placed on participants' right side specified the hand starting positions. A plate (28.5 cm to the right of the goal location and with a 13 cm diameter at 26 cm

from start position 1 in the participant's sagittal plane) specified the movement goal location. We tracked participants' fingertip movements using an Optotrak 3020 infrared tracking system. The Optotrak cameras were located to the left of the participants. To record index finger and thumb movement, sets of three infrared markers (forming a rigid body) were attached to the base of the participants' nails. The fingertip and tip of the thumb were calibrated in relation to the marker position, as participants grasped a wooden bar with a precision grip, placing their fingertips at two known locations on the bar.

***Stimuli***

**Experiment 1: Light objects made of wood.** Four differently shaped objects (defined as objects L, U, S and V; Figure 2b) each composed of 10 wooden (beech) cubes ($2.5^3$ $cm^3$), served as stimuli. Objects were fairly light with a mass of 97 g. Two of the objects featured cubes stacked on top of each other, whereas the other two objects were composed exclusively of cubes lying flat on the ground. The objects were presented to the participants at one of two orientations. Across orientations, object L was rotated by 180 degrees, objects U and V were rotated by 90 degrees, and object S was rotated by 55 degrees. Figure 2b shows the objects positioned as if viewed by a participant.

**Experiment 2: Heavy composite objects made of wood and brass.** For each of the 4 shapes from Experiment 1, we created 3 new objects (12 in total) to serve as stimuli for Experiment 2 (Figure 2c). Individual cubes were made of either wood or brass. The objects were composed of 5 cubes of each material, which made them fairly heavy with a mass of 716g. By reordering the sequence of wood and brass cubes, we shifted the location of each shape's CoM. For each shape we made one object in which brass and wooden cubes alternated with one another, and two bipartite objects, where the 5 brass cubes were connected to one another to make up one side of the object with the wooden cubes making up the other side. This configuration was also inverted, (i.e., wooden and brass cubes switched locations). All objects were presented at the same two orientations as Experiment 1.

**Object meshes.** Triangulated mesh replicas of all objects were created in Matlab; each cube face consisted of 128 triangles. To calibrate mesh orientation and position, we measured, using the Optotrack, four non planar points on each object at each orientation. We aligned the model to the same coordinate frame employed by the Optotrack using Procrustes analysis.

***Procedure***

Prior to each trial, participants placed thumb and index finger at a pre-specified starting location. In Experiment 1, two start locations were used (start 1 at 28 cm to the right of the chinrest in the participant's coronal plane and 9.5 cm forward in the sagittal plane; start 2 9 cm further to the right and 3 cm further forward, 23 cm from the center of the goal plate). Given that we observed no effect of starting position in our data, in Experiment 2 only the first starting location was employed. When the subject was at the correct start position, the experimenter placed one of the stimulus objects at the target location behind the opaque shutter screen. Each object could be presented at one of two orientations with respect to the participant. The experimenter could very precisely position each object at the correct location and orientation by aligning two small groves under each object with two small pins on the table surface.

Once both stimulus and participant were positioned correctly, a tone indicated the beginning of a trial, at which point the shutter window turned translucent. Participants were then required to pick up the object using only forefinger and thumb and place it at the goal location. Participants had 3 seconds to complete the task before the shutter window turned opaque. In Experiment 1, no instructions were given regarding how the objects had to be picked up. In Experiment 2, participants were instructed to keep the objects as level as possible.

Experiment 1 had sixteen conditions: two starting locations, four wooden objects of different shapes, each object presented at two orientations. Each participant repeated each condition five times (eighty trials per participant).

Experiment 2 had thirty-six conditions: twelve distinct objects (four shapes in three material configurations) presented at two orientations. Half of the participants handled only shapes L and

V, the other half handled shapes U and S. Each participant repeated each condition seven times (eighty-four trials per participant). In both experiments trial order was randomized. Following each trial, the experimenter visually inspected the movement traces to determine whether a grasp was successful or not. Grasps were deemed unsuccessful when the movement was too slow, when an object was dropped, or when tracking was lost. Unsuccessful grasps were marked as error trials, added to the randomization queue, and repeated. A total of 368 error trials (13.8% of trials from Experiment 1 and 13.9% from Experiment 2) were not analyzed.

***Training***

Each participant completed six practice trials (using a Styrofoam cylinder in Experiment 1, and by lifting random objects from the shapes not used in that participant's run in Experiment 2) prior to the experiment to give them a sense for how fast their movement should be in order to complete the entire movement within three seconds. Practice trial data were not used in analyses. Prior to Experiment 2, participants were familiarized with the relative weight of brass and wood using two rectangular cuboids of dimensions 12.5x2.5x2.5 cm, one of wood (50 g) and one of brass (670 g).

***Analyses***

All analyses were performed in Matlab version R2018a. Differences between group means were assessed via paired or unpaired t-tests, as appropriate. Values of $p<0.05$ were considered statistically significant.

**Contact points.** Contact points of both fingers with the object were determined as the fingertip coordinates at the time of first contact, projected onto the surface of the triangulated mesh models of the object. The time of contact with the object was determined using the methods developed by Schot et al (60) and previously described in Paulun et al (19).

**Grasp similarity.** We described each individual grasp $\vec{G}$ as a 6D vector of the x,y,z coordinates of the thumb and index finger contact points:

$$\vec{G} = [x_T, y_T, z_T, x_I, y_I, z_I]$$

To compute the similarity $S$ between two grasps $\vec{G_1}$ and $\vec{G_2}$, we first computed the Euclidian distance between the two 6D grasp vectors. We then divided this distance by the largest possible distance between two points on the specific object $D_{max}$, determined from the mesh models of the objects. Finally, similarity was defined as 1 minus the normalized grasp distance, times 100:

$$S = 100 * \left(1 - \frac{\| \vec{G_1} - \vec{G_2}, \|}{D_{max}}\right)$$

In this formulation, two identical grasps, which occupy the same point in a 6D space, will be 100% similar, whereas the two farthest possible grasps onto a specific object will be 0% similar. Within-subject grasp similarity was the average similarity between grasps from the same participant to the participant's own median grasp. Between-subject grasp similarity was the similarity between the median grasp of each participant and the median grasp across all other participants.

***Computational model***

The model takes as input 3D meshes of the stimuli and outputs a cost function describing the costs associated with every possible combination of finger and thumb position on the accessible surface locations of our objects (i.e., those not in contact with the table plane). First, we define the center of each triangle in the mesh as a potential contact point. Then, given all possible combinations of thumb and index finger contact points $\vec{CP_T} = [x_T, y_T, z_T]$; $\vec{CP_I} = [x_I, y_I, z_I]$, the surface normal at both contact points $\vec{n_T} = [x_T^n, y_T^n, z_T^n]$; $\vec{n_I} = [x_I^n, y_I^n, z_I^n]$, and the CoM of the object $\vec{CoM} = [x_{CoM}, y_{CoM}, z_{CoM}]$, the five penalty functions we combined into a computational model of grasp selection were defined as follows:

**Force closure.** For two-digit grasping, a grasp fulfills force closure when the grasp axis connecting thumb and index contact points lies within the friction cones resulting from the

friction coefficient between object and digits(14). A grasp perfectly fulfills force closure when the grasp axis is perfectly aligned with the vectors along which gripping forces are applied, which are the opposite of the contact-point surface normals. Therefore, we defined the force closure penalty function as the sum of the angular deviances (computed using the atan2 function) of the grasp axis from both force vectors $\overrightarrow{\boldsymbol{F_T}} = -\overrightarrow{\boldsymbol{n_T}}$; $\overrightarrow{\boldsymbol{F_I}} = -\overrightarrow{\boldsymbol{n_I}}$:

$$\boldsymbol{P_{FC}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) = \boldsymbol{atan2}\left(\left\|\overrightarrow{\boldsymbol{F_T}} \times \left(\overrightarrow{\boldsymbol{CP_I}} - \overrightarrow{\boldsymbol{CP_T}}\right)\right\|, \overrightarrow{\boldsymbol{F_T}} \cdot \left(\overrightarrow{\boldsymbol{CP_I}} - \overrightarrow{\boldsymbol{CP_T}}\right)\right)$$

$$+ \boldsymbol{atan2}\left(\left\|\overrightarrow{\boldsymbol{F_I}} \times \left(\overrightarrow{\boldsymbol{CP_T}} - \overrightarrow{\boldsymbol{CP_I}}\right)\right\|, \overrightarrow{\boldsymbol{F_I}} \cdot \left(\overrightarrow{\boldsymbol{CP_T}} - \overrightarrow{\boldsymbol{CP_I}}\right)\right)$$

**Torque.** If a force is applied at some position away from the CoM, the object will tend to rotate due to torque, given by the cross product of force vector and lever arm (the vector connecting CoM to the point of force application). Under the assumption that is possible to apply forces at the thumb and index contact points that counteract the force of gravity $\overrightarrow{\boldsymbol{F_g}}$, we can compute the total torque of a grip as the sum of torques exerted by each contact point. Therefore, we defined the torque penalty function as the magnitude of the total torque exerted by a grip:

$$\boldsymbol{P_T}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) = \left\|\left(\overrightarrow{\boldsymbol{CoM}} - \overrightarrow{\boldsymbol{CP_T}}\right) \times \overrightarrow{-\boldsymbol{F_g}} + \left(\overrightarrow{\boldsymbol{CoM}} - \overrightarrow{\boldsymbol{CP_I}}\right) \times \overrightarrow{-\boldsymbol{F_g}}\right\|$$

**Natural grasp axis.** Schot, Brenner and Smeets(21) have carefully mapped out how human participants grasp spheres placed at different positions throughout the peripersonal space, and provide a regression model that determines the naturally preferred posture of the arm when grasping a sphere. We input the configuration of our current experimental setup into the regression model developed by these authors, and found the natural grasp axis for our participants to be $\overrightarrow{\boldsymbol{NGA}} = [\boldsymbol{0.49\ 0.87\ 0}]$. We therefore defined the natural grasp axis penalty function as the angular deviance from this established natural grasp axis:

$$\boldsymbol{P_{NGA}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) = \boldsymbol{atan2}\left(\left\|\overrightarrow{\boldsymbol{NGA}} \times \left(\overrightarrow{\boldsymbol{CP_I}} - \overrightarrow{\boldsymbol{CP_T}}\right)\right\|, \overrightarrow{\boldsymbol{NGA}} \cdot \left(\overrightarrow{\boldsymbol{CP_I}} - \overrightarrow{\boldsymbol{CP_T}}\right)\right)$$

**Optimal grasp aperture for precision grip.** Cesari and Newell(23) have shown that, when free to employ any multi-digit grasp, human participants selected precision grip grasps only for cubes smaller than 2.5 cm in length. As cube size increases, humans progressively increase the

number of digits employed in a grasps. Therefore, since our participants were instructed only to employ precision grip grasps, we defined the optimal grasp aperture penalty function as 0 for grasp sizes smaller than 2.5 cm, and as a linearly increasing penalty for grasp sizes larger than 2.5 cm:

$$\boldsymbol{P_{OGA}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) = \begin{cases} 0, & if\ \left\|\overrightarrow{\boldsymbol{CP_I}} - \overrightarrow{\boldsymbol{CP_T}}\right\| < 25mm \\ \left\|\overrightarrow{\boldsymbol{CP_I}} - \overrightarrow{\boldsymbol{CP_T}}\right\| - 25, & if\ \left\|\overrightarrow{\boldsymbol{CP_I}} - \overrightarrow{\boldsymbol{CP_T}}\right\| > 25\ mm \end{cases}$$

**Object Visibility.** Under the assumption that humans are attempting to minimize the portion of the objects hidden from view by their hand, we defined the optimal visibility penalty function as the proportion of object still visible during each possible grasp. We first defined the line on the XZ plane that passes through the thumb and index finger contact points. We made the simplifying assumption that, given all possible surface points on the object $\boldsymbol{SP_{TOT}}$, the surface points $\boldsymbol{SP_{OCC}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right)$ that fall to the side of the line where the hand is located will be occluded. Therefore, the object visibility penalty function was defined as:

$$\boldsymbol{P_{OGA}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) = \frac{Length\left(\boldsymbol{SP_{OCC}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right)\right)}{Length(\boldsymbol{SP_{TOT}})}$$

**Overall grasp penalty function.** To obtain the overall grasp penalty function, each grasp penalty function was first normalized to the [0 1] range (i.e., across all possible grasps for each given object, independently of the other objects). Then, we took a weighted linear sum of the individual penalty functions, with all weights equal to 1:

$$\boldsymbol{P_O}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) = \boldsymbol{P_{FC}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) + \boldsymbol{P_T}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) + \boldsymbol{P_{NGA}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) +$$

$$\boldsymbol{P_{OGA}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) + \boldsymbol{P_{RT}}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right)$$

For display purposes this final function was normalized to the [0 1] range. The minima of this overall grasp penalty function represent the set of grasps that best satisfy the largest number of constraints at the same time.

**Model fitting.** In both Experiments 1 and 2, human participants executed repeated grasps to the same objects at each orientation. To fit the overall grasp penalty function to these human data, for each participant in each condition we first defined a human grasp penalty function $\boldsymbol{P_H}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}})$ in which all grasps selected by a participant onto an object were set to have 0 penalty, and all grasps that had not been selected were set to have a penalty of 1. Then, we fit the function:

$$\boldsymbol{P_{O,fit}}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}) = \sum_i C_i * \boldsymbol{P_i}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}})$$

to the human grasp penalty function. More specifically, we employed a nonlinear least-squares solver to search for the set of coefficients $C_i = [C_{FC}; C_T; C_{NGA}; C_{OGA}; C_{RT}]$ that minimized the function:

$$\boldsymbol{F(C_i)} = \sqrt{\boldsymbol{W}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}})} * \left[ \left( \sum_i C_i * \boldsymbol{P_i}\left(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}\right) \right) - \boldsymbol{P_H}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}) \right]$$

i.e. we searched for the set of coefficients for which $\boldsymbol{P_{O,fit}}$ best approximated the human grasp penalty function $\boldsymbol{P_H}$. The solver employed the trust-region-reflective algorithm; we set the lower and upper bounds of the coefficients to be 0 and 1, and 0.2 as the starting value for all coefficients. Critically, $\boldsymbol{W}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}})$ was a weight function which served to give equal weight to high and low penalty grasps in the human grasp penalty function, since the number of non-selected grasps with $\boldsymbol{P_H}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}) = 1$ vastly outnumbered the few selected grasps for which $\boldsymbol{P_H}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}) = 0$. Thus, for grasps where $\boldsymbol{P_H}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}) = 0$, $\boldsymbol{W}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}})$ was equal to the number of times the participant had selected that specific grasp. For grasps where $\boldsymbol{P_H}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}) = 1$ instead, $\boldsymbol{W}(\overrightarrow{\boldsymbol{CP_T}}, \overrightarrow{\boldsymbol{CP_I}}) = \frac{N_{G,selected}}{N_{G,non-selected}}$; where $N_{G,selected}$ was the total number of grasps performed by the participant onto the object, and $N_{G,non-selected}$ was the total number of non-selected grasps within the grasp manifold.

**Predicting Grasps.** The minima of both the equally weighted (non-fitted) and the fitted overall grasp penalty functions represent the set of grasps predicted to be optimal under the weighted linear combination of the five penalty functions included in our computational model. To visualize these predicted optimal grasps, we sampled them from the minima of the penalty functions. First, we removed all grasps with penalty values greater than the lower 0.1th percentile. The remaining grasps were therefore all optimal or near-optimal. From this subset, we then randomly selected (with replacement) a number of grasps equal to the number of grasps executed by the human participants. The probability with which any one grasp was selected was set to be 1 minus the grasp penalty, thus grasps with zero penalty had the highest probability of being selected. These sampled grasps can then be projected back onto the objects for visualization purposes (Figure 5a), or they can be directly compared to human grasp using the grasp similarity metric described above (Figure 5b,c).

**Data availability.** Data and analysis scripts will be made available from the Zenodo database (doi:xx.xxxx/zenodo.xxxxxxx upon publication).

**Acknowledgments.** The authors thank Dr. Karl Gegenfurtner for insightful feedback. This research was supported by the DFG (IRTG-1901: 'The Brain in Action' and SFB-TRR-135: 'Cardinal Mechanisms of Perception'), and an ERC Consolidator Award (ERC-2015-CoG-682859: 'SHAPE'). Guido Maiello was supported by a Marie-Skłodowska-Curie Actions Individual Fellowship (H2020-MSCA-IF-2017: 'VisualGrasping' Project ID: 793660).

## References

1. Levine S, Pastor P, Krizhevsky A, Ibarz J, Quillen D (2018) Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International Journal of Robotics Research* 37(4–5):421–436.

2. Smeets JBJ, Brenner E (1999) A New View on Grasping. *Motor Control* 3(3):237–271.

637 3. Rosenbaum DA, Meulenbroek RJG, Vaughan J, Elsinger C (1999) Approaching Grasping from
638 Different Perspectives. *Motor Control* 3(3):289–297.

639 4. Rosenbaum DA, Meulenbroek RGJ, Vaughan J, Jansen C (1999) Coordination of reaching and
640 grasping by capitalizing on obstacle avoidance and other constraints. *Experimental Brain Research*
641 128(1–2):92–100.

642 5. Smeets J, Brenner E (2001) Independent movements of the digits in grasping. *Experimental Brain*
643 *Research* 139(1):92–100.

644 6. Christopoulos VN, Schrater PR (2009) Grasping Objects with Environmentally Induced Position
645 Uncertainty. *PLoS Computational Biology* 5(10):e1000538.

646 7. Karok S, Newport R (2010) The continuous updating of grasp in response to dynamic changes in
647 object size, hand size and distractor proximity. *Neuropsychologia* 48(13):3891–3900.

648 8. Eloka O, Franz VH (2011) Effects of object shape on the visual guidance of action. *Vision Research*
649 51(8):925–931.

650 9. Volcic R, Domini F (2014) The visibility of contact points influences grasping movements.
651 *Experimental Brain Research* 232(9):2997–3005.

652 10. Volcic R, Domini F (2016) On-line visual control of grasping movements. *Experimental Brain*
653 *Research* 234(8):2165–2177.

654 11. Bozzacchi C, Volcic R, Domini F (2016) Grasping in absence of feedback: systematic biases endure
655 extensive training. *Experimental Brain Research* 234(1):255–265.

656 12. Kleinholdermann U, Franz VH, Gegenfurtner KR (2013) Human grasp point selection. *Journal of*
657 *Vision* 13(8):23–23.

658 13. Gilster R, Hesse C, Deubel H (2012) Contact points during multidigit grasping of geometric objects.
659 *Experimental Brain Research* 217(1):137–151.

660 14. Nguyen V-D (1988) Constructing Force- Closure Grasps. *The International Journal of Robotics*
661 *Research* 7(3):3–16.

662 15. Goodale MA, et al. (1994) Separate neural pathways for the visual analysis of object shape in
663 perception and prehension. *Current Biology* 4(7):604–610.

664 16. Lederman SJ, Wing AM (2003) Perceptual judgement, grasp point selection and object symmetry.
665 *Experimental Brain Research* 152(2):156–165.

666 17. Eastough D, Edwards MG (2006) Movement kinematics in prehension are affected by grasping
667 objects of different mass. *Experimental Brain Research* 176(1):193–198.

668 18. Lukos J, Ansuini C, Santello M (2007) Choice of Contact Points during Multidigit Grasping: Effect of
669 Predictability of Object Center of Mass Location. *Journal of Neuroscience* 27(14):3894–3903.

19. Paulun VC, Gegenfurtner KR, Goodale MA, Fleming RW (2016) Effects of material properties and object orientation on precision grip kinematics. *Experimental Brain Research* 234(8):2253–2265.

20. Roby-Brami A, Bennis N, Mokhtari M, Baraduc P (2000) Hand orientation for grasping depends on the direction of the reaching movement. *Brain Research* 869(1–2):121–129.

21. Schot WD, Brenner E, Smeets JBJ (2010) Posture of the arm when grasping spheres to place them elsewhere. *Experimental Brain Research* 204(2):163–171.

22. Voudouris D, Brenner E, Schot WD, Smeets JBJ (2010) Does planning a different trajectory influence the choice of grasping points? *Experimental Brain Research* 206(1):15–24.

23. Cesari P, Newell KM (1999) The scaling of human grip configurations. *Journal of Experimental Psychology: Human Perception and Performance* 25(4):927–935.

24. Huang HJ, Kram R, Ahmed AA (2012) Reduction of Metabolic Cost during Motor Learning of Arm Reaching Dynamics. *Journal of Neuroscience* 32(6):2182–2190.

25. Paulun VC, Kleinholdermann U, Gegenfurtner KR, Smeets JBJ, Brenner E (2014) Center or side: biases in selecting grasp points on small bars. *Experimental Brain Research* 232(7):2061–2072.

26. Bozzacchi C, Brenner E, Smeets JB, Volcic R, Domini F (2018) How removing visual information affects grasping movements. *Experimental Brain Research* 236(4):985–995.

27. Maiello G, Paulun VC, Klein LK, Fleming RW (2018) Object visibility, not energy expenditure, accounts for spatial biases in human grasp selection. *bioRxiv*. doi:10.1101/476101.

28. Volcic R, Kappers AML (2008) Allocentric and egocentric reference frames in the processing of three-dimensional haptic space. *Experimental Brain Research* 188(2):199–213.

29. Desanghere L, Marotta JJ (2015) The influence of object shape and center of mass on grasp and gaze. *Frontiers in Psychology* 6. doi:10.3389/fpsyg.2015.01537.

30. Glowania C, van Dam LCJ, Brenner E, Plaisier MA (2017) Smooth at one end and rough at the other: influence of object texture on grasping behaviour. *Experimental Brain Research* 235(9):2821–2827.

31. Cuijpers RH, Smeets JBJ, Brenner E (2004) On the Relation Between Object Shape and Grasping Kinematics. *Journal of Neurophysiology* 91(6):2598–2606.

32. Cuijpers RH, Brenner E, Smeets JBJ (2006) Grasping reveals visual misjudgements of shape. *Experimental Brain Research* 175(1):32–44.

33. Schettino LF, Adamovich SV, Poizner H (2003) Effects of object shape and visual feedback on hand configuration during grasping. *Experimental Brain Research* 151(2):158–166.

34. Chen Z, Saunders JA (2015) Online processing of shape information for control of grasping. *Experimental Brain Research* 233(11):3109–3124.

35. Mamassian P (1997) Prehension of objects oriented in three-dimensional space: *Experimental Brain Research* 114(2):235–245.

36. Rosenbaum DA, et al. (1990) Constraints for action selection: Overhand versus underhand grips. *Attention and Performance 13: Motor Representation and Control.* (Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US), pp 321–342.

37. Short MW, Cauraugh JH (1999) Precision hypothesis and the end-state comfort effect. *Acta Psychologica* 100(3):243–252.

38. Lee Hughes CM, Seegelke C, Schack T (2012) The Influence of Initial and Final Precision on Motor Planning: Individual Differences in End-State Comfort During Unimanual Grasping and Placing. *Journal of Motor Behavior* 44(3):195–201.

39. Castiello U (2005) The neuroscience of grasping. *Nature Reviews Neuroscience* 6(9):726–736.

40. Castiello U, Begliomini C (2008) The Cortical Control of Visually Guided Grasping. *The Neuroscientist* 14(2):157–170.

41. Janssen P, Scherberger H (2015) Visual Guidance in Control of Grasping. *Annual Review of Neuroscience* 38(1):69–86.

42. Theys T, Pani P, van Loon J, Goffin J, Janssen P (2012) Selectivity for Three-Dimensional Shape and Grasping-Related Activity in the Macaque Ventral Premotor Cortex. *Journal of Neuroscience* 32(35):12038–12050.

43. Murata A, et al. (1997) Object Representation in the Ventral Premotor Cortex (Area F5) of the Monkey. *Journal of Neurophysiology* 78(4):2226–2230.

44. Raos V, Umiltá M-A, Murata A, Fogassi L, Gallese V (2006) Functional Properties of Grasping-Related Neurons in the Ventral Premotor Area F5 of the Macaque Monkey. *Journal of Neurophysiology* 95(2):709–729.

45. Raos V, Umiltá M-A, Gallese V, Fogassi L (2004) Functional Properties of Grasping-Related Neurons in the Dorsal Premotor Area F2 of the Macaque Monkey. *Journal of Neurophysiology* 92(4):1990–2002.

46. Murata A, Gallese V, Luppino G, Kaseda M, Sakata H (2000) Selectivity for the Shape, Size, and Orientation of Objects for Grasping in Neurons of Monkey Parietal Area AIP. *Journal of Neurophysiology* 83(5):2580–2601.

47. Cant JS, Goodale MA (2011) Scratching Beneath the Surface: New Insights into the Functional Properties of the Lateral Occipital Area and Parahippocampal Place Area. *Journal of Neuroscience* 31(22):8248–8258.

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

736 49. Gallivan JP, Cant JS, Goodale MA, Flanagan JR (2014) Representation of Object Weight in Human
737 Ventral Visual Cortex. *Current Biology* 24(16):1866–1873.

738 50. Goda N, Tachibana A, Okazawa G, Komatsu H (2014) Representation of the Material Properties of
739 Objects in the Visual Cortex of Nonhuman Primates. *Journal of Neuroscience* 34(7):2660–2673.

740 51. Goda N, Yokoi I, Tachibana A, Minamimoto T, Komatsu H (2016) Crossmodal Association of Visual
741 and Haptic Material Properties of Objects in the Monkey Ventral Visual Cortex. *Current Biology*
742 26(7):928–934.

743 52. Borra E, et al. (2008) Cortical Connections of the Macaque Anterior Intraparietal (AIP) Area.
744 *Cerebral Cortex* 18(5):1094–1111.

745 53. Sakata H, Taira M, Murata A, Mine S (1995) Neural Mechanisms of Visual Guidance of Hand Action
746 in the Parietal Cortex of the Monkey. *Cerebral Cortex* 5(5):429–438.

747 54. Jeannerod M, Arbib MA, Rizzolatti G, Sakata H (1995) Grasping objects: the cortical mechanisms of
748 visuomotor transformation. *Trends in Neurosciences* 18(7):314–320.

749 55. Srivastava S, Orban GA, De Maziere PA, Janssen P (2009) A Distinct Representation of Three-
750 Dimensional Shape in Macaque Anterior Intraparietal Area: Fast, Metric, and Coarse. *Journal of*
751 *Neuroscience* 29(34):10613–10626.

752 56. Davare M, Rothwell JC, Lemon RN (2010) Causal Connectivity between the Human Anterior
753 Intraparietal Area and Premotor Cortex during Grasp. *Current Biology* 20(2):176–181.

754 57. Theys T, Romero MC, van Loon J, Janssen P (2015) Shape representations in the primate dorsal
755 visual stream. *Frontiers in Computational Neuroscience* 9. doi:10.3389/fncom.2015.00043.

756 58. Franz VH Optotrak Toolbox. *The Optotrak Toolbox: Control your Optotrak from within Matlab*.
757 Available at: http://www.ecogsci.cs.uni-tuebingen.de/OptotrakToolbox/.

758 59. Milgram P (1987) A spectacle-mounted liquid-crystal tachistoscope. *Behavior Research Methods,*
759 *Instruments, & Computers* 19(5):449–456.

760 60. Schot WD, Brenner E, Smeets JBJ (2010) Robust movement segmentation by combining multiple
761 sources of information. *Journal of Neuroscience Methods* 187(2):147–155.

762

763

764 Supplementary Information to

765 # *How humans grasp three-dimensional objects*

766 Lina K. Klein [a,†] , Guido Maiello [a,†,*], Vivian C. Paulun [a], Roland W. Fleming [a]

767 [a] Department of Experimental Psychology, Justus-Liebig University Giessen, Giessen 35394,
768 Germany

769 * Corresponding Author:

770 Guido Maiello

771 Department of Experimental Psychology, Justus-Liebig University Giessen, Otto-
772 Behaghel-Str.10F, Giessen 35394, Germany

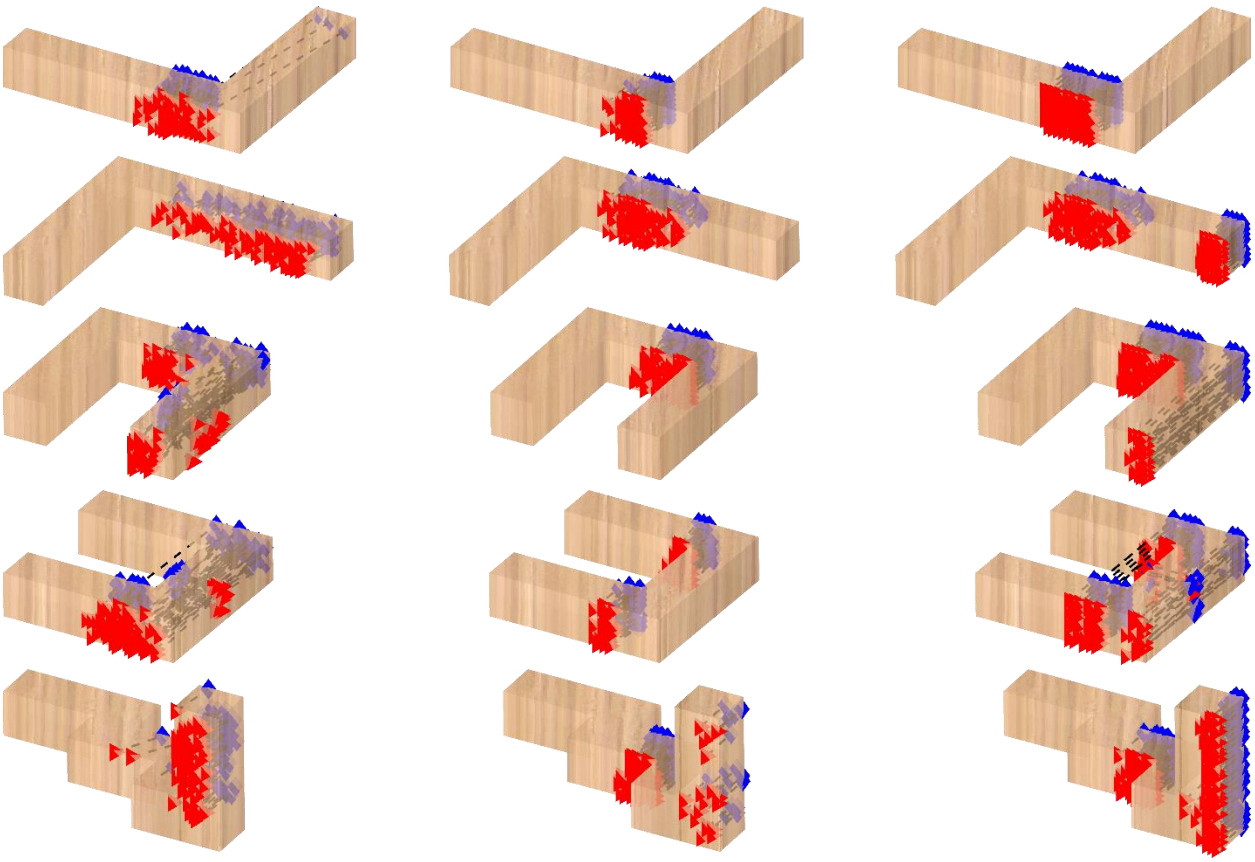
773 Email: guido_maiello@yahoo.it

774 [†] joint first authors; these authors contributed equally to this work

775

776 (a) Experiment 1

777 Human Unfitted Model Fitted Model



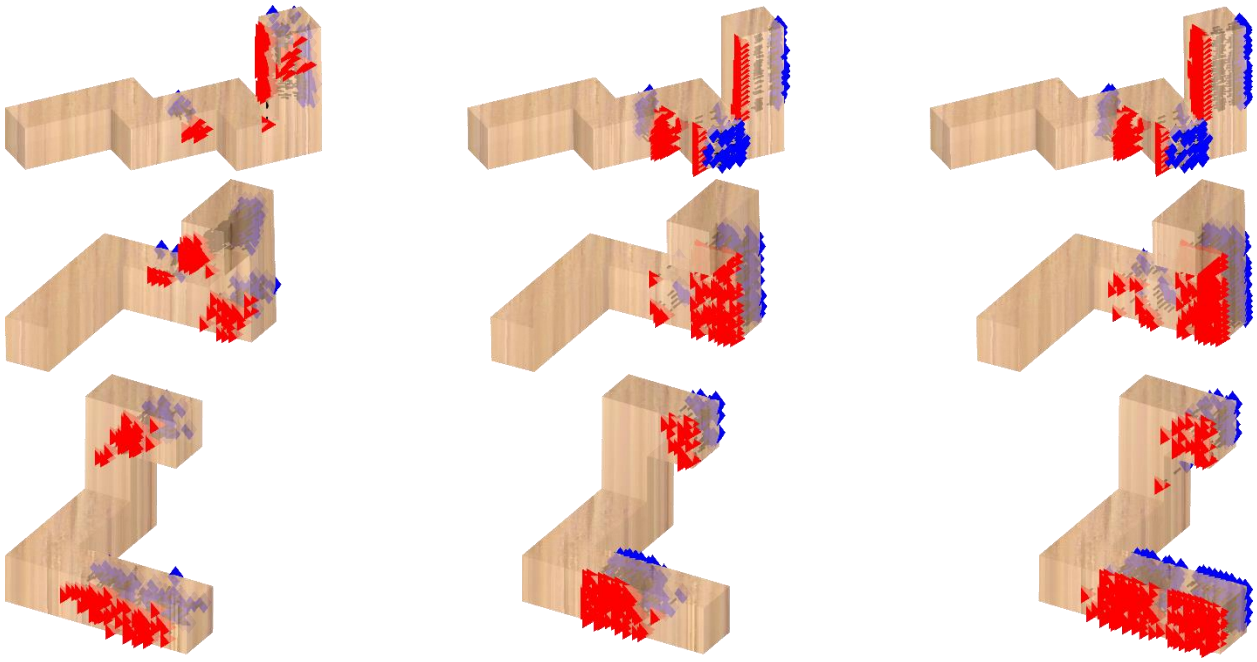
778

779

780

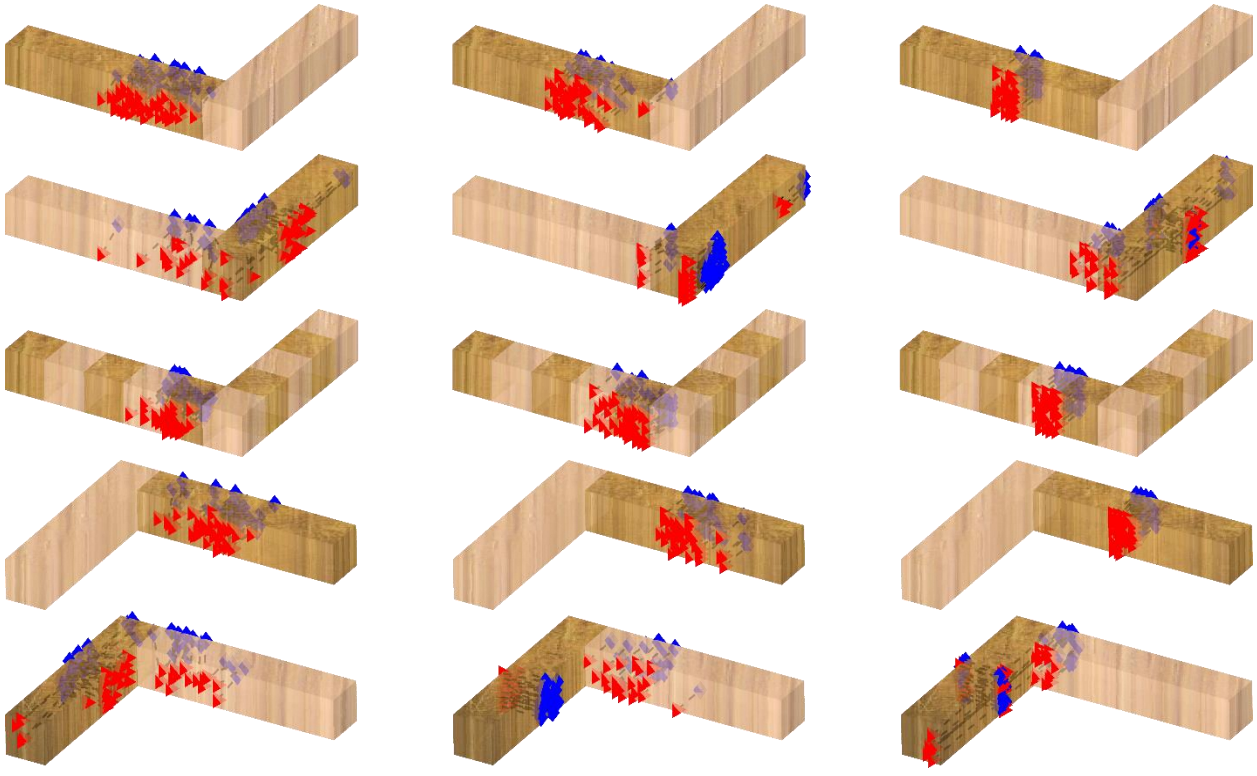
781

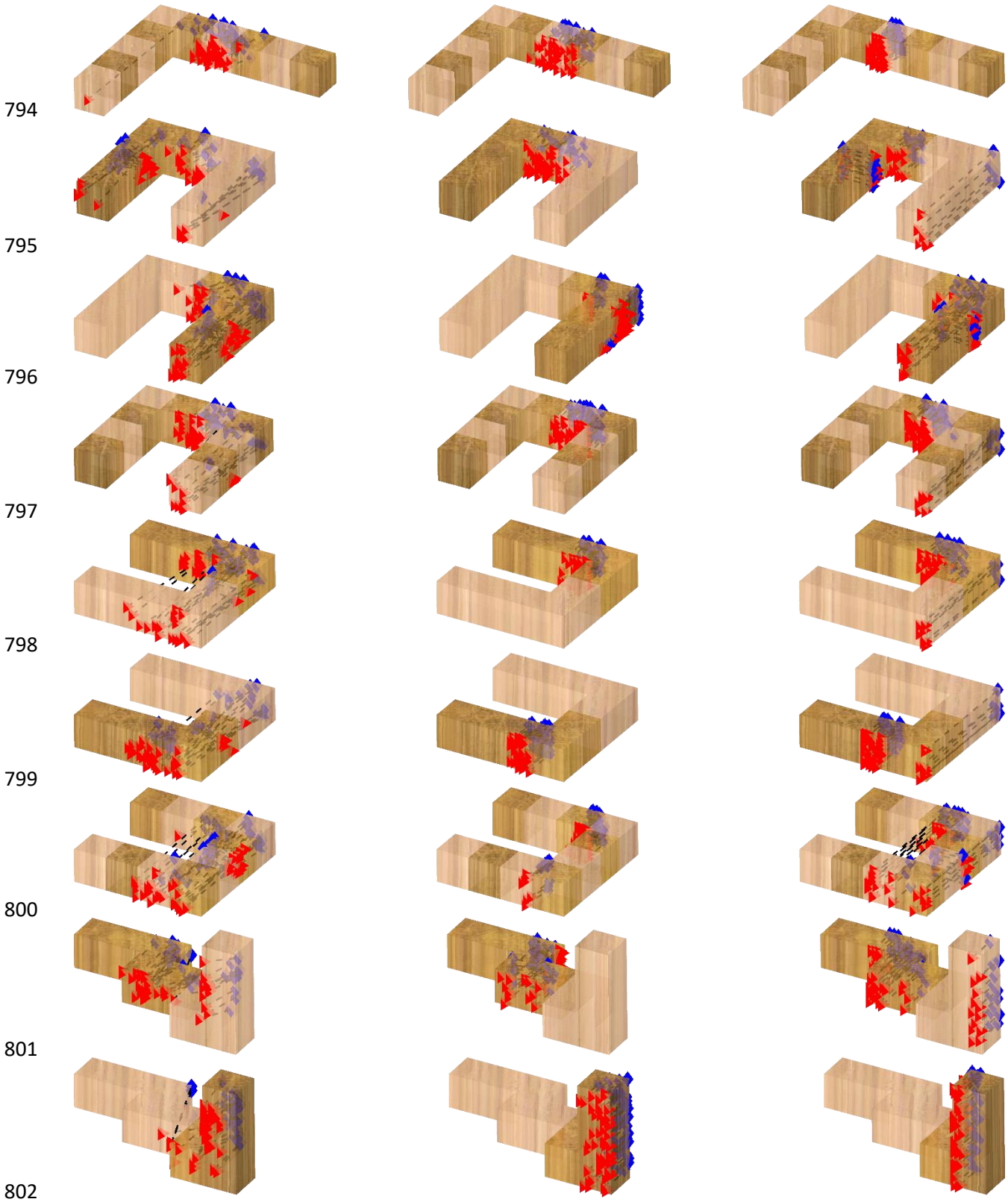
782

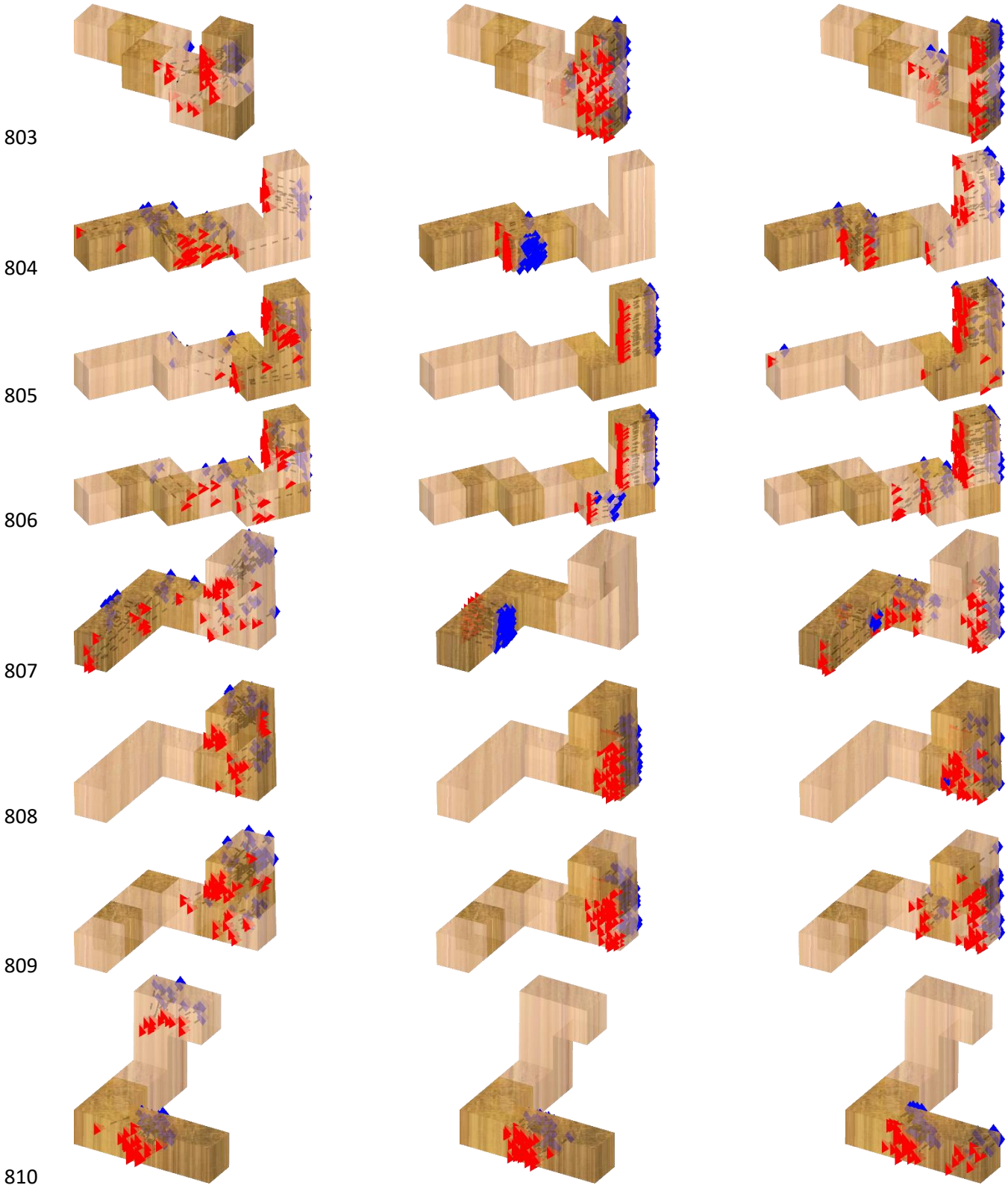


(b) Experiment 2

Human Unfitted Model Fitted Model







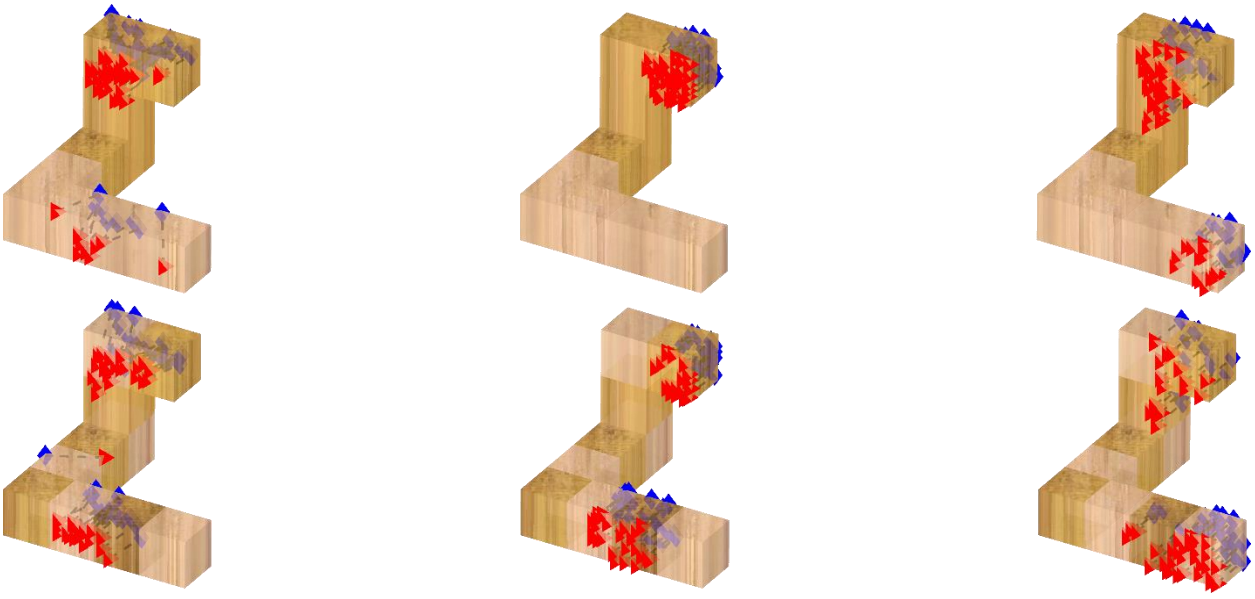


***Figure S1***. *Grasping patterns from human participants (left), unfitted model (middle), and fitted model (right). (a) Grasping patterns on wooden objects from Experiment 1. (b) Grasping patterns on mixed material objects from Experiment 2.*

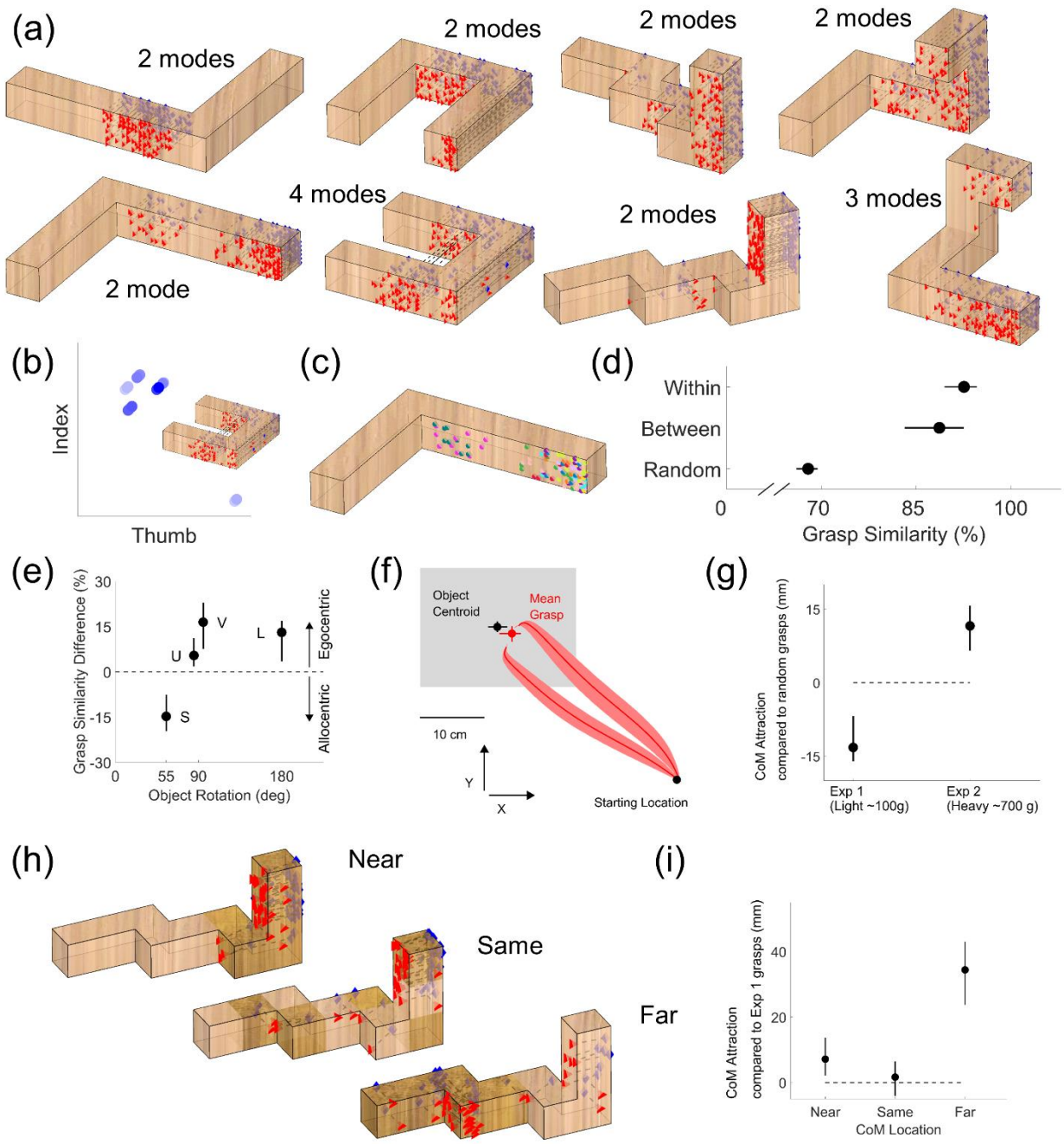



***Figure S2**. Pattern of empirical results recreated from simulating grasps from the fitted computational model. All panels are the same as in Figure 3 of the main manuscript, except that the data are simulated from the model. Only the grasp trajectories in panel (f) are from the human data, to highlight how the model correctly reproduces the biases in human grasping patterns.*