the left), alpha is −8 deg, and the grey circle shows the rigid rotation. We found that the perceived curvature increases monotonically as alpha goes from −45 to 45 (Bertamini & Smit 1998). Note that changes in alpha do not change the center of rotation, and that this effect is therefore inconsistent with the idea of internalised kinematic geometry. It suggests instead that motion tends to be orthogonal to the orientation of the object in the first frame.

Some observations on symmetric shapes are also important. In the case of a solid rectangle (Fig. 1b) there are not only the two paths that are always possible on the basis of Chasles' theorem (90 and 270 deg, respectively). There are two more paths identical to the first two except that alpha differs by 90 deg. Ignoring the longer paths, we still have a conflict between two possible solutions with the same angle of rotation. Both solutions can be seen, but the motion orthogonal to the orientation of the object (alpha=0) is seen more often by naive subjects. In this solution the object extends farther from the center of rotation, therefore we may be observing a difference in torque. If so, this would be an effect related to the physics of the event, not its kinematic geometry.

More axes of symmetry can be present in an object, such as in the case of an equilateral triangle (Fig. 1c). Kolers and Pomerantz (1971) have noticed that both rotation in the plane and rotation in depth can be seen in such cases (the depth solution being more likely when there is longer presentation time). What is important here is that a depth rotation of 180 deg is seen at least as often as a rotation in the plane of 60 deg. Surely this is a problem for an argument based on the simplicity of motion. We went even farther and tried quasi-symmetrical stimuli (Fig. 1d). Remarkably, motion in depth is seen even when the 60 deg rotation is a rigid motion, whereas a rotation in depth of 180 deg entails a shape change (one arm getting longer as the object moves).

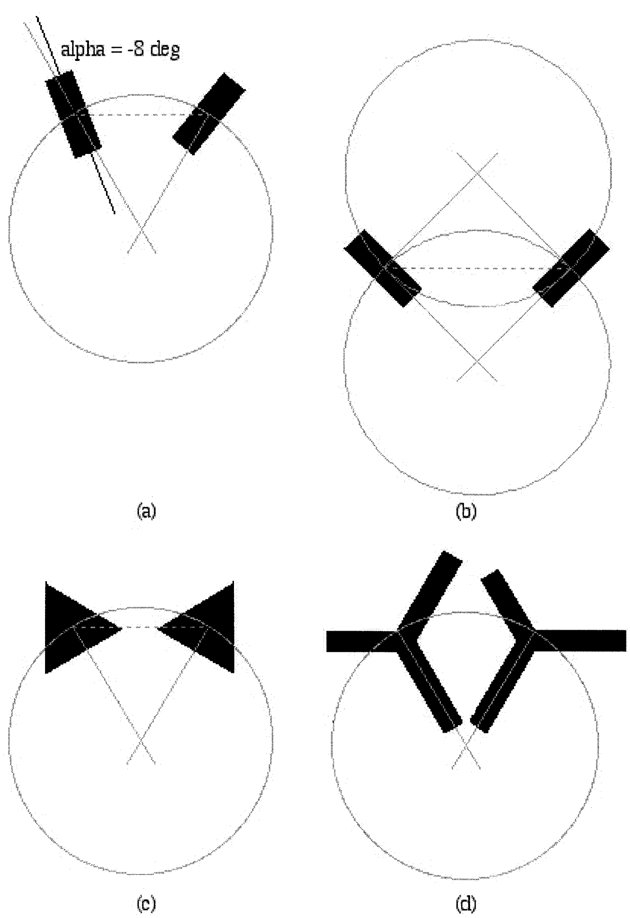

Figure 1 (Bertamini). To see the animations go to: http://www.liv.ac.uk/~marcob/todorovic.html

SHEPARD discusses the case where motion in depth by 180 deg is preferred to rotation by 180 deg and suggests that the reason depth rotation is preferred is because it is more consistent with retinal stimulation (i.e., if motion in the plane had taken place it would have been detected). I doubt that this could account for the case where 180 deg is compared to 60 deg, but it is a useful way of looking at the problem (i.e., what is the most likely motion given the available evidence), especially if we agree that apparent motion is a solution to poor temporal sampling (Watson & Ahumada 1983).

I hasten to acknowledge the noisiness of these data. Everybody looking at these displays will notice their inherent ambiguity. Percepts can and do change even for one individual over time. This multistability needs to be taken into account in any theory. I suggest that this multistability could be used constructively to study the way shape is represented. Taking the example of the equilateral triangle, the three axes of bilateral symmetry are identical from a geometrical point of view, but perceptually they are not. At any one time, one vertex is seen as the top and the opposite side as the base of the object. Such a chosen axis of orientation is important in determining the motion of the object (for other effects of shape on apparent motion, see McBeath et al. 1992). When the equilateral triangle is seen as oriented horizontally, the motion in depth (but not the rotation in the plane) is around a pivotal point along the axis of elongation. The importance of axes in constraining perceived motion is consistent with what SHEPARD is arguing, except that this does not mean that kinematic geometry has been internalised, it means that the representation of shape is not independent from the representation of motion.

We have recently found effects of pivot points in how motion is perceived, using random dot configurations (Bertamini & Proffitt 2000). These are all examples where the system assumes (or infers from spatial information) mechanical constraints on motion (Hoffman & Bennett 1986). Given the environment in which we live, this may be the best strategy. The only case in which mechanical constraints to motion do not exist are particle motions and they are not as common as extended object motion and joint motion. SHEPARD claims that physics would predict straight paths for the center of mass, but although this is true it misses out on the fact that given a certain shape not all motions are equally likely. Probably no tree that has ever fallen in a forest moves along a straight path, instead it rotates around a center at the base of its elongation.

Finally, as an aside, let me point out that the preference for motion orthogonal to object orientation is not a general effect. Werkhoven et al. (1990) have found quite the opposite result in short term apparent motion. This difference may be related to how the aperture problem affects the system at different scales. But this is another story.

"First, we assume a spherical cow . . . "

Lera Boroditsky[a] and Michael Ramscar[b]

[a]*Psychology Department, Stanford University, Stanford, CA, 94305–2130, USA;* [b]*School of Cognitive Science, University of Edinburgh, Edinburgh, EH8 9LW, Scotland.* **lera@psych.stanford.edu**
http://www-psych.stanford.edu/~lera/michael@dai.ed.ac.uk
http://www.dai.ed.ac.uk/homes/michael/

Abstract: There is an old joke about a theoretical physicist who was charged with figuring out how to increase the milk production of cows. Although many farmers, biologists, and psychologists had tried and failed to solve the problem before him, the physicist had no trouble coming up with a solution on the spot. "First," he began, "we assume a spherical cow . . . " [TENENBAUM & GRIFFITHS]

TENENBAUM & GRIFFITHS (henceforth T&G) present an ambitious attempt at a computational framework encompassing generalization, similarity, and categorization. Although it would seem elegant to account for all of similarity and/or categorization in a simple unitary framework, the phenomena in question are almost certainly far too complex and heterogeneous to allow this. A framework this general will inevitably fail to capture much of the intricacy and sophistication of human conceptual processing. That is, it may turn out to be a theory about spherical cows rather than cow-shaped ones.

T&G propose a model of similarity as generalization based on Bayesian inference. However, although T&G specify a framework (essentially, Bayes' rule and some ancillary equations), they fail to specify a procedure for generating, weighting, or constraining any of the input into this framework. At times, T&G base the representations in their hypothesis space on people's similarity judgments. It is hardly surprising that a model with people's similarity judgments built in can compute similarity. Further, the basis for T&G's claim that similarity is based on Bayesian generalization becomes unclear – in their model, generalization appears to be based on similarity and not the other way around. At present the framework relies solely on hand-coded and hand-tailored representations, while the few predictions it does make (relying on asymmetrical comparison and the size principle) are not borne out by data. We review just a few of the complications as illustrations below.

People's similarity judgments are based on a myriad of contextual, perceptual, and conceptual factors. In carrying out a comparison, people need to choose a way to represent the things to be compared as well as a strategy for comparing them. This means that a comparison between the same two items in different circumstances will yield different results. For example, in a replication of **T&G**'s study shown in the left panel of Figure 6 (with right-left position counterbalanced), 62% of our subjects picked the object-match (a) as most similar to the top example. But, if subjects were first given the example shown in the right panel of Figure 6 and then the question in the left panel, then only 33% picked the object-match. Changing how likely it was for people to notice and represent the relational structure of the stimuli had a dramatic effect on the results of the comparison. In another example, subjects were asked to say which of AXX or QJN was most similar to AHM (a problem structurally similar to T&G's in Fig. 6), and 43% chose QJN when the letters were presented in Chicago font (which makes all the letters look boxy). When the same letters were presented in Times font (which emphasized the pointy ends of the A's), only 17% chose QJN. Thus, even a trivial change in the perceptual properties of the stimuli can have a dramatic effect on how people choose to represent and compare the arrays.

Nothing inherent in **T&G**'s framework predicts these kinds of results. Although T&G's framework might allow for perceptual similarity, effects of context, and other factors to be coded into the hypothesis space, it is disappointing that it is these back-door (i.e., coded-in and not necessarily principled) elements, and not anything about the framework itself, that carry all of the explanatory power. Moreover, at times the specifics of the framework can even prevent the back-door solutions from working, even when these solutions are probably the psychologically correct ones. Consider the following example: When subjects were asked which of 1-911-ANALOGY or 1-208-BKSDEMG was most similar to 1-615-QFRLOWY, 75% of the subjects chose 1-208-BKSDEMG (chi^2=5.00, p <. 05) even though 1-911-ANALOGY shares 4 extra features with the base example, and the "1 in position 3, L in position 8, O in position 9, and Y in position 11" hypothesis is more than 72,000 times more restrictive than the "all different letters" hypothesis. Despite an advantage of more than 72,000 to 1, the size principle proposed by **T&G** as a new universal had no effect. We doubt that any one of our subjects even considered the "1 in position 3, L in position 8, O in position 9, and Y in position 11" hypothesis. Clearly the distinctive properties in 1-911-ANALOGY are responsible for the subjects' choices.

Although **T&G**'s model can discover distinctive features utilizing the size principle, it is limited to discovering the distinctive features of the base of the comparison (in T&G's framework, similarity is based on the intrinsically asymmetrical function of generalization, which depends only on the distinctive features of the base and not of the target). But for the subjects, the outcome of this problem depends on the distinctive features of the target (the opposite of what T&G predict). It seems unlikely, given the flexibility and sophistication of human thought, that all comparison processes will be bound by the asymmetrical properties of Bayesian inference. Further, if the model is extended to be able to perform bi-directional comparisons, how will it decide which of the computations to choose as the measure of similarity? Unless some principled way is specified, the model will be able to predict anything (and as such will explain nothing). It would appear that the model's predictions (asymmetrical comparison and the size principle) are not borne out by data. Rather, the hand-coded hypothesis space (a kind of a clairvoyant homunculus that can mysteriously assemble itself to fit any given occasion) carries most of the explanatory power.

Finally, we should evaluate any model not only on whether or not it can be falsified, but also, importantly, on its usefulness. How much does it add to our understanding of cognition? **T&G**'s model is only viable if we can somehow anticipate (and hand-code in) all the adjustments to the hypothesis space that will be required in any given situation (i.e., build in complete world knowledge). As such, the framework is either computationally unimplementable (if we can't build everything in) or psychologically uninformative (if we can).

A theory that applies equally well to all possible situations may apply poorly in each. This is especially true if generality requires us to disregard much of our hard-won understanding of the details of psychological processing. There is a vast literature documenting the complexity and diversity of representations and processes involved in similarity and categorization. The sheer variety of these psychological phenomena weighs heavily against any simple unitary account. Any such account can at best aspire to be a theory of spherical cows – elegant, but of little use in a world filled with cows that stubbornly insist on being cow-shaped.

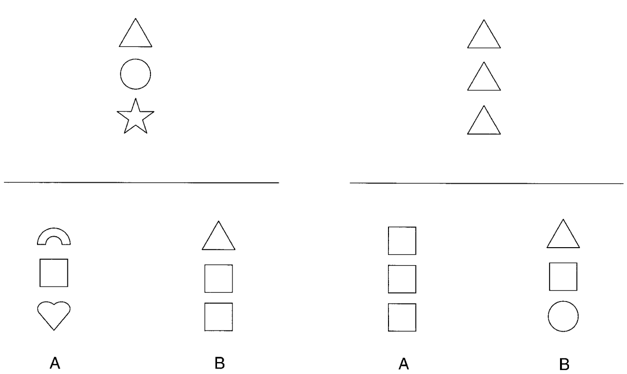

Figure 1 (Boroditsky & Ramscar). [This appears as Figure 6 in the article in this issue by Tenenbaum & Griffiths. The caption shown here is the caption written by T&G and appearing in T&G's article.] The relative weight of relations and primitive features depends on the size of the set of objects that they identify. Most observers choose B (the primitive feature match) as more similar to the top stimulus in the left panel, but choose A (the relational match) in the right panel, in part because the relation "all same shape" identifies a much smaller subset of objects than the relation "all different shapes."

ACKNOWLEDGMENTS
We would like to thank Alison Preston, Phillip Goff, Bradley Love, and Dedre Gentner for comments on earlier versions of this commentary.