

PUBLISHED VERSION

Lee, Michael D.

Extending Bayesian concept learning to deal with representational complexity and adaptation, *Behavioral and Brain Sciences*, 2001; 24 (4):685-686.

Copyright © 2001 Cambridge University Press

PERMISSIONS

<http://journals.cambridge.org/action/stream?pageId=4088&level=2#4408>

The right to post the definitive version of the contribution as published at Cambridge Journals Online (in PDF or HTML form) in the Institutional Repository of the institution in which they worked at the time the paper was first submitted, or (for appropriate journals) in PubMed Central or UK PubMed Central, no sooner than one year after first publication of the paper in the journal, subject to file availability and provided the posting includes a prominent statement of the full bibliographical details, a copyright notice in the name of the copyright holder (Cambridge University Press or the sponsoring Society, as appropriate), and a link to the online edition of the journal at Cambridge Journals Online. Inclusion of this definitive version after one year in Institutional Repositories outside of the institution in which the contributor worked at the time the paper was first submitted will be subject to the additional permission of Cambridge University Press (not to be unreasonably withheld).

10th December 2010

<http://hdl.handle.net/2440/32595>

locking on impact is observed for the modulation of muscle reflex responses and for the changes in overall hand impedance (the mechanical resistance to an imposed displacement, see Fig. 2). Remarkably, motor preparation of reflex responses and limb impedance is correctly timed on impact even when blindfolded subjects are alerted of ball release by an auditory cue but have no real-time information about TTC (Lacquaniti & Maioli 1989b). The hypothesis that an internal model of gravity is used by the brain to time catching actions has recently been tested in micro-gravity as well (McIntyre et al. 1999). Astronauts caught a ball projected from the ceiling at different, randomized speeds both on ground (1g) and in-flight (0g). Motor activity started too early at 0g, with time shifts in accord with the internal model hypothesis. Apparently, they did not believe their eyes that told them the ball was

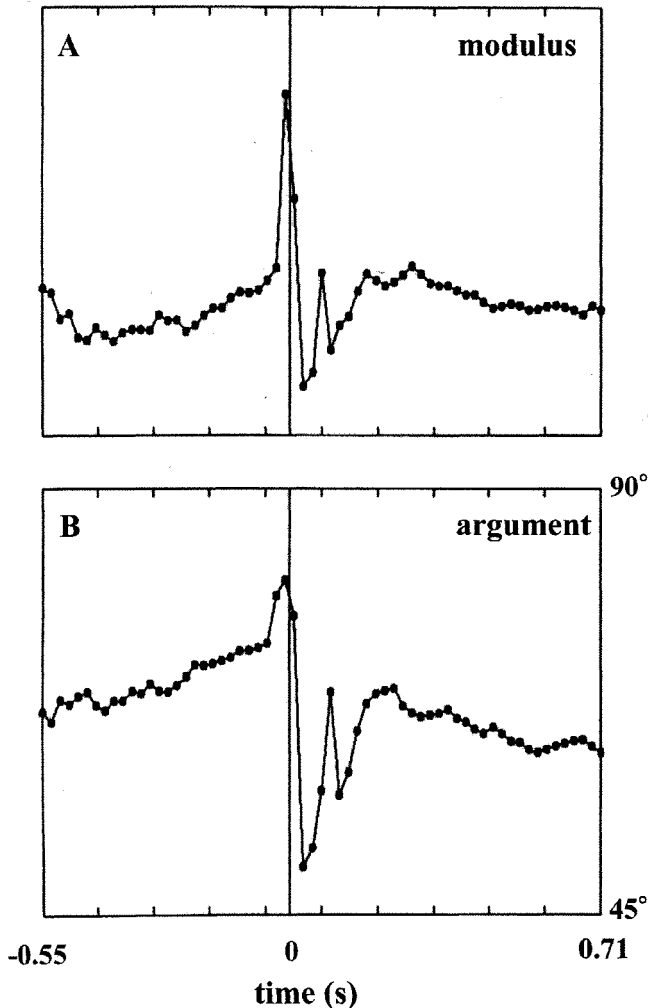


Figure 2 (Lacquaniti & Zago). Time course of the changes of end-point impedance during catching. Continuous, unpredictable perturbations were applied at the elbow joint by means of a torque motor, starting from ball release (time - 0.55 sec) through ball impact (time 0) and afterward. The time-varying values of stiffness and viscosity coefficients at the end-point were computed by cross-correlating input torque with output displacement. The modulus (arbitrary scale) and the argument of hand viscosity are plotted in A and B, respectively. A 0% argument corresponds to a horizontal vector pointing outward from the hand, whereas a 90% argument corresponds to a vertical, upward vector. Note that prior to ball impact, the magnitude of hand viscosity (and stiffness, not shown) increases significantly, while the direction of the viscosity vector rotates closer to the vertical, that is the direction of ball impact (modified from Lacquaniti et al. 1993b).

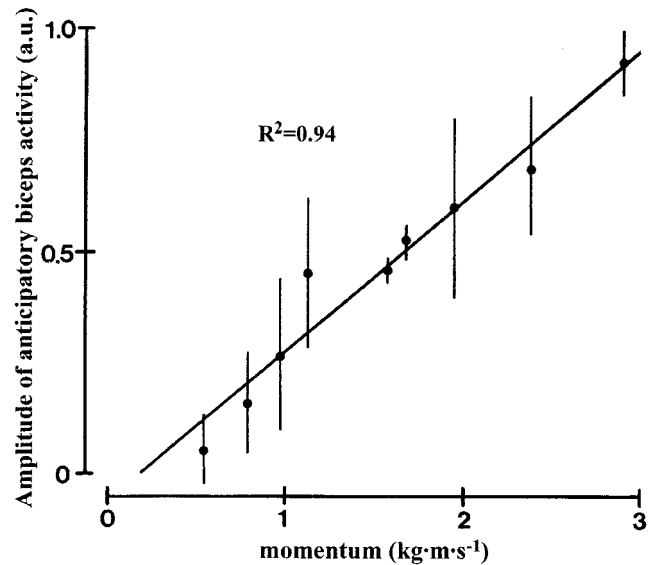


Figure 3 (Lacquaniti & Zago). Linear relation between the amplitude of biceps EMG anticipatory responses (mean value over the 50-msec interval preceding impact) and the momentum of the ball at impact time (modified from Lacquaniti & Maioli 1989a).

traveling at constant velocity, but they behaved as if the ball was still accelerated by gravity.

Catching studies also reveal that another dynamic parameter can be internalized, namely the predicted momentum at impact. Figure 3 shows that the amplitude of anticipatory muscle activity scales linearly with the expected momentum of the ball impact (Lacquaniti & Maioli 1989a). This was demonstrated using a factorial design, which involved the independent experimental manipulation of height of fall and mass of the ball. Thus, other kinematic or kinetic parameters could be excluded as putative control elements. In addition, it has been shown that, when the mass of the ball is unexpectedly changed, subjects scale their responses to the expected momentum.

In conclusion, we reviewed evidence that supports SHEPARD's hypothesis that during our evolutionary development we have internalized environmental regularities and constraints. In particular, we showed that physical laws may be internalized for our interaction with the environment even in cases in which they are not overtly exploited for perception and cognition. Moreover, the internal models of dynamics we have considered for the task of ball interception also satisfy BARLOW's criterion that the regularity must be turned to an advantage to have a biologically relevant value, as is well known to all fans of ball games.

Extending Bayesian concept learning to deal with representational complexity and adaptation

Michael D. Lee

Department of Psychology, University of Adelaide, SA 5008 Australia.
 michael.lee@psychology.adelaide.edu.au
<http://www.psychology.adelaide.edu.au/members/staff/michaellee/>

Abstract: While Tenenbaum and Griffiths impressively consolidate and extend Shepard's research in the areas of stimulus representation and generalization, there is a need for complexity measures to be developed to control the flexibility of their "hypothesis space" approach to representation. It may also be possible to extend their concept learning model to consider the fundamental issue of representational adaptation.
 [TENENBAUM & GRIFFITHS]

Two research areas in which Roger **SHEPARD** has made enormous contributions are stimulus representation (e.g., Shepard 1980) and stimulus generalization (e.g., Shepard 1987b). The Bayesian account of concept learning developed by **TENENBAUM & GRIF-FITHS (T&G)** addresses both of these areas, providing a unifying consolidation of Shepard's representational ideas, and a natural extension of the "consequential region" approach to modeling generalization. A number of challenges and problems, however, remain for future research.

On the representational front, **T&G** demonstrate that their approach to building representations is sufficiently flexible to accommodate spatial, featural, and a range of other established approaches. While **T&G** note that part of the attraction of Shepard's (1987b) theory is that it assumes well-defined representational structures, their number game demonstrates the need for richer representational possibilities. By modeling stimulus representations in terms of prior distributions across an unconstrained hypothesis space, **T&G** develop an approach that may be sufficiently sub-conceptual (Smolensky 1988) to act as a useful unifying framework.

The price of (representational) freedom, however, is eternal (complexity) vigilance. The representational flexibility of the hypothesis space approach demands that the complexity of the representations be controlled. In the absence of some form of Occam's Razor, there is a danger that arbitrary stimulus representations can be constructed to solve particular problems, without achieving the substantive interpretability, explanatory insight, and generalizability that is the hallmark of good modeling.

What is required is a method for imposing priors on a hypothesis space that satisfy representational constraints in a parsimonious way. Following **T&G**, it seems plausible that representational constraints could be internalized through evolution, or learned on the basis of interaction with the world. Any source of information that offers adaptive advantage provides a candidate for representational refinement. The important point is that the representational priors must accommodate the constraints at an appropriate level of generality. Representations fail to serve their adaptive purpose if they do not generalize, and do not allow what has been learned (or internalized) in the past to be brought to bear on present concerns.

T&G are certainly aware of this challenge, as their discussion of the origin of representational priors indicates. Their general notion of developing a "vocabulary for a variety of templates" to tackle the challenge is an intriguing and promising one. One of the fundamental tools needed to pursue this undertaking, however, is a mechanism for assessing the complexity of arbitrary hypothesis space representations, and **T&G** are comparatively silent on this issue.

Fortunately, there are grounds for optimism. The Bayesian framework adopted by **T&G** is well suited to addressing issues of model complexity (Kass & Raftery 1995), and there have been recent attempts to develop Bayesian complexity measures for multidimensional scaling, additive clustering, and other approaches to stimulus representation subsumed under the hypothesis space approach (Lee 1999). The additive clustering analysis (Lee 2001) is particularly promising in this regard, since it gives measures that are sensitive to the "functional form" component of representational complexity (Myung & Pitt 1997), as will surely be required for the general hypothesis space approach. Indeed, given the formal correspondence between **T&G**'s Bayesian model and Tversky's (1977) ratio model, and the close relationship of the ratio model to the contrast model that underpins additive clustering, some of the groundwork has already been laid.

In terms of stimulus generalization and concept learning, the model developed by **T&G** constitutes an impressive extension of Shepard's (1987b) approach, particularly through the introduction of the size principle. Their Bayesian formulation seems to capture important capabilities of human learning that are not obviously present in discriminative learning models such as **ALCOVE** (Kruschke 1992).

One issue that **T&G** do not substantially address is representa-

tional adaptation resulting from learning. A fundamental problem for any adaptive system with a memory is: how should established representations be modified on the basis of experience? The Bayesian account of concept learning involves the interaction of data-driven (perceptive) and knowledge-driven (apperceptive) components, and so is well placed to deal with this issue. Studies of learned categorical perception that measure the effects of concept learning on human mental representations (Goldstone et al. 2001) could provide one source of empirical data to guide theoretical development.

Ultimately, addressing the issue of adaptation requires an understanding of the way in which perceptive and apperceptive processes interact across different learning episodes and time scales. The Bayesian concept learning model modifies its representations to learn a particular concept from a small number of stimuli, but the permanence of these modifications is not clear. If a new concept is subsequently learned across the same stimulus domain, what is the effect of previous learning? Do the priors on the hypothesis space revert to their original state, or do they assume a different distribution that is partly influenced by the learned concept? In some cases, it seems likely that the representations will be unchanged. It would come as no surprise if human performance on repeated versions of the number game were shown to be independent of each other. For particularly salient concepts, or for conceptual relationships that are continually reinforced over time, however, there is a much stronger argument for change. On evolutionary time scales, the argument that representations have adapted in response to ancestral experience is compelling. Extending the Bayesian model of concept learning to balance the use of representations in learning with the use of learning in representation-building should be a focus of future research.

Finally, it may be worth some effort exploring the relationship between the Bayesian approach of **T&G**, and the "fast and frugal" approach to cognitive modeling (Gigerenzer & Todd 1999b). In discussing a related Bayesian model of prediction (Griffiths & Tenenbaum 2000), the same authors have argued that humans do not actually perform the Bayesian calculations specified by their model, but instead apply a simple heuristic that approximates the outcomes of these calculations. It would be interesting to know whether **T&G** hold the same view in relation to their model of concept learning and, if so, what sorts of heuristics they believe are likely to be involved.

ACKNOWLEDGMENTS

I wish to thank Daniel Navarro and Chris Woodruff for helpful comments.

Representation of basic kinds: Not a case of evolutionary internalization of universal regularities

Dennis Lomas

Department of Theory and Policy Studies, Ontario Institute for Studies in Education, Toronto, Ontario, M5S 1V6, Canada. dlomas@oise.utoronto.ca

Abstract: Shepard claims that "evolutionary internalization of universal regularities in the world" takes place. His position is interesting and seems plausible with regard to "default" motion detection and aspects of colour constancy which he addresses. However, his claim is not convincing with regard to object recognition.

[**SHEPARD**]

SHEPARD makes a convincing case for "evolutionary internalization of universal regularities in the world" with regard to "default" motion detection and to the aspects of colour constancy he addresses. His (provisional) attempt to apply the same principles to object recognition is not convincing. (I address the first five paragraphs of the sect. 1.10, "Formal characterization of generalization based on possible kinds.")