Available online at www.sciencedirect.com
ScienceDirect
Cognition 103 (2007) 460–472
COGNITION
www.elsevier.com/locate/COGNIT

Discussion

Rule following and rule use in the balance-scale task

Thomas R. Shultz *, Yoshio Takane

Department of Psychology, McGill University, Canada

Received 27 April 2006; revised 2 November 2006; accepted 7 December 2006

Abstract

Quinlan et al. [Quinlan, p., van der Mass, H., Jansen, B., Booij, O., & Rendell, M. (this issue). Re-thinking stages of cognitive development: An appraisal of connectionist models of the balance scale task. *Cognition*, doi:10.1016/j.cognition.2006.02.004] use Latent Class Analysis (LCA) to criticize a connectionist model of development on the balance-scale task, arguing that LCA shows that this model fails to capture a torque rule and exhibits rules that children do not. In this rejoinder we focus on the latter problem, noting the tendency of LCA to find small, unreliable, and difficult-to-interpret classes. This tendency is documented in network and synthetic simulations and in psychological research, and statistical reasons for finding such unreliable classes are discussed. We recommend that LCA should be used with care, and argue that its small and unreliable classes should be discounted. Further, we note that a preoccupation with diagnosing rules ignores important phenomena that rules do not account for. Finally, we conjecture that simple extensions of the network model should be able to achieve torque-rule performance.

Keywords: Connectionism; Rule learning; Balance scale; Latent class analysis; Rule-assessment method; Torque rule; Addition rule; Goodness-of-fit tests; Local independence assumption; Boundary estimates

* Corresponding author.
E-mail address: thomas.shultz@mcgill.ca (T.R. Shultz).

0010-0277/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2006.12.004

1. Introduction

A fundamental debate in cognitive science concerns the best theoretical account of knowledge representation, processing, and acquisition. Two main computational contenders have been the classic symbolic-rule account and the neurally inspired connectionist account. The classic view is that knowledge is represented in rules whose propositions refer to objects and events, that processing occurs as rules are selected and fired thus generating new symbolic propositions, and that knowledge is acquired by learning these symbolic rules. In many connectionist accounts, active knowledge is represented in fluctuating unit activations and long-term knowledge is represented as connections between units, processing occurs as activations are passed from one layer of units to another, and knowledge acquisition results from adjustment of connection weights. The symbolic view is sometimes referred to as *rule use*, and the connectionist view as *rule following*, to the extent that the environment affords regularities that a neural network can absorb.[1]

The use vs. following debate was joined by Quinlan, van der Maas, Jansen, Booij, and Rendell (this issue) in their critique of cascade-correlation connectionist models of development on the balance-scale task, one of the most frequently modeled tasks in developmental psychology. It is generally beneficial for a computational model to be examined from different perspectives than that of the original modelers. But if problems with a model are found, by either the original or secondary modelers, this need not trigger abandonment of the model. It is often more appropriate to determine whether problems can be fixed, particularly if the model offers useful insights, as cascade correlation has on several phenomena including conservation (Shultz, 1998, 2006), seriation (Mareschal & Shultz, 1999), transitive inference (Shultz & Vogel, 2004), integration of cues for moving objects (Buckingham & Shultz, 2000), pronoun acquisition (Oshima-Takane, Takane, & Shultz, 1999; Shultz, Buckingham, & Oshima-Takane, 1994), shift learning (Sirois & Shultz, 1998), learning of word stress (Shultz & Gerken, 2005), and habituation of infant attention to auditory (Shultz & Bale, 2001) and visual (Shultz & Cohen, 2004) information, in addition to balance-scale acquisition (Shultz et al., 1994).

In this rejoinder, we review available rule-detection methods for the balance scale, document and discuss the tendency of LCA methods to find small unreliable classes, underscore important balance-scale phenomena that rules cannot capture, and speculate about what might be required for neural networks to cover a torque rule.

2. Rule detection for the balance scale

Quinlan et al.'s critique relies on their use of a particular method for detecting rules known as Latent Class Analysis (LCA). We begin with a brief comparative

[1] A neural network functions the same whether the training environment is regular or not. But if the environment is regular enough to be described in rules, then a neural network might learn to behave as if it was following those rules, even though the rules are not explicitly represented as such within the network.