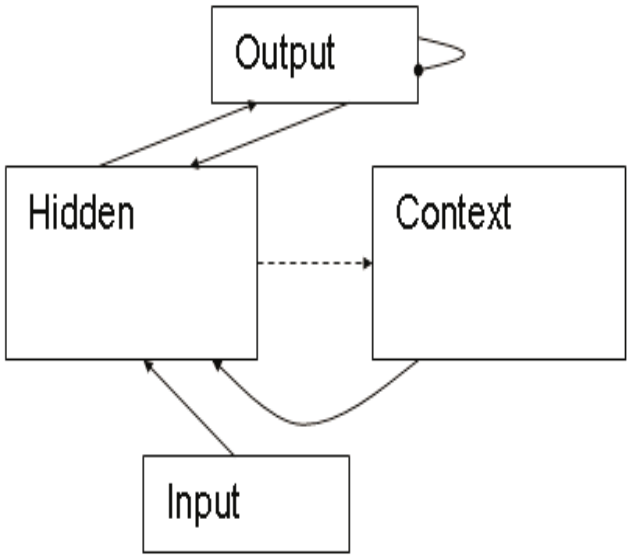



Figure 1 (Calvo Garzón). NCC network, with bidirectional connectivity and inhibitory competition, trained on a prediction task.

matrices, permits the network to alter the knowledge acquired in the weights by computing the difference between an initial phase where the networks activations are interpreted as its "expectation" of what's to happen, and a later phase in which the environment provides the output response to be taken as the teaching signal. Activation-based signals in a prediction task are not to be interpreted in Marcus's terms. The ecologically grounded prediction task of the networks does not incorporate universally open-ended rules. Unsupervised Hebbian learning, on the other hand, makes its contribution by representing in hidden space the first-order correlational structure of the data pool. Our NCC architecture delivers a correct syntactic interpretation of the infants' data. The data are accounted for without the positing of rule-fitting patterns of behavior (allegedly required to constrain novel data).

On development, where CC is graded as "better," the score may be made even more robust. A&L remark upon CC's anti-nativist stance on the nature/nurture debate. Marín et al. (2003) argue, in the context of poverty-of-stimulus arguments in Creole genesis, that CC eschews any form of nativism. Creole genesis, nativists contend, can only be explained by appealing to a Chomskian Universal Grammar (UG). Substratists contend that Creole genesis is influenced, crucially, by substratum languages. We show how the process by which a Pidgin develops into a Creole can be modelled by an SRN exposed to a dynamic (substratum-based) environment. In this way, an empiricist approach is able to account for Creole grammar as a by-product of general-purpose learning mechanisms. Connectionist theory, we argue, furnishes us with a (statistical) alternative to nativism. Taking into account that combined Hebbian and activation-phase learning drives SRN networks to a better performance on generalization than the backpropagation algorithm does (O'Reilly & Munakata 2000), a NCC replica of this simulation would further strengthen connectionist's stronghold on the development criterion.

Biologically plausible NCC would cast light as well upon other Newell Test criteria: Real-time Performance, where classical connectionism gets a "worse" grade, can be improved if we consider online dynamical coupling. In NCC models that do not depend on synchronous updates, it may be assumed, as A&L note, that "the time for a unit to reach full activation is a function of that activation" (sect. 5.2). Moreover, one-shot Hebbian learning (Rolls & Treves 1998), where a few event co-occurrences can contribute to fast recall, can also be seen as a motivation for not having to endorse a hybrid architecture. On the other hand, performance on the flexible behavior criterion would be enhanced as well. Notice that nonclassical, dynamical networks can compute any function to an arbitrary level of accuracy, while allowing for breakdowns in performance.

In general, I can see no good reason not to allow NCC to enter the decathlon. The best connectionist contestant should enter the competition, not a straw man (classical connectionism). It is usually argued that the endorsement of a symbolic-cum-subsymbolic stance would permit connectionism to remain at an appropriate level of realism (Palmer-Brown et al. 2002). However, "failure to acknowledge a symbolic level to thought" (target article, Abstract) cannot be interpreted as a weakness of connectionism when the score is revised as just described.

ACKNOWLEDGMENTS

The author was supported by a Ramón y Cajal research contract (Spanish Ministry of Science and Technology) during the preparation of this commentary. I thank Andy Clark, Eliana Colunga, and Javier Marín for helpful discussions.

## The Newell Test should commit to diagnosing dysfunctions

William J. Clancey

*Computational Sciences Division, MS 269–3, NASA Ames Research Center, Moffett Field, CA 94035.* **william.j.clancey@nasa.gov**
**http://bill.clancey.name**

**Abstract:** "Conceptual coordination" analysis bridges connectionism and symbolic approaches by positing a "process memory" by which categories are physically coordinated (as neural networks) in time. Focusing on dysfunctions and odd behaviors, like slips, reveals the function of consciousness, especially constructive processes that are often taken for granted, which are different from conventional programming constructs. Newell strongly endorsed identifying architectural limits; the heuristic of "diagnose unusual behaviors" will provide targets of opportunity that greatly strengthen the Newell Test.

Anderson & Lebiere's (A&L's) article evaluates cognitive theories by relating them to the criteria of functionality derived from Newell (1980; 1990). Suppose that the Newell Test (NT) has all the right categories, but still requires a significant architectural change for theoretical progress. I claim that "conceptual coordination" (CC) (Clancey 1999a) provides a better theory of memory, and that, without committing to explaining cognitive dysfunctions, NT would not provide sufficient heuristic guidance for leading in this direction.

Conceptual coordination (CC) hypothesizes that the store, retrieve, and copy memory mechanism is not how the brain works. Instead, all neural categorizations are activated, composed, and sequenced "in place," with the assumption that sufficient (latent) physical connections exist to enable necessary links to be formed (physically constructed) at run time (i.e., when a behavior or experience occurs). For example, if comprehending a natural language sentence requires that a noun phrase be incorporated in different ways, it is not moved or copied but is physically connected by activation of (perhaps heretofore unused) neural links. Effectively, Newell's "distal access" is accomplished by a capability to hold a categorization active and encapsulate it (like a pointer) so that it can be incorporated in different ways in a single construction. The no-copying constraint turns out to be extremely powerful for explaining a wide variety of odd behaviors, including speaking and typing slips, perceptual aspects of analogy formation, developmental "felt paths," multimodal discontinuity in dreams, and language comprehension limitations. CC thus specifies a cognitive architecture that bridges connectionist and symbolic concerns; and it relates well to the NT criteria for which ACT-R scores weakest – development, consciousness, language, and the brain.



https://doi.org/10.1017/S0140525X03250135 Published online by Cambridge University Press

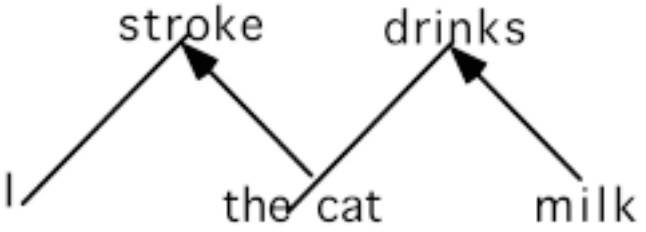



Figure 1 (Clancey). Unproblematic: "I stroke the cat that drinks the milk."

To illustrate, I provide a diagnostic analysis of an autistic phenomenon and then relate this back to how NT can be improved.

In CC analysis, a diagram notation is used to represent a behavior sequence, which corresponds in natural language to the conceptualization of a sentence. For example, according to Baron-Cohen (1996), an autistic child can conceptualize "I stroke the cat that drinks the milk." In one form of the CC notation, a slanting line to the right represents categorizations activated sequentially in time (e.g., "I – stroke" in Fig. 1). Another sequence may qualify a categorization (e.g., "the cat – drinks" qualifies "stroke"). This pattern of sequences with qualifying details forming compositions of sequences occurs throughout CC analysis. The essential idea in CC is to understand how categories (both perceptual and higher-order categorizations of sequences and compositions of them) are related in time to constitute conscious experience (Clancey1999a).

The challenge is to understand why an autistic child finds it problematic to conceptualize "I see the cat that sees the mouse." A traditional view is that the child lacks social understanding. But CC analysis suggests a mechanistic limitation in the child's ability to physically sequence and compose categories. Relating to other agents requires being able to construct a second-order conceptualization that relates the child's activity to the other agent's activity. Figure 2 shows the CC notation for the required construction.

The statement (the conceptualization being constructed) involves a triadic relation: I see the cat, the cat sees the mouse, and I see the mouse. There is one mouse that we are both seeing. Two "see" constructions are unified by identifying a detail (the mouse) as common to both. In effect, the child must conceive of a problem space (Clancey 1999a): A common categorization of an operand (mouse) enables categorization of multiple actions as being one action (seeing), an operator. Because the two actions are by different agents, accomplishing this identification integrates perspectives of *self* (what I am doing now) and *other* (what that object is doing now). Indeed, the conceptualization of agent appears to be inherent in this construction.

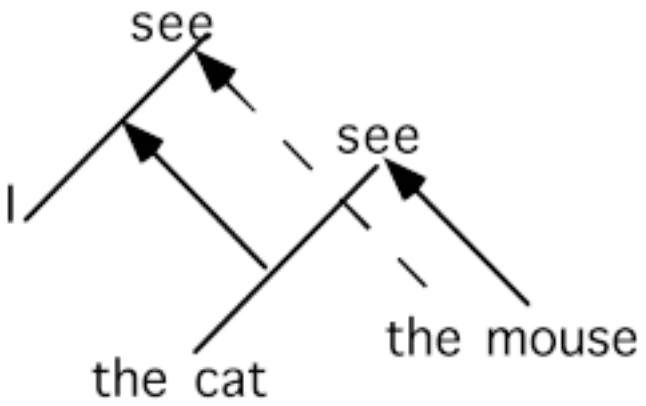



Figure 2 (Clancey). Problematic: "I see the cat that sees the mouse."

Put another way, two sequentially occurring conceptualizations (I see the cat; the cat sees the mouse) are held active and related: "I see the cat that sees the mouse" and "I see the mouse" become "I see that the cat sees the mouse" (i.e., the mouse that I am seeing). (The second-order relation is represented in Fig. 2 by the solid arrow below "I see.") Conceiving this relation is tantamount to conceiving what joint action is. Barresi and Moore (1996) characterize this as "integrating third and first person information" (p. 148), and contrast it with (Fig. 1) "embedding one third person representation in a separate first person frame" (p. 148). Related to Langacker's (1986) analysis, logical relations are not extra capabilities or meta "inference" capabilities, but generalizations of concrete accomplishments that arise through the capability to physically coordinate categories through identification, sequence, and composition in time. Mental operations are physical, subconscious processes, constrained by physical limits on how inclusion in new sequences can occur. The ability to hold two sequences active and relate them constitutes a certain kind of consciousness (e.g., not present in dreaming; Clancey 2000).

To summarize, the example requires relating sequential categorizations of seeing so that they become simultaneous; it exemplifies a second-order conceptualization of intentionality (my seeing is about your seeing; Clancey 1999b); and suggests that joint action requires being able to conceive the ideas we call operator and agent.

The pivotal heuristic in CC analysis is addressing *unusual* behaviors and experiences. These "targets of opportunity" appear to be de-emphasized by A&L's focus on normal behaviors "that people display on a day-to-day basis." For NT to provide heuristic guidance for discovering a theory like CC, grading for each criteria should include diagnosing unusual phenomena that everyone experiences (e.g., slips) and dysfunctions. For example, for the criteria of consciousness, we should direct theorization at explaining the phenomenology of dreaming, autism, compulsive-obsessive disorders, and the like. For natural language, include comprehension difficulties (e.g., subject relatives with center-embedded noun phrases; Clancey 1999a, Ch. 10). For development, explain how "felt paths" are constructed in children's learning (Ch. 5). For knowledge integration, explain slips (Ch. 6) and "seeing as" in analogy formation (Ch. 7). In this manner, learning in well-known architectures (e.g., MOPS, EPAM, SOAR) can be evaluated and the nature of problem spaces reformulated (Ch. 12).

The evolution criterion highlights the limitations of NT as stated. Rather than focusing on human evolution, this criterion should be about the evolution of cognition broadly construed, and hence should be inherently comparative across species (Clancey 1999b). Viewed this way, there is no "paucity of data," but rather a largely unexploited potential to make the study of animal cognition an integrated discipline with human problem solving. By including the heuristic "explain odd behavior" in the grading, we will naturally be guided to characterize and relate cognition in other primates, ravens, and the like. This is essential for relating "instinctive" mechanisms (e.g., weaving spider webs) to brain mechanisms, development, and learned higher-order categorizations (e.g., conceptualization of intentionality). A&L mention comparative considerations, but we should view this as a diagnostic problem, much as cognitive theories like ACT* have been used to explain students' different capabilities (Anderson et al. 1990). Furthermore, the research community should collect behaviors that have been heretofore ignored or poorly explained by computational theories and include them in the grading criteria.

Applying the Newell Test in this way – moving from the routine behaviors already handled more or less well, to diagnostic theories that relate aberrations to architectural variations – might bring symbolic and connectionist theories together and make the study of cognition a more mature science.



https://doi.org/10.1017/S0140525X03250135 Published online by Cambridge University Press