Doing without schema hierarchies: A recurrent connectionist approach to routine sequential action and its pathologies

Matthew Botvinick[1,2] and David C. Plaut[1]

[1]Department of Psychology, Carnegie Mellon University and
Center for the Neural Basis of Cognition, [2]Department of Psychiatry, University of Pittsburgh

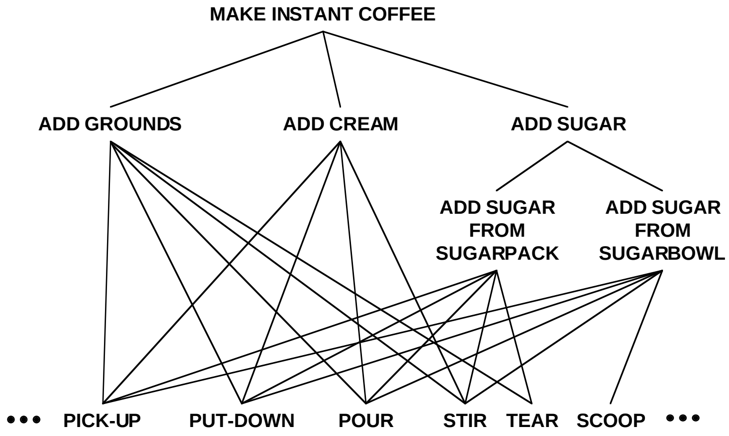
Presented at the Annual Meeting of the Cognitive Neuroscience Society
San Francisco, California, April 11, 2000

Abstract

In everyday tasks, selecting actions in the proper sequence requires a continuously updated representation of temporal context. Many existing models address this problem by positing a hierarchy of processing units, mirroring the roughly hierarchical structure of naturalistic tasks themselves. Such an approach has led to a number of difficulties, including a reliance on overly rigid sequencing mechanisms, an inability to account for context sensitivity in behavior, and a failure to address learning. We consider here an alternative framework, according to which the representation of temporal context is facilitated by recurrent connections within a network mapping from environmental inputs to actions. Applying this approach to a specific, and in many ways prototypical, everyday task (coffee-making), we examine its ability to account for several central characteristics of normal and impaired human performance. The model we consider learns to deal flexibly with a complex set of sequencing constraints, encoding contextual information at multiple time-scales within a single, distributed internal representation. Mildly degrading this context representation leads to errors resembling everyday "slips of action." More severe degradation leads to a pattern of disorganization resembling that observed in ideational apraxia or "action disorganization syndrome." Analysis of the model's function yields novel, testable predictions relevant to both normal and apractic performance. Taken together, the results obtained indicate that recurrent connectionist models may offer a useful framework for understanding routine sequential action.

The sequencing problem

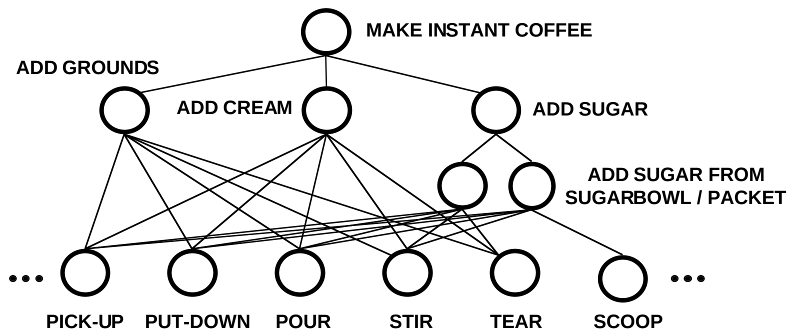
- Tempting hypothesis: Sequencing involves a chain of associations between successive responses.
- Problem: Any given action may occur in many contexts (e.g., action of playing given note on piano occurs in many pieces). Chaining is therefore insufficient (Lashley, 1951).
- Sequencing thus requires access to a richer representation of temporal context and sequential structure. Often termed a "schema."[1,4,6,11]
- A further complication: Structure of naturalistic sequences often involves multiple levels. For example, actions may fit together into subtasks, and subtasks into overall tasks[9,14]
- Frequent conclusion is that this calls for a hierarchy of schemas, pertaining to different levels of structure



Hierarchical models of action

A common approach

- Elements of task structure directly identified with individual processing elements (sometimes cast as implementation of schemas)[1,3,5,8,12]
- Hierarchical structure of task built directly into architecture



Successes

- Used to account for some aspects of:
  - Normal (error-free) sequencing in hierarchical domains[1,5,8,12]
  - “Slips of action:” Everyday errors made by normal subjects[1,12]
  - Action disorganization syndrome[6,14,15] (ADS): Variety of apraxia affecting sequencing in everyday routines (often following diffuse brain injury)[1]

Problems

- Limited account of learning; heavy reliance on built-in structure
- Reliance on inflexible or ad hoc sequencing mechanisms (e.g. reflex inhibition)
- Difficulty coping with quasi-hierarchical task domains, where details of subtask performance depend on larger task context

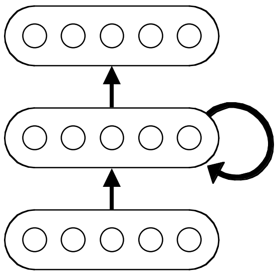
The present alternative

Basic tenets

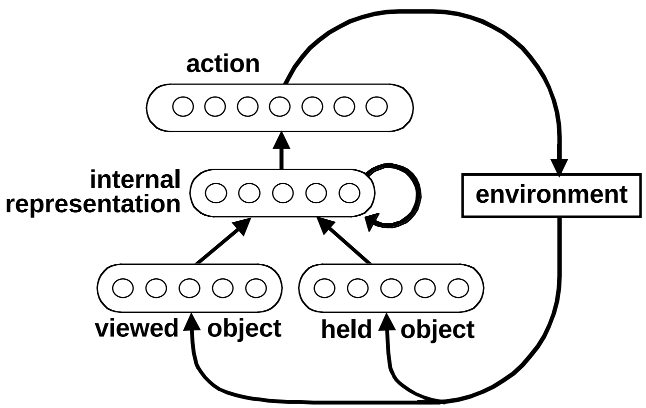
1. Schemas as emergent system properties, rather than explicit representations

2. Knowledge of sequences learned, rather than built into the architecture

- Approach based on simple recurrent network framework introduced by Elman[2] (right)
- Input mapped to output using learned, distributed internal representations
- Recurrent connections allow maintenance of information about temporal context



The model



- Simple recurrent network modified to model action on objects in environment
- Input layer represents features of 1) object viewed and 2) object held
- Output layer represents actions (including visual search, e.g. "locate-cup")
- Interpreting actions: viewed object = target; held object = implement
- Example:

  viewed: coffee, held: spoon, action: stir = "stir coffee using spoon"
- Perception-action loop: Inputs updated based on last action
- Modeling dysfunction: Random noise added to hidden layer at end of each processing cycle, degrading network's representation of temporal context

Target empirical phenomena

Normal behavior

- Ability to deal with hierarchically-structured tasks
- Ability to select actions/subtasks that may appear in multiple contexts
- Environmental cues alone often insufficient to guide action selection
- In some tasks, subtasks may be executed in variable order
- Some subtasks interchangeable
- Details of subtask performance often depend on context in which subtask is performed

Slips of action[10,11]

- Occur during distraction or preoccupation
- Occur at “decision points,” typically transition points between subtasks
- Usually take form of familiar subtask sequence, appearing intact but in wrong context
- Usually involve repetition or omission of a subtask, or lapse from one task into another
- Lapses typically involve shift from less frequent to more frequent task

Action disorganization syndrome[6,14,15]

- Fragmentation of sequential structure: With increasing ADS severity, increase in “independent actions” (actions lying outside a completed subtask)
- Characteristic error types: omissions, sequence errors (perseverations, reversals, anticipations), object and action substitutions, action additions, action quality errors
- Patients who make more errors commit a higher proportion of omission errors[15]

Methods

Task

- Coffee-making (4 versions)

  grounds → sugar (bowl) → cream → drink
  grounds → sugar (pack) → cream → drink
  grounds → cream → sugar (bowl) → drink
  grounds → cream → sugar (pack) → drink

- Tea-making (2 versions)

  teabag → sugar (bowl) → drink
  teabag → sugar (pack) → drink

- Background examples (265)

  inputs: *each possible scene (seen, held objects)*
  targets: *all physically-completable actions*

STEP	SEEN OBJECT	HELD OBJECT	ACTION
1	cup-1handle-clearliquid	nothing	locate-coffeepack
2	packet-foil-untorn	nothing	pickup
3	packet-foil-untorn	packet-foil-untorn	pullopen
4	packet-foil-torn	packet-foil-torn	locate-cup
5	cup-1handle-clearliquid	packet-foil-torn	pour
6	cup-1handle-brownliquid	packet-foil-torn	locate-spoon
7	spoon	packet-foil-torn	putdown
8	spoon	nothing	pickup
9	spoon	spoon	locate-cup
10	cup-1handle-brownliquid	spoon	stir
11	cup-1handle-brownliquid	spoon	locate-sugar
12	cup-2handles-lid	spoon	putdown
13	cup-2handles-lid	nothing	pinchlift
14	cup-2handles-sugar	lid	locate-spoon
15	spoon	lid	putdown
16	spoon	nothing	pickup
17	spoon	spoon	locate-sugarbowl
18	cup-2handles-sugar	spoon	scoop
19	cup-2handles-sugar	spoon-sugar	locate-cup
20	cup-1handle-brownliquid	spoon-sugar	pour
21	cup-1handle-brownliquid	spoon	stir
22	cup-1handle-brownliquid	spoon	locate-carton
23	carton-closed	spoon	putdown
24	carton-closed	nothing	pickup
25	carton-closed	carton-closed	peel-open
26	carton-open	carton-open	locate-cup
27	cup-1handle-brownliquid	carton-open	pour
28	cup-1handle-brownliquid-milk	carton-open	locate-spoon
29	spoon	carton-open	putdown
30	spoon	nothing	pickup
31	spoon	spoon	locate-cup
32	cup-1handle-brownliquid-milk	spoon	stir
33	cup-1handle-brownliquid-milk	spoon	putdown
34	cup-1handle-brownliquid-milk	nothing	pickup
35	cup-1handle-brownliquid-milk	cup-1handle-brownliquid-milk	sip
36	cup-1handle-brownliquid-milk	cup-1handle-brownliquid-milk	sip
37	cup-1handle-empty	cup-1handle-empty	saydone

Representations

VIEWED INPUT	HELD INPUT	ACTION
cup	cup	pickup
1handle	1handle	putdown
2handles	2handles	pour
lid	lid	peelopen
water	water	tearopen
brownliquid	brownliquid	pullopen
milk	milk	pinchlift
carton	carton	scoop
open	open	sip
closed	closed	stir
packet	packet	locate-cup
foil	foil	locate-sugar
paper	paper	locate-sugarbowl
torn	torn	locate-teabag
untorn	untorn	locate-coffeepack
spoon	spoon	locate-spoon
teabag	teabag	locate-carton
sugar	sugar	saydone
coffee-instruction	nothing	
tea-instruction		

Methods (continued)

Training

- Full training set included all coffee-making, tea-making, and background example sequences
- At each timestep, network predicted next action
- Weights updated using backpropagation through time, adapted to SRN framework
- Learning rule included term favoring small changes in hidden unit activation over larger changes (not necessary for results obtained).
- Training continued until correct action selected on every step (activation > .7; ~10,000 epochs)

Testing

- Initial input from coffee and tea sequences applied
- Most active output unit used to identify selected action
- Feedback from environment based on actions selected
- Degradation of temporal context: Same weights used, noise added to hidden unit activation at end of each cycle (Slips simulation: variance 0-0.1; ADS simulation: variance 0.1-0.5).

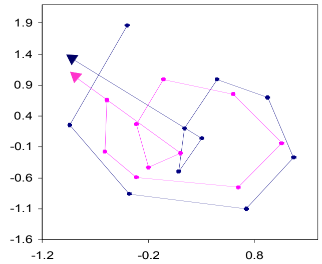
Evaluation of performance

- Normal performance: Sequences produced compared with sequences in training set
- Errors generated at low noise compared with slips of action (see target phenomena)
- Performance at higher levels of noise compared with behavior in ADS. Independent actions and specific error types quantified based on Schwartz et al. (1991)

Results (1): Normal behavior

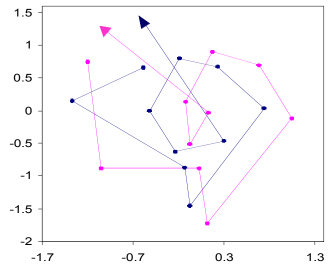
- At test, network spontaneously reproduced all sequences in training set, without errors
- System copes with hierarchically structured task without relying on hierarchical architecture
- Instead, hierarchical nature of task reflected in learned internal representations

Relationships among successive internal representations can be visualized using multidimensional scaling, which effectively projects the 50-dimensional internal representation onto two dimensions.



Coping with actions associated with multiple contexts

- Example: Adding sugar, as part of coffee vs. tea task
- Similar trajectories reflect identical stimulus-response mappings
- Difference between trajectories reflects maintenance of information concerning which task is being performed, crucial for guiding action selection at end of sugar sequence



Coping with flexible ordering constraints

- Example: Adding cream, before vs. after having added sugar
- Similar trajectories reflect identical stimulus-response mappings
- Difference between trajectories reflects maintenance of information concerning whether sugar has been added, needed for guiding action selection at end of cream sequence

Representing quasi-hierarchical structure

Consider the following scenario...

An office assistant has the job of making coffee for three different executives:

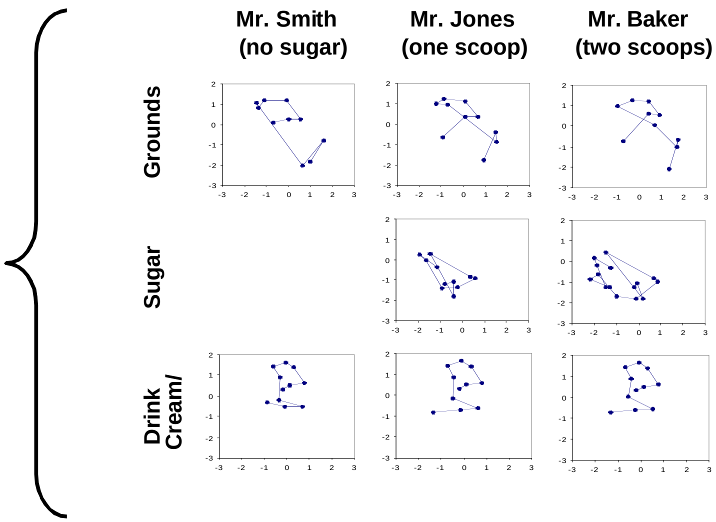
Mr. Smith (cream but no sugar)
Mr. Jones (cream and one scoop of sugar)
Mr. Baker (cream and two scoops of sugar)

Dilemmas for hierarchical approach

- Represent coffee-making as one “schema unit” or three?
- Represent sugar-adding as one schema unit or two?
- Basic problem: No way to simultaneously encode similarities and differences between tasks

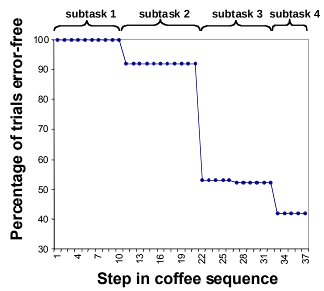
Network performance

- Network model learns to perform all sequences correctly
- Internal representations of related subtasks show “family resemblances”
- Use of distributed internal representation allows network to capture distinctive features of particular sequences, while still capturing similarities



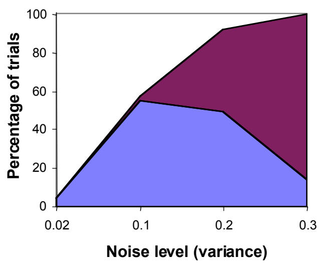
Results (2): Slips of action

Low levels of noise resulted in errors resembling everyday "slips of action"



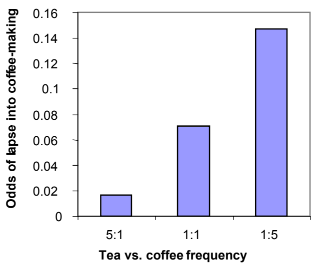
Errors occurred at "decision points"

- Survival plot shows percentage of trials still error-free at each step of coffee sequence
- Errors occurred primarily at transitions between subtasks (as reported for slips of action by Reason, 1990).



Errors took form of displaced but intact subtask sequences

- Blue region: Percentage of trials containing errors of subtask placement only. Purple: Other error trials
- At low noise, only subtask-placement errors occurred: Subtasks were repeated, omitted, and they "intruded" from other task (as in empirical data relating to slips of action; Reason, 1990).



Lapse errors showed effect of task frequency

- Case study: Lapse from tea- into coffee-making, reflected in intrusion of cream-adding sequence
- Occurred more often when tea-making infrequent relative to coffee-making during training (reproducing the frequency effect reported by Reason, 1990).

The etiology of slips

Slips occur when noise causes internal representation to resemble a pattern usually associated with a different situation

Example error:

sugar sequence → cream sequence →
sugar sequence repeated

Etiology:

Internal representations drift over course of cream sequence, coming to resemble those usually present when cream is being added *before* sugar

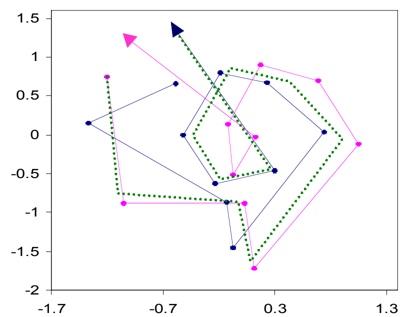
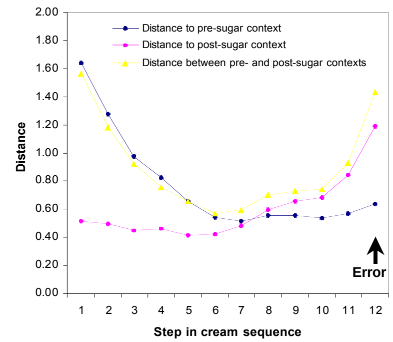
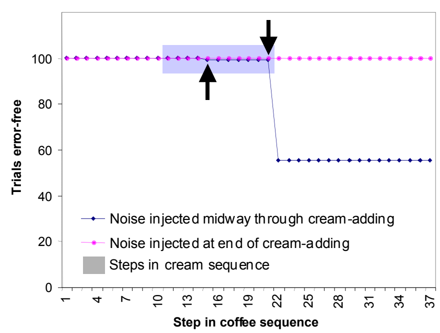


Illustration: During cream sequence, internal representations (green) drift away from those indicating sugar has been added (pink) toward those indicating sugar not yet added (blue).



Data: Distance of representations from usual post-sugar (pink) and pre-sugar (blue) patterns, over course of cream sequence.



- Even without noise, context information “compressed” toward middle of subasks (see yellow trace above-right)
- Implication: Noise injected toward middle of subtasks more likely to result in subsequent error than noise closer to end (left)
- Behavioral prediction: Distraction midway through a subtask should lead to errors at task end, and to more errors than noise injected later in subtask

Results (3): Action disorganization syndrome

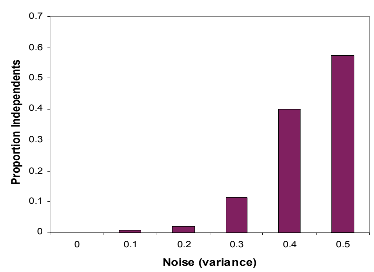
Higher levels of noise resulted in behavior resembling ADS

Sample of behavior

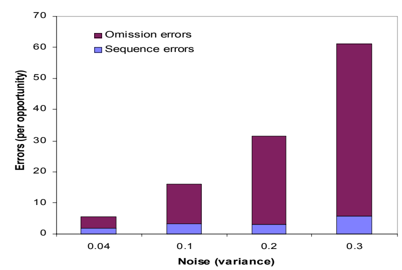
pickup cup
sip
putdown cup
pickup carton
peelopen carton
putdown carton
pinchlift [open] sugarbowl
putdown lid
pickup spoon
putdown spoon
pickup coffeepack
putdown coffeepack
pickup sugarbowl
putdown sugarbowl
pickup coffee-pack
pullopen coffee-pack
pour coffee-pack into cup
putdown coffeepack
pickup spoon
putdown spoon
pickup cup
sip
saydone

Model made same sorts of errors as ADS patients

Error type	Example
Omission	Sugar not added
Sequence:	
Anticipation	Pour cream without opening
Perseveration	Add cream, add sugar, add cream again
Reversal	Stir water then add grounds
Object substitution	Stir with coffee-pack
Gesture substitution	Pour gesture substituted for stir
Tool omission	Pour sugarbowl into cup
Action addition	Scoop sugar with, then put down, lid
Quality	Pour cream four times in a row

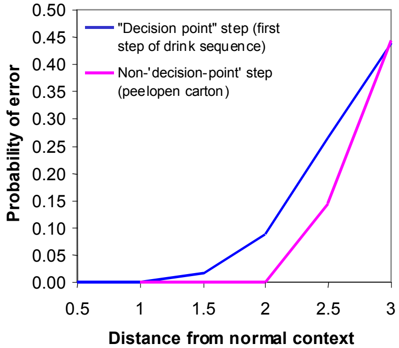


With increasing error rate, increasing fragmentation of sequential structure, increasing proportion of "independents" (as shown for ADS by Schwartz et al., 1991)



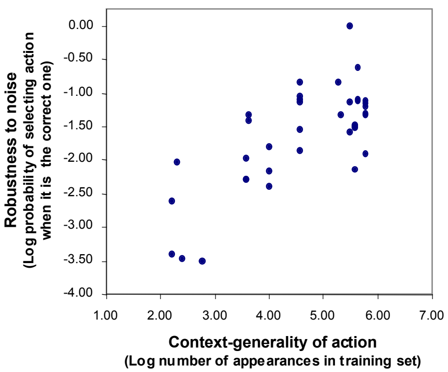
With increasing overall error rate, proportion of omissions rose, while proportion of sequence errors did not (as shown for ADS by Schwartz et al., 1998)

Slips and ADS: Points on a continuum



- As noise rises, errors occur first at “decision points,” later at other points
- Same mechanism for both: Noise causes internal representation to resemble one associated with different situation, different action
- Errors occur first at decision points because less distortion necessary to produce familiar context associated with a different action (Figure)
- Fits with idea that ADS represents exaggeration of action pathology present even in normal function (Schwartz et al. 1998)

Resorting to context-general actions



- With increasing noise, bias emerges toward selecting actions associated with large variety of contexts
- Result: Errors most frequent where correct action is a context-specific (Figure)
- In simulation most robust actions are *pickup, putdown*, and *locate.* Bias toward these actions at high noise fits with “toying behavior” and long periods of visual scanning observed in ADS[14]

Conclusions

- Recurrent networks appear to provide a plausible framework for understanding routine sequential action.
- According to this framework, action schemas inhere in emergent system dynamics rather than in discrete architectural elements. This allows the mechanisms needed to deal with complex sequencing constraints to develop through learning, rather than being built explicitly into the structure of the processing system, and enables the system to deal naturally with quasi-hierarchical sequential structure.
- The recurrent connecionist framework provides a parsimonious account of action pathology across the spectrum from slips of action to action disorganization syndrome, and makes a number of specific and testable predictions, providing the opportunity for future empirical tests of the approach.

References

1. Cooper, R. & Shallice, T. (In press). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology.*
2. Elman, J. L. (1990). Finding structure in time. *Cognitive Science,* **14,** 179-211.
3. Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton and E. Martin (Eds.), *Coding Processes in Human Memory.* Washington, D. C.: V. H. Winston & Sons.
4. Grafman, J. (1989). Plans, actions, and mental sets: Managerial knowledge units in the frontal lobe. In E. Perecman (Ed.), *Integrating theory and practice in neuropsychology.* Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
5. Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zock (Eds.), *Handbook of Spelling: Theory Process and Intervention.* Chichester, UK: John Wiley.
6. Humphreys, G. H. & Forde, E. M. E. (1998). Disordered action schema and action disorganization syndrome. *Cognitive Neuropsychology,* **15,** 771-881.
7. Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in Behavior.* New York: John Wiley & Sons.
8. MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills.* New York: Springer Verlag.
9. Miller, G. A., Galanter, E. & Pribram, K. H. (1960). *Plans and the Structure of Behavior.* New York: Holt, Rinehart and Winston.
10. Norman, D. A. (1981). Categorization of action slips. *Psychological Review,* **88,** 1-15.
11. Reason, J. T. (1990). *Human Error.* Cambridge, UK: Cambridge University Press.
12. Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science,* **6,** 1-36.
13. Schank, R. C., & Abelson, R. P. (1977). *Scripts, Plans and Understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
14. Schwartz, M. F., Edwards, E. S., Montomery, M., Palmer, C., & Mayer, N. H. (1991). The quantitative description of action disorganization after brain damage: A case study. *Cognitive Neuropsychology,* **8,** 381-414.
15. Schwartz, M. F. Montgomery, M. W., Buxbaum, L. J., Lee, S. S., Carew, T. G., Coslett, H. B., Ferraro, M., Fitzpatrick-DeSalme, E., Hart, T., & Mayer, N. (1998). Naturalistic action impairment in closed head injury. *Neuropsychology,* **12,** 13-28.