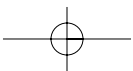
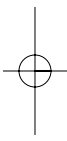
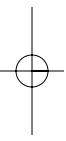


## **Section 6**

---

# **Cognitive rehabilitation theory**



## 22 For a theory of cognitive rehabilitation

---

### *Progress in the decade of the brain*

Argye E. Hillis

#### **Abstract**

A theory of cognitive rehabilitation should specify how change from a damaged state of cognitive processing can be modified into a normal, or more functional, state of cognitive processing. Such a theory should incorporate what is known about the cognitive representations and processes underlying normal cognition, how these are affected by brain damage, and how learning or modification of cognitive processing occurs. It is therefore argued that development of a useful theory of cognitive rehabilitation will require integrating advances from cognitive neuropsychology, experimental psychology, computational neuroscience, and molecular biology of the brain, as well as empirical evidence from various branches of rehabilitation. It is likely that such a theory will specify how behavioral rehabilitation strategies can be augmented by pharmacological agents.

Rehabilitation of cognitive impairments is among the most challenging and rewarding endeavors of clinician-scientists. Hence, it is no wonder that seemingly everyone wants to 'get in on the act.' In the past two decades, cognitive rehabilitation has been the focus of investigators and therapists in a wide range of disciplines: speech-language pathology, occupational therapy, clinical psychology, neuropsychology, experimental psychology, neurology, neuroscience, linguistics, education, neuroimaging, computationalism, and others. This cross interaction has been productive. For example, many speech language pathologists treating aphasia or other cognitive impairments caused by focal brain damage have found it useful to consider cognitive neuropsychological models of the cognitive processes underlying the task to be treated (see Chapey 2001; Coltheart, Chapter Brunsdon and Nickess, Chapter 2 this volume; Hillis 2002; Riddoch and Humphreys 1994a; Seron and DeLoche 1989, for examples). Several authors have argued that these models provide an essential first step to rehabilitation, in terms of identifying the components of each task that are impaired, and the components that are spared, allowing the clinician to capitalize on the spared components and to focus treatment or facilitation on the damaged components (Beeson and Rapczak 2002; Hillis 1993, 1994, 1998; Riddoch and Humphreys 1994b; Wilson and Patterson 1990). The focus of this chapter concerns the next step: How do we move beyond the demonstrated success in determining 'what to treat' to the critical issues of 'when, how, and how much to treat'?

We (Caramazza and Hillis 1993; see also Hillis 1993, 1998) have previously argued that the questions of when, how, and how much to treat cannot be answered on the basis of cognitive neuropsychological models. Such models have been developed to represent a theory of the normal cognitive processes that must be engaged in order to perform a task (e.g., reading a word). They are

silent with respect to which of these cognitive processes are most subject to change, or how changes might be brought about. We did not offer a theory of rehabilitation, but did lay out some criteria for what a theory of rehabilitation would entail.

In response to Caramazza and Hillis (1993), Baddeley (1993) argued that at least some of the crucial aspects of a theory of rehabilitation could be provided by learning theory and by connectionist models. Baddeley pointed out rehabilitation requires re-learning, so that theories of learning and memory and connectionist models of learning may be relevant to rehabilitation.

I am in full agreement that learning theory and connectionism have much to offer toward developing a theory of rehabilitation of cognitive impairments resulting from brain damage. But below I will point out the limitations of each of these domains, and will propose that further advances toward a theory of rehabilitation will be provided by neuroscience. That is, I will argue that a theory that can direct clinicians as to when, how, and how much to treat will depend on advancements in knowledge about how the human brain works, how it is damaged by stroke, trauma, and disease, and how it recovers from these injuries. The biology of the brain will need to be integrated with our psychological and computational models of normal cognitive processes. First, I will discuss the contributions and limitations of such psychological and computational theories, and then turn to the potential contributions of biology of the brain.

## Learning theory

Decades of experiments have provided a wealth of important principles about how people learn and retain new information: many of these principles have been based on the work of empiricists. We have learned much about operant and classical conditioning and shaping behaviors. Therapists incorporate random, intermittent reinforcement into their rehabilitation, based on evidence that these reinforcement schedules result in more durable training effects than constant reinforcement. Similarly, therapy is based on the principle that spaced practice is more effective than mass practice. There have also been recent illustrations of the effectiveness of 'errorless learning' (see Kessels and de Hann 2003 for review). Nevertheless, the limitations of these principles for cognitive rehabilitation are twofold. First, many of these principles are based on careful, well-documented, well-replicated observations, without concern as to the mechanisms of why they work (although we will return to issue of how and why they work, which has been elucidated by cellular level neuroscience). Without knowing the mechanisms, the principles cannot contribute much to a *theory* of rehabilitation. They only contribute to the *practice* of rehabilitation. Second, it is not clear that the damaged brain can always learn the same way as normal brains. For example, anoxic brain damage may cause substantial damage to the hippocampus and amygdala, structures that are crucial to learning new information. In such cases, intensive, spaced practice and appropriate reinforcement schedules, even in the appropriate modality, may not be sufficient for learning new facts. Also, brain injury may affect the chemistry of the brain in such a way that it can no longer respond to the rewards that would normally reinforce learning.

More recent learning theories have proposed mechanisms of laying down new memories, such as dependence on the phonological loop for learning through the auditory modality and dependence on the visual spatial sketchpad for learning visual information (Baddeley 1993). These theories recognize that brain damage can result in an inability to learn the way normal brains learn. Nevertheless, the constraints of such theories on rehabilitation are minimal (e.g., reliance on errorless learning, which might be chosen on the basis of other theories). Baddeley (1993) also suggests that in cases of impairment of the phonological loop, reliance on learning through the visual modality is crucial, while in cases of impairment of the visuospatial sketchpad reliance on learning through the auditory modality is essential. In other cases of impaired new learning, acquisition of new memories may depend on the

use of episodic long term memory – e.g., use of visual imagery mnemonics to learn people’s names (Wilson 1987) or reliance on implicit learning (Glisky *et al.* 1986). However, it is not clear how this approach differs from the cognitive neuropsychological approach of identifying the impaired and spared components of the cognitive processes underlying a task, and relying on the spared components (or focusing treatment on the impaired ones). Again, we see that cognitive models (in this case, of learning) help to identify what to treat more than how or when to treat.

## Connectionism

Computational models come in two main forms: those that learn (change the strength of their connections in response to input; e.g., parallel distributed models of McClelland and colleagues), and those in which connections strengths or ‘weights’ are specified (e.g., Dell 1986). The former are based on principles of Hebb (see [Hinton](#), McClelland and Rumelhart 1986, and Robertson, Chapter 23 this volume, for review). Hebbian learning has been summarized as, ‘Cells that fire together wire together.’ (author unknown; quoted from McClelland 2002). That is, the more often the firing of neuron A stimulates neuron B to fire, the more easily neuron A will be able to stimulate neuron B in the future. Connectionist models can capture this crucial feature of neurons – the more times a network is exposed to a particular stimulus, the stronger the connections that link the stimulus with the response become (and the more likely it becomes that the network will produce the same response to that stimulus in the future).

AU: Hinton et al 1986 – no entry in references

Furthermore, recent computational models have provided an account of the mechanisms of two complementary memory systems in the brain. It has long been recognized that the hippocampal system is crucial for laying down new information, but not for recall of remote memories, whereas the neocortex has the opposite role. A computational model described by McClelland and colleagues (1995) accounts for these two systems in the following way: new memories are laid down through changes in synapses in the hippocampal system, and these changes result in minor changes in the neocortex which support recent memories. Recent memories then decay, unless there is continued support from the hippocampal system. Continued support from the hippocampal system results in accumulated minor changes in the neocortex that eventually result in durable changes in synapses underlying remote memory. Computational models that have both rapid onset and fast decaying connections (like the rapid learning of new information by the hippocampal system) and gradual, interleaved, durable changes in connections (like the establishment of remote memories in the neocortex) that support one another can account for these learning systems.

Plaut (1996) also argued that connectionist models can not only simulate learning of new information and specific cognitive tasks (reading, naming), but can also reproduce some of the patterns of errors produced by neurologically impaired individuals when the network is ‘damaged’ (e.g., by adding noise, reducing connection weights, taking out a subset of connections, etc.). He went on to propose that connectionist models could be utilized in rehabilitation, and might provide a theory of rehabilitation. To illustrate, he reported that a network of his design learned a set of artificial associations more quickly when it was trained with atypical exemplars of a category than when it was trained with typical exemplars of the category.

AU: Kiran and Thompson 2001 – 2002 in references, which date

This prediction was empirically tested by [Kiran](#) and [Thompson](#) (2001), who showed that four aphasic patients with impaired lexical-semantics learned to name pictures more reliably when atypical exemplars were trained (e.g., ‘penguin’ as an exemplar of birds) than when typical exemplars (e.g., robin) were trained via semantic features. One interesting observation made by the authors was that one patient actually produced more semantic errors (e.g., robin → cardinal) after training, suggesting the possibility that the patient simply learned to say a category of names in response to that category of pictures (e.g., any bird name in response to any picture of a bird), and by chance, sometimes selected

the correct one. Furthermore, the difference in learning with atypical versus typical stimuli might have been due to the greater range of stimuli or semantic features employed when atypical stimuli were used than when typical stimuli were used. That is, there are more differences in features between penguins and chickens than between robins and sparrows. It is possible that utilizing equally similar atypical members (e.g. chickens and ducks and geese) and typical members would have yielded similar results.

Nevertheless, in this case it would appear that a connectionist model was useful in planning therapy. But the model did not really provide hints about how or when to treat, only a bit more detail on what to treat (atypical exemplars, or broader range of stimuli). However, another example described by McClelland and co-workers (McCandliss *et al.* 2002), indicates the possibility that such networks can shed light on how to treat as well.

McCandliss and colleagues (2002) described a connectionist model of learning to differentiate two very close phonemes, such as /r/ and /l/. They found that if the network was presented only with stimuli that are so close together that the stimuli elicit the same response, then the network could not learn the distinction. However, when the difference between stimuli was exaggerated, it could learn the distinction. Each error was followed by making the stimuli more distinct. After eight correct responses in a row, the distinction was made slightly more difficult. Finally, the network could make the difficult discrimination that it could not learn initially.

McCandliss and co-workers then tested this training model by teaching Japanese Americans to distinguish /r/ and /l/. The training was very intense (daily practice), comparable to the massive stimulation suggested by Hebbian learning. Results were mixed: the adults could learn just as the network had, beginning with exaggerated /r/ versus /l/ and progressing to incrementally harder distinctions (or regressing to easier ones with each error). And, as predicted, they did not learn the distinction when just intensively presented with the difficult (realistic) distinction. But what was unpredicted was that learning was considerably faster when the subjects were provided with feedback. In fact, they could learn the distinction just as quickly when only the difficult distinction was presented in intensive practice, as long as correct responses were followed by feedback.

Thus, connectionism provides some important insights into how people learn. It suggests that training should be intense (many repetitions of the stimuli are required), and nearly error free (to prevent incorrect connections from being strengthened). But it must allow enough errors or be sufficiently challenging to increase new connections.

The question remains: has connectionism provided the link between cognitive theory and the issue of when and how to treat? One key aspect of learning /r/ versus /l/ that was not predicted by the model was that learning was much faster once feedback was provided to the students. Furthermore, in the presence of feedback, learning was no faster when the exaggerated distinctions were presented first, compared to when the normal /l/ and /r/ sounds were presented first. This aspect of learning was not captured in the connectionist model; networks don't need feedback. But people do. McCandliss and colleagues report that the network could be modified to respond to feedback; nevertheless, the original network did not require it. Clearly, in some ways, people do not learn like computers do. We turn to crucial biological variables that distinguish network versus human learning, which may be the key to improving cognitive rehabilitation.

### **Biological theories: how computers are different from brains**

**Spatial organization** In computers, information is stored or represented in nodes and connections between nodes. There is need for the nodes to be organized such that nodes representing two adjacent spaces, or two tones along a continuum, are spatially adjacent in the network. In contrast, there is compelling evidence that there is spatial organization of information in the brain (see Kandel *et al.* 1995,

for review). There is retinotopic organization of visual cortex, such that two adjacent areas in the visual scene (or even an imagined scene) are represented in adjacent areas of cortex. Likewise, two similar tones are represented in auditory cortex closer to one another than are two dissimilar tones. That is, there is cochleotopic organization of human auditory cortex, although a network representing different tones would not need to have such localization. There is similar organization of the sensory cortex and motor cortex, representing the human body more or less in its physical shape. In fact, this localization of information in the brain may extend to higher levels of representation, such as multimodality representations of concepts, such fruits or animals. At least, we know that focal brain damage can selectively impair access to semantic representations in one category (e.g., animals) but not in others (e.g., Goodglass and Budin 1988; Hart *et al.* 1985; Hillis and Caramazza 1991; McCarthy and Warrington 1988). There have been attempts to account for such selective impairments without assuming any categorical organization of the brain, but these accounts do not satisfactorily explain highly selective deficits (Caramazza and Shelton 1998). In any case, it is clear that the brain has much more localization of function than is 'necessary' according to connectionist theory. It is also clear that the localization is modifiable, either by lesions that disrupt the cortical map or by change in input to the cortical map, or by experience – intense practice, rehabilitation, etc. (see Jenkins and Merzenich 1987; Merzenich *et al.* 1983; Xerri *et al.* 1996).

Au: – no entry in references

**Rate or intensity of stimulation** Another feature of brains that is not captured by connectionist models is the fact that human learning depends on the rate of stimulus presentation. Recent studies at the cellular level have provided evidence that learning occurs in the brain (as in computers) by change in the strength of connections between neurons. In the brain, this change in connection strength depends on rapid firing of a neuron that elicits firing of a connected neuron. This rapid string of effective firing of a neuron by another results in a change in the 'threshold' of the connection between these two neurons, such that second neuron will fire more easily in response to the first neuron than it did before. This process, called long-term potentiation (or long-term depression, when the connection is inhibitory), is thought to be the basis for human learning. Connectionist models that entail Hebbian learning capture many aspects of this 'learning'. That is, connection strength is changed by massive presentation of a stimulus that causes one node to activate a connected node. However, in most connectionist models, the rate of presentation or 'activation' is not important (although networks with rapid decay, as well as slow decay, connections could simulate a requirement for rapid presentation). In the animal (human or non-human) brain, the rate of presentation is an essential feature in determining long-term potentiation (LTP) or long-term depression (LTD). The results from brain experiments suggest that frequent, as well as massive, practice would be most effective in learning.

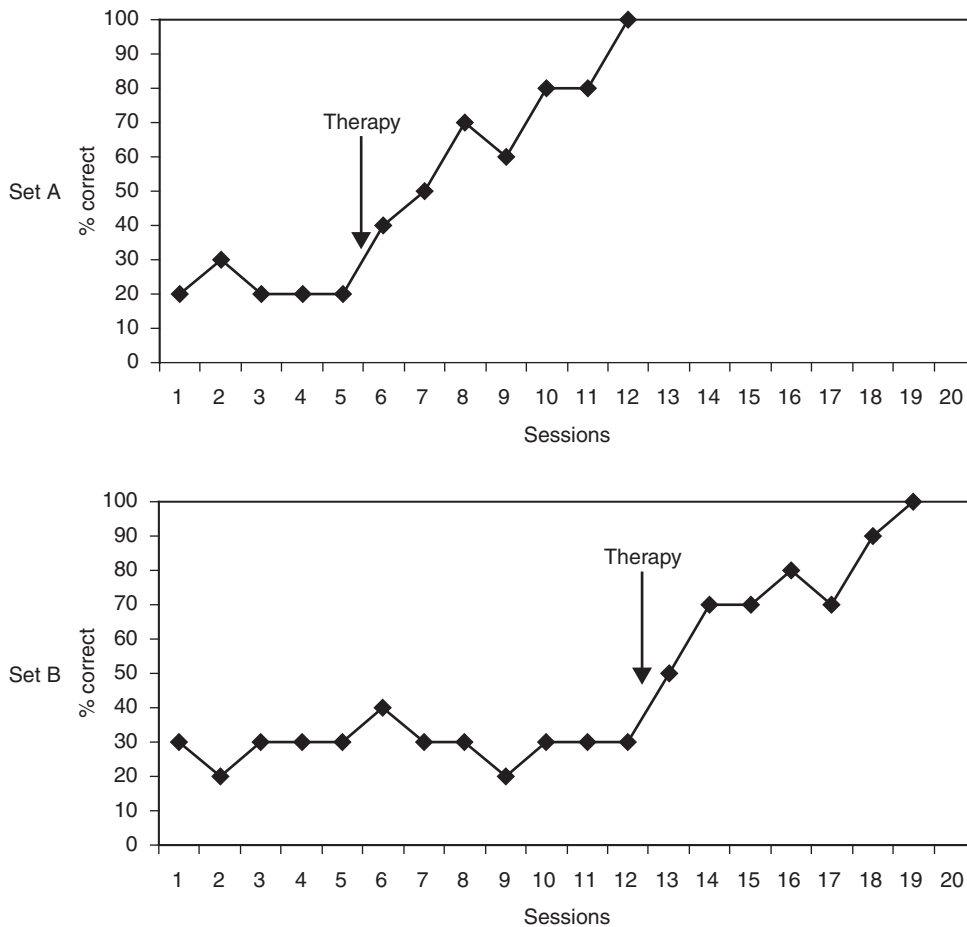
This principle of intense practice is borne out in our every day experiences in learning. We do not expect a child to become a proficient violinist by playing the violin once a week, even if he or she does it for many years. Rather, lengthy, daily practice results in the most rapid and effective learning of the violin. Likewise, improvements in a sport are often observed most readily after a week or so of intense practice (e.g. spring training), rather than a few months of weekly or biweekly practice. As a final example, learning of a language occurs most rapidly and effectively when there is 'language immersion' (e.g., moving to a foreign country) rather than through taking classes a few times per week.

Rehabilitation of language, cognitive, or motor function probably also depends on the same principles. That is, there is evidence that relearning of a function, such as access to spoken word form representations for word retrieval, occurs most rapidly (and perhaps only) when training occurs daily, rather than a few times per week. To illustrate, a young woman, H.G., with severe impairment in accessing both semantic representations and spoken word form representations for speech, following severe left frontotemporal and parietal brain damage, was trained in oral naming of fabrics (so that she could attain her goal of working in a fabric shop). Training involved eliciting the correct name in

276 | For a theory of cognitive rehabilitation

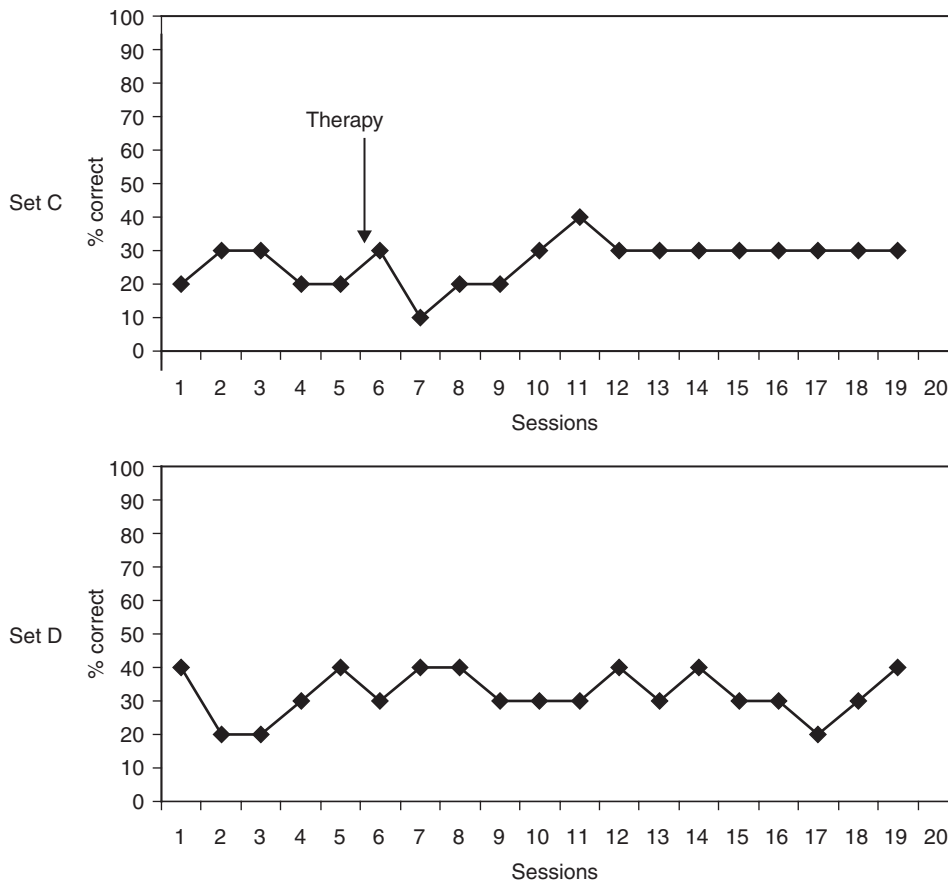
response to a fabric, through a series of cues, designed to gradually foster more independent responses. This therapy was effective in improving her naming when it was provided in daily, two-hour sessions, but not when the same therapy was provided in twice per week two hour sessions, as shown in Figures 22.1 and 22.2 (described in Hillis 1998). H.G. rapidly improved in naming Sets A and B when treatment was initiated five days per week (Figure 22.1). However, she failed to learn Sets C and D, which were matched in word frequency and length to Sets A and B, when the same treatment was applied just twice per week (Figure 22.2).

**Biochemical milieu in the brain** Also unlike computers, change in connection strength in the brain depends on chemical milieu. For example, LTD and LTP depend on the presence of particular neurotransmitters, which can be influenced by medications, reinforcement, motivation and other emotional states. Recent studies have shown that the presence of norepinephrine and acetylcholine together are essential for synaptic plasticity (LTD) (Kirkwood *et al.* 1999). Norepinephrine is released in states of excitement or challenge, and in response to rewards. The essential role of norepinephrine can account



**Figure 22.1**

Au: Pls  
provide  
legend



**Figure 22.2**

Au: Pls provide legend

for the remarkable effect of feedback during learning to discriminate /r/ versus /l/, as reported by McCandliss, McClelland and colleagues. This sort of feedback is not essential for learning in a connectionist network. Norepinephrine is also increased in response to certain medications, such as amphetamines. Recent studies, including small, double-blind randomized placebo-controlled trials, have demonstrated that administration of amphetamines along with language therapy resulted in more improvement in language in aphasic patients recovering from stroke, compared to language therapy alone (Walker-Batson 2000; Walker-Batson *et al.* 1991, 2001). Similar results have been reported for rehabilitation of motor skills after stroke (Walker-Batson *et al.* 1995).

The essential role of acetylcholine in synaptic plasticity is also borne out clinically. Neurons whose firing depends on acetylcholine degenerate in Alzheimer's disease (AD), a disease that severely affects memory and learning. Acetylcholine esterase inhibitors, which increase the amount of acetylcholine available to neurons, have been shown to improve cognition, or at least temporarily slow the rate of decline of cognition, in AD. These medications also may have a positive effect on recovery during language rehabilitation (Berthier *et al.* 2003).

Another medication that has been shown, at least in some studies (Albert *et al.* 1988; but see MacLennan *et al.* 1991), to improve outcome during language rehabilitation is bromocriptine,

Au: Walker-Batson et al 2001 – no entry in references  
Walker-Batson et al 1995 – no entry in references

a dopamine agonist. Dopamine is a neurotransmitter that has an important role in the experience of reward. Serotonin, a neurotransmitter that has a role in the experience of punishment and emotional state, may also influence the effectiveness of rehabilitation. Medications that increase available serotonin (e.g., paroxetine) or norepinephrine (e.g., nortriptyline) are important in treatment of depression, and may improve outcome in rehabilitation (Gillen *et al.* 2001). It is not clear whether the positive influence on rehabilitation is solely through reduction of depression (and the concomitant increase in motivation), or whether these medications also directly affect synaptic plasticity and thus, learning. Wilson (1997) observed that feelings, emotional state, and social and behavioral consequences have important influences on rehabilitation; the influence of these key neurotransmitters on synaptic plasticity may explain why.

Medications can also have negative effects on crucial neurotransmitters and on rehabilitation outcome (see Goldstein 1995, for review). For example, reserpine (an older medication used to treat high blood pressure) reduces the release of norepinephrine and epinephrine, and can inhibit effectiveness of rehabilitation. Haloperidol and other neuroleptic medications that block dopamine receptors (having the opposite effect of bromocriptine), can negatively affect rehabilitation outcome as well (Small 2002).

There are other important differences between the brain and computers that are difficult to quantify. For example, concurrent activities of the brain, such as distraction with bodily needs or emotional needs, can greatly influence learning. Such influences may account for why two individuals with apparently the same deficits do not equally respond to the same therapy (Caramazza and Hillis 1993).

**Stages of recovery** Finally, the brain recovers from damage through various stages of recovery (Hillis and Heidler 2002). The first, acute stage during the first few days after onset stroke or brain injury, probably depends on tissue recovery, through restoration of blood flow to the brain, resolution of edema, restoration of membrane integrity, and normalization of ionic balance. For example, we (Hillis and Heidler 2002) found that all patients who had impaired word comprehension at Day 1 of stroke, and who recovered word comprehension by Day 3 of stroke, showed reperfusion (restored blood flow) to Wernicke's area by Day 3, as demonstrated by repeated magnetic resonance perfusion imaging. Patients who failed to recover word comprehension by Day 3 showed persistent hypoperfusion, or low blood flow, in Wernicke's area.

Intermediate recovery, starting days to weeks after stroke, and continuing for months or possibly years, involves variable degrees of reorganization of structure/function relationships. This reorganization can be shown by serial functional imaging (fMRI or PET scanning), which shows changes during recovery in areas of the brain that are 'activated' during a specific cognitive task (Cappa *et al.* 1997; Heiss *et al.* 1997, 1999; Karbe *et al.* 1995; Leff *et al.* 2002; Musso *et al.* 1999; Ohyama *et al.* 1996; Thiel *et al.* 2001; Thompson 2000; Thulborn *et al.* 1999; Warburton *et al.* 1999; Weiller *et al.* 1995; Weiller 2000). Reorganization has also been demonstrated through single cell recording experiments in monkeys, who show changes in cortical maps of motor and sensory function, demonstrated by individual neurons' responses to stimulation, after peripheral or central lesions (e.g., Merzenich *et al.* 1983). This reorganization depends on relearning through synaptic plasticity (LTP and LTD), as influenced by the availability of neurotransmitters during intense practice or stimulation. The availability of neurotransmitters, in turn, depends crucially on the correct chemical milieu, such as medications, and on the individual's mood and responses to reinforcements.

In the chronic stage (months to many years after brain lesion), recovery of cognitive skills may depend on learning new ways of carrying out the task. For example, patient S.J.D. (Hillis and Caramazza 1995) was unable to write verbs, although she could say verbs and could write both nouns and verbs at five years after her stroke involving the entire perisylvian language cortex. S.J.D. learned to write verbs by self-cuing the written form by converting the first sound of the oral form of the verb

Au:  
Goldstein  
1995 –  
1998 in  
references,  
which date  
is correct?

Au: – no  
entry in  
references

to a letter. This self-cuing strategy involved relearning of sound-to-print conversion rules. On the other hand, H.G. (Hillis 1998) was unable to *say* content words, although she had unimpaired articulation and other motor speech skills and was able to *write* content words spontaneously (the modality effect opposite to S.J.D.'s). She needed to relearn print-to-sound conversion rules in order to convert from the written to the spoken form of the content word. However, converting the first few letters did not result in effective self-cuing, analogous to S.J.D.'s self-cuing. Instead, she needed to relearn the pronunciation of each individual content word. To illustrate, H.G. pronounced the word *one* as 'own', and the abbreviation *Mrs.* as [mers]. These words had to be misspelled for her as the 'regularized' version of the word in order to elicit the correct pronunciation. That is, she had to learn that *one* is pronounced like *wun*, and *Mrs* as *missuz*. The relearning by both S.J.D. and H.G. would also have required new synapses, established through LTP and LTD, and would thus have been influenced by intensity of training and the neurochemical milieu.

It is very likely that different interventions are required for different stages of rehabilitation. That is, in the acute stage (the first few days after stroke or brain injury) medical and surgical interventions to restore blood flow (e.g., thrombolytics, revascularization, induced blood pressure elevation) in stroke, and perhaps neuroprotective medications in both stroke and traumatic brain injury, may be most effective in restoring tissue viability. In the intermediate stage, weeks after stroke, intense practice, perhaps combined with appropriate medications, may be best for inducing reorganization. In the chronic stage, behavioral approaches such as shaping and conditioning with appropriate rewards (+/- medications to increase the effectiveness of the reward system) may be most effective in teaching compensatory strategies and laying down new pathways to access previously established cognitive processes.

## Conclusions

We are just beginning to have the elements with which a theory of rehabilitation can be built. These elements include, at the very least:

1. A model of the cognitive processes underlying the task to be relearned, and of the processes required to relearn it (relying on advances in cognitive neuropsychology and computational science);
2. A hypothesis about the network of neural structures and mechanisms underlying these cognitive processes (relying on advances in neuroimaging, single cell recording in primates, and other components of cognitive neuroscience);
3. A proposal about how change in these neural elements, such as synaptic plasticity, takes place (relying on recent advances in basic neuroscience), and
4. An informed theory of how neuroplasticity is influenced by medications and reinforcements (relying on recent advances in neuroscience, integrated with clinical neurology, psychiatry, and psychology).

It is likely that the actual strategies of therapy that arise from such a theory will not differ substantially from those that have been discovered empirically in behavioral and cognitive approaches, although the intensity of treatment and combined use of cognitive rehabilitation and pharmacotherapy might be ramped up. Moreover, such a theory would shed light on why and how and when rehabilitation is most effective, and would thus allow clinicians to apply their strategies in the most effective times and frequencies, and to the most appropriate individuals.

