CRITICAL REVIEW

What is cognitive reserve? Theory and research application of the reserve concept

YAAKOV STERN
Cognitive Neuroscience Division, G.H. Sergievsky Center, The Taub Institute, and Departments of Neurology, Psychiatry, and Psychology, Columbia University College of Physicians and Surgeons

(Received August 22, 2000; Revised February 26, 2001; Accepted February 28, 2001)

Abstract

The idea of reserve against brain damage stems from the repeated observation that there does not appear to be a direct relationship between the degree of brain pathology or brain damage and the clinical manifestation of that damage. This paper attempts to develop a coherent theoretical account of reserve. One convenient subdivision of reserve models revolves around whether they envision reserve as a passive process, such as in brain reserve or threshold, or see the brain as actively attempting to cope with or compensate for pathology, as in cognitive reserve. Cognitive reserve may be based on more efficient utilization of brain networks or of enhanced ability to recruit alternate brain networks as needed. A distinction is suggested between reserve, the ability to optimize or maximize normal performance, and compensation, an attempt to maximize performance in the face of brain damage by using brain structures or networks not engaged when the brain is not damaged. Epidemiologic and imaging data that help to develop and support the concept of reserve are presented. (JINS, 2002, 8, 448-460.)

Keywords: Functional imaging, Alzheimer’s disease, Compensation, Brain damage, Epidemiology

INTRODUCTION

The idea of reserve against brain damage stems from the repeated observation that there does not appear to be a direct relationship between the degree of brain pathology or brain damage and the clinical manifestation of that damage. For example, Katzman et al. (1989) described 10 cases of cognitively normal elders who were discovered to have advanced Alzheimer’s disease (AD) pathology in their brains at death. They speculated these women did not express the clinical features of AD because their brains were larger than average. Similarly, most clinicians are aware of the fact that a stroke of a given magnitude can produce profound impairment in one patient and while having minimal effect on another. Something must account for the disjunction between the degree of brain damage and its outcome, and the concept of reserve has been proposed to serve this purpose.

There have been many attempts to produce a coherent theoretical account of reserve. This paper will attempt to review and synthesize concepts that have been suggested, such as threshold, compensation, neuronal or brain reserve, and cognitive reserve. Many of these terms have been used interchangeably by previous authors, and they have not had well-accepted definitions. Specific definitions will be offered for these concepts that attempt to capture potential theoretical distinctions between them.

The concept of reserve should be relevant to any situation where the brain sustains injury. In addition, it will be argued that the concept of reserve should be extended to encompass variation in healthy individuals’ performance, particularly when they must perform at their maximum capacity. Nevertheless, many of the concrete examples will be framed around AD, with the implicit assumption that the discussion has implications for brain damage in general. AD has some unique advantages for examining disease-induced changes in brain function. AD pathology affects cortical circuitry that subserves a wide range of cognitive functions, and its pathology is more likely than conditions...
such as stroke to affect similar anatomic sites across subjects, allowing better generalization. AD is also slowly but inexorably progressive, providing a more sensitive indicator of the severity of brain insult required before cognitive networks change. On the other hand, the potential for adaptation of recovery might vary between slowly progressive and acute pathologies, so studies of AD may not always have direct implications for studies of other conditions.

Finally, it should be noted that this paper is not intended to be a comprehensive review of all of the literature relevant to the concept of reserve. For example, there is a large body of work investigating the concept of reserve in the context of HIV-related cognitive functioning that will not be addressed here (Basso & Bornstein, 2000; Pereda et al., 2000; Satz et al., 1993; Starace et al., 1998; Stern et al., 1996). Rather, work has been selected that helps to develop and support the ideas that will be presented.

**DEFINING RESERVE**

One convenient, although not entirely accurate, subdivision of reserve models revolves around whether they envision reserve as a passive process, or see the brain as actively attempting to cope with or compensate for pathology. In passive models, reserve is defined in terms of the amount of damage that can be sustained before reaching a threshold for clinical expression. In the active models, reserve revolves around differences in how the task is processed. These two approaches are not mutually exclusive. Ultimately, some combination of these two approaches might best describe the empirical observations that have prompted us to develop the concept of reserve.

**Passive Models: Brain Reserve or Threshold**

Many investigators have proposed passive models including Katzman (1993; *brain reserve*) and Mortimer et al. (1981; *neuronal reserve*). This type of model has also long been implicitly adopted by most clinicians. The threshold model, critically reviewed by Satz (1993), is one of the best articulated passive models. The threshold model revolves around the construct of *brain reserve capacity* (BRC). While BRC is a hypothetical construct, concrete examples of BRC might include brain size or synapse count. The model recognizes that there are individual differences in BRC. It also presupposes that there is a critical threshold of BRC. Once BRC is depleted past this threshold, specific clinical or functional deficits emerge.

This formulation, illustrated in Figure 1 (derived from Satz, 1993b), is sufficient to account for many clinical observations. Assume that two patients have two different amounts of BRC. A lesion of a particular size might result in a clinical deficit in a person with less BRC (Patient 2), because it exceeds the threshold of brain damage sufficient to produce that deficit. However, an individual with greater BRC (Patient 1) could remain unaffected, because this threshold is not exceeded. Thus, more BRC can be considered protective factor, while less BRC would impart vulnerability. An apparently intact individual with pre-existing brain damage can tolerate less new brain damage than another individual without this underlying pathology: the pre-existing damage reduces the amount of remaining BRC, so the new lesion is sufficient to exceed the functional impairment cutoff.

Many observations about the prevalence and incidence of AD are consistent with the threshold model. Figure 2 (based on Katzman, 1993) illustrates that the progression of AD pathology and the clinical expression of AD can be discontinuous. AD pathology probably begins to develop many years before the disease is expressed clinically and...
slowly becomes more severe. The threshold model would assume that when synapses are depleted beyond some critical point the initial symptoms of dementia will appear. At some point after this, depletion will result in symptoms of sufficient severity to allow the diagnosis of AD. The arrows surrounding the point in the figure where clinical symptoms appear denote the fact that there are individual differences in reserve capacity, and these differences result in later or earlier expression of clinical symptoms. In patients with more reserve, synapse loss must be more severe before clinical symptoms appear and the symptoms appear later. Conversely, symptoms would appear earlier in a patient with less reserve.

The threshold approach can be extended to account for more than just differences in the onset of a clinical outcome. Because reserve mediates between the pathology and its clinical outcome, the level of reserve should also influence the severity of clinical symptoms after the threshold for their appearance has occurred. Here, at any level of disease pathology, patients with more reserve evidence more mild levels of clinical severity.

Active Models

The active models of reserve suggest that the brain actively attempts to compensate for brain damage. I suggest that in its active form, there can be at least two types of reserve. The first is cognitive reserve. This could take the form of using brain networks or cognitive paradigms that are less susceptible to disruption. I propose that this type of reserve is a normal process used by healthy individuals when coping with task demands. The second is compensation: using brain structures or networks not normally used by individuals with intact brains in order to compensate for brain damage.

Cognitive Reserve

Cognitive reserve parallels the concept of brain reserve in that it is a potential mechanism for coping with brain damage. In the threshold model, the reserve capacity typically consists of additional synapses or an increased number of redundant neuronal networks. Cognitive reserve focuses more on the “software.” This could consist of the ability of the cognitive paradigm underlying a task to sustain disruption and still operate effectively. Alternately, this could consist of the ability to use alternate paradigms to approach a problem when the more standard approach is no longer operational.

The concept of cognitive reserve provides a ready explanation for why many studies have demonstrated that higher...
levels of intelligence, and of educational and occupational attainment are good predictors of which individuals can sustain greater brain damage before demonstrating functional deficit. Rather than positing that these individuals’ brains are grossly anatomically different than those with less reserve (e.g., they have more synapses), the cognitive reserve hypothesis posits that they process tasks in a more efficient manner.

The concept of cognitive reserve also differs from the passive threshold approach in other important ways. Recall that in the passive model, individuals may have different levels of BRC and a lesion of the same size is sufficient to deplete BRC below the critical threshold in some individuals but not others (see Figure 1). The cognitive reserve model is illustrated in Figure 4. Here the 2 patients have the same amount of BRC (again, let’s say, the same number of synapses). However, Patient 1 has more cognitive reserve than Patient 2, in that Patient 1 uses more efficient processing mechanisms. As a result, Patient 1 can tolerate a larger lesion than Patient 2 before functional impairment is apparent. Thus, the cognitive reserve model does not assume that there is some fixed cut-off or threshold at which functional impairment will occur. The critical threshold differs from one person to the next, depending on how efficient or resilient the “software” is in using the remaining neural substrate. Putting it another way, the threshold approach supposes that the person with more BRC has more to lose before they reach some clinical cut-point. The cognitive reserve hypothesis focuses less on what is lost and more on what is left. In the case of AD, one individual may begin to express clinical features when synapses are depleted to a particular number, while an individual with more cognitive reserve may be able to operate effectively with the same number of synapses.

From a strict point of view, the differences in cognitive processing envisioned by the cognitive reserve model must also have a physiologic basis, in that the brain must ultimately mediate all cognitive function. The difference is in terms of the level of analysis. Presumably, the physiologic variability subsumed by cognitive reserve is at the level of variability in synaptic organization, or in relative utilization of specific brain regions. Thus cognitive reserve implies anatomic variability at the level of brain networks, while brain reserve implies differences in the quantity of available neural substrate.

The cognitive reserve model also does not assume that a specific type of brain damage will have the same effect in each person. Because of individual variability in how they cope with brain damage, the same amount of damage will have different effects on different people, even if BRC is held constant.

A proposed definition of cognitive reserve is: the ability to optimize or maximize performance through differential recruitment of brain networks, which perhaps reflect the use of alternate cognitive strategies. Since the changes in brain recruitment associated with reserve are a normal response to increased task demands, this definition suggests that cognitive reserve is present in both healthy individuals and those with brain damage, and is reflected in the modulation of the same brain networks. In essence, an individual who uses a brain network more efficiently, or is more capable of calling up alternate brain networks or cognitive strategies in response to increased demand may have more cognitive reserve. The definition encompasses two possibilities, differences in recruitment of the same network, and differential ability to recruit alternate networks. These possibilities will be discussed in turn. This discussion is extremely speculative, although some evidence to support these speculative lines will be reviewed.

More efficient use of brain networks

This idea is based on studies of how normal individuals respond as tasks are made increasingly difficult and of individual differences in task performance. The rough parallel here is that, in effect, brain damage acts to increase task difficulty. Several functional imaging studies suggest that a common response to increasing task difficulty in normal individuals is increased activation of areas involved in an easier version of the task and/or the recruitment of additional brain areas (Grady et al., 1996; Grasby et al., 1994; Gur et al., 1988; Rypma et al., 1999). There are also individual differences in how this additional recruitment occurs. For any level of task difficulty, more skilled individuals typically show less task-related recruitment than less skilled individuals. When the more skilled individual exerts herself maximally, she can perform better than the less skilled one. This increased efficiency and larger dynamic range can be considered reserve. If brain damage is considered a form of demand (similar to increased task difficulty), then a person with more cognitive reserve would be able to cope with more brain damage and still maintain effective functioning. Often, increased task difficulty results in the recruitment of additional brain areas or networks. We might

Fig. 4. The cognitive reserve model. Two patients have the same amount of brain reserve. However, Patient 1 has more cognitive reserve than Patient 2, in that Patient 1 uses more efficient processing mechanisms. As a result, Patient 1 can tolerate a larger lesion than Patient 2 before functional impairment is apparent.
speculate that, in a person with more reserve, this additional recruitment would occur at a higher difficulty level.

**Differential ability to recruit alternate networks**

This possibility is more speculative, but is consistent with the concept of cognitive reserve. The point is simply that a person with more reserve might be able to call on a larger array of alternate networks for solving the problem at hand. As a concrete example, a trained mathematician might be able to solve a mathematics problem many different ways, while a less experienced individual might have only one possible solution strategy available. The mathematician would have more flexibility in solving the problem if any particular solution strategy was precluded. This built in redundancy would permit greater resilience in the face of brain damage.

These two ideas about reserve might form a heuristic framework for designing studies about cognitive reserve. Studies can be aimed at behavior in unimpaired individuals, taking advantage of inherent inter-individual differences in skills or intelligence. Predictions can then be made about how these individual differences might affect response to brain damage.

**Compensation**

The term cognitive reserve can be limited to the variability seen in non-brain damaged individuals, which distinguishes it from compensation, which might be reserved for a specific response to brain damage. This distinction emerges from the consideration of findings in functional imaging studies that compared task-related activation in impaired and unimpaired groups with a more critical eye. For example, several functional neuroimaging studies comparing task-related activation in AD patients and controls found more marked and extensive activation in the patients (Becker et al., 1996; Deutsch et al., 1993; Grady et al., 1993). These findings have been interpreted as evidence that the patients compensated for AD pathology. That is, since pathology impaired the patients’ ability to mediate the task through the same brain network used by the controls, the patients compensated by engaging alternate brain areas during task performance. One may ask whether this observed “compensation,” as the investigators called it, is the same thing as cognitive reserve. If compensation truly represents a change that is induced by brain damage, then it might be important to distinguish between compensation and cognitive reserve. This distinction has not been commonly used in the reserve literature. However, recent functional imaging studies are often carefully designed to ensure that observed group differences are not simply a function of task difficulty.

Compensation is thus first defined in the negative, in that it cannot be simply a normal response to difficulty. In addition, the term compensation implies an attempt to maximize performance in the face of brain damage by using brain structures or networks not engaged when the brain is not damaged.

One of the theoretical reasons why discriminating between reserve and compensation might be important is that it helps critically evaluate the results of studies comparing impaired and unimpaired populations. For example, one study (Becker et al., 1996) reported a comparison of PET rCBF in AD patients and elderly controls performing a verbal list-learning task. Three list-lengths, of one, three, and eight were used. In the eight-word task (compared to the three-word task), patients showed decreased activation of the lateral frontal cortex relative to controls. However, dorsolateral prefrontal cortex and areas surrounding the angular gyrus were more active than in the controls. The authors suggested that this may represent a response to neuropathologic changes that is specific to AD patients. In the proposed classification scheme, this would be considered compensation. Later, using another analytic technique, the same authors concluded that both patients and controls used the same underlying network to mediate the memory task (Herbster et al., 1996). They reported that observed group differences they originally reported were a result of differential activation of this network, probably because the task was more difficult for the AD patients. In the proposed classification scheme, this would be considered cognitive reserve. In assessing the brain’s response to damage, it clearly will be important to know which responses are within the range of normal behavior and which only occur in the presence of pathology.

Disentangling compensation and reserve presents a specific experimental design problem. Studies must use tasks that allow for systematic manipulation of task difficulty. Ideally, task difficulty should be equated across individuals, not just subject groups. Once task difficulty is equated, group differences in patterns and levels of functional activation are more likely to represent compensation, and not reserve.

Stern et al. (2000) tried to determine whether the pathology of Alzheimer’s disease (AD) alters the brain networks subserving performance on a memory task, while carefully controlling for task difficulty. H$_2$[O]$^15$PET was used to measure regional cerebral blood flow in patients and healthy elders during the performance of a verbal recognition task. Task difficulty was matched across participants by adjusting the size of the list that each subject had to remember such that each subject’s recognition accuracy was 75%. In the healthy elders, a network of brain areas involving left anterior cingulate, anterior insula, and left basal ganglia was activated during task performance. Higher study list size (SLS) was associated with increased recruitment of this network, indicating that this network was associated with task performance and that subjects who could recruit the network to a greater degree could perform the task better. Only 3 AD patients also expressed this network in a similar manner. This network used by the controls and a minority of AD patients may mediate reserve, in that it...
appears to be recruited to cope with the demands presented by the activation task, and differential recruitment of the network is directly related to the ability to perform the task. Individuals who are able to activate this network to a greater degree may have more reserve against brain damage.

The remaining 11 AD patients recruited a different network during task performance, consisting of left posterior temporal cortex, calcarine cortex, posterior cingulate, and the vermis. Again, in these patients, higher SLS was associated with increased activation of this network. Thus the majority of AD patients did not use the same network as controls to mediate task performance, but rather used an alternate network. The healthy elders also expressed this network during task performance, but it did not mediate their performance of the task, as indicated by the lack of correlation between their expression of this network and their SLS.

Whether the patients’ use of the alternate network represents compensation is a matter of definition. The alternate network was used by patients in the place of the normal network in an attempt to mediate task demands, suggesting that its use was compensatory. However, this network was also activated by the healthy elders, indicating that it was not unique to the patients. If the term compensation is reserved for the use of a novel network that emerges in response to pathology, then the alternate network does not meet this criterion. On the other hand, the role played by the alternate network differed in patients and controls, in that it appeared to be mediating the ability to achieve larger study list sizes in the patients but not the healthy elders. This novel use of the network may arise out of the inability to use the standard network, and thus may be considered compensation. Studies such as this one may lead to better understanding of the brain mechanisms underlying compensation and reserve.

Similar considerations of compensation and reserve apply when examining other issues, such as recovery from stroke. Some aspects of recovery simply rely on resolution of factors such as edema. More interesting from our perspective is recovery of function that results from alterations in the brain networks that underlie specific behavior. Thus it is important to demonstrate that the patient is really using a novel brain network, as opposed to a degraded version of the “normal” network.

**RESEARCH APPROACHES TO RESERVE**

Research investigating reserve must focus on three components: brain damage, clinical expression of the brain damage, and the theoretical mediation of reserve. Of course, the key question to be answered is what actually mediates the relationship between brain damage and its clinical outcome. Most typically, study designs attempt to establish some operational definition for at least two of the three components. By measuring these two components they attempt to make inferences about the third. This theme will be developed further in the discussion below, which reviews some operational definitions that have been used for each component of the model, and describes some representative studies that have used these definitions.

**Indices of Pathology**

The problem of assessing the nature and degree of brain damage is a familiar one in neuropsychological studies. The optimal measure of brain damage would be some anatomic index of that damage. In some cases, these indices can be relatively direct, even without resorting to post mortem studies. In stroke, for example, direct measures of the volume of stroke, in combination with some consideration of their location, might be useful. Most often, we must resort to proxy measures of pathology. In studies of head trauma, for example, we have no direct measure of neuronal damage, but clinical indices have been established that appear to capture that severity to some degree. These include measures of the severity of the head trauma itself or measures of its sequela, including the duration of loss of consciousness. In AD, there is no existing direct measure of pathologic severity. In addition, we run a danger of confusing measures of outcome with measures of pathologic severity; because they are outcomes, clinical measures of disease severity, such as mental status or activities of daily living scales cannot be used to estimate the pathologic burden. An optimal approach to this problem is clinicopathologic studies, where postmortem indices of pathology, such as synapse loss, or amyloid burden, are related to aspects of the clinical presentation observed during life. Several studies of reserve have used this challenging approach. In particular, Snowdon et al. have demonstrated a relation between measures of linguistic ability acquired at an early age, and the presence of AD pathology noted post mortem (Snowdon et al., 1996).

An alternate approach is to develop a proxy measure for pathology that can be used during life. Some studies of AD have used the characteristic reduction in parietotemporal and frontal perfusion and metabolism seen at rest in AD (Prohovnik et al., 1988) as an index of the severity of AD pathology. The perfusion deficit correlates with disease severity and increases with disease progression (Foster et al., 1984) and its distribution overlaps with the cortical areas with the greatest density of histopathological abnormalities (Brun & Englund, 1981; Pearson et al., 1985; Rogers & Morrison, 1985). Because this perfusion pattern is not unique to AD (Schapiro et al., 1993), the degree of CBF deficit might be used as a marker of the severity of AD pathology, but not necessarily as a diagnostic indicator.

One functional imaging study found that, in patients matched for overall severity of dementia, the parietotemporal flow deficit was greater in those with more years of education (Stern et al., 1992). This observation was confirmed in a later PET study (Alexander et al., 1997). After controlling for clinical dementia severity, higher education was correlated with reduced cerebral metabolism in prefrontal, premotor, and left superior parietal association areas.
These studies support the idea that although pathology was more advanced in patients with higher education, the clinical manifestations of the disease were comparable to those in patients with lower education and less pathology. Presumably the patients with more education had more cognitive reserve.

Because of the general difficulty in ascertaining or quantifying pathology, many research designs do not attempt to do so. In general, if a research design specifies two of the three components, the third can be inferred. Thus often studies measure clinical outcomes while specifying some index of reserve, and then attempt to make inferences about underlying pathology. One such study (Stern et al., 1995a) matched AD patients for clinical severity and followed them prospectively. AD patients with greater education or occupational attainment died sooner than those with less attainment. Here, the proxy for reserve is the level of educational and occupational attainment. The outcome is death, which is more likely when AD pathology is more advanced. The observed relation between the level of reserve and mortality implies that at any level of assessed clinical severity, the underlying pathology of AD is more advanced in patients with more educational or occupational attainment. This would result in shorter duration of diagnosed disease before death. A recent study did not replicate this finding (Geerlings et al., 1997), but a follow-up study by the same group, using patients with more advanced dementia, did (Geerlings et al., 1999).

Outcome Measures

One of the key outcomes in the study of reserve is the presence or absence of some clinical entity. For example in AD, many studies have attempted to determine whether there is a relation between some measure of reserve, such as education, and the prevalence or incidence of AD. Many studies have observed higher prevalence of AD in individuals with lower education (Bonaiuto et al., 1990; Callahan et al., 1996; The Canadian Study of Health and Aging, 1994; Glatt et al., 1996; Gurland et al., 1995; Hill et al., 1993; Korczyn et al., 1991; Mortel et al., 1995; Ott et al., 1995; Prencipe et al., 1996; Sulkava et al., 1985; Zhang et al., 1990). Note that the assumption here is that since education is associated with reserve against the expression of AD pathology, AD should be less prevalent in individuals with higher education. These studies do not directly measure AD pathology, and assume that its prevalence is relatively equal across education groups.

A major weakness of prevalence studies for the study of reserve is the potential confounding of the determination of the outcome with the measure of reserve. For example, the diagnosis of dementia relies on the presence of cognitive deficit. Individuals with lower educational attainment may simply perform worse on the psychometric tests used to identify these deficits, while those with higher education perform better. It is possible to eliminate some of this diagnostic bias with incidence studies, where the outcome is the new diagnosis of dementia in a previously non-demented individual. If all subjects originally have the same diagnostic evaluation and are judged to be nondemented, the diagnosis of incident dementia at followup necessarily implies major decline from initial performance. This minimizes the chance of misdiagnosing a nondemented individual who could never have passed the diagnostic tests. Several groups have reported that the relative risk of incident dementia was increased in subjects with low education (Evans et al., 1993; Letenneur et al., 1994; Stern et al., 1994; White et al., 1994). Other prevalence (Beard et al., 1992; Fratiglioni et al., 1991) and incidence (Cobb et al., 1995) studies have not found an education effect for AD. Between-study differences are probably a function of differences in study samples. For example, in contrast to studies with positive findings, the Cobb et al. (1995) study had only a small percentage of subjects who did not complete grade school (8.1%), perhaps limiting power to observe an educational effect. Also, lower educational attainment was strongly confounded with increased age in that study.

The incidence studies suggest that higher prevalence of dementia in the low education/occupation group is not simply a result of detection bias, because all subjects pass the screens and are rated as nondemented at least once. Still the possibility of bias exists even in incidence studies. It may have been more difficult to detect new dementia in the high education and occupation groups because of lowered sensitivity of neuropsychological tests in these groups. The Stern et al. (1994) study tried two approaches to address this issue. First, subjects with "questionable dementia" were eliminated from analyses. This ensured a substantial change in performance for the diagnosis of dementia. They also evaluated the validity of the dementia diagnoses by investigating functional decline. Functional scores in the newly demented patients from both high and low education groups declined significantly from baseline values, while those in the nondemented groups remained stable. To the extent that the diagnosis of dementia corresponds to changes in performance that disrupt daily activities, the possibility of detection bias is minimized.

The epidemiologic studies above used a dichotomous outcome, the presence or absence of AD. Neuropsychologists excel in developing and applying measures that ascertain the effects of brain damage. Standard measures of cognition, both global and specific can be used as outcome measures in the study of reserve. A great advantage of continuous outcome measures is that they provide more statistical power than dichotomous variables.

The value of a continuous outcome measure is demonstrated by a study in which AD patients were matched for clinical severity and performance on a memory test (Stern et al., 1999). Patients with higher educational and occupational attainment showed more rapid decline in their memory functioning. A similar relationship between educational attainment and rate of decline in memory scores was also noted in another study (Teri et al., 1995). A cognitive-reserve-based explanation for these findings would be that, because patients with higher educational and occupational attain-
ment have more cognitive reserve, more pathology is required before memory begins to be affected. However, AD pathology progresses independently of educational and occupational attainment, and when pathology becomes very severe there is no longer a substrate for cognitive reserve to come into play. Thus, the severity of AD pathology at the initiation of memory deficit varies as a function of reserve, but the level of pathology associated with severe clinical dysfunction does not vary as a function of reserve. The result is a shorter time between the initiation of memory loss and severe memory disability in patients with higher educational and occupational attainment. Note that this interpretation relies on the assumption that AD pathology progresses independently of reserve. Thus, differential rates of clinical progression can provide insight into how reserve may mediate the relation between pathology and outcome. This approach does not rely a direct measure of pathology but rather on the progressive nature of the pathology.

In contrast to the findings in AD, several studies of normal aging have found more rapid cognitive decline in individuals with lower educational attainment (Albert et al., 1995; Butler et al., 1996; Farmer et al., 1995). Similarly, lower education has been associated with greater risk of functional decline in nondemented individuals (Snowdon et al., 1989). These findings, in healthy individuals, suggest that reserve is allowing them to cope more successfully with age-related changes.

It is also important to consider noncognitive outcomes that may be mediated by reserve. Among these are changes in day-to-day function, which are assessed by measures of basic and instrumental activities of daily living as well by other indices of function such as vocational measures. In addition, it is important to consider whether reserve may mediate the affective consequences of brain damage, such as depression. One epidemiologic study of AD is instructive in this regard (Geerlings et al., 2000). The authors found a reduced incidence of AD in individuals with higher versus lower educational attainment. However, the presence of depression was predictive of incident dementia only in the higher education group. They reasoned that cognitive reserve allowed individuals with more education to cope with AD pathology longer, thus delaying the cognitive symptoms of AD. However, the reserve was not successful in coping with another outcome of AD pathology, depression. Thus individuals with higher education were more likely to manifest an early depression as an early sign of the disease.

Measures of Reserve

The selection of research measures to represent reserve is dependent on the investigator’s theoretical concept of what reserve is. When selecting a proxy for reserve, one must keep in mind that its true mode of action may not be in accord with the theoretical reason for its selection.

For advocates of the idea of brain reserve, anatomic measures such as brain size, head circumference, synaptic count, or dendritic branching are effective measures of reserve. A direct test of the brain reserve hypothesis is to determine whether larger head or brain size is associated with reduced prevalence, risk, or severity of dementia. Several studies have shown such a relationship (Aksari & Stoppe, 1996; Graves et al., 1996; Mori et al., 1997; Schofield et al., 1997). Schofield et al. (1997) conducted a dementia prevalence study in 649 community-dwelling elders. Head circumference was measured in a standardized fashion. Analyses controlled for age, education, ethnicity, and height. Women in the lowest head circumference quintile were 2.9 times more likely to have AD than those in the upper four quintiles. Similarly, men in the lowest quintile were 2.3 times more likely to have AD. These findings suggest that individuals with larger brains may have more reserve against AD pathology.

Even in passive models such as threshold models, it is clear that there are demographic features that serve as proxies for reserve. These include measures of socioeconomic status, such as income or occupational attainment. Educational attainment has also been a widely used proxy for reserve, probably because it is relatively easy to ascertain. Finally, specific measured attributes have been used as indices of reserve, including literacy, IQ, and measures of specific cognitive functions.

Great care must be taken not to confound measures of reserve with measures of outcome. For example, many studies have evaluated the relation between IQ or education and the prevalence or incidence of AD. In these studies, as explained above, the possibility that education or IQ simply confounds the diagnosis of AD must be addressed. Kittner (1986) suggested adjusting for education when screening for dementia, in order to avoid ascertainment bias. In a dissenting view, Berkman (1986) suggested that we must remain open to the view that “educational level and/or socioeconomic behavior correlated with it may be genuine risk factors for senile dementia and are worthy of scientific exploration in their own right.” Simply controlling for education during assessment will not supply the solution to this question.

In 1981, Gurland wrote:

It is still an open matter whether there is an important sociocultural contribution to the prevalence of Alzheimer’s and other forms of dementia occurring in the senium, but the evidence now available is sufficiently intriguing to warrant further study of the issue.

Similar considerations were raised by other investigators (Mortimer, 1988). The reason that education or SES might serve as proxies for reserve is an important issue for research. Lower SES is associated with increased risk for toxic or environmental exposures, nutritional deficiencies, or perinatal insult (Katzman, 1993; Mortimer & Graves, 1993). Thus, some studies have noted an association between education and vascular or alcohol-related dementia but not AD (Cobb et al., 1995; Del Ser et al., 1997; Frattigioni et al., 1991), suggesting that education may simply be a proxy for other factors or exposures that mediate risk. Some
studies have controlled for conditions observable in adults, such as stroke (Stern et al., 1994) and still found relationships between education and dementia. Still, the potential interactions between proxies for reserve and clinical outcomes requires careful investigation. In particular, the important issue of the effect of perinatal or early childhood exposures will require prospective studies that follow children from before birth.

Education might also be a marker for innate intelligence, which may in turn be genetically based or a function of exposures. Some studies suggest that an estimate of IQ, or premorbid IQ might actually be a more powerful measure of reserve in some cases (Albert & Teresi, 1999; Alexander et al., 1997). One study had the unique opportunity to evaluate the relation between mental ability, as assessed in 1932 at age 11, and the incidence of dementia (Whalley et al., 2000). It found that mental ability scores were significantly lower in children who developed late-onset dementia when compared to those who remained nondemented.

Still, education, or other life experiences, may impart reserve over and above that obtained from innate intelligence. Studies have demonstrated separate or synergistic effects for higher educational and occupational attainment, suggesting that each of these life experiences contributed independently to increased reserve (Evans et al., 1993; Mortel et al., 1995; Rocca et al., 1990; Stern et al., 1994, 1995b). Occupational attainment also served as a measure of reserve in CBF studies of AD patients (Stern et al., 1995b). Thus several variables that are descriptive of life experiences might influence the cognitive reserve in the same way that education does. In addition to occupational attainment, these may include leisure activity (Bickel & Cooper, 2000; Fabrigoule et al., 1995; Hultsch et al., 1999; Kondo et al., 1994), and literacy (Manly et al., 1999). Also, recent analyses in our group suggest that, after controlling for educational level and measures of intelligence, Spanish-speaking individuals who learned to speak English have reduced risk of incident dementia when compared to those who never learned to speak English. This might suggest that the experience of acquiring another language imparts reserve.

ASSUMPTIONS OF BOTH ACTIVE AND PASSIVE MODELS: CAUTIONS FOR DESIGNING AND INTERPRETING RESEARCH

Research investigating reserve tends to treat the three components of the reserve model as if they are independent. This assumption is probably not true. As discussed above, many of our proxies for reserve are not independent of the outcome measures. This section reviews other areas of potential interaction between components of the reserve model.

A straightforward explication of either the threshold or cognitive reserve model assumes independence between the source of brain damage and reserve. For example, when exploring reserve in AD, it is convenient to begin by assuming that the progression of AD pathology is independent of the number of neurons a person has or their level of education. This simplifying assumption may not be valid, however. Some investigators (Friedland, 1993; Swaab et al., 1998) have suggested several mechanisms by which chronic neuronal activation stemming from educational exposure or other sources might actually be protective against the development of AD pathology. Also, the genetic determinants that influence reserve factors such as brain size, intelligence, or memory capacity might also influence the advent of the brain pathology. For example, APOE influences the timing of the onset of AD, but also has been related to the rate of change in memory capacity in nondemented elders (O’Hara et al., 1998). Alternately, one study found that linguistic ability evidenced in a writing sample produced at age 22 was predictive of later AD, controlling for education (Snowdon et al., 1996). This might suggest that lowered linguistic capacity, even at that early age, was due to existing pathology.

A related assumption is that the BRC remains relatively constant. For brain reserve, this simply means that there are a fixed number of neurons or synapses to lose. For cognitive reserve this means that the brain substrate remains stable, while the cognitive functions which that substrate mediates can vary. Again, this simplifying assumption may not be valid. It has been suggested that education may stimulate increased synaptic growth in the developing infant or child (Katzman, 1993), similar to that seen in animals reared in an enriched environment (Diamond, 1988). Increased dendritic branching has been noted in individuals with more education (Jacobs et al., 1993), although the causal direction of this relationship cannot be determined. Recently, it has been recognized that the adult brain is continuously generating new, functioning, neurons (Gould et al., 1999). Further, animal studies suggest that aspects of life experience, including enriched environment (Kemperman et al., 1997) and exercise (van Praag et al., 1999) can increase the amount of new neurons that are generated in mature animals. Thus some variable ostensibly associated with cognitive reserve, such as education, may dynamically influence the underlying neural substrate.

DIFFERENTIAL IMPLICATIONS OF ACTIVE AND PASSIVE MODELS

The active and passive models of reserve can often lead to the same predictions, so it is useful to consider the differential implications of the two types of models. As stated, the brain reserve, passive, models rely on actual anatomic differences to determine who has more or less reserve. Thus, there is no ready explanation in these models for why educational and occupational attainment, or IQ, should impart reserve other than to rely on the assumption that these experiences must modify brain anatomy in some way. The cognitive reserve hypothesis does not rely on gross differences in brain anatomy.
Some of the research findings described above are more easily explained using an active reserve model. For example, two functional imaging studies described above found that, in patients matched for overall severity of dementia, the parietotemporal flow deficit was greater in those with more years of education. These studies suggest that although pathology was more advanced in patients with higher education, the clinical manifestations of the disease were comparable to those in patients with lower education and less pathology. Given no gross difference in brain size, passive models do not have a ready explanation for why the patients with more education can tolerate more pathology. Some studies described above have also shown that AD patients with higher educational attainment have more rapid memory decline and die sooner. There is no ready explanation for this observation using a passive model. In fact, the explanation put forward above using an active reserve model relies on the assumption that “brain reserve” does not differ as a function of education, and essentially posits that there comes a point when pathology becomes so severe that there is no longer a substrate for cognitive reserve to come into play. Thus the active and passive models can produce different predictions and can differ in their ability to explain observed phenomena.

CONCLUSION

Cognitive reserve is a rich concept that has great heuristic value for research. While reserve is basically a simple concept, upon consideration there can be many layers of theoretical complexity. Consideration of the concept of reserve suggests that it cannot be considered as a unidimensional entity. While they overlap to some degree, the concepts of brain reserve and cognitive reserve may produce different predictions about the impact of brain pathology on function. Further, the differentiation of reserve and compensation may have practical utility particularly when attempting to study reserve using functional imaging.

Careful thought must go into translating theories about reserve into a research design. Also, we must resist the urge to invoke reserve reflexively to explain any anomalous result. Still, the concept of reserve does have great salience for the investigation of variability in individual performance and for understanding how the brain responds to challenge and pathology.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institutes on Aging (RO1 AG14671).

REFERENCES

Cognitive reserve


