DEMENTIA AS A MAJOR PUBLIC HEALTH CONCERN: INTELLIGENCE TESTING REVISITED

PATRICIA STOCKTON

Submitted for PhD

The London School of Economics and Political Science

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

In 1976 it was proposed that senile dementia, a potential affliction of old age, be redefined as Alzheimer's disease, a rare diagnosis previously assigned to presenile dementia occurring in middle life. In response to a "public culture" generated by those caring for the afflicted, together with leaders of the biomedical reserch community, substantial financing has been allocated by the U.S. Congress to the National Institutes of Health, for investigation of senile dementia redefined as a "dread disease". This has funded studies in the neurosciences, and a range of epidemiological and high technology diagnostic investigations for which psychiatry developed a "case-finding" method.

The "cognitive paradigm" for dementia was conceived by American psychiatry within a now dominant "biological" model which imputes physical causation to mental disorders and stresses "objectivity" in diagnosis. This has legitimated the use of "mental test" instruments based upon, or validated against, "intelligence tests" developed by psychologists for quantification of "intelligence" now redefined as "cognition".

In a study funded by the National Institute of Mental Health, three cognitive assessment instruments were administered to a sample of individuals aged 60-93 with a broad range of educational experience across the age spectrum. Education rather than age was found to be the most significant predictor of test results for each instrument, and when the tests were repeated a marked "learning effect" was detected among those with the least education and lowest baseline scores. However, the identification of low education as a predictor, albeit less powerful than age for "cognitive impairment" indicative of dementia in other investigations has now been interpreted as a "risk factor" rather than a confounding variable and now enters into genetic mental testing models.

Negative stereotyping of "old age", strongly associated with images of "senility", and "burden of ageing" economic arguments have therefore been reinforced by the dissemination of prevalence estimates from epidemiological studies conducted in communities in which there is an inverse correlation between age and education. In the meantime, basic scientists have failed to discriminate precisely between neuropathological changes indicative of "disease" and those of "normal ageing" or to establish a functional link between such changes and dementia behaviour *in vivo.* In consequence the legitimating rationale for public financing of the "Alzheimer's enterprise", i.e. "clinical benefit" remains elusive.

TABLE OF CONTENTS

TABLES

FIGURES

Chapter 1

AGEING, SOCIAL POLICY AND THE MEDICAL ENTERPRISE

"This disease will have a very big impact on our nation" The Washington Post, November 10, 1989

Dementia, the "loss of the mind" has been recognised and portrayed in literature over the centuries as a distressing potential affliction of old age, eliciting pity, mockery, and sometimes compassion, but until recent years, little or no attention from the medical profession. Senile dementia had been variously defined in classifications of mental disorders as "acquired imbecility", Kraepelin, 1906 xvii and the "psychosis" i.e. madness of old age, DSM-II, 1968 p24 and some suffering from the condition, particularly among the poor, came under the nominal care of the medical profession in insane asylums or their successor institutions. Scull, 1989 p249 The results of those few epidemiological studies reported prior to the 1970s, none of which were conducted in the United States, indicated that the incidence increases with age, with prevalence estimates in the population aged 65 and over ranging from 2.5 percent in one Norwegian community Bremer, 1951 to 4.6 percent in Newcastle, England. Kay et al, 1964 p153

In 1989, however, the following headline appeared on the front page of one of the most influential newspapers in the United States: **"Alzheimer's Toll Heavier: Illness May Affect 4 Million in U.S."**. The Washington Post, November 10, 1989 [1] The article went on to report the results of an epidemiological study published that day in the Journal of the American Medical Association. Evans et al, 1989 The newspaper described the medical journal

---

[1] A similar headline appeared on the front page of another prestigious daily newspaper: "Study finds Alzheimer's Disease afflicts more than was estimated". New York Times, November 10, 1989

findings as "the clearest evidence to date of the extent of a disease that until recently was poorly understood and often overlooked by physicians". It stated that the "rate of disease rises far more dramatically with age than had been suspected ..... 3 percent of those age 65 to 74 had the disease, but 18.7 percent of those between 75 and 84 had it, as did more than 47 percent of those age 85 and over". The study had been carried out in a "working-class neighbourhood of Boston" where "82 percent of the people in the community over 65" were given "brief memory tests", and a proportion of all those first interviewed received "more detailed psychiatric exams and laboratory tests". It was further noted that the research had been conducted by a "team from Harvard Medical School". The director of the study, a physician, had been "sobered by these findings", and had made the following statement: "Demographic trends make it clear that this disease will have a very big impact on our nation in the next 50 years".[The Washington Post, November 10, 1989]

The newspaper went on to report that the "National Institute on Aging estimates that within 50 years as many as 14 million Americans could have the disease", defined in the article as a "degenerative illness of the nervous system with no known cause, cure or treatment". Executives from the national Alzheimer's Association were quoted as responding to the study with the observations that "the financial and public health implications are remarkable", and "..... the impact could eventually affect us all". Finally, there was a quotation from an editorial that accompanied the study in the same medical journal.[Larson, 1989] This stated that "these new findings are likely to be disturbing to the general public, given the widespread fear of Alzheimer's disease".[The Washington Post, November 10, 1989]

When studies by academic and private institutions, or government agencies, relating to defence, economic growth, international relations, or other affairs perceived to be of national importance, are reported in the Washington Post, the typical response is analysis of the assumptions underlying the study and comment on the implications of the findings in editorial columns, frequently written *ad hoc* by those with relevant knowledge of the subject. In this instance, however, there was no appraisal of the report on the editorial pages and no printed reaction from medical doctors, demographers, epidemiologists, economists or members of other disciplines in the many academic and "biomedical" institutions of Metropolitan Washington. Nor was there any published reaction from the "lay" readership for whom it was apparently unnecessary to explain that the topic of the report was senile dementia, redefined as "Alzheimer's disease".

The social construction of reality

Berger and Luckmann posited that the "root proposition" for their 1966 "treatise on the sociology of knowledge" derived from the proposal by Marx that "man's consciousness is determined by his social being". They argued that what passes for "knowledge" in a society, regardless of its ultimate validity or invalidity as "judged by whatever criteria", is "socially constructed". Insofar as all human "knowledge" is developed, transmitted and maintained in social situations, "the sociology of knowledge must seek to understand the processes by which this is done (so that) a taken-for-granted 'reality' congeals for the man in the street".[Berger & Luckmann, 1966 p3-6]

They further argued that society is itself constructed both as "objective" and "subjective" reality. As "subjective reality", society derives from the process of socialisation or the internalisation of the objective "social world".[Berger & Luckmann, 1966 p130] Objective reality lies in the legitimation of "institutions", which are the products of "habitualised" human activity, and may only be adequately understood in the context of the historical processes from which they are derived.[Berger & Luckmann, 1966 p54-5] Legitimation, built upon the use of language, is the process of "explaining and justifying", ascribing "cognitive validity" to the institutional order, and "normative dignity" to its practical imperatives.[Berger & Luckmann, 1966 p93]

They also posited that different, but potentially overlapping levels of legitimation may be distinguished, ranging from the linguistic "objectification" of simple human experience, to the level containing explicit theories by which an institutional sector is legitimated and perpetuated in terms of a differentiated body of knowledge, entrusted to specialised personnel to transmit through "formalised initiation procedures". The latter pertains to relatively autonomous "subuniverses of meaning", based upon the development of role-specific bodies of knowledge, which may be socially structured by criteria such as sex, age, or occupation.[Berger & Luckmann, 1966 pp93-5] Not only does an increasing number and complexity of subuniverses make them increasingly inaccessible to outsiders, but they raise the potential for the "institutional order" or for some part of it to be "apprehended as a non-human facticity", through the "reification of social reality". Reification implies that man is capable of forgetting his own authorship of the human world and that the "dialectic between man, the producer, and his products is lost to consciousness". In this way the products of human activity are apprehended as if they were "facts of nature, results of cosmic laws, or manifestations of divine will".[Berger &

Luckmann, 1966 pp88-9

They also proposed that if the institutional order is to be taken for granted in its totality as a meaningful whole, it must be legitimated by placement in a "symbolic universe" which provides a comprehensive integration of all discrete institutional processes. Symbolic universes were defined as "bodies of theoretical tradition", referring to realities other than those of everyday experience, i.e. as "social products with a history".[Berger & Luckmann, p95] While not proposing an evolutionary scheme, different conceptual machineries of universe maintenance could be historically identified, of which the most conspicuous are mythology, theology, philosophy and science.[Berger & Luckmann, 1966 p110] At the same time both institutions and symbolic universes are legitimated, maintained or changed by living individuals who have concrete social locations, social interests and their own "definitions of reality", described as having "self-fulfilling potency".[Berger & Luckmann, 1966 p128]

A paradigm for enquiry

Berger and Luckmann generated a theoretical perspective on the sociology of knowledge which has influenced the analysis of institutions and policies on both sides of the Atlantic since its publication. In a 1973 paper on social policy and social change, Carrier and Kendall directed attention to the question of "whose definitions of reality are embodied in welfare legislation, and whose particular concepts of society and its social problems are involved in the policy-making process?"[Carrier & Kendall, 1973 p221] Silverman, in developing a sociological framework for the study of organisations, rejected theories "reifying" the concept of institutions in favour of the view of "social reality as socially constructed, socially sustained and socially changed".[Silverman, 1970 p5] Estes acknowledged the influence of Berger and Luckmann in proposing a paradigm for the study of ageing in America: "knowledge is socially generated ...... It may be accepted as factually legitimate, based upon empirical demonstrations of proof or upon the judgements of proclaimed experts and authorities who possess status and power...... Although socially generated, such knowledge and expert opinion take on the character of objective reality, regardless of inherent validity".[Estes, 1979 p6] Freidson subtitled his analysis of the medical profession in the United States, "a study of the sociology of applied knowledge" and invoked the work of Berger and Luckmann in arguing that "like law and religion, the profession of medicine uses normative criteria to pick out what it is interested in, and its work constitutes a social reality that is distinct from (and on occasion virtually independent of) physical reality".[Freidson, 1973 p206]

Both the original treatise on the "social construction of reality", and the application and extension of its concepts by other analysts, afford a model for exploration of the "constructs" of the Washington Post report described above. Ostensibly, accredited medical researchers had conducted a scientifically valid study and discovered that a "disease", which a decade earlier was rarely diagnosed in medical practice, and totally unknown to the population at large, could now be identified in epidemic proportions in the older population to add to the "widespread fear" which had already been generated in the general public. The "disease" had been detected by "brief memory tests" conducted among elderly people in a working-class neighbourhood of one American city. Based upon demographic projections, a "reified" "national institute" had extrapolated the prevalence estimates from this small population cluster to present and future generations of a country of 250 million people, and national economic and public health "implications" had been attributed to these extrapolations.

In his work on the "structure of scientific revolutions", Kuhn proposed that "scientific knowledge" derives from the adoption of changing "paradigms of inquiry", of which a "formative ingredient" is "the beliefs espoused by a given scientific community at a given time".[Kuhn, 1962 p4] Many questions may be raised regarding the methodology of the Boston study and the generalisability of its findings, but more fundamentally it may be asked how and why the investigation was conceived at that particular "given time". Furthermore, why were the "beliefs" or "definitions of reality" of those who formulated the paradigm so widely shared that the results, and projections based upon them, could be accepted without challenge? It may be argued that the answer lies both in what may be accepted as "factually legitimate", i.e. declining mortality, increasing life-expectancy, and growth in the proportion of older people in the population, but also in the social constructs which have been built around these phenomena.

From this perspective it may be proposed that the newspaper report and the journal article on which it was based are representative of the intellectual output of a "medical enterprise", the institutions and funding for which derive from the health care policies which have been implemented in the United States, not only on behalf of the "elderly", a category which is itself a social construct, but for the advancement of "science" in the understanding of disease. It may be further posited that the "paradigm of enquiry" underlying the epidemiological study of Alzheimer's disease, was generated less by the "beliefs espoused by the scientific community" than the power of the medical profession to subvert and promote these policies in their own self-interests. At the same time it may be argued that expenditures on behalf of this and other medical enterprises, and the publicity surrounding their output, serve to reinforce the arguments of those projecting population ageing as an economic "burden" for society. On the other hand, proponents of the "burden of ageing" concept tend to propose reducing or eliminating state retirement benefits, whereas the solution suggested by the medical community to "physical" evidence of the negative effects of ageing, is further public investment in its own clinical and research activities.

Ideology, social policy and population ageing

Marx defined "ideology" as "ideas serving as weapons for social interests", a concept adapted by Berger & Luckmann as a "definition of reality attached to a concrete power interest". The distinctiveness of ideology is that "the same overall universe is interpreted in different ways". An ideology may be taken on by a group because of specific theoretical elements that are conducive to its interests, but ideologies that are irrelevant to their interests may be adopted by other groups. The "pluralistic" nature of modern societies means that there is "a shared core universe" and "different partial universes coexisting in a state of mutual accommodation" so that "outright conflict between ideologies has been replaced by varying degrees of tolerance or even co-operation".Berger & Luckmann, 1966 p123-5

It may be argued that in the United States there is a dominant "ideology" that has defined, more consistently than in other industrial countries, the "reality" within which social policy has been conceived and implemented. The ideology is that of *laissez-faire* capitalism, i.e. individualism, family responsibility and the free-market. Furthermore it may be proposed that the "economy" and other constructs or "products of human activity" with which it is associated are frequently apprehended as if they are "facts of nature, or the results of cosmic laws", i.e. neutral or value free in their impact. At the same time, the ideology of the market is combined with a shared "core universe" of democratic pluralism generating a perception that the "public interest" derives from accommodation between competing interest groups, based on an assumption of equality of class, status and power. In consequence, concepts such as "collective responsibility" have, historically, been absent from political debate.

John Myles, in his work on the "political economy of public pensions", argued that the transformation of "old age" into a social category to which one gains access by virtue of reaching a specified age derives from the institution of a set of age-based income entitlements administered by the state, i.e. the public pension.[Myles, 1989 p2] At the same time he proposed that it is "startling" to observe that in all the "post-War capitalist democracies" the state has become the major source of income in old age, since intuitively this would not be predicted from the principles of social organisation upon which they were founded.[Myles, 1989 p1] On the other hand, retirement ages and relative benefit levels vary from country to country, and he proposed that since the "the social, legal and political constituency we call the elderly was created and given form by social, political and economic forces", this socially constructed "category" is therefore continually subject to these same forces.[Myles, 1984 p120][2]

With waves of immigration from other countries, racial divisions, vast natural resources and the geographic potential for internal migration, there is little evidence of working class solidarity in the United States to support the arguments of Navarro that "class conflict" has been instrumental in achieving the predominant social reforms that were passed almost exclusively in the 1930s and 1960s.[3] The Social Security Act of 1935, the only "universal" social policy, marked the adoption of the "retirement principle" at a later stage than in most other industrial countries, and the legislation was passed in the face of overwhelming evidence of market failure with a depressed economy and massive unemployment. Myles has argued that passage of social insurance legislation under the Roosevelt administration was legitimated as a "means for

---

[2] In support of his argument, Myles has noted that in 1981 contrasting policies were espoused in France and in the United States in the face of high rates of inflation and unemployment, and low economic growth. In France a Keynesian approach was adopted: the retirement age was lowered and pensions were increased to remove older workers from the work force and stimulate demand.

In the United States, by contrast, it was proposed that public funding of retirement income was "the cause rather than a cure" for economic stagnation and exactly opposite policies were proposed, resulting in compromise legislation for a gradual increase in the retirement age beginning in the year 2000. Myles argued that with "a stroke of the legislative brush" the future size of the elderly population in the United States was reduced from earlier projections, while in France it was immediately increased.[Myles, 1989 p5-6]

[3] Navarro has argued that in all capitalist societies including the United States, the "working class and its conflict with the capitalist class - a conflict that takes place in all political, social and civil spheres of society - have produced a series of reforms that can be cumulative and can determine the transformation and even the breakdown of the capitalist system". He further argued that whether the working classes achieved "universalism" depended primarily on "how powerful they were versus the capitalist classes that opposed both universalism and the expansion of benefits".[Navarro, 1991 p588-90]

rationalising labour inputs", i.e. increasing economic efficiency by "removing older workers from the industrial labour force".Myles,1991 p298 [4]

In the 1960s, however, in a period marked by the African-American civil rights movement and anti-war protests, a range of social legislation was passed by the U.S. Congress for which there was little legitimation within the dominant ideology. Among the policies enacted, which in combination formed what has been described pejoratively by critics as a "welfare state for the elderly",Myles, 1989 p6 were provisions for a range of social services in the Older Americans' Act, and government subsidised medical insurance under Medicare (Title XVIII of the Social Security Act), 1965. There were also increases in state retirement benefits and in amendments to the Social Security Act passed in 1972, these were indexed to the rate of inflation.Myles, 1991 p300 Myles has proposed that these improvements in social insurance derived from a "temporary shift" in the balance of "class power" and that in order to redress this perceived loss of power in the political process, corporate America "nurtured scholarship and research", based on the *laissez-faire* model, to legitimate reversal of the 1972 policy. The intellectual production of this exercise was the "social construction" of ageing as an "economic

---

[4] Intervention in the free market is legitimated by the concept of "Pareto optimality". The "welfare economic" model supports the market by arguing that for every Pareto optimal point on a production possibility frontier there is a state of perfectly competitive equilibrium. The *a priori* assumptions for perfect competition are rationality, perfect information on the part of both producers and consumers, and no externalities or economies of scale. Where these do not hold, or there is inequitable distribution of income according to the norms of society, arguments may be made for government intervention in the market by way of regulation, subsidy or provision to promote equity and efficiency. Child labour was not excluded from the market until passage of the National Labour Standards Act, 1938.

burden".Myles, 1991 p302-5 [5]

Myles has tended to emphasise the positive benefits of state pensions in providing a level of income security, and giving workers a valued "right to cease working before wearing out".Myles, 1991 p304 Other analysts such as Estes in the United States, and Townsend in the United Kingdom have stressed the negative effects of national insurance and other social policies for the older population perceiving provisions such as a statutory retirement age as disempowering and creating a situation of "structured dependency" for those to whom it applies.Estes, 1979, Townsend, 1981 At the same time, Walker, in a discussion of the international adoption of the "burden of ageing" argument to legitimate "anti-welfare policies" proposed that the theory has generated "an increasingly prevalent assumption that societal ageing should be a source of ..... remorse rather than pride".Walker, 1990 p378

At the end of the 19th Century, life expectancy at birth in the United States was 48 years for men and 51 for women. Declining mortality during the next half-century resulted in a 37 percent increase in longevity for men to 66 years, and a 40 percent increase for women to 72 years. Since 1950, life expectancy has continued to increase, but at a slower rate and with higher gains for women than for men. By 1990, life expectancy had risen to 72 years for men and 79 for women. Similar changes occurred

---

[5] Myles has described the establishment and funding by the business community of policy institutes such as the American Enterprise Institute and the National Bureau of Economic Research to develop economic models to legitimate the reversal of the 1972 amendments to the Social Security Act. The director of the latter, Martin Feldstein became chief economic advisor to President Reagan in 1982.Myles, 1991 p303 The validity of the influential econometric models developed have been subject to continuing debate and evaluation (see: Barr N. and Whynes D. Current Issues in the Economics of Welfare, 1993, pp53-4, 142-9, and 164).

Table 1.1

LIFE EXPECTANCY AT BIRTH FOR MEN AND WOMEN: SELECTED COUNTRIES, 1900, 1950, 1990

	Circa 1900		Circa 1950		Circa 1990	
	MEN	WOMEN	MEN	WOMEN	MEN	WOMEN
United States	48.3	51.1	66.0	71.7	72.1	79.0
England & Wales	46.4	50.1	66.2	71.1	73.3	79.2
Denmark	51.6	54.8	68.9	71.5	72.6	78.8
Norway	52.3	55.8	70.3	73.8	73.3	80.8
Sweden	52.8	55.3	69.9	72.6	74.7	80.7
The Netherlands	48.6	51.2	70.3	72.6	74.2	81.1
France	45.3	48.7	63.7	69.4	73.4	81.9
Australia	53.2	56.8	66.7	71.8	73.5	79.8
Japan	42.8	44.3	59.6	63.1	76.4	82.1

*Data in this table were derived from a variety of country–specific data that have been assembled into an international database housed and maintained by the U.S. Bureau of the Census.

Source: United States Bureau of the Census. Center for International Research, Washington, D.C.

in all industrialised nations during this period with a relative narrowing of differences between countries in mortality rates and thus in life expectancy (see Table 1.1).[Kinsella, 1991 p1197S]

In the early 20th Century, decreases in mortality rates were highest among infants and younger age groups, but by the 1940's they were spread more evenly across the age spectrum. During the past three decades, rates have declined at a slower rate, and the decline has been relatively higher among older age groups.[Kinsella, 1991 p1198S] In the United States between 1960 and 1990 there was a 19 percent increase in longevity for women at age 65, and a 17 percent increase for men, compared with an overall 8 percent improvement in life expectancy at birth during the same period. By 1990, predictions of years of life remaining at age 65 were 19.0 for women in the United States and 15.0 for men (see Table 1.2).

Mortality experience has been similar in all economically advanced countries but there have been differences between the United States and Europe in rates of population growth and in the age structure of the population, attributable to differences in levels of fertility and immigration. In 1901 less than 5 percent of the population of England and Wales was aged 65 or older, but by 1970 13 percent were in older age groups, a percentage that had risen to nearly 16 by 1990. In the United States, by contrast, 4 percent of the United States population was aged 65 and over in 1900, rising to 11 percent by 1980, and 13 percent by 1990, an increase in absolute numbers from 3 to 30 million. It is the ageing of the "baby-boom generation", born between 1945 and 1965, that will result in the highest rate of growth in the population aged 65 and over beginning after 2010 (see Table 1.3). At the same time, as a consequence of declining mortality

Table 1.2

LIFE EXPECTANCY AT BIRTH & AT AGE 65, AND PERCENTAGE INCREASE PER DECADE BY SEX. UNITED STATES, SELECTED YEARS: 1900–1990

	AT BIRTH			
	MEN		WOMEN	
YEAR		Percent Increase		Percent Increase
1900	46.3	--	48.3	--
1950	65.6	(41.7)	71.1	(47.2)
1960	66.6	(1.5)	73.1	(2.8)
1970	67.1	(0.8)	74.8	(2.3)
1980	70.0	(4.3)	77.4	(3.5)
1990	71.8	(2.6)	78.8	(1.8)
	AT AGE 65			
1900	11.5	--	12.2	--
1950	12.8	(11.3)	15.0	(12.3)
1960	12.8	(0.0)	15.8	(5.3)
1970	13.0	(1.6)	17.0	(7.6)
1980	14.1	(8.5)	18.3	(7.6)
1990	15.1	(7.1)	18.9	(3.3)

Source: National Center for Health Statistics: HEALTH, UNITED STATES, 1992 Hyattsville, MD, Public Health Service, 1993.

Table 1.3

TOTAL POPULATION AND PERCENTAGE IN OLDER AGE GROUPS: CENSUS DATA & FUTURE PROJECTIONS, SELECTED COUNTRIES, 1900–2025

		Total Population	Percentage Aged:		
Country	Year	(thousands)	65 & over	75 & over	80 & over
United States	1900	75,793	4.0	1.2	0.5
	1950	151,325	8.1	2.6	1.1
	1970	205,051	9.8	3.7	1.8
	1990	250,410	12.6	5.3	2.8
	2010	282,547	13.9	6.5	4.0
	2025	298,253	20.0	8.4	4.8
United Kingdom	1900	32,528	4.7	1.4	0.5
	1950	43,758	11.0	3.6	1.5
	1970	55,632	12.8	4.5	2.2
	1990	57,418	15.7	6.9	3.7
	2010	59,617	17.1	8.3	5.1
	2025	60,032	21.5	10.8	6.3
Japan	1920*	55,963	5.3	1.3	0.5
	1950	83,200	4.9	1.3	0.5
	1970	104,331	7.1	2.1	0.9
	1990	123,567	11.8	4.8	2.4
	2010	130,480	21.3	10.0	5.7
	2025	125,279	26.7	15.2	9.3

*The first Population Census was conducted in Japan in 1920.

Source: United States Bureau of the Census. Center for International Research, Washington, D.C., 1993, and Statistics Bureau, Management and Coordination Agency, Government of Japan.

at more advanced ages, demographic projections indicate that 20 percent or more of the population of all similarly developed countries will be aged 65 and over by 2025, with 5 to 9 percent aged 80 and over.[6]

It may be argued that there is supporting evidence for the "burden of ageing" argument in the demographic changes which have occurred and are empirically verifiable from census data, as well as in future projections which are social constructs based upon assumptions regarding continuing trends in mortality, fertility and immigration. On the other hand, Estes, from the perspective of the "structured dependency" of "old age" has addressed both the dominance of the market ideology underlying proposals to reduce or eliminate retirement income, a direct benefit accounting for the highest proportion of public expenditures on behalf of the elderly,[Estes, 1982 p577] and also the distribution of benefits from the supply of goods and services under other government policies of which the older population are the ostensible beneficiaries.

In 1979, Estes proposed that the "major problems faced by the elderly in the United States are .... ones that are socially constructed". She argued that policies developed to remedy normative conceptions of social problems and priorities, even when supported in principle by the older population, are subverted by individuals and groups with greater resources of class, status and power to influence policy choices in their own self-interests. The term used by her to describe the outcome of the exercise of such influences was the "ageing enterprise", said to comprise physicians, hospitals, insurance

---

[6] It should be noted, however, that if all projections are realised a smaller proportion of the total population of the United States will be in older age groups in the next century than in Japan or the United Kingdom in the next century (see Table 1.3) or many other European nations.

companies, national and local bureaucracies, and other professionals and institutions that "serve the aged in one capacity or another".[Estes, 1979 p1-2] Estes posited that the policies and procedures that had evolved had accommodated to the interests of the "agencies and professionals that make up the ageing network", while disempowering and fostering dependency in those for whom they were devised.[Estes, 1979 p24-5]

Weber, in a discussion of legitimate authority and bureaucracy proposed that "the capitalist entrepreneur is the only type" who has been able to maintain "at least relative immunity from subjection to the control of rational bureaucratic knowledge", while the rest of the population has tended to be organised in "corporate groups which are inevitably subject to bureaucratic control".[Weber, 1947 p28] The particular focus of Estes' 1979 work was the network of bureaucratic structures established to implement the coordination of social services under the Older Americans Act. On the other hand, while Estes may argue that the procedures devised and implemented within state agencies benefit those employed within them rather than serving the "needs of the elderly" for which the expenditures have been allocated, those who are employed within such "enterprises" are not entrepreneurs in the sense of being free from bureaucratic control, and deriving direct profits from the growth of the older population.

In a 1971 discussion of the "political integration of capitalism" in the United States, Birnbaum argued that the "public sector" serves as a means for providing "an infrastructure without which capitalist accumulation could not continue". Within this infrastructure he included "essential services" such as roads and other forms of transport, but also support for private industry as provided in defense contracts, describing the latter as a "remarkable instance of the socialisation of loss and the privatisation of profit".[Birnbaum, 1971 p383] It may be argued that an equally remarkable example is to be found in the medical profession which has received vast government investments in capital equipment, education and training, and research activities. At the same time, by subscribing to the ideology of the free market while exercising monopoly professional power American medicine has maintained the status of its members as independent entrepreneurs while subverting government policies in its own self-interests. With public funding of the institutional infrastructure exempting its members from assuming the real investment risks of entrepreneurialism, it has been able to construct a vast "ageing enterprise" based upon the indemnity insurance model which was, under the influence of the profession, adopted for the Medicare programme. Furthermore, the profits of the enterprise accrue both from market expansion with population ageing, and from the legitimacy accorded the profession, as an "autonomous subuniverse",[Berger & Luckmann, 1966 p93] to construct "disease specific" sub-enterprises based on the promotion of innovative techniques or the "discovery" of new diseases, activities which are in turn supported by public investment in research on behalf of "medical science".

The Medicare enterprise

It may be posited that health services show failure of all the preconditions of the competitive market, but although there has been state intervention in the supply of "public goods" such as clean water and sanitation, in contrast with other "capitalist democracies", there has been no legislation to secure "universal" provision of personal medical care in the United States. Rather, its financing subscribes to the "consumer sovereignty or private insurance model". Entry to the market for the majority of the population is achieved as a "fringe benefit" of employment, the cost of which is subsidised by the government as a tax-deductible business expense.[Schieber et al, 1992 p12]

Culyer, in a discussion of current issues in welfare economics argued that the "mostly private" market in the United States is both "unnecessarily costly and inequitable, with some 50 million citizens (20 percent of the population) either uninsured or underinsured".[Culyer, 1993 p153] The uninsured include the unemployed and those in part-time and low paid occupations indicating that there is, implicitly, a "human capital" argument in the provision of medical care to those of working age and their dependents. The system is costly in that an analysis comparing 24 industrialised member countries of the Organisation for Economic Cooperation and Development demonstrated that America had the highest ratio of medical care spending to gross domestic product (GDP), and the fastest rate of growth in these expenditures between 1980 and 1990. Furthermore, the returns, as defined by indicators such as life expectancy and infant mortality, compare unfavourably with Japan and other countries with a little over half the ratio of medical care to GDP expenditures.[Schieber et al, 1992 p12]

It has been argued that the primary risk to the efficiency of indemnity medical insurance is that of "moral hazard" or over-consumption by the consumer who receives care at a marginal cost which is less than the true marginal cost of provision. High medical costs might therefore be attributed to over-consumption in the exercise of "consumer sovereignty" by those with market access. To constrain such behaviour out-of-pocket expenditures including "copayments", "deductibles" and "coinsurance" are therefore introduced as disincentives on the demand side. On the other hand in a classic essay by Arrow it was proposed that "uncertainty" is a primary characteristic of the medical care market, and that information and advice becomes the "commodity" purchased from the medical profession.[Arrow, 1963 p18] In an open-ended financing system there are therefore no effective disincentives to "moral hazard" on the supply side, i.e. the creation of demand by performing "unnecessary" procedures and other forms of "encouragement" to over-consume. Strategies such as reducing the fees allowed for services, tend to act as incentives to increase supply to offset lower earnings.[Culyer, 1993 p162-4] Nevertheless, the dominant but inefficient private indemnity insurance model was adopted to provide medical care to the older American population.

Ehrenreich and Ehrenreich in a discussion of health care and social control, detailed methods adopted by the medical profession in the 1960s to discourage support among the "lay public", i.e. their patients, to the proposed Medicare legislation.[Ehrenreich & Ehrenreich, 1974 p36] Stevens, in her 1971 discussion of American medicine, documented the opposition of the leaders of the profession in the post-War period to a succession of proposals to introduce either national health insurance, labelled "socialised medicine", or non-means-tested government subsidised care for the elderly.[Stevens, 1971 p432-8] She therefore posited that with passage of Medicare, strongly supported by the older population, "the public emerged as the dominant decision maker (over) the vested interests of organised medicine", interpreting it as a breakthrough that would lead to a universal system.[Stevens, 1971 p442] On the other hand, an "open-ended" government commitment was made for the payment of "customary" and "reasonable fees" to the suppliers of services for those enrolled in the Medicare programme and contracts were awarded to the private insurance industry to negotiate fees and administer benefits.[Stevens, 1971 p448-53] Furthermore, regulations were written to maintain clinical freedom, and attempts to introduce national health insurance with its potential for bureaucratic regulation and budgetary control have been resisted to this day.[7]

Stevens noted that in the 1950s life expectancy was increasing, raising the prevalence rates of conditions such as heart disease, cancer and cataract, but less than half this expanding potential market of older people had medical insurance to purchase the industry output,[Stevens, 1971 p433] the supply and sophistication of which had increased with support from government investment in the hospital infrastructure.[8] She argued that Medicare legislation provided "benefits for expensive health services", i.e. hospital investigations and treatments, thus expanding the market for tertiary care and protecting the middle class from economic loss due to medical bills".[Stevens, 1971 p442] On the other hand there were no provisions for financing primary and preventive care which might potentially reduce the demand for such services. It may therefore be posited that having failed to prevent its passage, the law was written to accommodate to professionally-

---

[7] The final Medicare law declared that "there would be no supervision or control by any federal office or employee over the practice of medicine, or the manner in which medical services are provided", and an initial provision for a "national medical review committee" was repealed in 1967 under pressure from the American Medical Association.[Stevens, 1971 p458-9]

[8] The 1946 Hospital Survey and Construction Act (commonly known as the Hill-Burton Act) disbursed $3.7 billion in federal government funds and generated $9.1 billion in state and local financing between 1947 and 1971, with no requirement for the many hospitals which were constructed with these funds or the doctors who practiced in them to provide care to low income or uninsured populations.[Starr, 1982 p349-51]

perceived market imperfections, i.e. the inability of the elderly to meet the costs of acute and high technology care.

Starr, in his work on the "social transformation of American medicine" described the outcome of the introduction of Medicare in 1967 as a "bonanza" for the profession, and a "giveaway to the providers".[Starr, 1982 p370, p382] National data demonstrate that the population enrolled in the programme grew by 63 percent between 1967 and 1990 (from 19.5 to 30.9 million), but the average annual "payment per enrollee" increased by more than 1,200 percent (from $217 to $2,869).[National Center for Health Statistics, 1993 p186] In a 1991 report by the Health Care Financing Administration, the agency responsible for oversight of the Medicare programme, rising costs were attributed to a combination of "economywide inflation", "increases in use and intensity of services", and "medical price inflation". Those related to the latter two components have consistently outpaced the national rate of inflation, and have been resistant to measures such as "reductions in payments to physicians (and) in payments for diagnostic laboratory tests and some overpriced surgical procedures".[Levit, 1991 pp39-40] At the same time it may be proposed that the continued growth in the "use and intensity of services" derives both from the endemic incentive for "moral hazard" and from the construction of innovative sub-enterprises, and strategies to expand the market by specialties which benefit less from the indemnity system and consequently from Medicare.

The science of medicine, specialisation, and competitive power

From the perspective of Berger and Luckmann, the medical profession may be construed as an "autonomous subuniverse of meaning" with a role specific stock of knowledge that is "esoteric as against the common stock of knowledge". Berger & Luckmann, 1966 p85-6 At the same time, it may be posited that compared with their peer groups in other developed countries, the profession in the United States has achieved what is arguably a unique status as demonstrated by its independent entrepreneurial status and the preservation of the market model for the purchase of medical care. It may be further argued that other distinguishing and related characteristics of the American profession *vis a vis* many peer groups overseas are the lack of hierarchical structure as between "specialists" and "generalists", and the dominance of hospital-based and independent "fee for service" practice, and that these differences derive from the "standardisation" of medical education by the profession as a postgraduate doctoral degree in the early 20th century.[9]

Berger and Luckmann posited that the definitions of reality of "science" represent

---

[9] The "standardisation" of medical education came with the implementation of the recommendations of the 1910 Flexner Report, commissioned by the Council on Medical Education of the American Medical Association. The rationale given for the educational reforms, which also included the elimination of the majority of medical schools then in existence, was the "great discrepancy" that had opened up between "medical science and medical education". Starr, 1982 p116-8

Stevens, in her analysis of American medicine noted the "appalling conditions" and lack of educational resources in many of the commercial medical colleges then in existence. Nevertheless their closure was recommended without regard to geographic distribution, the socio-economic characteristics of the communities they served or their potential for improvement. She proposed that the Flexner Report and its adoption had a "stultifying social effect" by forcing the closure of many institutions established to give "apprenticeship" medical training to women and minorities, and thereafter excluding these groups and others with low income from the profession. Furthermore, the reforms failed "to produce physicians for different purposes", i.e. doctors for both urban and rural areas, and generalists as well as specialists. Stevens, 1971 p67 the "secularisation and sophistication" of universe maintenance compared with earlier "conceptualisations of the cosmos" as invoked by mythology or theology. "Science" removes "knowledge as such from the world" so that lay members of society no longer know how their universe is to be conceptually maintained, merely "who the specialists of universe-maintenance are presumed to be".[Berger & Luckmann, 1966 pp110-12] Freidson, in his 1973 work on the sociology of medicine, argued that (American) "medicine's position today is akin to that of state religions yesterday", with an officially approved monopoly to define health and illness and to treat the latter.[Freidson, 1973 p5] He further proposed that "public belief" in the "scientific foundation" of medical practice and a "secure and practical technology" was essential to the achievement of professional "autonomy" and to gaining the "confidence of the lay clientele".[Freidson, 1973 p21-2] Similarly, Starr, in his study of the "social transformation of American medicine" argued that medicine has achieved an unequalled position in the social structure, combining "broad cultural authority, economic power and political influence".[Starr, 1982 p6] He proposed that "authority" involves a "surrender of private judgement" and reflects the probability that "particular definitions of reality and judgements of meaning and value will prevail as valid and true".[Starr, 1982 pp12-16] The profession had gained this special "claim" to authority, based in part on its "bonds with modern science", the privileged status of scientific knowledge in the "hierarchy of belief", and the role of doctors as "intermediaries between science and private experience".[Starr, 1982 p4-6] It may therefore be proposed that the power of American medicine as an "autonomous subuniverse" within the "symbolic universe", or "hierarchy of institutions", was achieved by legitimation of its place within the dominant "universe-maintaining" conceptual machinery of "science" and that in educating its members to be "scientists", the status of "medical science" was, by association, thereby ascribed to its practice. It may further be proposed that the professional structure, payment system and the power to reject or subvert government legislation designed to provide greater access to medical care derives from a pervasive "belief" in "science" as the determinant of the actions of its practitioners.

It may be posited that the underlying thesis of Stevens' 1971 work on American medicine was its failure to serve the "public interest", which she attributed in large part to the fact that in 1971 some seventy to eighty percent of American doctors were defined as "full-time specialists".[Stevens, 1971 p3] Intuitively it might be assumed that the specialisation process, which gained momentum in the years after the "standardisation" of medical education, and the dispersion of its members among a large number of specialty groups would result in a highly fragmented profession.[10] On the other hand Stevens proposed that by compromise and by agreeing to limit medical practice to the chosen specialty, the establishment of "specialty boards" was a "dynamic process of status recognition which became a system of immutable vested interests".[Stevens, 1971 p243] Similarly, Starr has argued that the growth of specialisation resulted in "an informal control system in medical practice". There was a shift from dependence on clients to dependence on colleagues within a referral network, which promoted a change in the profession from a competitive

---

[10] After the Flexner reforms were implemented, sixteen "specialty boards" were established in the inter-War years including the American Board of Ophthalmology in 1917, the American Board of Psychiatry and Neurology in 1934 and the American Board of Neurological Surgery in 1940.[Stevens, 1971 p542]

to a "corporate orientation", and enabled physicians to assert their "long-run collective interests over their short-run individual interests".[Starr, 1982 p229-30] At the same time it may be posited that there were and continue to be status differences, and that the "corporate orientation" which generated political homogeneity also legitimated the development of a payment system congruent with the practice patterns and self-interests of higher status specialties.

Starr has proposed that specialisation gives producers partial relief from competition and enables them to benefit from whatever comparative advantage they may enjoy. At the same time "the specialist typically gives up those services offering the lowest return and concentrates on those offering the highest", and in medicine these tend to be services performed in hospitals because of the savings in time to the doctor, and the "standard of higher fees for procedures that are or were at one time complex".[Starr, 1982 p76-7] Starr also observed that "surgery enjoyed a spectacular rise in prestige and accomplishments" in the late 19th century following the introduction of anaesthesia, antiseptics and aseptic operating techniques.[Starr, 1982 p156] It may therefore be argued that the predominance of hospital-based practice and the fee-for-service system in American medicine derives from the model which yielded the highest returns to the prestigious practitioners of surgery as professional status and monopoly power were institutionalised. On the other hand although the collective power of the profession has been employed to maintain the dominance of the model, as demonstrated by its adoption in the Medicare programme, it may be argued that it is a system which discriminates against those specialties which are primarily "consulting" and lack either personal technical skills or access to technologies, the application of which are separately and highly rewarded.

A 1992 survey published by the Center for Health Policy Research of the American Medical Association showed the median net income for surgeons, and for members of the associated specialties of anaesthesia and radiology to be $200,000 per year compared with $120,000 for internal medicine, $107,000 for psychiatry and $93,000 for general/family practice.[Gonzalez, 1994 p18] Furthermore, surgeons had a high share of the Medicare market, treating 31 per 100,000 enrollees, whereas the proportion receiving psychiatric care was too small to be separately recorded.[Gonzalez, 1994 p4] It may therefore be proposed that for the "consulting" specialties there is an incentive to increase income by expanding the referral network for the detection of "new" diseases, and that a legitimating infrastructure for innovation and diversification in clinical practice has been provided to the academic leaders of both the technical and consulting specialties by public investment in research to extend the frontiers of scientific medicine.

The "national medical research enterprise"[Strickland, 1972 p210]

The first National Institute of Health evolved from the bacteriological laboratory of the U.S. Public Health Service, and was established by federal legislation in 1930 for "study, investigation and research in the fundamental problems of the diseases of man".[Strickland, 1972 p28] Strickland, in a review of medical research policy in the United States, proposed that in planning for peacetime, successful collaboration during the 2nd World War between the "government and the private sector, especially the universities", promoted the passage of the Public Health Service Act, 1944. This authorised the National Institutes of Health (NIH) to "pay for research to be performed by universities, hospitals, laboratories, and other public or private institutions".[11] In the debate one Congressional leader argued that "the expenditure of more funds in medical research would lead to a longer life and better health for the people of the country".[Strickland, 1972 p18-19]

Prior to the 1940s financing for medical research in the United States came primarily from charities such as the American Cancer Society, dedicated to raising funds for "fighting" specific diseases, and support from "corporate capitalism".[Starr, 1982 p227] The latter derived from the Rockefeller, Carnegie and other private foundations, which in 1940 gave $4.7 million for medical research compared with a total budget for the NIH of $700,000.[Strickland, 1972 p27] By the late 1960s, however, the number of separate institutes of the NIH had increased to ten, some established *ad hoc* by Act of the U.S. Congress, others at the initiative of the U.S. Surgeon General,[Strickland, 1972 p192] for investigation of a continuum of conditions perceived to be amenable to resolution by investment in medical research, including cancer, mental health, stroke, child health and heart disease. At the same time the budget grew from $2.5 million in 1945 to $1.6 billion by 1970, the NIH director distributing the funding evenly between the institutes.[Rettig, 1977 p53] On the other hand, although the NIH had over 13,000 full-time employees in 1968, some eighty percent of its budget was allocated for "extra-mural research", primarily to academic medical institutions.[Stevens, 1971 p358-60]

Stevens argued that NIH awards to external investigators provided a major subsidy for the "nominally independent medical schools", and by 1955 one-third of their income

---

[11] The second institute established was the National Cancer Institute in 1937, the legislation for which set a precedent for permitting the award of grants for extramural research.[Stevens, 1972 p358]

was from research grants, predominantly from the federal government which thereby became "essential to the existence of the schools".[12] She proposed that faculty members in the post-War period responded to an "apparently never-ending stream of federal wealth", which was channelled to the NIH to be reallocated "through a complicated, Byzantine system of national peer-review groups and advisory committees". Furthermore, within the medical schools the subsidy "created a system of powerful baronies built around specific research projects and interests".[Stevens, 1971 p359-60] Starr noted that "to a remarkable degree, control over research was ceded to the scientific community (and) approval of grant applications as well as basic policy issues rested with panels of nongovernmental scientists". Medical researchers who presented proposals to the NIH and received government grants "enjoyed autonomy within the constraints of professional competition",[Starr, 1982 p343] i.e. without public oversight of the distribution and use of funding or the outcome of the research endeavours supported.

Arrow in 1963 argued that the free market will tend to underinvest in the production of "information" by way of research "since the cost of reproduction is usually much less than that of production".[Arrow, 1963 p18] Abel-Smith similarly proposed that an "externalities" argument may be made for collective subsidy of medical research since the potential

---

[12] Stevens proposed that medical schools in the 1950s accepted the "back door of research subsidy"[Stevens, 1971 p359] as an alternative to direct federal aid. Similarly, Starr has noted that the American Medical Association opposed government proposals for grants and scholarships for medical education in the 1940s and 50s, as setting a "dangerous precedent".[Starr, 1982 p347] Both Stevens and Starr agree that such aid would have entailed requirements to increase the number of students and to award scholarships for those with low income thereby potentially reducing the "social homogeneity" and prestige of the profession as well as control over its size.

returns to investment in new methods for the prevention or cure of disease may have widespread benefits, and support from "voluntary subscriptions" does not represent the full "value to society".[Abel-Smith, 1976 p34] It may therefore be posited that government funding for medical research was legitimated as a collective good for the promotion of "public health", a legitimacy which was reinforced by the place of American medicine in the "symbolic universe" of science. On the other hand it may be argued that in common with the Medicare programme, the 1930 legislative remit and the 1944 policy were subverted by the medical profession, the NIH providing the infrastructure for entrepreneurialism in academic medicine by individual researchers on behalf of their own careers, the institutions with which they were associated, and their own medical specialties. It may be further proposed that in expropriating the NIH agenda in its own self-interests the profession has invoked emotive rather than scientific arguments, and has exercised collective power in opposition to any perceived threat to autonomy or attempts to restrict the growth in resources allocated to the "national medical research enterprise" while seeking to expand the share of resources allocated for extra-mural research.

Strickland has suggested that there were, from the beginning, philosophical differences in interpreting the language of the 1930 statute (see above) which established the NIH, between those who supported research into "fundamental problems" and those who sought to emphasise the "diseases of man".[Strickland, 1972 p28] It was argued by proponents of the latter approach, one that was frequently endorsed by politicians, that basic research in the biological sciences is pursued over the longer term with unpredictable results whereas more immediate benefits to society may be derived from clinical investigations into the development and evaluation of new technologies for the treatment of established disease.[13] On the other hand, the focus of the NIH intra-mural research has tended to be on basic sciences such as microbiology, while the agenda for investigations by the extra-mural clientel has been self-selected by the academic representatives of medical specialties, and funding primarily supports "applied research" or the development and evaluation of new diagnostic and therapeutic interventions, both medical and surgical. In contrast with the Medicare programme, however, for which an open-ended commitment of government funds was achieved by the profession, the NIH budget is reviewed annually by the U.S. Congress. It may therefore be posited that rather than philosophical differences with regard to research priorities, the "disease" emphasis was defined both by the "realities" of the medical profession as a whole, and by competition between different specialties in seeking funding for extending the scope of their particular sphere in the interests of academic advancement within their institutions and the practice of their clinical colleagues. Furthermore, the interests of academic entrepreneurialism were combined with professional ambivalence with regard to the NIH as a "bureaucratic" government agency.

On the one hand it may be argued that the reified "National Institutes of Health", which employed members from a range of scientific disciplines for the conduct of basic research, embodied concepts such as "progress in the fight against disease", and thereby assigned scientific validity to the studies for which it served as a financial conduit.

---

[13] One Congressman speaking in support of reallocating resources in favour of clinical research is quoted as asking the rhetorical question, "Whoever died of micro-biology?". Strickland, 1972 p192

These studies then legitimated innovation and diversification in the supply of services to the market while subsidising medical schools and the activities of the academic faculties.[14] On the other hand the internal activities reduced the pool of resources available for external entrepreneurialism. Furthermore, it may be posited that there are preventive implications from research in fields such as immunology and bacteriology, which tend to be a focus of studies undertaken by the NIH, and that these may be perceived as antithetical to the expansionary self-interests of the profession.

Starr has documented the historical opposition of private medicine to measures which might pose any economic threat, even those for preventing the spread of diseases such as tuberculosis, when undertaken by state agencies.[Starr, 1982 p181-5] [15] On the other hand, Stevens has noted that where government funding via the NIH resulted in the successful development of new techniques, such as a test for the early detection of cancer, there was no challenge to doctors to change medical institutions to promote greater public access to the outcome of the research and the benefits therefore accrued to the profession.[Stevens,

---

[14] A recent example was published in the Washington Post under the headline: "NIH Backs Operation for Strokes". The article stated that, based on a study conducted by neurosurgeons at 39 hospitals, officials at the National Institutes of Health were recommending "surgery to unclog the major arteries supplying blood to the brain" as a "preventive strategy" for strokes. It was reported that "the number of people who might benefit from the operation (costing $15,000 per operation) could number in the tens of thousands".[The Washington Post, October 1, 1994]

[15] Starr has argued that "devotion to medicine in pursuit of health", reinforced by successes such as a vaccine for polio-myelitis, reinforced public support for financing of medical research in the post-War period.[Starr, 1982 p346-7] On the other hand, research for the Sabine and Salk polio vaccines was funded by voluntary contributions, and their distribution by the U.S. Department of Health in the 1950s was undertaken only at the insistence of members of Congress, not the medical profession.[Strickland, 1972 p248]

1971 p352 It may therefore be posited that arguments to reallocate resources from the intra-mural programme of the NIH, in favour of extra-mural research, served both to restrict the scope of its investigations and also to limit the growth and power of the institution, thereby eliminating its potential for implementing overtly "public health" policies by disseminating the output of the basic science or the clinical studies it supported. At the same time Strickland noted that it was argued that the benefits of the "research enterprise" would be dispersed by what he described as "a kind of trickle-down theory" even without universal access to medical care. Strickland, 1972 p259 [16]

The "dread disease" model

Strickland entitled his review of medical research policy, "Politics, science and dread disease", the latter referring to the "symbolic language" Berger & Luckmann, 1966 p40 constructed by medical academics to persuade members of Congress in the annual budget review of the merits of increasing research funding for the total NIH budget with additional increments for specific selected conditions, a process in which "lay advocates" were recruited to give testimony. These lay advocates included individuals suffering from specific illnesses, paid representatives of powerful voluntary societies, and distinguished doctors speaking on behalf of "citizens' committees" supporting the need to sponsor further developments in heart surgery, etc. Strickland, 1972 p142-3

Rettig in a discussion of the National Cancer Act, 1971, which allocated $1.5 billion

---

[16] In support of this argument both Starr Starr, 1982 p343 and Strickland have argued that politicians were persuaded by organised medicine that ever increasing government investment in medical research was an effective alternative to national health insurance. Strickland offers the following quotation by Congressman Melvin Laird: "Medical research is the best kind of health insurance the American people could have". Strickland, 1972 p213

in dedicated research funds for the disease, argued that the "lay lobby" which was formed and promoted the "conquest of cancer" campaign in 1969 was seeking primarily to restore the *status quo ante* of NIH funding for medical research, the budget for which was increased below the rate of inflation in 1968, for the first time in the post-War period.[Rettig, 1977 p10] On the other hand, beginning in the 1960s the financial management and merit of extra-mural research supported by the NIH also came under critical scrutiny in a series of reports by a Congressional committee and there were requests for greater accountability from academic investigators and their institutions.[17] Furthermore, in the debate over the legislation, it was initially proposed that responsibility for cancer research be assigned to a new agency directly accountable to the U.S. President, and outside the jurisdiction of the NIH, which was criticised in the national press for its emphasis on basic science to determine "causes" rather than applied research for treatment of disease.[Rettig, 1977 p138-40]

It may be argued that cancer is the prototypical "dread disease",[18] having received research funding from the federal government from as early as 1922, and being the first

---

[17] The U.S. Congress, Committee on Government Operations: The Administration of Grants by the NIH, 1962 was the first in a series of critical reports produced under the chairmanship of Congressman L.H. Fountain and termed the Fountain Reports.[Strickland, 1972 p312] Stevens has noted that the 1962 report, based on a detailed audit, revealed "misuse of grants, excessive payments, purchase of unnecessary equipment, and inadequate administrative review".[Stevens, 1971 p368]

[18] "The time has come in America when the same kind of concentrated effort that split the atom and took man to the moon should be turned toward conquering this dread disease. Let us make a national commitment to achieve this goal.[President Richard M. Nixon, The State of the Union, January 22, 1971]

condition to merit the establishment of a separate national research institute in 1937 (see Footnote 11). At the same time "cancer" appears in many forms affecting different types of body cells and all organs of the body and many if not most surgical and medical specialties are engaged in its treatment. It may therefore be proposed that the "conquest of cancer" campaign was promoted to offset proposed reductions in funding for medical research, and that the disease was chosen as a focus because of the high probability of attracting research funding and the potential for disseminating the proceeds among a wide range of specialties. It may be further proposed that the campaign was undertaken in the context of a perceived threat to autonomy from requests for more public oversight of the finances disbursed to the research community (see Footnote 17), and that the counter threat to the integrity of the NIH posed by proposals to remove cancer research from its ambit served as a warning against intrusion into the "autonomous subuniverse".[19] On the other hand, the 1971 Act undermined the principle of relative parity in the allocation of resources between the various institutes and introduced competition for a diminished share, while at the same time introducing a model for securing dedicated funding from Congress for other academic research enterprises.

The National Institute on Aging and the "paradigm of enquiry"

From the perspective of the proposition by Kuhn that "scientific knowledge" derives

---

[19] The Dean of Johns Hopkins Medical School and President of the Association of American Medical Colleges, in response to requests for greater accountability for the outcome of federally supported extra-mural research, argued that "it is virtually impossible to put a dollar value on medical advances".[Turner, 1966 p111] With regard to proposals for more detailed accounting of grant expenditures he argued that "down that road lies the pall of mediocrity; the sterility of the managed human spirit".[Turner, 1966 p116]

from the adoption of changing "paradigms of inquiry", of which a "formative ingredient" is the "beliefs espoused by a given scientific community at a given time",$^{\text{Kuhn, 1962 p4}}$ it was argued, earlier in this discussion, that the epidemiological study conducted in East Boston, from which prevalence estimates of Alzheimer's disease were disseminated to the national media, was conducted in the context of the growth of the proportion of older people in the population, of the health care policies which have been implemented on their behalf, and of the power of the medical profession to subvert such policies in their own self-interests (see above). It may be further argued that the specific "paradigm of inquiry" in which the East Boston study was conceived was the "dread disease" model, constructed by the medical community in the interests of sustained public funding for academic entrepreneurialism, while warding off any perceived threat to "professional autonomy" from the NIH as an agency of government, including that posed by the establishment of the National Institute on Aging under the Research on Aging Act, 1974.

Stevens has described the process by which the American medical specialties "carved out the boundaries of its field", defined its content, and decided on acceptable modes of practice and behaviour" as being "at root a political process, arising from the relative successes of interest groups".$^{\text{Stevens, 1971 p218-9}}$ On the other hand it may be argued that "specialisation" may also be perceived as the dissection of the human being into its component parts by specialist groups competing for exclusive jurisdiction and relief from competition. From this perspective it may be argued that "holistic" medicine as pursued by general practitioners is discouraged and assigned low status by American medicine because of its negative impact on the market for specialist practice. Furthermore, although the term "geriatrics" has been attributed to an American physician, Nascher, in the first decade of the 20th century,[Fennell et al, 1988 p39] no medical specialty comparable with "paediatrics" has achieved legitimacy in the United States.[20] Lockett, in a review of "ageing, politics and research", documented strong opposition, both from the profession and from the director of the NIH, dating back to the 1950s, when "the NIH had more money than it could spend", to proposals for a national institute for ageing studies.[Lockett, 1983 p39-40] In response to those who advocated specific investment in "ageing" research it was argued that the existing institutes were addressing the problems of old age including cancer, heart disease and stroke, and "ergo the NIH is gerontology".[Lockett, 1983 p39] [21] On the other hand Lockett has posited that the "cost of Medicare", which beginning in 1967 appeared with the "medical research component" in the budget of the Department of Health Education and Welfare, was a deciding factor in the debate over legislation for the National Institute on Aging, with Congress accepting the argument that greater public investment in "biomedical research could help prevent old age disabilities and save taxpayers' money".[Lockett, 1985 p182] To this end, and in spite of political opposition,[22] a "multi-disciplinary" Congressional mandate was given to the new

---

[20] In 1988 the American Board of Internal Medicine and the American Board of Family Practice conducted the first joint examination for a certificate of competence in geriatric medicine.

[21] It was reported that the NIH director would not admit "that there was such a thing as gerontology and geriatrics".[Lockett, 1983 p39]

[22] President Nixon vetoed an Act passed in 1972 to establish the National Institute on Aging. He signed the second Act into law in 1974 at a time when he and members of his administration were under investigation by Congress for criminal activities just two months before he was forced to resign from office.[Lockett, 1983 p152]

institute to develop "a plan to coordinate and promote research into the biological, medical, psychological, social, educational and economic aspects of ageing".[Lockett, 1983 p4-5] [23]

Thus it may be posited that the "dread disease" paradigm was invoked in the context of a lack of perceived legitimacy for the concept of an institution dedicated to: the "holistic study" of ageing; the threat to professional entrepreneurialism posed by attempts to curb the rate of increase in public spending on the Medicare programme; and the mandate for inclusion of the "social sciences" in the "national medical research enterprise". It may be further posited that the results of East Boston epidemiological study represent the output of a further subversion of public policy in the successful construction of an "Alzheimer's enterprise".

---

[23] In 1938 an international symposium took place in Washington, D.C., the proceedings of which were published under the title: "Problems of Ageing: biological and medical aspects".[Cowdry, 1939] In the introduction it was proposed that the purpose was to address the "process of ageing" and what is arguably a valid and still unresolved investigative issue, i.e. "the possibility of distinguishing between the cumulative but physiological involutions that inevitably take place in all individuals as they grow older, and pathological changes that occur in ageing individuals as the result of adverse environmental conditions".[Cowdry, 1939 p xiii]

Lockett notes that the symposium was called by those who anticipated "potential social disaster by the end of the century" from population ageing.[Lockett, 1983 p9] A "Club for Research on Ageing" was formed by some of the participants who subsequently voted in 1945 to establish the Gerontological Society as a non-profit organisation.[Lockett, 1983 p17-18] She proposed that the "multidisciplinary voices of gerontologists", advocating national investment in a research agenda similar to that proposed at the 1938 meeting, were heard in the debate over the Research on Aging Act, 1974.

Chapter 2

THE ALZHEIMER'S ENTERPRISE

The redefinition of senile dementia

In 1907, Alois Alzheimer, a German neurologist, published the medical history of a woman with rapidly progressive dementing disease who was first examined by him when she was 51, and who died some four and a half years later. Based on the post-mortem pathological findings in her brain Alzheimer argued that "all in all, we have before us evidence for a specific disease process". He further argued that "it behoves us not to be satisfied with attempts to force .... clinically unclear observations to fit one of the disease categories familiar to us".[Alzheimer, 1907 p1-2] For seven decades thereafter his name was associated with a new "disease category", i.e. idiopathic "pre-senile dementia" occurring in a few unfortunate individuals at relatively young ages.

In 1974, at a medical symposium, it was formally proposed by a neurologist, Robert Katzman and T.B. Karasu, a psychiatrist, that "we should drop the term 'senile dementia' and include these cases under the diagnosis of Alzheimer's disease".[1] The rationale for this proposal was that "senile dementia and Alzheimer's disease are progressive dementias in which the pathological findings are identical ..... neither the clinician, the neuropathologist nor the electron microscopist can distinguish between the two disorders except by the age of the patient..... It is time to drop the arbitrary age

---

[1] The proposal by Katzman and Karasu was made in a paper presented in 1974 at the Houston Neurological Symposium, sponsored by the University of Texas Health Science Center at Houston, Texas. The proceedings of the conference were published the following year in: Fields, W.S. (ed.), Neurological and Sensory Disorders in the Elderly, 1975.

distinction and adopt the single designation, Alzheimer's disease".[Katzman & Karasu, 1975, p106-7]

In 1976, a medical journal editorial was published by Katzman entitled: "The prevalence and malignancy of Alzheimer's disease: a major killer". This summarised the arguments first presented at the 1974 symposium, including the proposition, which remains unsubstantiated to this day, that under the broadened definition incorporating individuals of all ages with dementing disease for which no specific aetiology can be determined, "Alzheimer's disease may rank as the fourth or fifth most common cause of death in the United States". The use of the term "malignant", which also appeared in the earlier presentation and publication, he justified by arguing that "patients with senile dementia" show a "marked decrease in life expectancy, depending on the age of onset of symptoms".[Katzman, 1976 p217-18]

"The Alzheimer's disease movement"[Fox, 1989]

Fox has documented a sequence of events initiated by Katzman in 1976 which he termed "the rise of the Alzheimer's disease movement". Based on a letter sent by Katzman to Donald Tower, director of the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS), together with the editorial described above, Tower "agreed that dementia was a public health problem" and proposed a conference to review the current status of research in the field.[Fox, 1989 p76] In 1977 the conference which was entitled "Alzheimer's disease - senile dementia and related disorders" was held at the National Institutes of Health. It was organised as a collaborative venture between the NINCDS, the National Institute of Mental Health and the National Institute on Aging, with the objective of providing information for "programming and allocation of resources" and also "to stimulate the interests of participants as well as others in the problem".[Fox, 1989 p77] In the same year Robert Butler, a psychiatrist who was appointed in 1976 to be the first director of the National Institute on Aging (NIA), adopted Alzheimer's disease as the "disease-based" research focus for the new institution.[Fox, 1989 p77-8] Subsequently, in a 1979 "research agenda" prepared by the National Institutes of Health, and published by the U.S. Department of Health, Education and Welfare, the study of "Alzheimer's disease and related dementias" was identified as one of its "priority areas". The arguments advanced for its inclusion by leaders of the NIA were that the disease "had major public policy implications and .... was a scientifically promising research area".[Fox, 1989 p90]

Based upon personal communication with Fox, Butler is quoted as stating that in order for the condition to be accepted as a "national priority" thereby securing government funding for its investigation, it was necessary to "make Alzheimer's disease a household word" and that the director focused his efforts both on "representing the NIA to Congress and in developing a public constituency to be an advocate for the Alzheimer's disease research cause".[Fox, 1989 p79-82] In pursuit of the latter objective, Butler and Katzman undertook to organise representatives of "grass-roots" self-help groups, which had been formed by caregivers and relatives of those with dementing disease and in a meeting held in 1979 at the National Institutes of Health a national "Alzheimer's Disease and Related Disorders Association" was established. As reported by Fox, Katzman became co-chairman of the Medical and Scientific Advisory Board and a member of the Executive Committee, and successfully advocated that efforts be targeted towards "public education and scientific research" in the field of Alzheimer's disease rather than to brain injuries or other conditions associated with signs of dementia in order not to "dissipate any potential resources we might accumulate".[Fox, 1989 p84] [2] A "consulting firm" was employed by the organisation to gain access to individual senators and congressmen,[Fox, 1989 p85] and a "plethora of (Congressional) hearings documenting the devastating effects of Alzheimer's disease, as well as the need for increased research funds" were conducted between 1980 and 1986.[Fox, 1989 p91] In addition an "information office" was established at the National Institute on Aging to provide briefings and written materials to the national media about Alzheimer's disease and any "major scientific accomplishments .... in terms that lay people could understand".[Fox, 1989 p89]

Prior to the mid-1970s Alzheimer's disease was a condition rarely diagnosed in clinical practice and a name unfamiliar to the general public or to their elected representatives but in what may be construed as the successful achievement of the strategy articulated by Butler, i.e. "to make Alzheimer's disease a household word" as a means to establishing it as a "national priority" for research funding, within a decade the name of the long-deceased German doctor was to appear in newspaper headlines,[The Washington Post, November 10, 1989, The New York Times, November 10, 1989] and public financing from the U.S. federal budget for the investigation of Alzheimer's disease and related disorders at the National Institute on Aging, rose from $1.5 million in 1977 to $80 million in 1989. This

---

[2] The national group was entitled successively the Alzheimer's Disease and Related Disorders Association, the Alzheimer's Disease Society, and is now the Alzheimer's Association.[Fox, 1989 pp80-2]

increase was largely attributable to specific legislation including the Health Research Extension Act, 1985 and the Alzheimer's Disease and Related Dementias Services Research Act, 1986. Under the 1985 Act (Public Law 99-158) Congress approved funding for ten "centres of excellence" across the country to "coordinate basic, clinical, social, and behavioural studies of Alzheimer's disease, create national standards for diagnosis, and establish a well-defined clinical population for future studies.[Fox, 1989 p94-6] Furthermore, the NIA has served as a conduit for financing investigation of Alzheimer's disease not merely by its own intramural and extramural researchers who are the primary beneficiaries but also by other constituencies of the NIH. By 1995 the estimated annual budget for Alzheimer's research was some $305 million of which $216 million was allocated for NIA studies with $46 million for those sponsored by the National Institute of Neurological and Communicative Disorders and Stroke, $25 million for the National Institute of Mental Health, and lesser sums for other institutions (see Table 2.1).

Theoretical perspectives on Alzheimer's disease

A range of sociological theories have been advanced for the study of ageing and interpretation of phenomena with which it is associated.[Bond et al, 1993 p31-50] These include: the structural "conflict" perspective deriving from the Marxist model of the differential distribution of power and authority within capitalist economies, a model which underlies the "political economy of old age" and the theory of "structured dependency" as articulated by Estes,[Estes, 1979, Estes, 1982] Myles,[Myles, 1984] Townsend,[Townsend, 1981] Walker,[Walker, 1981, Walker, 1990] and others (see Chapter 1); structural functionalist "consensus" theories which

Table 2.1

NATIONAL INSTITUTES OF HEALTH
ALZHEIMER'S DISEASE
(Dollars in thousands)

	1992 Actual	1993 Actual	1994 Actual	1995 Estimate	1996 Estimate
National Institute on Aging	$197,080	$202,744	$212,797	$216,360	$220,700
National Cancer Institute	501	593	564	486	494
National Heart, Lung, & Blood Institute	2,043	1,954	2,076	2,135	2,200
National Institute of Dental Research	367	0	0	0	0
National Institute of Diabetes & Digestive & Kidney Diseases	2,000	2,100	2,200	2,300	2,400
National Institute of Neurological Disorders & Stroke	38,844	42,266	44,460	45,883	47,349
National Institute of Allergy & Infectious Diseases	1,558	2,364	1,585	1,631	1,664
National Institute of Child Health and Human Development	713	266	392	400	400
National Eye Institute	115	1,444	1,335	1,381	1,427

Source: National Institute on Aging, Budget Office, 1995

Table 2.1 (contd.)

Alzheimer's Disease (con't)

	1992 Actual	1993 Actual	1994 Actual	1995 Estimate	1996 Estimate
National Institute of Environmental Health Sciences.......	424	97	356	356	369
National Institute of Arthritis & Musculoskeletal & Skin Diseases..........	1,758	1,625	1,434	1,500	1,550
National Institute on Deafness & Other Communication Disorders................	1,022	1,438	1,585	1,660	1,687
National Institute of Mental Health.....................	26,990	24,185	24,390	25,166	25,887
National Institute on Alcohol Abuse & Alcoholism.............	1,027	823	554	600	600
National Center for Research Resources..............	1,194	1,911	1,204	1,447	1,564
National Institute of Nursing Research................	3,188	3,369	3,122	3,230	3,370
Fogarty International Center....................	43	60	50	55	200
Total, NIH..............	278,867	287,239	298,104	304,590	313,857

Source: National Institute on Aging, Budget Office, 1995 focus primarily upon the role of older people within society; the unstructured sociological and social psychology perspectives of "symbolical interactionism" for analysis of the actions of individuals and groups; and "critical social science" theory which adopts an historical perspective on social reality to examine "the conditions under which our taken-for-granted ideas are created and developed" and to challenge the assumption that overall benefit to society necessarily derives from advances in science and technology.[Bond et al, 1993 p39-40]

Gubrium, in analysing the results of his observational study conducted among support groups affiliated with the Alzheimer's Disease and Related Disorders Association (ADRDA) adopted a further sociological perspective, i.e. that of ethnomethodology and phenomenology, interpreting the social world in terms of the meaning given by individuals to their own experiences and "biography". In his analysis Gubrium rejected the social constructionist model of "Alzheimer's disease" as employed by Fox (see above). He argued that "there is no need to resort to a gratuitous and difficult to prove theory of diagnostic machination to understand the growth of disease in general and Alzheimer's in particular" and rejected the notion that "disease as an entity and in its details is concertedly brought into existence because those who have special interests in it need it in order to apply their wares and make good on their concern, expert or otherwise". He posits that "Alzheimer's disease is more fundamentally understood as a cognitive, not a social movement" representing the "need for elaboration" of those concerned, either as professionals or as a labour of love or as a combination of both in attempting "to do something about the suffering believed to be caused by Alzheimer's".

He further posits that "the elaborations produced are integral parts of the related activity" whether or not they are underwritten by the force of vested interests".[Gubrium, 1986 p69-71]

It may be argued that Gubrium's critique of the constructionist model is persuasive insofar as it relates to the experiences and responses of family members and "semi-professionals" including nurses, recreation therapists and family counsellors concerned with the daily care of those with senile dementia. On the other hand his discussion does not explore the institutional, political and economic context within which this distressing condition was redefined as "disease", the initiating role of the medical profession in its "elaboration" to their peers and to the lay public including the U.S. Congress, nor the chronology of these activities. For example Robert Katzman is introduced into Gubrium's 1985 treatise as a "neurologist" who "has pointed out ..... Alzheimer's presenile dementia and senile dementia as essentially the same disorder, on both clinical and pathologic bases"[Gubrium, 1986 p7]; Robert Butler enters the narrative as "former director of the National Institute on Aging and a psychiatrist",[Gubrium, 1986 p34] and the ADRDA is first described as a "consumer-oriented source of written description" whose efforts since its "establishment in 1979" have resulted in part in "a growing public documentation of the disease".[Gubrium, 1986 p23-4] He later posits that "with the emergence of the Alzheimer's disease organisation, the cultural apparatus has come into place ..... to form and fill the public side of the disease". The "basic facts of the realities at stake" were sorted out "by the serious scientists and practitioners" and they were "hardly propagandistic", but the ADRDA became the "focal apparatus for formulating, embellishing and publicly diffusing ..... the concrete personal experiences of its members and spokespersons" as a "public culture".[Gubrium, 1986 p112-3] Further into his discussion Gubrium proposes that a "reified" National Institutes of Health "was persuaded to call a meeting" at which the ADRDA was founded, citing an observation that differences between delegates from the various support groups was resolved by "the voices of wisdom" of "such distinguished educators as Dr. Katzman ..... Dr. Butler and Dr. Tower of NIH".[Gubrium, 1986 p144] Therefore, Gubrium appears either to discount or to be unaware that, as well documented by Fox, the ADRDA did not "emerge" but was established by Katzman and Butler who sought out and invited representatives of self-help and support groups from around the country to the 1979 meeting at the NIH, and that Katzman, as co-chairman of the Medical and Scientific Advisory Board and member of the Executive Committee sought successfully to influence the organisation's priorities and agenda.[Fox, 1989 p82-5]

Gubrium argues that it "is the fashion in social sciences to describe realities as the products of concerted intention, even conspiracy" although he himself acknowledges that "there may be a measure of concerted intention and conspiracy in all domains of science and industry",[Gubrium, 1986 p206] and also that there are "clear elements" of a "social movement" in the ADRDA's "attempt to 'orchestrate' awareness and reform" on caregivers' behalf.[Gubrium, 1986 p71] It may be proposed, however, that it is not necessary to suggest that "diseases are made in Washington by medical interests" in order to perceive the outcome of the activities of individuals and groups who share a common "concern" as reflecting and institutionalising the realities of those with greater power and authority, and that a decade after Gubrium's study was published "Alzheimer's disease" also lends itself to analysis within both social constructionist and critical theory.

The "elaboration of disease" or a research "enterprise"?

From a phenomenological perspective it may be argued that Robert Katzman, an academic neurologist, having perceived congruence between the clinical and neuropathological signs of Alzheimer's disease and senile dementia, interpreted his observations as evidence of a continuum of "disease" which contrasted with the existing diagnostic dichotomy between early onset "pathology" and "normal" age-related change, and therefore proposed that the age distinction between the two conditions be dropped. In his 1976 editorial in a journal of neurology he elaborated on his "belief" that "senile as well as pre-senile forms of Alzheimer (sic) are a single disease" and argued that clinical, neuro-pathological and epidemiological investigations were indicated. His use of the "symbolic" word "malignancy" and his conjectures with regard to Alzheimer's disease under its expanded definition as a "major killer" and the "fourth or fifth most common cause of death in the United States",[Katzman, 1976 p217-8] may be perceived as a means of generating concern and interest among members of his own and related specialties in a condition previously largely ignored by the medical profession.

The original proposal to redefine senile dementia as disease was made jointly by Katzman, the neurologist, together with Karasu, a psychiatrist,[Katzman & Karasu, 1975 p106-7] and intuitively it might be assumed that for a "brain/behaviour disease",[Gubrium, 1986 p48] which is classified in medical nomenclature as a "mental illness" either the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) whose director Donald Tower was first approached by Katzman with what he perceived to be a "public health problem",[Fox, 1989 p76] or the National Institute of Mental Health would be appropriate institutions for financing its investigation. However, Robert Butler, the first director of the National Institute on Aging was himself a psychiatrist and his decision to adopt Alzheimer's as a "disease-based" research focus for the new institution may be interpreted from the perspective of the realities of his own professional "biography" and the difficulties he confronted with his appointment. These latter included: attempts by Congress, beginning in the late 1960s, to restrain growth in public expenditures including those for the National Institutes of Health; opposition to the enabling legislation for the National Institute on Aging from the Association of American Medical Colleges, representing the "external" clientel of the NIH, and also from the "intra-mural" community of primarily basic scientists, an opposition which may be interpreted as a self-interested response from both groups to sharing scarce resources with an increasing number of institutions (see Chapter 1). Furthermore, as previously observed, the holistic specialty of "geriatrics" has never been legitimated by the profession in the United States and the concept of a "national institute for ageing studies", first proposed in the 1950s "when the NIH had more money than it could spend", was opposed by its director who argued that the conditions of old age were already under investigation by existing institutes dedicated to research into cancer, heart disease, diabetes, etc.[Lockett, 1983 p38-9]

It may therefore be proposed that, based upon his own experience which combined psychiatry with gerontology, senile dementia was already a "concern" to Butler,[3] and

---

[3] In 1975 Robert Butler had published a book entitled "Why Survive?: Being Old in America, New York, Harper & Row, 1975" which was critical of public policies and perceptions of ageing. Fox argues that it was this work that commended him to members of the appointment board of the National Institute on Aging.[Fox, 1989 p74]

that stimulated by the interest of Katzman in a disorder which is unequivocally associated with advancing age he perceived it to be a legitimating focus for the National Institute on Aging since dementing disease had previously attracted limited research interest and its investigation would not duplicate the efforts of other biomedical researchers.[4] Moreover, the legitimacy of his institution might be further enhanced by securing earmarked funding from the U.S. Congress for study of Alzheimer's disease, rather than competing for resources with other institutions of the NIH. To this end the definition of senile dementia as a "distinct disease entity separate from the varied experiences of normal ageing"[Gubrium, 1986 p3] accorded with the preference of congressmen for funding applied research into specific "diseases of man" rather than microbiological and related basic science studies.[Strickland, 1972 p192] Moreover, a precedent had been established in the 1970s for securing funding additional to the global NIH budget with passage of the National Cancer Act, 1971,[Rettig, 1977] (see Chapter 1).

Strickland in his discussion of "politics, science and dread disease", observes that the U.S. Congress has tended to be responsive to "lay advocacy" by individuals suffering from specific illnesses, and representatives of voluntary associations organised on their behalf, by increasing resources via the NIH not merely for study of cancer which is the prototypical "dread disease", but also for neurological diseases, open-heart surgery, and

---

[4] Fox notes that research into dementing disease was supported by the National Institute for Neurological and Communicative Disorders and Stroke before it was adopted by Butler for the National Institute on Aging but there is no evidence that it was a research "priority".[Fox, 1989 p93]

other conditions and technologies.[Strickland, 1972 p142-3] [5] Since no organisation existed to represent the interests of those with dementing disease who are incapable of representing themselves, the initiative of Butler and Katzman in forming a national association to encourage a "public culture", rooted in the experiences of those who have known or are undergoing the "tribulations" of caring for the afflicted may be interpreted as an expression of their "concern", based on the reality of the need to compete for medical research resources with other interest groups, a reality which must be continually reinforced. Gubrium, from his observations of the ADRDA and its members writes that "those concerned are reminded repeatedly that Alzheimer's is a disease for which there is no cure ...... (and) medical researchers, service providers, and the ADRDA itself, remind patients, caregivers and others that there is only one source of hope for a cure and that will be provided when it comes, by medical research". To this end the ADRDA serves as an advocate on behalf of the research community for federal government resources and also raises and distributes private funds with its own "research grant program".[Gubrium, 1986 p105-6] On the other hand it may be unnecessary for the concerned to be "reminded repeatedly" that hope for a "cure" lies in research in order to be motivated to generate funding for the biomedical community. In a later paper, Gubrium

---

[5] Cancer was the first non-communicable disease to attract public funding for research in pursuit of a "cure" for what was described by one U.S. senator in the 1920s as a "loathsome, deadly and insatiate monster".[Strickland, 1972 p1] The National Cancer Institute, was the first of the National Institutes of Health dedicated to a specific condition, the 1937 legislation for its establishment setting a precedent in providing for the distribution of public funds for extramural research (see Chapter 1).[Stevens, 1971 p358] Conquest of the "dread disease" was also the "crusade" initiated by members of the medical profession which succeeded in subverting efforts by legislators in the late 1960s to restrain the growth in public expenditures including those for the NIH with passage of the National Cancer Act, 1971.[Rettig, 1977 p277]

distinguishes between what he defines as the "ameliorative mode" which "structures" the disease experience in the developmental logic of "doing something about it", and the "tribulation mode" which "destructures" the illness in "frustration over the lack of perceived order in what is happening".[Gubrium, 1987 p16-17] From this perspective it may be argued that the "ameliorative mode" finds expression in support for research and that this is congruent with a pervasive "belief" in the potential of medical science which finds expression in deference to the authority of the profession in determining policies and priorities in relation to health care and the investigation of "disease".[6]

Gubrium posits that the ADRDA "seeks funds for research because there are

---

[6] As previously discussed (see Chapter 1), Starr argues that American medicine has achieved an unequalled position in the social structure, combining "broad cultural authority, economic power and political influence". He proposes that authority reflects the probability that "particular definitions of reality and judgments of meaning will prevail as valid and true" and that the profession achieved its authority based in part on its "bonds with modern science" and the privileged status of scientific knowledge in the hierarchy of belief".[Starr, 1982 p4-6] Thus it was proposed that legitimation of its place within the "universe maintaining" conceptual machinery of "science",[Berger & Luckmann, 1966 p110-12] had been institutionalised in the National Institutes of Health (NIH), agencies of the U.S. Public Health Service. On the other hand, while there has been sustained support for public investment in medical research in the post-War period, fueled by evidence that "the magic of science and money had worked" as with the successful development of vaccines for poliomyelitis, albeit with financing from charitable contributions,[Starr, 1982 p347] some have questioned the lack of public accountability and the distribution of benefits derived from some eighty percent of the NIH budget which is reallocated to academic medical centres, the recipients of which "enjoy autonomy within the constraints of professional competition".[Starr, 1983 p343] Stevens observed that where NIH grants resulted in the successful development of new diagnostic or other techniques doctors were not challenged to promote greater public access,[Stevens, 1972 p352] and Strickland defined the NIH as the "national medical research enterprise", the "public health" benefits of which would be dispersed through "a kind of trickle-down theory" in the absence of universal health insurance.[Strickland, 1972 p259] Therefore whatever progress may be achieved in the understanding and management of disease, public investment in medical science may also be perceived as reinforcing the power and authority of American medicine to promote or subvert policies related to health care in the entrepreneurial self-interests of its members, as in the institutionalisation of the "market" model for health care in the Medicare programme.

apparent pathologies to investigate. As a result, what is revealed are the disease's concrete lesions and operating systems, whose prevalence demands resources for further investigation".[Gubrium, 1986 p207] However, although fund raising for research by family support groups may be interpreted as "ameliorating the disease experience",[Gubrium, 1987 p16] in the absence of a "cure" there continue to exist the immediate and ongoing financial, physical and emotional needs of those afflicted and those who care for them, but there is little evidence that the "concern" of the "serious scientists and practitioners" extends to elaborating the disease from the perspective those currently experiencing its tribulations. On the contrary, as documented by Fox, when the ADRDA was first established a "broader focus" was sought by those with "personal experiences of caring for relatives with cognitive disabilities and a desire to obtain assistance for people with brain impairments and their caregivers", but "as a researcher, Katzman advocated targeting scarce resources that would benefit the cause of biomedical research",[Fox, 1989 p85] and whatever may have been achieved by the activities of the ADRDA in promoting service provision at the state and local level, federal government policy and financing has since been directed to "research".[7] Furthermore it has been posited by a former

---

[7] In 1987 the research office of the U.S. Congress published a report entitled "Losing a Million Minds: Confronting the Tragedy of Alzheimer's Disease and Other Dementias" in which testimony given by representatives of the ADRDA is cited. The document was produced out of "Congressional concern with the plight of those suffering from dementing disease",[U.S. Congress, 1987 piii] and the report itemises a list of "federal policy priorities" of which the first was "support for biomedical research", but "financing long-term care" was also included.[U.S. Congress, 1987 p5] "Federal policy options" for assisting with long-term care included: "making dementia a presumptive cause of disability for Medicare", i.e. extending eligibility below the age of 65 for early-onset dementia, and modifying the "Medicare benefit package", which is directed primarily to acute and tertiary care, to include "adult day care, personal care, chore services, attendant care".[U.S. Congress, 1987 p467] None of these policy options have ever been implemented.

President of the Gerontological Society of America that the focus on Alzheimer's studies has diverted resources away from the adoption of a broader national agenda in the field of ageing research,[Adelman, 1995 p526-32] (see below). He concurs with Fox in interpreting the activities "orchestrated" by Katzman and Butler as a "movement", one which is, arguably, congruent with the constructionist model of "moral entrepreneurship".

In a discussion which he termed the "social construction of illness" Freidson argued that as a "consulting" profession, "medicine is committed to treating rather than merely defining and studying man's ills" and that it therefore has "a mission of active intervention guided by what, in whatever time and place it exists, it believes to be ill in the world". The profession does not merely treat the illnesses laymen take to it, but also seeks to discover illness of which laymen may not be aware, one of the greatest ambitions of the doctor being to discover and describe a "new" disease or syndrome and to be immortalised by having his name used to identify the disease. He further argued that "insofar as illness is defined as something bad - to be eradicated or contained - medicine plays the role of moral entrepreneur". Freidson posited that there is a division of labour in this "moral entrepreneurship": when in the course of practice a doctor finds illness of which an individual is unaware, his task is "modest and unassuming". On the other hand the "major moral entrepreneurs" in medicine are those seeking to influence public opinion and political policy in different ways. These include public spokesmen for the organised profession or its specialties, seeking "to alert the public to the important dangers of a given disease or the virtues of a given kind of health"; representatives of medical institutions or organisations whose activities are primarily political and are directed towards implementing "measures designed to improve what they see to be the public health"; and "special lay interest groups sometimes led by physicians but always including at least one prominent physician, which crusade against the menace of its own specially chosen disease". The latter groups seek public "attention and resources" with objectives that include: applying the label of disease "to conditions not considered illnesses before"; "removing the stigma of some diseases by changing their labels", and redefining an illness so that it moves in the public mind from "chronic, serious or incurable to ... curable or controllable".[Freidson, 1973 pp252-4]

It may be argued that all the types and objectives of "major" moral entrepreneurship described by Freidson are identifiable in the redefinition of senile dementia as Alzheimer's disease and the activities which were subsequently pursued: Katzman fulfilled the ambition of the physician to "discover and describe a new disease", although he did not assign his own name to the redefinition of one by which an earlier physician had already been "immortalised". He and Robert Butler were, however, active in both founding and leading a lay interest group in a crusade against the "menace" of this new illness. Changing the label of "senile dementia" to that of "disease" and promoting discussion and a "public culture" may be interpreted as actions designed to diminish if not remove the stigma. Furthermore, the adoption of Alzheimer's disease by the National Institute on Aging (NIA) as a research focus for the institution and the advocacy of the ADRDA under the guidance of Katzman and Butler may be perceived as a process in which the condition moved in the public mind from "incurable" to "curable or controllable", a move that found expression in legislation enacted by the U.S. Congress to allocate resources for its study. Therefore, on the one hand the success achieved by Katzman and Butler may be perceived as deriving from activities undertaken by "serious scientists and practitioners" which reflected "the basic realities at stake". Gubrium, 1986 p113 or Alternatively, the selection of Alzheimer's disease as a focus for the NIA and the raising of dedicated funding for its investigation may be interpreted as an expropriation by the medical profession of its multi-disciplinary remit[8] in favour of one "disease" of ageing for which neuropathological studies may yield hope of a "cure" but which, in turn, legitimate a range of clinical diagnostic and epidemiological research conducted to "quantify" the problem and to seek "causal" associations within a positivist paradigm which is compatible with basic and clinical science perspectives.

Alzheimer's disease as a "physical reality" and "neuroscientific" research

In pursuing his discussion of the social construction of illness Freidson proposed that the medical profession in the course of achieving a monopoly over the content and practice of its work has gained "well nigh exclusive jurisdiction over determining what illness is ... with the authority to label one person's complaint an illness, and another's complaint not." He further argued that like law and religion, the other traditionally

---

[8] As previously discussed (see Chapter 1) the enabling legislation for the National Institute on Aging (NIA) was passed in the face of protracted opposition from interest groups within the academic and clinical medical communities but with strong advocacy from other groups including the Gerontological Society of America. Lockett proposes that the deciding factor in the debate was the argument that "biomedical research could help prevent old age disabilities and save taxpayers' money", i.e. from the open-ended financing of the Medicare programme. Lockett, 1983 p182 To this end the NIA was established with a mandate to "coordinate and promote research into the biological, medical, psychological, social, educational and economic aspects of ageing". Lockett, 1983 p5

autonomous professions, medicine uses normative criteria to select its fields of interest, so that its work "constitutes a social reality that is distinct from, and on occasion virtually independent of, physical reality".[Freidson, 1973 pp205-6] Alois Alzheimer first linked early onset dementia and *post mortem* pathology as evidence of a "specific disease process",[Alzheimer, 1907 p1-2] and therefore Alzheimer's disease may not be construed merely as a medically constructed "social reality". However, although the redefinition of senile dementia as Alzheimer's disease, based upon perceived congruence between pathological and clinical findings for the two disorders, has enabled those "concerned" whether as researchers or caregivers to elaborate on "senility" as "disease" rather than "normal ageing", Gubrium observes that "the factual status of Alzheimer's as a distinct entity separate from the varied experiences of normal ageing is not obvious".[Gubrium, 1986 p3]

There is evidence that from the beginning pathologists were not convinced that the "tangled bundle of fibrils" and "miliary foci" described by Alzheimer,[Alzheimer, 1907 p2] (later termed neurofibrillary tangles and senile plaques), were unique to the condition associated with his name. Kraepelin, an "organic" psychiatrist and director of the institute in which Alzheimer practiced, wrote in a treatise on "senile and pre-senile dementias" that it was an open question as to "whether we are dealing with a singular disease or only with a stage of a different disease", noting points of similarity in the clinical and pathological signs associated with both conditions.[Kraepelin, 1910 p2] In 1968, six years before Katzman proposed the redefinition of senile dementia as Alzheimer's disease, the results of a longitudinal study conducted by Blessed et al. under the sponsorship of the United Kingdom Medical Research Council were published and in the introduction to the paper the authors discussed the physical changes in the brain associated with dementia, as described both by Alzheimer and by others, dating back as far as 1892. They argued that there were difficulties in interpreting these reports since "far from being specific, such changes could be demonstrated in the brains of well-preserved old people". Furthermore, others who had compared clinical signs with autopsy results had noted that "senile dementia could occur without the presence of senile plaques".[Blessed et al., 1968 p797]

Blessed and colleagues sought to resolve the "brain/behaviour" relationship by matching post-mortem findings with *in vivo* "dementia scale" ratings, based on the assessment of behavioural and intellectual phenomena, from a sample of older people with signs of dementing disease and also with autopsy samples from an additional group with no similar evidence of impairment. They found a "general tendency for functional incapacity to be correlated with the number of senile plaques", but although the "association between measures of dementia and mean plaque count was generally close, in a few cases some degree of discrepancy between the two measures was noted." They therefore raised the possibility that plaque count, i.e. observable anatomical change, was a crude "quantitative" measure of "qualitative differences ... inaccessible to present methods of examination". They also noted in their discussion the results of then recent electron-microscopy studies of the previously unvisualised anatomy of the brain.[Kidd, 1964] [Terry et al, 1964] These indicated that plaque formation is secondary to degenerative nerve cell processes and suggested that differences between demented and "well-preserved" old people may lie in a different rate of progression of "one and the same process".[Blessed et al, [1968 p808] A longitudinal study published in 1988 by Crystal et al. used electron-microscopic techniques but advanced little on the earlier findings. Written under the sub-title: "non-demented subjects with pathologically confirmed Alzheimer's disease", the group compared dementia measures "with *post mortem* neuropathology, including senile plaque and neurofibrillary tangle counts" and found changes characteristic of the "disease" in the brains of six out of nine non-demented subjects. Furthermore, although all those with a large number of cortical neurofibrillary tangles were demented, seven of nineteen demented subjects had "numerous senile plaques but no senile tangles". The group concluded that for subjects aged over 75, "quantitative pathological criteria for Alzheimer's disease" did not distinguish between demented and non-demented individuals.[Crystal et al, 1988 p1682-7] In a recent autopsy study by Nagy et al. of the brains of "49 prospectively evaluated patients with Alzheimer's disease" quantitative analyses were performed to determine the "relative roles of plaques and tangles" in the dementing process using three different sets of published pathological criteria. Their results differed somewhat from Crystal et al. in finding dementia to be consistently related to a combination of plaques and tangles. They distinguished between and highlighted the "irrelevance" of "amyloid plaques", i.e. random deposits of starch-like protein, to the stage of the condition *in vivo* and the "relevance" of "neuritic plaques", in which amyloid protein combines with the debris of degenerative nerve cells, the highest correlations of neurofibrillary tangles being with neuritic plaque densities. The investigators also identified variations in pathological changes between different lobes of the brain but "found no pathological measure which correlated with the age of patients at death".

They observed, however, that the contribution made by the study to refining pathological diagnostic criteria was limited by the lack of inclusion of "undemented" subjects.[Nagy et al, 1995 p21, 29] Recently, McLoughlin and Lovestone proposed that advances in understanding the molecular biology of "amyloid precursor protein" and its "neurotoxic potential" may demonstrate a functional link to explain the apparent lack of anatomical correlation between plaques and tangles and the variations in and between study results.[McLoughlin & Lovestone, 1994 p439] Therefore, in spite of refinements in analysing the microanatomy of the ageing and/or dementing brain over the past thirty years and in understanding the molecular and physiological changes with which they are associated, there are unresolved dilemmas with regard to the pathogenesis of dementia and in consequence therapeutic or preventive measures remain elusive.

In 1977 the research group of which Blessed was a member published evidence from a study in which post-mortem tissue samples from both "normal patients" and "patients with depression and dementia" were examined. This revealed neurochemical changes in the brains of those with dementia which were related to the degree of senile plaque formation. It was proposed that the findings were evidence for encouraging investigation of "brain biochemistry" which might provide "a basis for therapeutic regimens".[Perry et al, 1977 p189] In a follow-up paper, evidence for the model, defined as the "neurotransmitter" theory of senile dementia, was enlarged upon and a recommendation was made for further research to explore the potential for "neuro-pharmacological manipulation",[Perry, et al, 1978 p1459] i.e. the development of drugs to compensate for the biochemical changes or imbalance found to be associated with dementia. The specific neurotransmitter identified as deficient by Perry et al. was acetylcholine but, as observed by Wurtman in 1985, deficits in other transmitters were later detected indicating that neuron loss is not selectively restricted to those releasing one specific agent. Wurtman discussed this and other conceptual models for Alzheimer's disease including: the "infectious agent model" derived from studies of Creutzfeld-Jacob disease; the "toxin model" based upon finding high concentrations of aluminium in neurofibrillary tangles; the "genetic model" in which disease may relate to an inborn error of metabolism or to the inherent vulnerability of the individual to environmental factors; and the "abnormal protein" or "amyloid deposition" model. He suggested that just as a range of similar models were developed and found to have validity in the genesis of cancer none could be entirely discounted.[Wurtman, 1985 p62-74] However, McLoughlin and Lovestone in 1994 proposed that biological and genetic research indicate that "amyloid deposition is likely to be fundamental in all familial and sporadic cases of Alzheimer's disease" and that the "amyloid cascade hypothesis" in which there is genetically determined triggering of neurotoxic protein deposits,[Hardy & Higgins, 1992 p184-5] "remains the best model of the pathogenic process". On the other hand, they observe that speculations with regard to potential "clinical benefits" are "almost destined to prove inaccurate". Furthermore, although genetic findings may improve "diagnostic accuracy", developing treatments based upon molecular biology is a "distant prospect".[McLoughlin & Lovestone, 1994 p439-40]

Alzheimer's disease was selected by the director the National Institute on Aging as "a scientifically promising research area" in the year in which the "neurotransmitter" theory of senile dementia was disseminated, and the efforts undertaken in conjunction with Robert Katzman to achieve funding for its investigation included the establishment of the ADRDA as a lay advocacy group.[Fox, 1989 p90] As previously discussed (see above), Gubrium observed that patients, caregivers and others are repeatedly reminded by medical researchers, service providers and the ADRDA that "the only hope for a cure will be provided, when it comes by medical research".[Gubrium, 1986 p106] However, in the same section he also proposed that physicians frequently perceive caregivers as "overly concerned with medical breakthroughs and cures".[Gubrium, 1986 p92] It may be argued that from the constructionist perspective it was members of the medical profession in the role of "moral entrepreneurs" who raised expectations by moving senile dementia "in the public mind from chronic, serious or incurable to ..... curable or controllable",[Freidson, 1973 p254] and nearly twenty years after Alzheimer's disease was adopted as a focus for the NIA those "concerned" as patients and caregivers in the succeeding decades have had little alternative to placing their hopes in scientific breakthroughs even though effective treatments remain a "distant prospect".[9] In the meantime, while neuroscientists relate the potential for therapy to "diagnostic accuracy" in terms of genetics and microbiology, based primarily on research conducted outside the United States, the American medical profession has been supported in clinical and epidemiological studies which legitimate the

---

[9] At the time of writing, one "neurotransmitter" drug has been approved by the U.S. Food and Drug Administration for those diagnosed with "Alzheimer's disease". This is Cognex (tacrine) based on the cholinesterase inhibitor tetra-hydroaminoacridine. One double-blind, placebo-controlled clinical trial indicated a very small effect for the drug in reducing the rate of decline in the dementing process.[Davis et al, 1992] However many subjects show no benefit from the drug, there are significant side effects including nausea and vomiting and liver dysfunction, and one Australian study recommended that "tacrine not be used widely in the treatment of dementia".[Davies et al, 1989 p164]

imputation of an incurable disease, for which the "functional link" has not been fully established between behaviour and pathological changes in the brain, or between the latter and those of "normal ageing", to an increasing proportion of older people.

The diagnostic enterprise

Robert Katzman's proposal to "drop the term senile dementia" and adopt the single designation "Alzheimer's disease" was made originally in 1974, in conjunction with T.B. Karasu, a psychiatrist,[10] some three years prior to publication of the "neuro-transmitter" theory of dementia with its therapeutic potential (see above). In the Katzman and Karasu paper it was asserted that "senile dementia is a malignant disease"[Katzman & Karasu, 1975 p105] a proposition that, based upon the elimination of any age distinction in diagnosis, was reiterated by Katzman in his 1976 editorial on "the malignancy of Alzheimer's disease", [Katzman, 1976 p217] (see above). In the first publication Katzman and his co-author elaborated on the analogy between dementia and the disease synonymous with malignancy, i.e. cancer, observing that "our colleagues in oncology, faced with a malignant disorder, show little reluctance to proceed with much more difficult and dangerous tests than those used by neurologists and psychiatrists". Therefore in addition to "the electro-encephalogram, skull films and brain scan" they advocated the use of a range of "invasive" diagnostic measures including "the lumbar puncture, the infusion manometric test, the isotope cisternogram and the pneumo-encephalogram". The rationale for these investigations was to discriminate between conditions which may become manifested by evidence of dementia, with the objective of detecting some potentially treatable or

---

[10] See footnote 1.

"reversible" disorder such as a non-malignant brain tumour.[Katzman & Karasu, 1975 p128] The paper by Katzman and Karasu was entitled "the differential diagnosis of dementia" and the related concepts of "differential diagnosis" and "reversible dementias" have been extensively invoked in clinical practice and research in the ensuing years. In 1980, for example, the recommendations of a multi-specialty task force established by the National Institute on Aging to consider "treatment possibilities for mental impairment in the elderly" were published which listed sixty-one "reversible causes" to be considered and investigated for all subjects with dementia.[The National Institute on Aging, 1980 p260-2] [11] In 1987, a "Consensus Development Conference on the Differential Diagnosis of Dementing Disease" was sponsored jointly by the National Institute on Aging (NIA), the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS), and the National Institute of Mental Health. In the "consensus statement" which was disseminated at the end of the conference, a wide range of testing was again advocated for all subjects and there was discussion of the use of new techniques including magnetic resonance imaging (MRI) and measurement of regional blood flow and metabolism with positron emission tomography (PET) and single photon emission computed tomography (SPECT) for "improvement in accuracy of diagnosis".[National Institutes of Health, 1987 p6-7] In addition, in 1984, the report of a task force was published which was established by the

---

[11] The report proposed that the "aged brain is extremely sensitive to the internal environment" and that "almost any disorder that alters this environment - cardiac, pulmonary, renal or hepatic failure, endocrine disorders, water and electrolyte disturbances, anoxia, anaemia, infections, nutritional deficiencies, and hypothermia or hyperthermia", together with lesions in the brain, "various therapeutic and non-therapeutic chemical substances", and "anaesthesia or surgery" may all induce "impairment".[The National Institute on Aging, 1980 p260-2]

NINCDS and the Alzheimer's Disease and Related Disorders Association (ADRDA), under the auspices of the U.S. Department of Health and Human Services, to develop "clinical criteria" specifically for the diagnosis of Alzheimer's disease. The task force, of which Robert Katzman was a member, listed a range of indicators for a "probable", "possible" and "definite" diagnosis, the latter category being "confirmed" by "histopathologic evidence obtained from a biopsy or autopsy", an assertion which may be questioned in view of the neuroscientific findings (see above). The primary physical criteria for the diagnosis were the "absence of systemic disorders or other brain disease" that might account for signs of dementia, and an extensive list of conditions was designated for "exclusion" by "examination of body fluids and non-neural tissues" as well as advanced technological investigation of the brain.[McKhann et al, 1984 p939-41]

Johnson in a discussion of "professions and power" argued that professional status may be analysed as a process in which occupational groups engage in an ideological struggle to legitimate a series of "claims". Among "practitioners who personally confront laymen as an essential part of their work and therefore need to have their expertise taken for granted", the "claim" to autonomy or independence as a pre-condition for fulfilling obligations to consumers, gives pre-eminence to "the diagnostic relationship". This relationship then serves as a control mechanism, the diagnosticians assuming an authoritative role *vis a vis* their clients, while subordinating emergent and potentially competing providers to their control and direction. Thus, by legitimating the diagnostic relationship and invoking it to maintain pre-eminence, related occupations fail to achieve "autonomy" and may therefore carry out treatment only in respect of a prior diagnosis by a physician.Johnson, 1972 p57-8

Freidson proposed that "institutionalised expertise", derived from "demonstrable scientific achievement", has legitimated the exclusive jurisdiction achieved by the medical profession over determining what constitutes illness or disease, and that the jurisdiction thereby established extends far beyond its capacity to "cure".Freidson, 1973 p251 It may also be proposed that "scientific achievement" in medicine has frequently related to diagnostic rather than "curative" technologies developed by scientists and "inventors", not all of whom are or were members of the medical profession, but whose work reinforces physician authority.[12] [13] Furthermore, in spite of advances in the understanding and management of disease, it may be argued that for some specialties such as neurology, dedicated to conditions affecting the brain, spinal cord and other structures of the nervous system, therapeutic capabilities continue to be exceeded by the authority imbued in the "diagnostic relationship". Therefore on the one hand, from Gubrium's perspective the introduction of the concepts of "differential diagnosis" and "reversible dementias" by Katzman and Karasu, representatives of neurology and psychiatry, may be interpreted

---

[12] Kuhn noted that X-rays were discovered "through accident" by a physicist, Roentgen, the type of discovery "that occurs more frequently than the impersonal standards of scientific reporting allow us to realise".Kuhn, 1962 p57

[13] Starr posited that the introduction of diagnostic aids in the 19th century such as the stethoscope, the laryngoscope to inspect the throat, the ophthalmoscope to examine the eye, all strengthened the authority of the doctor, reduced dependence on the patient's statement of symptoms, and "increased the asymmetry of information". The addition of technologies such as the X-ray and the electro-cardiograph, which generated data on a patient's condition "seemingly independent of the physician's as well as the patient's subjective judgment", and made possible the presentation of the results of these examinations to other members of the profession promoted the claim to "objective judgement" in the diagnostic process.Starr, 1982 p136-7

as expressions of "concern" with understanding a condition not previously well investigated, the diagnostic testing of which brings reassurance that "something is being done" by applying scientific methods to those afflicted. On the other hand, promotion of the same concepts may also be perceived as an exercise in "expansionary control" by clinical medicine over a new clientel.

In a 1974 discussion of American health care and social control it was proposed by Ehrenreich and Ehrenreich that medicine has both "exclusionary" and "expansionary" dimensions, representing two different forms of control. "Disciplinary control" includes adoption by the medical profession of attitudes and behaviours which discourage those unable to pay from seeking treatment, and results in the geographic isolation in institutions of those such as the "insane" with conditions which are not amenable to conventional medical interventions. In contrast with these exclusionary methods, "co-optative" or expansionary control is exercised over those with market access, by means such as expanding the definitions of illness, promoting the use of medical services and technologies, and fostering dependency on the system.[Ehrenreich & Ehrenreich, 1974 pp28-9] Katzman and Karasu's initiative in encouraging aggressive diagnostic investigation of those with senile dementia, a group which it may be argued had traditionally been subject to "exclusionary control", did not occur until after the retired population achieved universal access to the market via the Medicare programme in the late 1960s and it is unclear that the clinical course of senile dementia/Alzheimer's disease is indistinguishable from treatable conditions either with or without the employment of high technology.

It has been widely argued that the indemnity medical insurance model which pays for each physician consultation with additional fees for all procedures undertaken incurs the risk of "moral hazard" on both the demand and the supply side,[Culyer, 1993 p162-3] and that its adoption for Medicare has been a "giveaway to the providers",[Starr, 1982 p382] promoting increases in "the use and intensity of services"[Levit, 1991 p39-40] (see Chapter 1). Gubrium acknowledges that with the high cost of medical care questions have been raised with regard to "the widespread use of expensive diagnostic tools" but he defers to medical authorities who argue that "the cost of a CAT scan is less than a one-days stay in hospital and does serve to detect a number of reversible and/or treatable dementias ..... which can present as dementia (*sic*)".[Gubrium, 1986 p11] On the other hand, as documented in the clinical criteria for Alzheimer's disease, it is characterised by an "absence" of physical signs and symptoms such as visual disturbances, incoordination, paralysis or loss of sensation associated with other conditions which may present as dementia (see above).[McKhann et al, 1984 p940] Furthermore, in the absence of brain tumours, changes in the brain revealed by technologies such as CAT scans and magnetic resonance imaging are not necessarily distinguishable from "normal ageing",[Jacoby et al, 1980 p256, Buonanno, 1987 p58] and even where other conditions which enter into the differential diagnosis such as Creutzfeld-Jacob and Huntington's disease are detected, few are treatable or reversible.

The various task forces of the National Institutes of Health acknowledge the relative lack of therapeutic potential associated with physical evidence derived from extensive clinical testing, the cumulative cost of which for any subject may far exceed a one-day stay in hospital, since the vast majority of those subjected to a full range of diagnostic investigations will be ascribed a diagnosis of senile dementia/Alzheimer's disease,[National

Institutes of Health, 1987 p4 or another of the "irreversible dementias".National Institute on Aging, 1980 p260 Therefore, although such measures may bring at least temporary hope to patients and caregivers that the latest technologies will reveal some "curable" condition, it may be argued that the benefits accrue primarily to the clinical medical community from the consultations and testing conducted within a network of specialist referrals in what may be perceived as a "diagnostic enterprise" which includes economic incentives for lowering the threshold for testing, the intensity of which is, arguably, legitimated by the concept of "malignancy". On the other hand, although it is "uncertainty" that has been proposed by the various task forces established by institutions within the NIH as the rationale for extensive clinical investigations, an uncertainty which analysts such as McLoughlin and Lovestone argue will be resolved by genetic and molecular research,McLoughlin & Lovestone, 1994 p439-40 (see above) nevertheless this has not impeded the establishment of diagnostic criteria for determining prevalence estimates of Alzheimer's disease or related phenomena in community studies.

Documenting the "silent epidemic"

Gubrium observed in 1986 that "formal descriptions of Alzheimer's disease are commonly introduced with alarm. ...... The demographic urgency of the Alzheimer's disease problem is perhaps best captured by one of the most frequently used slogans of the related movement's promotional literature: the 'silent epidemic'. The slogan ..... informs us of two alarming facts. One is that Alzheimer's disease is something that is not fixed but is spreading in the population; the disease problem is getting worse. The other is that until now its enormity has gone unnoticed". He argues that "though some persons may be concertedly rhetorical in presenting the facts of Alzheimer's disease, others are seriously concerned with what they believe to be the realities at stake".[Gubrium, 1986 p34-5] However, with the publication in 1989 of the results of a study initiated and funded by the National Institute on Aging (NIA) and their dissemination by its "information office" to the national media who reported the findings under headlines such as "**Alzheimer's Toll Heavier: Illness May Affect 4 Million in U.S.**" the "alarming facts" of the "realities at stake" were confirmed, executives of the ADRDA being quoted as observing that "the financial and health implications are remarkable" and "the impact could eventually affect us all"[The Washington Post, November 10, 1989] (see Chapter 1).

Epidemiology derives from a recognition in the 17th century of the value of "routinely collected data in providing information about human illness" but the range of epidemiological studies and data sources have extended far beyond the collection of statistics on births and deaths to include disease-specific episodes of ill health as documented by doctors and hospitals, intervention studies or "clinical trials" to determine the risks and benefits of therapies, and population "screening" for disease.[Hennekens et al, 1987 p3-5] Writing in 1964 on the "uses of epidemiology", J.N. Morris argued that "medicine is social science as well as human biology, and the epidemiological is the main method of studying the social aspects of health and disease". Epidemiology provides "intelligence" for the health services: the nature and relative size of problems are described and an indication given of their importance to the community. However, he also observed that in moving beyond simple numerical counts of "vital statistics" "morbidity means what it is defined to mean".[Morris, 1964 p277 & p40] Similarly, in a recent work on the study of ageing, Bond et al. argue that epidemiological studies are necessary for "the planning of services and policy making", but that "it is obviously important to have sound diagnostic criteria". On the other hand they also suggest that such research "can of itself improve the instruments used to detect disease or to measure change over time" and that it can "enable the identification of 'risk factors'".[Bond et al, 1993 p41] Bond's discussion referred specifically to the diagnosis of dementia, the clinical syndrome indicative of Alzheimer's and, by definition, other dementing diseases.

The study from which data on unexpectedly high rates of Alzheimer's disease were derived was entitled "Established Populations for Epidemiologic Studies of the Elderly (EPESE)", designed as a longitudinal investigation of individuals aged 65 and over living in three, and later four, communities in the United States "specifically to produce estimates of the prevalence and incidence of chronic conditions, impairments and disabilities with their associated risk factors".[Cornoni-Huntley et al, 1986 p1] The method combined a lengthy questionnaire to determine self-reported disease, disability, and health risks from smoking and obesity, together with testing or "screening", for high blood pressure, depression and "mental status".[Cornoni-Huntley et al, 1986 p130] The assessments of mental status from one of the study sites, East Boston, an ethnically homogeneous "working class neighbourhood" in which the study instruments were translated for those "who usually spoke Italian", were used to select cases for further "neuro-psychological evaluation", a brief review of medical history and blood tests in order to make a diagnosis of Alzheimer's disease. No attempt was made to assess the extent to which individuals were "impaired" by the "disease" since it was "difficult to apply this criterion in a uniform, meaningful way".[Evans et al, 1989 p2552-3] In view of the failure of neuroscientists to establish a functional link between the brain and behaviour, and the extensive clinical testing advocated in order to exclude other conditions from the diagnosis, questions may be raised as to how morbidity was "defined" in this study which detected Alzheimer's disease in forty-seven percent of those aged 85 and over; whether the diagnostic criteria were "sound", and whether the instruments used in the screening process had been validated and "refined". Furthermore, the EPESE study conducted by the Epidemiology, Demography and Biometry Programme of the NIA may be perceived as largely duplicative of the research of other government institutions.

It may be proposed that the funding of major community investigations was a legitimate initiative by an institute for ageing studies with a multi-disciplinary mandate from the U.S. Congress (see Footnote 8). However, data similar to those sought in the EPESE study were available from other federal agencies: the National Center for Health Statistics conducts regular household surveys from which are derived age-specific prevalence estimates of disease and physical impairments together with data on nutrition, smoking, and the use of, and expenditures for medical care; and the Health Care Financing Administration generates extensive data from the Medicare programme on illness for which older people receive testing and treatment as well as funding other cross-sectional and longitudinal surveys.[14] In 1983, two physician administrators at the

---

[14] Data on chronic disease and disability, based upon household interviews were made available from the following sources in the 1980s:

The National Center for Health Statistics:

National Health Interview Survey: 1979-81, 1982-84, 1985-87, and 1986 Functional Limitation Supplement

NIA used prevalence estimates of chronic disease and disability in the older population prepared by the National Center for Health Statistics to refute the arguments of James Fries in support of the theory of the "plasticity of ageing", i.e. that the onset of chronic illness may be postponed until more advanced ages by health promotion and changes in lifestyle.[15] Schneider and Brody argued against such "seductive predictions" proposing that in the future "chronic diseases will probably occupy a larger proportion of our life span", observing that "Alzheimer's disease" is "one of the major chronic disease of ageing".[Schneider & Brody, 1983 p855] However, they presented no American data on senile dementia/Alzheimer's disease and it may be argued that there was a specific rationale for seeking to determine the prevalence of the condition chosen as the legitimating research focus for their institution, the results of which in East Boston were extrapolated to the entire U.S. elderly population. On the other hand, the dementia syndrome indicative of Alzheimer's disease is a "mental disorder" the criteria for which are defined by psychiatry and in 1980 the National Institute of Mental Health itself initiated an extensive

---

Longitudinal Study on Aging, 1984 and 1986

The Health Care Financing Administration:

National Long Term Care Survey, 1982, 1984, 1989.

[15] In a 1980 discussion of "ageing, natural death, and the compression of morbidity", Fries argued that although life expectancy had increased throughout this century, there was evidence from cellular studies to indicate that the human life span is relatively "fixed". One the other hand he also argued that "the practical focus on health improvement over the next decades must be on chronic instead of acute disease, on morbidity not mortality, on quality of life rather than its duration, and on postponement rather than cure". He further proposed that with increasing "rectangularisation of the survival curve", i.e. continuing reductions in "premature death", and the deferred onset of diseases associated with mortality in old age by preventive care and lifestyle change, "high level medical technology applied at the end of the natural life span epitomises the absurd".[Fries, 1980 p130-5]

epidemiological study of the adult population which included "screening" for "organic mental disorders" (see Chapter 3).

In 1987, Richard C. Adelman, then President of the Gerontological Society of America (GSA), an organisation that had long advocated the establishment of a multi-disciplinary institute for the study of ageing, and which awarded a fellowship to Gubrium to undertake his phenomenological study of Alzheimer's disease,[Gubrium, 1986 pix] argued that a progressively increasing share of the NIA budget had become "devoted to diseases of the elderly", while support for "research on biological, behavioural and social phenomena that are unrelated to disease has hardly grown at all". This he proposed marked a "blatant shift in priority" that was in conflict with "the language and intent with which Congress established the National Institute on Aging".[Adelman, 1987 pB1-2] In 1995, Adelman, a biochemist, pursued his argument, observing that priorities had not changed and that "the NIA invests a disproportionately large share of its resources in research on Alzheimer's disease at the expense of other interests of the broader scientific community in gerontology". He acknowledged that Congressional funding for the investigation of senile dementia/Alzheimer's disease is earmarked for that purpose, but argued that the "Alzheimer's movement" deriving from the commitment and activities of the leadership of the NIA and the advocacy of the ADRDA has limited the potential for securing funding for a wider range of studies of the ageing process and that the medical profession and its "physician scientists" are the primary beneficiaries of the public investment in its study.[Adelman, 1995 p526] Estes and Binney in a 1991 paper supported Adelman in positing that "the primary rationale for, and thrust of, the NIA's research agenda has contributed to the usurpation of all applied problems in ageing by medicine" and ageing problems which cannot be defined in medical or clinical terms, or directly linked to them, are unlikely to be accorded significance by the NIA "review committees" who decide on the merit of the research proposals submitted for funding. In consequence those social sciences, those that are "quantitative" and "positivist" are favoured over those that are qualitative or related to "system level problems".[Estes & Binney, 1991 p125-6]

Gubrium acknowledges in a later work on "the politics of field research", that "there is no neutral money" and grant-giving organisations whether the National Institutes of Health, pharmaceutical companies or other public or private sources, "will seek to channel research in particular directions".[Silverman & Gubrium, 1989 p2] The funding for investigation of Alzheimer's disease has supported studies in the "neurosciences", the field in which there is an ongoing quest for a "cure", and also in the clinical and social sciences initiated and conducted by the various institutions of the National Institutes of Health and driven by what Gubrium in 1986 observed to the "facts of apparently alarming proportions".[Gubrium, 1986 p207] On the other hand one may question the "realities at stake" and the "facts" which are generated within a paradigm legitimated by a specialty, i.e. psychiatry, which has consistently defined and redefined its own socially constructed disorders and which now supports extensive medical and other research interests in what Adelman describes as the "Alzheimerisation of Ageing".[Adelman, 1995 p526]

Chapter 3

THE PSYCHIATRIC ENTERPRISE

Senile dementia and changing definitions of "mental disease"

"The head that will become a skull is already empty. Madness is the *deja vu* of death".[Foucault, 1965 p16] In his discursive work on "madness and civilisation" Foucault proposes many definitions of madness but here the image he evokes is that of senile dementia. Dementia, the Latin word "madness", with the literal meaning "loss of the mind", has become equated in medical terminology and in the popular imagination with a particular type of madness, one of the potential afflictions of old age which has been recognised and depicted in literature over the centuries. Kitwood, a clinical psychologist, in a discussion of the "dialectics of dementia" argued that of all the changes associated with ageing, perhaps none is more distressing to contemplate than the loss of mind, of "self", of "personhood". He proposes that dementia engenders a "malignant social psychology" towards those afflicted the dynamic for which may derive from time constraints and lack of inter-subjective insight among caregivers; a tendency to stereotype and withdraw respect from the sufferer as a "non-person", and also to maintain a "psychological distance", a defence against anxiety triggered by the knowledge that the same fate may befall anyone: for "to be with a confused elderly person is to be reminded painfully of one's own possible future".[Kitwood, 1990 p184-5] Dementia may also be conceptualised from a sociological perspective as "deviancy", the "property of deviance" being conferred in phenomenological analysis on things perceived to be anomalous to "rules of description and classification" by which the social world is constituted.[Downes & Rock, 1982 p172-4] However, "rules of description and classification" have been defined for many "unmanageable cases" within the nomenclature of mental illness.

It was in the research institute of Emil Kraepelin (1855-1926) that Alois Alzheimer described his eponymous "disease",[Alzheimer, 1907] and it is Alzheimer's mentor who has been credited with developing the first "classification of mental diseases".[1] Kraepelin defined two main categories of mental disease: "varieties of insanity" and "varieties of imbecility". Among the former he included "melancholia", "maniacal depressive conditions", "dementia praecox" (renamed schizophrenia in 1911), and "moral insanity", the latter containing subcategories of "congenital", "morbid personalities", and "morbid criminals and vagabonds". In the Kraepelin typology "senile dementia" was categorised as "acquired imbecility" as opposed to "congenital imbecility-idiocy".[Kraepelin, 1906 p xvii]

In 1952 the American Psychiatric Association published the first Diagnostic and Statistical Manual of Mental Disorders (DSM). The foreword noted that the purpose was to develop a uniform nomenclature to replace "a polyglot of diagnostic labels and systems, effectively blocking communication and the collection of medical statistics".[DSM, 1952 pv] Earlier classifications devised in the United States had been developed piecemeal by different medical institutions for diseases, both physical and mental, and the purpose of the DSM was to develop a "suitable diagnosis for every case seen by a psychiatrist". During the 2nd World War psychiatrists were using "a system of naming developed primarily for the needs and caseloads of public mental hospitals" and were operating

---

[1] Kraepelin was Professor of Psychiatry in the University of Munich. The Classification of Mental Diseases was contained in a volume of his "Lectures on Clinical Psychiatry" edited in 1906 by Thomas Johnstone, M.D. Edin., M.R.C.P., Lond.

"within the limits of a nomenclature which was specifically not designed for 90% of the cases handled". "Relatively minor personality disturbances, which became of importance only in the military setting had to be classified as 'Psychopathic Personality'...... (and) the 'psychoneurotic label' had to be applied to men reacting briefly with neurotic symptoms to considerable stress".[DSM, 1952 p vi-vii]

Since 1952 there have been four revisions of the DSM, in 1968,[DSM-II, 1968] 1980,[DSM-III, 1980] 1987,[DSM-III-R, 1987] and 1994,[DSM-IV,1994] and with each edition diagnostic categories, the "diseases" themselves, and the symptoms associated with them have appeared, disappeared, or been redefined. In 1952 "vagrancy" and "vagabondage" had ceased to be "varieties of insanity" as in Kraepelin's classification but were included as "supplementary terms" or symptoms.[DSM, 1952 p122] In 1968 all mention of "vagrancy" had disappeared, but "homosexuality" became a "diagnosis" under "Personality Disorders and Certain Other Non-Psychotic Mental Disorders".[DSM-II, 1968] By 1980, "homosexuality" had been dropped, but other categories and conditions were introduced, for example: "disorders of impulse control" which included "pathological gambling" and "kleptomania"; "psychosexual dysfunctions", and "tobacco dependence".[DSM-III, 1980 p380-6] The DSM increased in size from one hundred to nine hundred pages between 1952 and 1994, and the latest edition includes "pre-menstrual dysphoric disorder" as well as "caffeine intoxication" and "caffeine withdrawal".[DSM-IV, 1994 p213-4] For children there are separate diagnostic criteria for "mathematics disorder", "reading disorder" and "disorder of written expression", "e.g. writing grammatically correct sentences and organised paragraphs."[DSM IV, 1994 p46-53]

Senile dementia has itself undergone changes both in nomenclature and in its diagnostic criteria. In 1952 the condition was categorised as a "chronic brain disorder", and defined as "chronic brain syndrome associated with senile brain disease", to be distinguished from "chronic brain syndrome associated with other disturbances of metabolism, growth or nutrition", of which one was "Alzheimer's disease".[DSM, 1952 p22] In 1968 the condition reverted to the designation "senile dementia" and together with "pre-senile dementia", which included "Alzheimer's disease", was categorised under "psychoses associated with organic brain syndromes".[DSM-II, 1968 p24] In 1980, however, "dementia" without distinction between senile and pre-senile onset was categorised under "organic mental syndromes and disorders" in a greatly expanded section. Among "aetiological factors" it was stated that "primary degenerative dementia of the Alzheimer type is the most common".[DSM-III, 1980 p110]

Kraepelin described senile dementia as a progressive "mental disease" of which the primary characteristics at various stages were "disturbance of the faculty of attention to new impressions - strange fancies - presbyophrenia - senile catatonia".[Kraepelin, 1906 p xiv] In the first two editions of the DSM the diagnostic definition of the condition was determined to be: "self-centredness, difficulty in assimilating new experiences, and childish emotionality". Deterioration from the mental state at which the diagnosis was made could be "minimal or progress to a vegetative existence".[DSM, 1952 p18 DSM-II, 1968 p24] By 1980, however, a new conceptual model for senile dementia had been developed in which the prior emphasis on "personality" characteristics was discounted (see Figure 3.1 below).[DSM-III, 1980 p111] Furthermore, all terms associating the generic condition "dementia" with "insanity" or "madness", either in its definition or its classification, for example "psychosis" and "presbyophrenia", were discarded. In a revision of the DSM-III published in 1987 the new model was itself modified. Evidence of impairment in "social or occupational functioning", the primary diagnostic criterion in the 1980 edition, was discontinued seven years later and the major determinant of dementia became "memory impairment" which was expanded to include both long and short-term memory (see Figure 3.2). The other criteria remained essentially the same, but new subcategories of "mild", "moderate" and "severe" dementia were introduced.[DSM-III R, 1987 p107]

The proposal to redesignate "senile dementia" as "Alzheimer's disease" by eliminating age distinctions in the diagnostic process, based upon the argument that the neuropathology and clinical course of the two conditions were indistinguishable, was made by a neurologist, Robert Katzman, representing the specialty with competitive control over the investigation and treatment of conditions affecting the physical space of the "brain", together with T. B. Karasu, a psychiatrist representing the medical "science of the mind".[Katzman & Karasu, 1975 p106] However, as previously discussed, in spite of a there were problems associated with an ostensibly simple redefinition which in spite of international research efforts remain unresolved, e.g. the establishment of a functional link between behavioural evidence of Alzheimer's disease/senile dementia and *post mortem* physical changes in the brain; the overlap of such changes with those identified in "normal ageing", and the determination of their aetiology,[Blessed et al, 1968 p808, Crystal et al, 1988 p1682, McLoughlin & Lovestone, 1994 p439-40] (see Chapter 2). On the other hand, research has not been confined to the neurosciences from which a "cure" may derive and an extensive range of related investigations have been conducted by other disciplines and medical specialties. In the United States these may be analysed in the context of the institutions and public financing made available by way of the Medicare programme and the National Institutes of Health and their potential for clinical and academic entrepreneurialism (see Chapter 1). However, successful efforts to secure funding for research into Alzheimer's disease were led by the first director of the National Institute on Aging they have benefitted many specialties and disciplines in the biomedical community (see Table 2.1.) and it is psychiatry via its own institutions, i.e. the American Psychiatric Association and the National Institute of Mental Health which formulated and legitimated the "case finding" method for the "disease".

It may therefore be posited that the change in nomenclature from "senile dementia" to "dementia" in the 1980 DSM-III,[DSM-III, 1980 p107] and the adoption of a new model for one which had varied little in previous classifications reviewed for this discussion, were conceptualised in support of the detection of "probable" and "possible" cases of Alzheimer's disease in clinical practice, in diagnostic research, and in epidemiological studies.[McKhann, 1984 p939] It may be further posited that the DSM-III model and its 1987 revision were constructed by a specialty which has consistently expanded and redefined its activities, as demonstrated by the growth in size of the DSM, while adopting different aetiological paradigms for the phenomena it defines as "mental illness" in response to factors including: changing market conditions for its services and the need for professional satisfaction; the influence of lay interests; and also the pursuit of scientific legitimacy within the medical profession.

Mental illness as deviance, stigma, and senile dementia

Thomas Szasz, a longtime and prolific critic of his own specialty, in a work first published in 1962, described psychiatry as a "pseudo-medical enterprise".[Szasz, 1972 p14] He argued that there is, as a rule, little disagreement among scientists concerning basic physiological, biochemical, or physical theories", whereas in religion and politics "we expect to find conflicting systems and ideologies". He therefore posited that since "psychiatry is characterised by a plethora of diverse, competing, and often mutually exclusive theories and practices ..... (it is) more like religion and politics than science".[Szasz, 1972 p96] It may be argued that the "diverse, competing, and often mutually exclusive theories and practices" to which Szasz referred derive from the differing "realities" of groups who became affiliated within the specialty including: neuro-psychiatric researchers such as Kraepelin and his pupil Alzheimer who sought resolution of the Cartesian dilemma of the duality of mind and body in post-mortem dissection and analysis of the brains of those diagnosed with "varieties of insanity";[Kraepelin, 1906 p xvii] Sigmund Freud (1856-1939), also a neurologist, and his disciples who reverted to "introspection" in the light of failure to establish a physical aetiology for the majority of mental "diseases";[2] and those described by Scull in an historical perspective on

---

[2] Syphilis was the 19th century archetypal paradigm for a systemic disease which in its late stages was often associated with insanity and with post-mortem lesions in the brain. However, not all 19th century medical academics focused on the organ of the brain as the genesis of "madness". Benjamin Rush (1746-1813) Dean of the University of Pennsylvania Medical School in Philadelphia, the "founder of American psychiatry", inferred "madness to be seated primarily in the blood vessels", based on the rapid pulse rates of most of those physically confined.[Szasz, 1970 p139] Freud is cited as proposing in 1925 that evidence would be found of either infections or "toxic substances" in the aetiology of all types of mental "disorders".[Szasz, 1979 p10]

Anglo/American psychiatry as "mad doctors" who exercised political power, reinforced by "ingenious metaphysical arguments dressed in the trappings of science", to gain control of the charitable, private, and state institutions established for the care of the insane in Britain and the United States.[3] [Scull, 1989 p140] Based upon his perception of the role of such groups in "the manufacture of madness" Szasz posited that psychiatrists are not concerned with "mental illnesses and their treatments" but with the "personal, social and ethical problems of living" and that "mental illness is a myth".[Szasz, 1972 p269]

There is evidence that institutions for the confinement of "madness", described by Scull as "the most solitary of afflictions to the people who experience it, but .... the most social of maladies to those who observe its effects",[Scull, 1989 p8] predate its legal sanctioning as a "medical disease", some by many centuries.[4] It may therefore be argued that "mental illness" may not be construed merely as a social construct but that with diverse origins, and changes in nomenclature and practice, a range of theoretical frameworks lend themselves to the analysis including the political economy perspective adopted by

---

[3] Scull proposes that doctors in Britain became interested in the "lucrative business" of private madhouses in the 18th century and thereafter secured legislation in the 1774 Madhouse Act to require "medical certification of insanity" before confinement.[Scull, 1989 p147] Nineteenth century legislation was secured to exclude the Quakers or other lay groups from control over the care provided in their own institutions and superintendents of state asylums often set up private homes.[Scull, 1989 p284] Scull demonstrates that developments in the asylum movement and in psychiatry in general have been remarkably similar in Britain and the United States.

[4] For example the St Mary of Bethlehem Hospital (Bedlam) was founded in London in the early 15th century for the segregation and confinement of a small number of lunatics.[Scull, 1989 p221] Foucault argues that the "great confinement of the insane" began in France in the 17th century with the founding of the *Hopital General.*[Foucault, 1965 p39] The York Retreat was founded by Quakers in the late 18th century and introduced the concept of "moral treatment" rather than simple physical confinement and restraint.[Scull, 1989 p81]

Scull and an array of "labelling", "social control" and other theories which comprise the sociology of deviance. Szasz has been described as holding to a "kind of *laissez faire* individualism",Ingleby, 1983 p174 and although functionalist, symbolic interactionist, phenomenological and alternative interpretations have been proffered for the analysis of "rule breaking" and responses to it there is overlap between apparently divergent philosophies. Downes and Rock propose that in a sense phenomenology "is a reaction to others' reactions, a consciousness of others' consciousness, a knowledge of others' knowledge" but its interpretation of deviance is intimately implicated in a conception of rules of description and classification and that in this respect "functionalism, structuralism and phenomenology are indistinguishable at points".Downes & Rock, 1982 p167-172

In a 1964 discussion of "definitions of health and illness" Talcott Parsons proposed "illness" to be "deviant behaviour" that is "partially and conditionally legitimated", i.e. the sick person is not held to be responsible, but is required to recognise that to be ill is "inherently undesirable", has an "obligation to seek competent help" and to co-operate fully with the therapeutic agency" in order to achieve recovery.Parsons, 1964 p274-5 Parsons placed his analysis of the "health-illness complex among other mechanisms of social control" in the context of an emphasis on "health" in the "American value system".Parsons, 1964 p277 [5] Berger and Luckmann argue that legitimacy has a "cognitive as well as a normative element", i.e. "knowledge" precedes "values" both in the legitimation of institutions, and in the integration of discrete institutional processes within a "symbolic

---

[5] Missing from Parson's discussion is an analysis of who defines "illness" or the parameters for "being ill", nor why, given the esteem ascribed "health", there had been a failure to provide universal insurance to enable the sick to meet their obligations.

universe" which represents the most comprehensive level of legitimation. Deviance from institutional norms is then "co-extensive with the limits of theoretical ambition and ingenuity on the part of the legitimators, the officially accredited definers of reality".[Berger and Luckmann, 1966 p96-7] In reviewing the work of Parsons from the perspective of Berger and Luckmann's treatise on the sociology of knowledge Freidson proposed that "the rise to social prominence of a value such as health is inseparable from the rise of a vehicle for the value - an organised body of workers who claim jurisdiction over the value".[Freidson, 1973 p251] He argued that "illness" may be analysed both as "biological deviance" from the normal or desirable and also "as a form of social deviance the treatment of which is the licensed domain of the medical profession".[Freidson, 1973 p212] He further argued that illness as a "biophysical state" involves changes in bone, tissue, vital fluids etc. in both human and other living organisms and may therefore be defined as "disease". On the other hand, "when a physician diagnoses a human condition as illness, he changes the man's behaviour by the diagnosis: a social state is added to a biophysical state by assigning the meaning of illness to disease". It is in this sense that "the physician creates illness ...... and that illness is a kind of social deviance analytically and empirically distinct from mere disease".[Freidson, 1973 p223] He therefore sought to analyse not merely the behaviour of those to whom the label "illness" has been assigned but also the "diagnosis behaviour" and "treatment behaviour" of those who impute this form of deviance.[Freidson, 1973 p216]

Based upon the analysis by Freidson, and in the context of the sociology of knowledge, the expanding nomenclature of "mental illness" may be perceived as reflecting the realities of the profession which has been officially accredited in defining or assigning meaning to both "health" and "illness" and has exercised "ambition and ingenuity" in extending its jurisdiction beyond "biological deviance" by ascribing medical labels to what Freidson has termed "disapproved behaviour".[Freidson, 1973 p248] He proposes that since the profession "has a mission of active intervention" in seeking out and treating that which it defines as "illness", imputing a "social and therefore moral meaning" to human conditions, the physician assumes the role of "moral entrepreneur" when he defines as "illness" behaviours such as "homosexuality" and "kleptomania" (see above), in just the same sense as a minister of religion who defines them as "sins".[6] Such claims by the medical profession then lead to the "illogical conclusion" that these behaviours are illnesses and also to the "assumption that (they are) properly managed only by physicians".[Freidson, 1973 p251-3] On the other hand, the jurisdictional expansion argument or what Ingleby terms the "hypothesis of medical imperialism"[Ingleby, 1983 p175] does not fully explain the selection and later deletion of conditions such as "vagrancy" from the classification, or the changes in nomenclature, and variations in the diagnostic criteria for detection of the same condition, i.e. "senile dementia" in the different editions of the DSM.[DSM II, 1968 p24 DSM-III, 1980 p111, DSM-IIIR, 1987 p107]

Berger and Luckmann argue that in being the "historical products of human activity, all socially constructed universes change, and change is brought about by the concrete actions of human beings".[Berger & Luckmann, 1966 p116] Therefore although medicine may be

---

[6] Freidson's proposal that medical practice may be perceived as "moral entrepreneurship" is not restricted to "mental" illness but is particularly apposite to this discussion. He argues that there is an "irreducible moral judgement in the designation of illness as such, a judgement the character of which is frequently overlooked because of the virtually universal consensus that exists about the undesirability of much of what is labelled illness".[Freidson, 1973 p252]

defined as an "autonomous subuniverse" with "role specific knowledge that (is) esoteric as against the common stock of knowledge"[Berger & Luckmann, 1966 p85-6] it is not independent of the society in which it exists nor immune to the influence of "rival definitions of reality" which are decided upon in the sphere of rival social interests and become translated into competing theories.[Berger & Luckmann, 1966 p120] They also argue that "the relationship between ideas and their sustaining social processes is always a dialectical one ..... theories are concocted to legitimate existing social institutions and social institutions are changed in order to bring them into conformity with existing theories".[Berger and Luckmann, 1966 p128] From this perspective it may be argued that medical definitions of "deviance" have varied in part to accommodate to theories underlying changes in social institutions and also to shifts in the allocation of resources with which such changes may be associated, i.e. from funding "institutional" care in workhouses and asylums in the 19th century to the establishment of an infrastructure via the welfare state for "community" care of "vagrants". Thus it may be argued that "moral insanity", with sub-categories of "congenital" and "morbid criminals and vagabonds" in the Kraepelin typology (see above), was constructed by medicine in response to the public funding provided for institutional control of deviant behaviours in one era and abandoned when rival definitions of reality prevailed and resources were diverted to alternative social programmes, new "mental disorders" then being constructed for generation of an alternative market to compensate for the loss of such resources.[7] Such an alternative

---

[7] In support of this argument, Andrew Scull proposed that 19th century industrialisation required the workforce to "internalise new attitudes and responses to discipline themselves". This was translated into "faith in the capacity for human improvement by social and

was found with in adopting the "psychogenic" or "functional" aetiological theory of mental disorders, generating a community-based clientel without evidence of "insanity" for treatment of "the personal, social and ethical problems of living".Szasz, 1972 p269 [8]

"Mental illness" may therefore be analysed both in relation to the authority of the medical profession to define and redefine those conditions which it has encompassed within its ambit, and also to the self-interested accommodation of its members to changes in social institutions. On the other hand, just as the profession is not immune to the effects of the legitimation of rival definitions of reality in social policy, Freidson argues that the social meaning of illness is not entirely supplied or controlled by the physician.Freidson, 1973 p260 In support of this argument he proffers examples of the expansion of medical jurisdiction in response to the activities of "lay interest groups"

---

environmental manipulation" in a variety of settings which may be regarded as "techniques of social discipline" and included factories, workhouses, schools, prisons and asylums.Scull, 1989 p91-2 He argued that the medical profession exercised political power to gain control over charitable, private and state institutions established for the control of insanity and then proceeded to increase the patient population by extending the boundaries of mental pathology, based on "morality" rather than "madness" and on prediction, determined by criteria such as heredity, rather than observation.Scull, 1989 p28

In a discussion of the "decarceration" of the mentally ill in the 1950s and 60s he proposed that this process may be perceived as an humanitarian response to the conditions which came to prevail in the huge "custodial warehouses" the size of which accommodated to expanded definitions of madness; as evidence of medical advances in the management of insanity, or as a recognition by governments that with the development of the welfare state, the "community" provided a cheaper alternative to institutional care.Scull, 1976 p191-4

[8] Szasz posited that Freud "invented the neuroses to justify calling conversation and confession 'psychoanalysis'", which he regarded as "a form of medical treatment", thereby "having a class of new diseases on his hands".Szasz, 1979 p38 He also argued that although terms such as "psychoanalysis" had never been registered as "trademarks", psychiatrists had been "amazingly successful in cornering and controlling the market for these 'products'".Szasz, 1979 p4

seeking the "application of the label illness to conditions not considered illness before", for example "alcoholism", proposing that "the movement to reinterpret human deviance as illness has its roots in humanitarianism" deriving from a quest for legitimacy for "disapproved behaviours" and the substitution of "treatment" for the retribution or atonement sought for crime or sin.[Freidson, 1973 p250-4] However, the effects of substituting the "label" from one deviant category for another are not necessarily positive.

Parsons argued that illness assigns legitimacy to an "inherently undesirable" or deviant state in that a sick person "cannot be held responsible for the incapacity" which justifies assumption of the "sick role",[Parsons, 1964 p274] but the Parsonian legitimation of "illness" is both temporary and conditional upon seeking therapy and co-operating in achieving recovery. Such legitimacy therefore applies primarily to acute illnesses and/or those which may be successfully treated by the medical profession. Freidson suggested that "blame" is not necessarily imputed to the chronically ill or permanently impaired but that legitimacy is again conditional upon striving for self-improvement, even though individuals are "incurable in the absolute sense", and also upon "limiting demands for privileges", i.e. there are limits to the amount and type of deviant behaviour. He therefore argued that the "illness" label may "discourage punitive reactions (but) does not discourage condemnatory reactions" and that a condition may be medically legitimated but "socially stigmatised".[Freidson, 1973 p231-5]

In 1963 Erving Goffman posited that "stigma", the mark of disgrace, incurs the imputation of a range of "discrediting" attributes to one "undesirable" characteristic and the ascription of a "virtual social identity" that is distinct from "actual social identity".

He defined three types of stigma: "tribal" relating to race, nation, and religion; "abominations of the body" or physical deformities; and "blemishes of individual character", inferred from a known record of imprisonment, alcoholism, homosexuality, "mental disorder", etc. By definition we (the normals) believe the person with stigma "is not quite human" and the stigmatised individual tends to hold the same beliefs about identity that we do: "the standards incorporated from the wider society equip him to be intimately alive to to what others see as his failing".Goffman, 1963 p2-7 Others in the symbolic interactionist school have elucidated on "mental disorders" themselves as outcomes of the "orderly production of deviance and deviants" with interdependence between those who identify deviance, in this instance psychiatrists, and the "rule breakers" who are persuaded to accept "labels" and comply with therapeutic authority. The deviant gains insight and a new dependent definition of "self" and those who award the new identity are rewarded with professional satisfaction and economic benefits in expanding the client base.Downes & Rock, 1983 p151-3 [9]

Berger and Luckmann argue that "the self" as a subjectively and objectively recognised "identity" can only be understood in the social context in which it is shaped",Berger & Luckmann, 1965 p50 and those to whom a "stigmatised identity" has been ascribed are "imprisoned in the objective reality of society". They further argue that maximal success in the production of "socially predefined" identities occurs in societies with a

---

[9] Similarly Scott discusses the selection of "eligible" blind individuals, i.e. the near-blind, the employable, and those without other handicaps, and their socialisation by reinforcing their conceptions of "self" as "helpless, docile, dependent, or incapacitated", in the interests of the prestige and fund-raising prospects of the "blindness" agencies and workers.Scott, 1969 p16-19

"simple division of labour and minimal distribution of knowledge" but when a group that has been "socially pre-defined as a profiled type" is sufficiently large and durable to serve as a "plausibility structure for different definitions of reality" counter-identities may arise within the profiled group and, as the distribution of knowledge becomes more complex, the possibility exists for choosing between such alternative "reality definitions".[Berger & Luckmann, 1966 p165-7] However, although the socially predefined identity may, arguably, have been successfully rejected by some such as those with "abominations of the body" who have adopted an identity other than that of "cripple", Berger and Luckmann also argue that "subjective identity" is a "precarious entity". It may therefore be posited that when "the 'sane' apprehension of oneself as possessor of a definite, stable and socially recognised identity"[Berger & Luckmann, 1966 p100] has been threatened if not negated by a label which defines one as a "mental cripple", the potential for identifying with others so defined for adoption of a counter-identity is diminished by the diagnosis itself. Nor, although Freidson proposes it to be among the objectives of lay support for moral entrepreneurship,[Freidson, 1973 p254] will stigma be eradicated simply by new "linguistic designations" which once incorporated into "the social stock of knowledge"[Berger & Luckmann, 1966 p68] become imbued with the same "social meaning", for example changing the label from "drunkenness" to "alcoholism" or "senile dementia" to "Alzheimer's disease".

Elaborating upon Parson's proposition that "illness" as a "social state" is "partially and conditionally legitimated", incurring an "obligation to seek competent help",[Parsons, 1964 p274] Freidson argued that the highest level of legitimation is accorded to acute conditions and those for which there is effective medical therapy (see above), and it may be posited that stigma seldom if ever attaches to illness that does not persist. Therefore it is by the successful development of preventive or therapeutic measures that the mark of disgrace is removed from previously stigmatised medical conditions, i.e. by the early treatment of leprosy before there is physical evidence that individuals are "lepers", rather than by changing the label to Hansen's disease, or by group rejection in the "leper colony" of a pre-defined identity. It may therefore be posited that lay support for the extension of medical jurisdiction derives fundamentally from the perception that a "problem" exists and that in combination with "public belief" in the profession's competence and skill, based upon demonstrable scientific achievement, it is amenable to medical therapeutics. However, this "strengthening of a professionalised control institution that can remove from laymen the right to evaluate their own behaviour and that of their fellows",[Freidson, 1973 p250] will legitimate the exercise of diagnostic authority and the formulation of criteria which may extend the problem beyond the group it was designed to "cure".

Gubrium, in discussing the generation of a "public culture of Alzheimer's disease", led by the national association of lay support groups, elaborates on the "tribulations" experienced by caregivers and its adoption of "slogans", including "the silent epidemic" and "the disease of the century", which "attempt to alert the public to the Alzheimer's disease problem .... the tragic way that the disease affects those concerned".[Gubrium, 1986 p119] The association also "serves the scientific legitimation of the war against Alzheimer's disease",[Gubrium, 1986 p134] successfully supporting the biomedical community in this legitimation as demonstrated by the public financing allocated for its investigation (see Chapter 2 and Table 2.1). On the other hand, it may be argued that significant negative effects may devolve from a "public culture" surrounding a disease, i.e. senile dementia, the victims of which share the double jeopardy of stigma attaching to a condition which Kitwood argues engenders a "malignant social psychology"[Kitwood, 1990 p184-5] and the "symbolic stigmatisation" of negative stereotyping of old age which is strongly associated with and reinforced by images of the disease itself.[Featherstone & Hepworth, 1993 p308-13]

There is little evidence that those who sought to "ameliorate" the disease experience by supporting the medical community in "doing something about it"[Gubrium, 1987 p17] ever advocated or anticipated extending the Alzheimer's diagnosis to anyone not overtly "a diseased person".[Gubrium, 1986 p91] Nevertheless, it may be argued that this expansionary potential was legitimated by American psychiatry in abandoning the categorisation of "senile dementia" as "insanity" or "psychosis" in 1980, and in formulating a new diagnostic model, one which was redefined in 1987 to include "mild", "moderate" and "severe" evidence of the "organic mental disorder".[DSM-III, 1980 p111, DSM-III-R, 1987 p107] On the one hand it may be argued that the models for in detecting evidence of "dementia" may be perceived as "case-finding" methods for clinical and epidemiological studies conducted by psychiatry and other specialties in response to the funding of research into Alzheimer's disease. However, all DSM-III models were formulated to achieve greater "objectivity" in psychiatric diagnosis in what may be interpreted as a response to "labelling" theories of mental disorders (see below). Furthermore it may be argued that they were conceived within a diagnostic and therapeutic paradigm for which the successful outcome of Alzheimer's disease research would proffer support and hence greater scientific legitimacy for the specialty within the medical hierarchy.

The "crisis of legitimacy" and the "biological paradigm"

Berger and Luckmann propose that "a body of knowledge, once it is raised to the level of a relatively autonomous subuniverse of meaning has the capacity to act back upon the collectivity that has produced it" and that the autonomy of the "subuniverse" makes for "special problems of legitimation *vis a vis* both outsiders and insiders". They argue that "it is not enough to set up an esoteric subuniverse of medicine" but that the lay public must be convinced that it is "right and beneficial, and the medical fraternity must be held to the standards of the subuniverse", the latter being kept from "quackery", i.e. "from stepping outside the medical universe in thought or action, both by external controls and by a body of professional knowledge that offers them 'scientific proof' of the folly ..... of such deviance".[Berger & Luckmann, 1966 p87-8]

Psychiatry lies within the autonomous collectivity of medicine but it may be argued that it has deviated from the institutional norms of the fraternity by failing to offer "scientific proof", based upon what Freidson terms the "standards of verifiability and reliability",[Freidson, 1973 p210] for phenomena now termed "mental disorders", a designation described by the American Psychiatric Association itself as "a concept" for which there is "no definition (that) adequately specifies precise boundaries".[DSM-III-R, 1987 p xxii] Therefore, although some lay groups may be convinced that its activities are "right and beneficial", there have been consistent challenges to the legitimacy of psychiatric paradigms and practices both from "outsiders and insiders" including critiques of American psychiatry by foreign "insiders".[Kendell, 1975, Lipowski, 1989, Guimon, 1989]

Scull described the specialty as having "enjoyed a perpetually marginal and unenviable position in the social division of labour ...... but a step away from a profound crisis of legitimacy".[Scull, 1989 p22] It may be argued, however, that the status of psychiatry reflects not merely what Scull terms its "persistently dismal therapeutic capacities" which, arguably applies to other specialties such as neurology and others dealing with chronic and degenerative diseases, but more significantly the lack of verifiable "conceptions of aetiology" for the phenomena which it has, over time selected and later deleted from its changing nomenclature. Furthermore, following the failure of "organic" psychiatrists such as Kraepelin to find a consistent link between "deviant" behaviour and brain pathology the specialty was further marginalised in shifting its dominant paradigm from the "brain" to the "mind" and to Freudian and related theories, based on the concept of "unconscious forces which in turn are the results of instinctual drives and early experiences",[Szasz, 1972 p22] a model that although intuitively powerful cannot be tested, verified or "refuted" thereby failing to meet the Popperian criterion for "scientific method".[Popper, 1964 p151] However, Guimon, a Spanish psychiatrist in a paper entitled the "biases of psychiatric diagnosis" posited that since the late 1960s American psychiatry has adopted a new "reductionist" model which has been promoted by a "new generation of biological psychiatrists" or "Neokraepelinians".[Guimon, 1989 p33]

Szasz had previously observed that with the introduction of new classes of tranquillisers in the 1950s and their administration to the "insane" in the "decarceration" of the state asylums,[Scull, 1976] inferences had been made from the "therapeutic intervention" to the nature of the disease, based on the argument that because the drugs affect "psychotic behaviour" in "ways that many people consider to be desirable" the outcome of their use is evidence of "abnormal functioning of the brain".[Szasz, 1979 p91-2] Guimon proposed that the "biological paradigm" for mental disorders is based both on the "availability" of an expanding pharmacopoeia and the detection of "alleged biological markers", or abnormal levels of neurochemicals in the brain associated with conditions including anxiety and depression. His use of the adjective "alleged" relates to the lack of evidence as to whether "biological markers" are the by-product or the stimulus of mental states for which specific drugs such as anti-depressants have been developed. Their use to change the neurophysiological functioning of the brain, like that of non-specific sedatives, is therefore also based in part upon imputation. However, the new model infers "scientific legitimacy" for the conceptualisation of "insanity" as well as "disapproved behaviours" and states of mind as "organic diseases" and is congruent with the "neuro-transmitter" theory of dementia with its potential for "neuropharmacological manipulation",[Perry et al, 1978 p1459] an avenue which has been pursued in Alzheimer's disease research since the 1970s but without current success (see Chapter 2).

Kendell, a British psychiatrist writing in 1975, defended the concept of "mental illness" even without the identification of physical lesions in the brain, but noted that whereas a century earlier the specialty was concerned only with "madness", the notion of its "proper role" had expanded steadily "particularly in North America". He therefore proposed that attempts to discredit the concept by Szasz and others, might be a reaction to the "absurd claims we have made that all unhappiness and all undesirable behaviour are manifestations of mental illness". He further suggested that psychiatrists were straying "outside our proper boundary" and there were "other people" whose "concepts" and "training" were more appropriate as sources of help for the "woes of mankind".[Kendell, 1975 p314] However, the "biological" or "neuro-pharmacological" model offers legitimation for the designation of both "madness" and "all unhappiness" not merely as "illness" but as "organic disease", and may be perceived as a vindication of the "humanitarian" movement's quest for "treatment" as an alternative to "retribution"[Freidson, 1973 p254] since causation may now be attributed to "a chemical imbalance in the brain".[Lipowski, 1989 p252] On the other hand, it has been argued by critics including Szasz that therapy consists in substituting chemical dependence and control for other forms of restraint. Furthermore, others such as Guimon and Lipowski, a Canadian psychiatrist, have considered the political and economic context within which the model has been adopted.

Lipowski posits that the pendulum has swung periodically from "extreme biologism" or "mindless" psychiatry to a "one-sided emphasis on psychodynamics and social issues" or what has been termed "brainless" psychiatry, and that the specialty has recently "rediscovered the brain with a vengeance".[Lipowski, 1989 p252] However, he proposes that the immediate motivation for adoption of the biological model in the United States was "political and economic concerns", i.e. unless psychiatrists drew closer to medicine "they would lose their medical identity as well as their patients .... to non-medical mental health practitioners".[Lipowski, 1989 p250-2] The model may therefore be perceived as affording greater legitimacy for the specialty while strengthening professional control with an armamentarium of conventional medical therapeutics, since maintaining legal authority over the prescribing of drugs is a means by which the medical profession asserts dominance over the competition, which in the field of "mental illness" includes clinical psychologists, lay psychotherapists, social workers, self-help groups and others.[10] [11] Guimon also describes inferences made within the model as "biologism", defined by him as the "unwarranted extension of biological explanations to psychosocial issues ..... motivated (in part) by the wish to conceal the socio-political factors that encourage the use of certain forms of treatment or bias choice of research topics".[Guimon, 1989 p34-5]

Neither Guimon nor Lipowski reject the basic assumptions of the biological paradigm but the latter argues that "to assume ..... biochemical processes underlie mental activity and behaviour does not imply they are the causal agents but rather constitute mediating mechanisms" influenced by information inputs received from the body and the environment and by the subjective meaning of that information for the individual. He also proposes that "at the present time neuroscience offers precious little that we could apply in our daily work with patients". Lipowski is largely reacting to what her perceives a trend towards a "purely scientific and technological approach to man, to medicine and to psychiatry".[Lipowski, 1989 p250-3] Guimon, however, acknowledges the development of "effective drugs" and that "the biological model has led to important

---

[10] As previously discussed (see Chapter 2), Johnson argues that the "diagnostic relationship" is given pre-eminence by occupational groups seeking to legitimate their claim to professional autonomy, and that this relationship serves as a control mechanism *vis a vis* their clients and other potentially competing health groups.[Johnson, 1972 p57-8] However, Medawar proposes that in Britain doctors had defined themselves as a "profession" rather than as "traders in health" by uniting to gain monopoly control of the profitable trade in opium, the premier psychotropic drug of the 19th century.[Medawar, 1992 p23]

[11] From the perspective of Lipowski the pharmacological model may also be construed as a means for American psychiatry to respond to incentives in the indemnity insurance system by increasing the "throughput" of patients under a structure which excludes long-term counselling or "psychotherapy", paying only for a limited number of consultations for the diagnosis and treatment of discrete episodes of either "mental" or "physical illness" (see Chapter 1).

discoveries". On the other hand, he observes that it has benefitted both "patients and the pharmaceutical industry" and argues that "diagnoses" have led to a search for drugs while new drugs have themselves "forced the creation of, or at least an over emphasis on certain 'diagnoses'", e.g. "panic disorder" which entered the nomenclature in 1980.

Scull proposes that during the period of the asylum movement psychiatry abandoned "references to real behaviour in favour of surmises concerning future behaviour", arrogating to itself a "margin of interpretation whose bounds are no longer discernible", and that it "was able to extend the boundaries of mental pathology to encompass marginally deviant affective symptoms and make a plausible case for the reality of partial sanity".[Scull, 1989 p248] Guimon argues that the Neokraepelinians have demonstrated an "over-enthusiasm" which is to those with "historical knowledge" reminiscent of the "19th Century Organicist School" and focuses his critique specifically on the invocation of both physical and statistical science in expanding the nomenclature of mental illness and claiming validity for diagnoses constructed within the biological paradigm, i.e. they have shown a tendency to "elevate symptoms to diseases" while stressing a "scientific approach to diagnosis" as in the 1980 American diagnostic manual.[Guimon, 1989 p33-35]

The DSM-III and scientific objectivity

In 1980 the American Psychiatric Association (APA), published the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (1980) (DSM-III). The introduction observes that prior to its publication "the clinician (was) largely on his or her own in defining the content and boundaries of diagnostic categories" and that it "includes such new features as diagnostic criteria, ... much expanded descriptions of the disorders, and many additional categories (some with newly-coined names)", but excludes "several time-honoured categories". It also notes that the manual "reflects an increased commitment in our field to reliance on data as the basis for understanding mental disorders".[DSM-III, 1980 p1-8]

In 1970, Robert L. Spitzer, chairman of the Task Force on Nomenclature and Statistics for the DSM-III, published an article on the development of the Psychiatric Status Schedule (PSS), a "standardised" questionnaire that was more comprehensive in its coverage of "psychopathological signs and symptoms" than numerous rating scales, inventories and forms that had previously been developed primarily for "the evaluation of in-patients". Therefore although the PSS instrument was designed to "improve the research value of clinical judgements ..... based on data collected during a psychiatric interview" it was designed to be suitable for "evaluation of the psychopathology and role functioning of both patients and non-patients".[Spitzer et al, 1970 p41] It was reported that both "validity" and "reliability" testing was conducted, but although the methodology used for the former was non-specific, the statistical techniques used to determine "reliability" for "symptom scale" values were described, reported upon,[Spitzer et al, 1970 p48-50] and adopted in constructing the DSM-III diagnostic models, i.e. prior to its publication, the National Institute of Mental Health funded "field trials" to measure "inter-rater reliability" for the newly defined DSM-III disorders, employing methods similar to those used in developing the PSS, i.e. the correlational relationship was determined between the results of "pairs of clinicians making independent diagnostic judgments of patients".[DSM-III, 1980 p5] It may therefore be argued that the DSM-III marked the successful outcome of reasonable efforts by Spitzer and others to introduce "scientific" objectivity into the psychiatric diagnostic process, to answer both internal and external challenges to a specialty whose clinicians had been "largely on (their) own in defining the content and boundaries of diagnostic categories",[DSM-III, 1980 p8] and in most instances without physical signs in support of clinical judgment. On the other hand it may also be argued that there is potentially a greater negative impact from "defined criteria", an "increased commitment in (the) field to reliance on data as the basis for understanding mental disorders" and a "scientific approach to diagnosis"[Guimon, 1989 p33] than the benefits derived from eliminating perceptions of psychiatry as a discipline which engages in subjective or arbitrary labelling based upon a model for which there is a lack of consistent neuroscientific evidence.

Guimon posits that psychiatric diagnosis and classification "involves a loss of information, the magnitude of which is determined by the use to which the diagnosis is put and the assumptions and theoretical orientation of the clinicians involved in the exercise". He proposes that the DSM-III was developed in the context of the Neokraepelinian paradigm which "holds that psychiatric disturbances are real illnesses" and that underlying this model is the imputation of a correlational relationship between physical indices, known or unknown, and behavioural evidence, which in the DSM-III is "measured" in the number and severity of "symptoms" of "major depression", "panic disorder", or other defined conditions.[Guimon, 1989 p33-5] However, the DSM-III claims to be "generally atheoretical with regard to aetiology", its authors arguing that "inclusion of aetiological theories would be an obstacle to use of the manual by clinicians of varying theoretical orientations" but it is also stated in the introduction that in addition to the "organic disorders" for which "patho-physiological processes" have been identified or are "presumed", "specific biological aetiologies" or an "interplay" of biological, psychological and social factors will "undoubtedly" be found for other mental disorders of "unknown aetiology".DSM-III, 1980 p6-7 It may therefore be proposed that there is evidence in support of Guimon's contention that the assumptions of the DSM-III "task force" were defined by the potential explanatory power of the new paradigm. On the other hand, he is selective in his rejection of its validity and applicability to the diagnostic process.

Guimon argues persuasively that "psycho-pathological phenomena" are not susceptible to measurement in the same sense as the testing of blood, blood pressure etc., and that "the standards of 'objectivity' in measuring and description which are demanded by statistical analysis have not yet been achieved". From his own orientation, which is implicitly based in Freudian psycho-dynamic theory, he rejects the selection of "phenomena susceptible to observation and description" as proposed "symptoms" or diagnostic criteria in the DSM-III, to the exclusion of phenomena registered through "intuition" or "subjective interpretation".[12] However, although Guimon challenges the validity of the attempt by those formulating the manual both to "deal with diagnosis on an empirical basis" and to impute an association between "psycho-pathology" and physical processes he argues that when studying "organic brain disorders" researchers "can afford to look for hard physical indices".Guimon, 1989 p34-6

---

[12] Guimon does not discuss the fact that the DSM-III diagnostic models were themselves subjectively defined by the "task force" members and "subjective interpretation" was not eliminated from the "symptoms" itemised in the manual which include "sexually seductive appearance", "indiscriminate socialising" and "impaired judgment".DSM-III, 1987 p521-3

Lipowski, who interpreted "biochemical processes" underlying mental activity and behaviour not as "causal agents" but rather as "mediating" between information received from the environment and its subjective meaning (see above), argued that biological psychiatry "confuses the distinction between aetiology and correlation". However, while discounting the current value of "neuroscientific progress" to clinical practice in psychiatry he did concede its potential to "discover" a "cure" for the dementias.[Lipowski, 1989 p252-3] It may therefore be proposed that critiques of the neurophysiological model have been applied predominantly to those mental states which have been defined as "functional" or "psychogenic" and not to the "organic" disorders, but as acknowledged in the introduction to the DSM-III, even for these conditions it may be necessary to "presume" causation by "pathophysiological processes".[DSM-III, 1980 p6] However, when "Kraepelinian" studies, such as those conducted by Blessed et al. in 1968 and Crystal et al. in 1988, have sought to establish a correlation between autopsy findings in the brain and behavioural evidence of dementia *in vivo*, "discrepancies" have been found.[Blessed et al, 1968 p797 Crystal et al p1682] Furthermore the aetiology or pathogenesis of changes identified in those with clinical evidence of dementia has yet to be determined and in consequence the dilemma of the "functional" link between the "brain" and the "mind" remains unresolved,[McLoughlin & Lovestone, 1994 p439] (see Chapter 2). It is, however, from the "organic disease" model supplemented by evidence of "desirable" responses to psychotherapeutic drugs and statistical evidence of "reliability" between diagnosticians that scientific legitimacy has been imputed for the new paradigm and its diagnostic models.

Many years ago, Szasz proposed that the concept of mental illness undermines the principle of personal responsibility, inviting the treatment of individuals as "irresponsible patients",Szasz, 1972 p269 and it may argued that the development of "objectively" determined and "reliable" diagnostic models within the biological paradigm reinforces this proposition: the individual is not held to be responsible for the "disease" but its control is abdicated to the profession. Moreover, with the professional imputation of "physical" causation, the potential for implying "misdiagnosis" or "inappropriate labelling" may be rendered more difficult for the laity, encouraging more acceptance of diagnoses which are, arguably, rendered less stigmatising. On the other hand, it was earlier proposed that it is by the successful development of preventive or therapeutic measures that the mark of disgrace is removed and in the absence of a "cure" stigma will not necessarily be diminished. However, the correlational diagnostic model implies legitimacy for ascribing diagnoses at an early or sub-threshold stage of the disease, an expansionary potential will that will be realised with general acceptance of the scientific validity of the paradigm.

Guimon argues that "changes in psychiatric fashion in the U.S.A. have had international impact".Guimon, 1989 p33 However, in 1974 Spitzer compared data from several studies which showed mean rates of agreement between psychiatrists in arriving at a diagnosis for some eighteen different disorders, ranging from a low of 0.24 to a high of 0.77. In particular there were "strikingly large diagnostic differences" between psychiatrists of different nationalities, American doctors having much higher rates of detection of "psychopathology".Spitzer et al, 1974 p344-5 [13] Uniform and therefore higher rates

---

[13] Videotape recordings of a small number of psychiatric interviews shown to a large number of British, Canadian and American psychiatrists elicited in one case a diagnosis of "schizophrenia" from 2 percent of British psychiatrists compared with 69 percent in the United

of diagnosis of mental disorders might be predicted from international use of the DSM-III and its "objective" diagnostic criteria but as demonstrated by Guimon and Lipowski, there is resistance to the underlying assumptions of the biological paradigm and its therapeutic implications by some foreign psychiatrists. On the other hand their critiques relate largely to "psychogenic" and not "organic" disorders, and intuitively it might be assumed that increased rates of diagnosis, deriving in part from early "detection" would be among those "discredited" both by the "negative stereotyping" of old age[Featherstone et al, 1993 p308] and the highest prevalence of stigmatising "organic disease", i.e. "dementia" for which a new model was also formulated in DSM-III.

Dementia: the new diagnostic paradigm and its differential diagnosis

In the DSM-III, "senile dementia" was redefined as "dementia", of which "primary degenerative dementia of the Alzheimer type is the most common", and was categorised as an "organic mental disorder" rather than "psychosis associated with organic brain syndromes" as in the DSM-II.[DSM-II, 1968 p24] An extensive description of the predominant features of the condition was given which were summarised as specific "diagnostic criteria" but the definition in the narrative description and the primary criterion were essentially the same, i.e. "a loss of intellectual abilities of sufficient severity to interfere with social and occupational functioning". The other criteria were "memory impairment" and at least one other identifiable "symptom" from a list which included "impaired judgment" and "personality change" together with "disturbances of higher cortical function" as determined by activities including "constructional difficulty" or items that

---

States and an intermediate percentage in Canada.[Spitzer et al, 1974 p345]

Table 3.1

DSM-III Criteria for Dementia

A. A loss of intellectual abilities of sufficient severity to interfere with social or occupational functioning.

B. Memory impairment

C. At least **one** of the following:

1. Impairment of abstract thinking as manifested by concrete interpretation of proverbs, inability to find similarities and differences between related words, difficulty in defining words and concepts, and other similar tasks.

2. Impaired judgement

3. Other disturbances of higher cortical function, such as aphasia (disorder of language due to brain dysfunction), apraxia (inability to carry out motor activities despite intact comprehension and motor function), agnosia (failure to recognise or identify objects despite intact sensory function), "constructional difficulty" (e.g. inability to copy three-dimensional figures, assemble blocks, or arrange sticks in specific designs)

4. Personality change, i.e. alteration or accentuation of premorbid traits

DSM-III, 1980 p111

may be defined as "puzzle solving".[DSM-III, 1980 p107-111]

Although narrative descriptions of senile dementia had remained largely unchanged from Kraepelin's 1906 classification until 1980, and even though the newly constructed DSM-III dementia model was included in "inter-rater reliability" studies conducted prior to publication of the manual,[DSM-III, 1980 p470] the definition of dementia and the diagnostic model became subject to "modification" within seven years with the publication of the DSM-III-R. By 1987, a "loss of intellectual abilities of sufficient severity to interfere with social and occupational functioning" was no longer the primary criterion for diagnosis of "dementia". The primary "symptoms" became: "demonstrable evidence of impairment in short and long-term memory" and one from the previous list which included "higher cortical function" items including "constructional difficulty". A third criterion was the specification that the first two items should interfere "significantly" with "work, usual social activities or relationships with others". In addition, sub-categories were introduced for classifying evidence of dementia as "mild, moderate or severe", in what may be interpreted as the imputation of a correlation between the level of objectively determined "impairment" and pathological changes in the brain (see Figure 3.2).[DSM-IIIR, 1987 p111] In both the DSM-III and DSM-III-R "physical" diagnostic criteria were also specified, i.e. "evidence from the history, physical examination, or laboratory tests of a specific organic factor that is aetiologically related to the disturbance". In the absence of such evidence, however, "an organic factor necessary for the development of the syndrome can be presumed".[DSM-III, 1980 p112, DSM-IIIR, 1987 p107]

Guimon, in his critique of the DSM-III models argues that "psychiatric diagnosis and classification involves a loss of information, the magnitude of which is determined by the use to which the diagnosis is put and the assumptions and theoretical orientation of the clinicians involved in the exercise".[Guimon, 1989 p33] On the other hand, while rejecting the potential for relating "objective" assessments of "psycho-pathological" phenomena to underlying physical disease processes does not reject the underlying assumptions of the model as they relate to mental disorders classified as "organic",[Guimon, 1989 p33-35] (see above). However, Berrios, another European psychiatrist expresses similar concerns with regard

Table 3.2

DSM-III-R Criteria for Dementia

A. Demonstrable evidence of impairment in short- and long-term memory. An individual is unable to learn new information, or to remember things well-known in the past, such as birthdates or addresses.

B. At least **one** of the following:

1. Impairment in abstract thinking, as shown by inability to find similarities and differences between related words, and difficulty in defining words and concepts.
2. Impaired judgement, as indicated by inability to plan or cope with financial, family or other issues.
3. Other disturbances of higher cortical function, such as aphasia (disorder of language); apraxia (inability to carry out motor activities despite intact motor function); agnosia (failure to recognise or identify objects despite intact sensory function); "constructional difficulty" (e.g. inability to copy three-dimensional figures).
4. Personality change, i.e. alteration or accentuation of premorbid traits.

C. The disturbance in A and B significantly interferes with work or usual social activities or relationships with others.

Criteria for severity of Dementia:

Mild: Work or social activities are significantly impaired but capacity for independent living remains with adequate personal hygiene and relatively intact judgment.

Moderate: Independent living is hazardous and some degree of supervision is necessary.

Severe: Activities of daily living are so impaired that continual supervision is required, e.g. unable to maintain minimal personal hygiene.

DSMIII-R, 1987 p107

to the DSM-III model for "dementia" and in particular its DSM-III-R modification, perceiving it to be narrowly focussed and lacking power to discriminate between dementia and other mental disorders of old age.

Those relatively few psychiatrists who have studied the mental health of the elderly during the past thirty to forty years consistently found "depression" or "affective disorders" to be the most prevalent diagnosis, with an onset often associated with physical illness and bereavement, both of which have a high incidence in late life.[Roth et al, 1952 p79, Kay et al, 1964 p150, Gurland, 1982 p11-26 Copeland et al, 1987 p468] Furthermore, depression in the elderly has been found to present as "pseudodementia" with evidence of difficulty concentrating, memory loss, self-neglect, and psychomotor retardation or agitation, signs and symptoms which are also characteristic of dementing disease.[Roth et al, 1952 p68-9, Roth, 1976 p80-1, Blazer, 1993 p165-6] In addition, and closely related to the prevalence of chronic illness among older people is multiple medication use, both prescribed and non-prescribed, which may induce pseudodementia as well as other "neuro-psychiatric" phenomena including mild states of "delirium"[14] or clouding of consciousness.[Burks, 1979 p69-78, Blazer, 1983 p31-5] Berrios argues that the new model which defines dementia "within a very narrow cognitive compass" cannot discriminate reliably in "cross-sectional diagnosis" between "true dementia and its behavioural phenocopies, e.g. depressive pseudodementia, quiet delirium etc." thereby undermining the complexity of the diagnostic process.[Berrios, 1989 p11]

He proposes that dementia is characterised by "major disintegration of psychological organisation" which is "likely to involve systems such as perception, motility,

---

[14] Delirium is further discussed in Chapter 5.

personality, emotional experience and volition". Furthermore, like Kitwood who suggests that deleterious effects may be engendered by a "malignant social psychology" towards those with dementia,[Kitwood, 1990 p184] Berrios does not discount "psychodynamic" influences in its aetiology. He argues, however, that "personality" and other symptoms which were, historically, "part of the definition of dementia" are "complex and difficult to operationalise" and that an underlying rationale for the DSM-III and DSM-III-R was the development of "diagnostic instruments" for research purposes.[Berrios, 1989 p11-14] However, instruments had been developed by others who sought both to take account of the multidimensional nature of the dementia syndrome and to distinguish it from other mental disorders. For example, when Blessed compared "quantitative measures" with post-mortem findings in subjects with and without evidence of senile dementia the Dementia Scale developed for the study consisted not merely of items related to intellectual functioning, "memory" and "concentration", but also "performance of everyday activities", and changes in "habits", "personality" and "interests".[Blessed et al, 1968 p808-9] In addition, the Geriatric Mental State Schedule (GMS), developed in the 1970s covers symptoms of all psychiatric conditions prevalent in old age and seeks information on physical illness, with attention to "hearing" and "disabilities of speech and vision" which may confound the reliability of diagnoses in the elderly.[Copeland et al, 1976 p441] Furthermore, the GMS was used in the development of AGECAT, a computerised diagnostic system designed to assist in epidemiological research,[Copeland et al, 1986 p89-99] a purpose for which it has been employed.[Copeland et al, 1987 p466-473, Blessed et al, 1991 p193-8]

Therefore it may be argued that others had addressed the problem of achieving greater reliability in psychiatric diagnosis in old age while acknowledging the complexities of the process. On the other hand, it may also be argued that none of the above instruments meet standards for total scientific "objectivity": changes in "personality", "habits" etc. in the Blessed Dementia Scale are reported by a relative or other informant; the GMS is a "semi-structured" psychiatric interview [Copeland et al, 1976 p439] and AGECAT requires ranking of symptoms by the interviewer for sorting by computer into diagnostic "clusters".[Copeland, 1987 p467] The DSM-III-R model, described by Berrios as the "cognitive paradigm" for dementia, in which the primary, but subjectively determined, DSM-III diagnostic criterion, i.e. a loss of intellectual abilities of sufficient severity to interfere with social or occupational functioning, was displaced by "memory impairment", may be perceived as an attempt to enhance "reliability". However, it may also be interpreted as a retrospective legitimation for the instrument adopted to determine the prevalence of "organic" mental disorders in a major epidemiological study undertaken by American psychiatry in which the DSM-III diagnostic models were "operationalised".

The National Institute of Mental Health and quantitative legitimacy

Guimon, in his critique of what he perceived to be a "Neokraepelinian biological paradigm" argued that in view of the "excessive success" achieved by a small group who have a high proportion of citations in the literature, there would seem to exist an "Invisible College" who, through their political power facilitate views and certain trends".[Guimon, 1989 p34] It may be proposed that insofar as it exists the "college" comprises leaders of the mutually supporting institutions of the American Psychiatric Association (APA) and the National Institute of Mental Health (NIMH) deriving power both from professional authority and from resources allocated to psychiatry via the national Institutes of Health, the "national medical research enterprise",Strickland, 1972 p216 which enable the specialty to document the reality of its diagnostic constructs in epidemiological studies, the results of which, when disseminated, may serve to reinforce their perceived validity and that of the paradigm within which they were conceived.

The leaders of the APA observed that the rationale for compiling the first Diagnostic and Statistical Manual of Mental Disorders (DSM) was the inadequacy of the existing nomenclature of mental illness, "developed primarily for the needs and caseloads of public mental hospitals"DSM, 1952 pv-vii (see above). However, Starr argues that psychiatry "emerged from World War II with an enhanced public image" deriving from: the rejection of more than 1 million men for military service "because of mental and neurological disorders"; the diagnosis of 850,000 "psycho-neurotic cases" during the 2nd World War; and "unexpected success in the treatment of psychiatric patients".[15] Furthermore, "these statistics (were presented) as measures of America's great unmet need for psychiatric services" in support of the Mental Health Act, 1946 which founded the National Institute of Mental Health (NIMH).Starr, 1982 p344-5 It may therefore be

---

[15] The introduction to the DSM acknowledges that the diagnoses assigned to military recruits were "specifically not designed for 90% of the cases handled" and that "minor personality disturbances" had to be "classified as "psychopathic personality" and the "'psycho-neurotic label' had to be applied to men reacting briefly with neurotic symptoms to considerable stress", the latter, arguably, accounting for the "unexpected success" in recovery rates.

Szasz in a work on the "manufacture of madness" discussed the influence of William C. Menninger, chief psychiatrist for the military during the 2nd World War, noting that nearly 2 million out of 15 million individuals were rejected for military service and that 37 percent of all medical discharges from the armed services were "on grounds of neuropsychiatric disability". He argued that "if a psychiatrist's greatness is measured by the number of people he 'diagnoses' as mentally ill then William Menninger was a very great psychiatrist indeed."Szasz, 1970 p38-9 proposed that "scientific" legitimacy was derived from "statistics" and that numerical data gathering was subsequently institutionalised by American psychiatry in the title of its "diagnostic and statistical manual" and in the early establishment of the "biometrics branch" at the NIMH.[DSM, 1952 px] However, as previously described, prior to the publication of the DSM-III there were no "defined criteria" for the statistical evidence collected.[DSM-III, 1980 p1]

In 1980, the year in which the DSM-III was published, the National Institute of Mental Health (NIMH) funded the Epidemiologic Catchment Area Study (ECA), a survey of approximately 20,000 individuals in five cities in the United States, conducted to determine the lifetime and current prevalence of mental disorders, both "treated and untreated". The conditions detected by the questionnaire assessment instrument employed in the study, i.e. the "Diagnostic Interview Schedule" (DIS), were those which had "evolved into the official nomenclature of the DSM-III".[Regier et al, 1991 p1-9] The DIS did "not attempt to cover all 122 adult diagnoses in DSM-III". For example, "disorders difficult to identify in the absence of a medical history or examination ..... were omitted".[Leaf et al, 1991 p13] Given the problems in distinguishing between dementia as a psychiatric diagnosis in the elderly and its "behavioural phenocopies",[Berrios, 1989 p11] and given the high prevalence of chronic physical conditions in old age and the effects of medications taken for them, intuitively it might be assumed that dementia or related conditions would be excluded from the DIS. However, "the ECA investigators felt it important to consider organic mental disorders" and to this end "cognitive impairment" was considered "the closest approximation".[George et al, 1991 p291-4] Furthermore, in contrast with the "yes/no" answers elicited to determine whether respondents had experienced symptoms of other DSM-III mental disorders, participants were asked to respond to a series of "mental tests".

Senile dementia, arguably a prototypical organic mental disorder, had appeared in the nomenclature of mental diseases from the time of Kraepelin (see above). However, although the National Institute of Mental Health (NIMH) was established in 1946, it was not until after the condition, redefined as Alzheimer's disease, had been adopted by the National Institute on Aging as its own research focus that a Mental Disorders of Ageing Research Branch was established by the NIMH. The DSM-III model for "dementia", its operationalisation in mental testing in the ECA study, and the interest in research into mental disorders in old age may therefore be interpreted in the context of the attempt by American psychiatry to achieve legitimacy for the "objective" correlational diagnostic model conceived within the "biological paradigm" for mental disorders, but also of the resources allocated by the U.S. Congress to the biomedical community for the investigation of a condition disregarded by the specialty within whose ambit it had fallen. On the other hand, the diagnostic mental testing model adopted for the detection of senile dementia and other organic disorders on behalf of its own research activities and those of other medical specialties and disciplines was one which had been largely rejected by psychology, the social science within which it had been conceived and elaborated.

Chapter 4

THE MENTAL TESTING ENTERPRISE

"Mental tests" and the legitimating "paradigms of inquiry"

In 1884 Francis Galton (1822-1911) invited visitors to the International Health Exhibition at the South Kensington Museum to pay threepence to enter the "Anthropometric Laboratory". Skull measurements, as a proxy for brain size, were taken, and entrants were asked to subject themselves to a series of tests of memory, reaction time and sensory acuity. These assessments conducted on "established principles" were designed to support the theory of "natural ability", based upon a perceived association between the "efficiency" of the central nervous system and "intelligence". More than nine thousand men and women were sufficiently intrigued to pay for these first large-scale "mental tests".[Fancher,1985 p41] Galton acknowledged the limitations of these observations as indices of "intelligence", describing them as designed to "obtain a general knowledge of the capacities of man by sinking shafts, as it were, at a few critical points".[1]

In spite of the inability of Galton or others including Alfred Binet (1857-1911) in France and Lewis Terman (1877-1956) in the United States, described by Fancher as "the intelligence men", to arrive at a consistent definition of the concept they were seeking to quantify, "mental tests" were institutionalised within state educational systems over the next decades, and 1.75 million conscripts to the United States Army for the First World

---

[1] From Galton's appendix to: Cattell J.M. Mental tests and measurements. Mind 15:373-81, 1890.

War were subjected to and classified by such assessments.Fancher, 1985 pp117-49, Kevles, 1968 p565-81 In response to the success of an international enterprise which included not only "the whole American army" but his own testing of 30,000 candidates for the Civil Service in Britain, Charles Spearman (1863-1945), a disciple of Galton, wrote in 1923: "signs are not lacking that, before many years, the measuring will have been extended to the entire civilized population of the earth, and perhaps to much even of the uncivilized".Spearman, 1923 p2

Foucault, in his work "*Surveiller et punir*" argued that "the growth of a capitalist economy gave rise to the specific modality of disciplinary power, whose general formulas, techniques of submitting forces and bodies, in short 'political anatomy' could be operated in the most diverse political regimes, apparatuses and institutions".Foucault, 1979 p221 At one extreme of Foucault's conceptualisation of a "great carceral continuum" is the "enclosed institution" as represented by the prison or "house of correction" and by the insane asylum. At the centre lies the "means of correct training" as embodied in the military, but extending to the school and hospital, with the disciplinary power of "hierarchical observation, normalising judgement and .... the examination.Foucault, 1979 p184 At the other extreme is "panopticism",[2] a "more subtle discipline of generalised surveillance outside the formally egalitarian juridical framework, but existing in the "systems of micro-power that are essentially non-egalitarian and asymmetrical that we

---

[2] Jeremy Bentham, in a series of letters published in 1791, devised the principles of the "Panopticon" as a means of improving prison discipline. Although proposed as an architectural structure, Foucault perceived an analogy between the efficiency and inescapability of "panoptic" surveillance and the non-institutional disciplinary systems and methods of the "human sciences".Foucault, 1979 p195-205

call the disciplines". Foucault argues that the methods developed within the disciplinary technologies, i.e. the "tests, interviews, interrogations and consultations" became the "methodology for the empirical sciences". He further argues that in adapting these technologies to the sciences of clinical medicine, psychiatry, psychology, criminology, education etc. the mechanism of "objectification could become the instrument of subjection", with mutual reinforcement between the "formation of knowledge and the increase of power".[Foucault, 1979 p222]

Scull endorses Foucault's insight into the relationship between power and knowledge, and into the systems of control which may derive from ostensibly "benevolent" interventions but argues that Foucault offers no "systematic" discussion of the political or economic context in which the "disciplinary technologies" were constructed.[Scull, 1989 p14-19] Scull posits that a central achievement of industrialisation was the imposition upon the workforce of a system of domination based upon the internalisation of new attitudes and responses in the promotion of self-discipline. He further argues that from the success of this "form of compulsion" evolved a perception that the character of the workforce was no longer "fixed", but was "plastic and amenable to improvement through appropriate management and training" in factories, schools, prisons, and other "structurally similar techniques of social discipline".[Scull, 1989 p91-2] It may therefore be argued that the "systems of micro-power" and "objectification" were conceived within the paradigm of classical economic theory with its rationalist conception of the social order and focus on the individual as the only significant unit of analysis.[Abrams, 1968 p8] It may be further argued that the institutions of "social discipline", rather than developing methods of objectification and surveillance, became "laboratories" for the employment of techniques constructed by the "human sciences" within a positivist model of human intellect and social behaviour compatible with political economy but with implications beyond "appropriate management and training".

Berger and Luckmann posit that "the relationship between ideas and their sustaining social processes is always a dialectical one ..... theories are concocted to legitimate existing social institutions and social institutions are changed in order to bring them into conformity with existing theories". Berger & Luckmann, 1966 p128 They also posit that any "significant theme (which) spans spheres of reality may be defined as a symbol, and the linguistic mode by which such transcendance is achieved may be called symbolic language" with the potential for construction of "immense edifices of symbolic representation that appear to tower over the reality of everyday life". Berger & Luckmann, 1966 p40 It may be proposed that in the mid-19th century a linguistic edifice deriving from the "symbol system" of "science" was constructed upon the Darwinian concept of "natural selection", i.e. "survival of the fittest implies multiplication of the fittest".[3] It may be further proposed that this linguistic construct became not merely an "objectively real element in everyday life" if not a "doctrine (which) becomes an ideology", Berger & Luckmann, 1966 p124 but that Herbert Spencer articulated a biological paradigm for human society already implicit in the evolving social sciences which legitimated the attribution of

---

[3] Charles Darwin (1809-1882) published his treatise "On the Origin of Species by Natural Selection" in 1859. In 1865 Herbert Spencer (1820-1903) wrote in "The Principles of Biology": "This survival of the fittest which I have here sought to express in mechanical terms is that which Mr Darwin has called 'natural selection' or the preservation of favoured races in the struggle for life". Herbert Spencer, 1865 Chapter 12, p3

inequality in social institutions to physical or "hereditary" differences between individuals. At the same time, based upon a "medical" model, it afforded a rationale for interventions when based upon "scientific" evidence produced by techniques for identifying both the "fittest" and the "unfit".[4]

Berger and Luckmann, in describing "ideology" as "a definition of reality attached to a concrete power interest", proposed that "the distinctiveness of ideology is that the same overall universe is interpreted in different ways, depending on concrete vested interests within the society in question". Berger & Luckmann, 1966 p124 From this perspective it may be argued that many different individuals and groups were accommodated by and co-opted in support of the mutually reinforcing ideologies of *laissez faire* capitalism and the 19th century biological paradigm from which derived the "mental testing enterprise". These included philosophers contemplating the scientific method of empirical research in relation to human society; reformers with a Benthamite vision of "statistics" as the "science of social reform"; Eyler, 1979 p20 mathematicians who applied their skills to advancing techniques for evaluating quantitative data, and researchers in "individual psychology" such as Alfred Binet whose interest in age-specific mental testing originated in observing the development of his own children. Fancher, 1985 p59-60

In 1860, five years before Spencer's treatise and twenty years before Galton began

---

[4] The Sixth Annual Report of the Council of the Journal of the Statistical Society of London, 1840 contains the following passage:

"By this cultivation of statistics of all sorts only can we arrive at a knowledge of the physiology of societies, and comprehend the paroxysms of disease which they sometimes exhibit in a state of violence, or the exhilaration of health, which displays itself in time of peace. Empirical treatment of symptoms, without this knowledge, must be as vain in its effects upon the body politic as upon the human frame". Eyler, 1979 p20 to "sink shafts" to measure the "capacities of man", it was proposed by the British delegate to the International Statistics Congress that an effort be made "to measure the population (of European countries) collecting data not only on stature, weight and strength, but on intelligence".[5] The speaker, Dr. William Farr, superintendent of the Statistical Department of the General Register Office for forty years and a pioneer in the use of epidemiological data to achieve sanitary reforms was concerned with the "racial" consequences of his work in increasing the potential survival of the "unfit" a definition of reality promoted as eugenic science by Galton.[Eyler, 1979 p154-5] [6] It may be argued, however, that while seeking to limit state intervention, the extensive "anthropometry" advocated by Farr could be introduced only via publicly financed institutions, ie. Foucault's "means of correct training".[Foucault, 1979 p184] In consequence it was with the institutionalisation of state education that legitimation was achieved by psychologists for applying the technologies of objectification and classification to "intelligence", from which evidence could be derived in support of "benevolent" interventions for "co-optative control" of those determined to be the "fittest" and the exercise of "exclusionary control" over the "unfit".[7] On the other hand, in spite of the expansionary ambition of Spearman

---

[5] William Farr was the leading British delegate to the International Statistical Congress which held nine sessions in various European capitals between 1853 and 1876.[Eyler, 1979 p34]

[6] Francis Galton founded the Eugenics Society and in 1904 established a chair in Eugenics at University College, London.[Eyler, 1979 p208]

[7] There is evidence that in Germany, France, Britain and individual states in the U.S., legislation was passed to "discover" all mentally retarded children.[Stern, 1914 p7] In Britain the Mental Deficiency Act of 1913 laid upon the local authorities the duty of documenting all mentally defective children in their areas between the ages of seven and sixteen. This resulted in rates ranging from 0.73 to 16.14 per 1000 children.[Armstrong, 1983 p64] The "exclusionary control" exerted over such children ranged from segregation in special schools to commitment to insane

for "measuring" the "population of the earth", Spearman, 1923 p2 social scientists did not achieve sustained or universal legitimacy for quantifying the concept in the adult population. In in the past fifteen years, however, a mental testing enterprise has been constructed within the Neokraepelinian biological paradigm for mental disorders (see Chapter 3) which has, realised if not exceeded the expansionary anthropometric ambitions of Farr in the 19th century and Spearman, his colleagues and mentors in the early 20th century (see below).

Within the 19th century social-Darwinian biological paradigm, it was the "intelligence" of youth, a concept which was arguably synonymous with "rationality"; the perceived correlation of this construct with "natural ability", and its cumulative impact on the competitive equilibrium of the free market which legitimated the activities of the "intelligence men". In the late 20th century, however, it is the "problem of ageing" which confronts the capitalist economies (see Chapter 1), a perceived threat which has, it may be proposed, served to promote universal acceptance of the model for dementia formulated by American psychiatry within the biological paradigm for "mental disorders", and which has legitimated "panoptic surveillance" of a different demographic group, i.e. the elderly. On the other hand, the "methodology of the empirical sciences" which was developed within the earlier enterprise has been appropriated in generating data on "intelligence" in old age, redefined as "cognition". This new enterprise is in turn is based on a model which is now imbued with the diagnostic authority of medicine, but

---

asylums. Furthermore several states in the U.S. passed laws for involuntary sterilisation of the "feeble minded". Pickens, 1968 p21, Fancher, 1985 p114 which had been largely rejected with the elaboration of knowledge by psychology, the human science within which "mental testing" was first conceived and "operationalised".

Professional autonomy and "applied" psychology

Freidson argues that "a profession attains and maintains its position by virtue of the protection and patronage of some elite segment of society which has been persuaded that there is some special value in its work. Its position is thus secured by the political and economic influence of the elite which sponsors it - an influence that drives competing occupations out of the same area of work, that discourages others by virtue of the competitive advantages conferred on the chosen occupation, and that requires still others to be subordinated to the profession". However, "if a profession's work comes to have little relationship to the knowledge and values of its society it may have difficulty surviving. The profession's privileged position is given by, not seized from society, and it may be allowed to lapse or may even be taken away". Freidson also distinguishes between those which he describes as the "consulting or practicing professions" and the "scholarly or learned professions". He argues that "unlike science and scholarship which create and elaborate the formal knowledge of a civilisation, practicing professions have the task of applying that knowledge to everyday life" and must therefore gain the confidence of a "lay clientel" in order to survive. He further argues that in order to maintain their autonomous position in society some linkage must be sustained between the work of practicing professions and that which "the layman considers desirable and appropriate".[Freidson, 1973 p72-4] Based upon this argument it may be proposed that "intelligence psychology" achieved a brief period of "autonomy" as a consulting profession but that by failing to gain the confidence of the laity for extending its measuring beyond childhood its privileged position was "allowed to lapse".

Pickens, discussing the "progressive movement" in the United States in the early 20th century, one that promoted "social science" within the biological and eugenicist paradigm as a guide to "reform matters", argues that its proponents "had tremendous influence upon American social, political and academic life".Pickens, 1968 p4 It may therefore be posited that legitimacy for mental testing of U.S. Army recruits in 1917, in another Foucaultian institution for "correct training", derived from the authority accorded "intelligence" practitioners under the movement's sponsorship. Furthermore the data generated was successfully employed in "reform matters", i.e. in support of "exclusionary control" legislation.[8] On the other hand, Kevles documents a lack of congruence between the "knowledge and values" of the military and the activities of the "mental meddlers" whose contemporary critics argued that the army had been made "a laboratory for their own purposes".Kevles, 1968 p574 [9] Furthermore, the results of the

---

[8] The 1917 U.S. Army tests were conducted under the direction of Robert Yerkes (1876-1956) chairman of the American Psychological Association who argued that "psychology has achieved a position which will enable it to substantially help to win the war and shorten the necessary period of conflict".Kevles, 1968 p571

The results of employing "scientific methods", to test both English and non-English speaking conscripts, literate and illiterate, were used as evidence of the low "native intelligence" of recent immigrants from Eastern European and Mediterranean countries. In the 1924 Immigration Act the U.S. Congress set minimal quotas for residents of these countries to enter the United States in order to prevent a further lowering in "national mean intelligence".Fancher, 1985 p128

[9] Kevles in a paper on "psychologists and the military in World War 1" notes that a much less ambitious psychology project was also legitimated in a scheme to assist in the "selection of officers by rating personality traits".Kevles, 1968 p569 This was evaluated as successful and the Committee on the Classification of Personnel, was incorporated into the military organisation

enterprise, which indicated that "the average soldier was nearly a moron",[Pickens, 1968 p152] attracted critical evaluation both from the "lay" public and from social scientists on both sides of the Atlantic.[10] It may therefore be proposed that the "intelligence men" lost their independence as a "practicing profession" in failing to convince the layman that extending the enterprise to the general adult population was "desirable and appropriate", and while psychology continued and diversified as an academic discipline with competing schools or paradigms of inquiry, applied "intelligence psychology" came largely under the sponsorship of medicine and the specialty with which it had shared a congruent organic or biological model of mental functioning, i.e. psychiatry.[11]

In 1939 David Wechsler (1896-1981) published the first edition of the Wechsler Adult Intelligence Scale (WAIS), a test battery that is still in use.[Lezak, 1983 p241, Chase et al, 1984 p1244-7] He noted in the narrative that the medical profession had been inclined to restrict the psychologist to "apersonal psychometrics", as in the group testing of school children. His role as chief psychologist in the psychiatric division of a New York teaching hospital

---

at the end of the War whereas the intelligence testing enterprise was abolished.[Kevles, 1968 p579]

[10] Fancher cites the example of a series of articles written by Walter Lippman, a distinguished journalist, who proposed that the standards applied to the army tests must have been "ludicrously inappropriate".[Fancher, 1985 p130] In a lecture given at the London School of Economics in 1923 on "The measurement of progress in public health", Arthur Newsholme argued that "assuming the drafted men are a fair sample of the entire (American) population of approximately 100 million (the results) justify scepticism as to the significance of the tests employed in this gigantic series".[Newsholme, 1923 p200]

[11] William Stern (1871-1938), writing on the "psychological methods of testing intelligence" in 1914 argued that "adult abnormal individuals form the chief material of the psychiatrists" who had devised "whole series or systems of examinations (but) by far the greater portion of them take on the character of questions and qualitative tests rather than that of quantitatively gradable tests". In consequence, "physicians turn to psychologists to secure normal series for the various psychiatrical tests of intelligence".[Stern, 1914 p6-7]

was that of "intelligence diagnosis" from the results of which he listed WAIS scores associated with different specific mental disorders.[12] However, he also observed that "even more enlightened psychiatrists look upon an intelligence examination as they do upon a (blood) test; when an I.Q. does not come up to expectations they feel privileged to disregard it",Wechsler, 1939 p48-9 i.e. applied "intelligence" psychology had become a "paramedical occupation".

Medicine, "paramedical occupations" and the "scholarly professions"

As previously discussed, Paul Starr, in his treatise on the "social transformation of American medicine" argued that medicine has achieved an unequalled position in the social structure combining "broad cultural authority, economic power and political influence", and that its special claim to authority was based in part upon its "bonds with modern science" and the privileged status of scientific knowledge in the hierarchy of belief",Starr, 1982 p4-6 (see Chapter 1). Similarly, Freidson posited that "medicine's position today is akin to that of state religions yesterday" and that this status derived both from a "secure and practical technology", which had been essential to achieving the confidence of the lay clientel, and from "public belief" in the "scientific foundation" of medical practice.Freidson, 1973 p21-2 He argued that such legitimation had enabled medicine to extend its jurisdiction beyond any "capacity to cure" not merely into those pursuits he defined as "moral entrepreneurship", i.e. seeking out that to which it has assigned the social

---

[12] In each edition of the WAIS, Wechsler published test scores for "normal" subjects and for those with a changing nomenclature of mental disorders. In 1941 these included "schizophrenia", "adolescent sociopaths", "sex delinquency", "neurotics" and "mental defectives".Wechsler, 1941 p149-50 In 1958 "sex delinquency" had disappeared but "anxiety states" were introduced.Wechsler, 1958 p171-3

meaning of "deviancy" or "illness",[Freidson, 1973 p251-3] (see Chapters 2 & 3) but also to control over groups defined as the "paramedical occupations".[Freidson, 1973 p48]

Freidson proposes that "in a way unparalleled in any other industry, the physician controls and influences his field and all who venture near it. Whereas the division of labour in non-professionalised fields is ordered by historical accident, economic and political power, competition and purely functional inter-dependence, the division of labour surrounding the highly professionalised activity of healing is ordered by the politically supported dominant profession." On the other hand, "the services of many other occupations are useful to the physician ..... even if dangerous to his monopoly. The solution to the problem is not, as may be the case with direct competitors, to drive them out of practice but rather to promote state control over those occupations so that medicine may supervise or direct their activities".[Freidson, 1973 p47-9] It may be argued, however, that the power of medicine extends beyond the "paramedical occupations" which may be "useful to the physician" and are subject to licensure, e.g. "clinical psychologists", to include members of the nominally independent "scholarly or learned" professions who are recruited in support of the "national medical research enterprise", the resources and agenda of which are controlled by the dominant profession.

Earlier in this discussion it was proposed that medicine had successfully restricted the allocation of resources to the intra-mural programme pursued primarily by biochemists, bacteriologists etc. within the National Institutes of Health (NIH). This was based on the argument that basic research is pursued over the longer term with unpredictable results, whereas more immediate benefits to society derive from clinical investigations of new technologies for treatment of established disease, the research focus of medical school faculty members (see Chapter 1). It was further proposed that this argument was constructed not merely in the interests of academic entrepreneurialism, but to limit the power of the NIH as a government agency, and to reduce potential threats to professional autonomy or economic self-interest arising from new discoveries and their dissemination as "public goods". In contrast, opportunities for social scientists to participate in "biomedical" research have expanded in recent years both within the NIH and in extra-mural studies but critics argue that these have been narrowly focussed.

In 1988 the President of the Gerontological Society of America, a basic scientists, argued that a progressively increasing share of the budget of the National Institute on Aging (NIA) with its multi-disciplinary mandate had been devoted to "diseases of the elderly" at the expense of "research on biological, behavioural and social phenomena that are unrelated to disease",[Adelman, 1988 pB1] (see Chapter 2). Estes and Binney in a paper entitled the "biomedicalisation of ageing", similarly argued that the NIA research agenda has been "usurped" by medicine and that where resources are allocated for social science research those disciplines which are "quantitative" and "positivist" are favoured over those which are "qualitative" or focused on "system level problems".[Estes & Binney, 1991 p125-6] It may be proposed, however, that the quantitative sciences of "biostatistics", epidemiology and demography are favoured not merely in research sponsored by the NIA but also by the National Institute of Mental Health on behalf of psychiatry the medical specialty which has adopted an "increasing reliance on data in the understanding of mental disorders",[DSM-III, 1980 p1] (see Chapter 3). In consequence these same disciplines have been employed in the new "mental testing enterprise" legitimated by psychiatry in the search for senile dementia redefined as Alzheimer's disease, the pursuit of which by the NIA has been perceived by Adelman as the "Alzheimerisation of ageing".[Adelman, 1995]

It may therefore be argued that just as the medical profession controls, influences, and selectively employs the services of the "paramedical occupations" feeling "privileged to disregard",[Wechsler, 1939 p48-9] or by extension to appropriate knowledge which is useful to it, it is similarly selective with regard to the "scholarly professions", rejecting those which are "qualitative" and therefore working within or elaborating paradigms of inquiry which may be inimical to its interests while embracing those whose members will serve as technicians in generating data within its own diagnostic parameters. It may be further argued that in the new mental testing enterprise the medical profession has appropriated the services of disciplines which employ the methodologies formulated in "proving" and "quantifying" the 19th century construct of "intelligence", while choosing to disregard the body knowledge that has since been elaborated by different schools of psychology, including both those subscribing to the paradigm of "learned behaviour" and the paramedical descendants of the "intelligence men" who came to reject the one-dimensional concept of intellectual ability and its susceptibility to "psychometrics".

Quantitative methodologies and the reification of constructs

It was earlier proposed that American psychiatry had derived legitimacy, as demonstrated by the passage of legislation to establish the National Institute of Mental Health in 1946, from the generation of "quantitative data" on its "successes" with the military, based in large part on the "inadequacy" of the nomenclature of mental disorders during the 2nd World War and the application of "insanity" labels to "men reacting briefly with neurotic symptoms to considerable stress".[DSM, 1952 pv-vi] In the past twenty-five years, however, psychiatry has focussed on more than simple tabulations of the number of subjectively determined or "qualitative" diagnoses by seeking to "standardise" its mental illness constructs in questionnaire instruments such as the Psychiatric Status Schedule (PSS),[Spitzer et al, 1970] and in the formulation of "objective" diagnostic criteria as in the DSM-III. Furthermore, it has sought to establish "reliability" for these instruments and diagnostic criteria by evaluating the results of their "operationalisation" by "pairs of clinicians making independent diagnostic judgments of patients".[DSM-III, 1980 p5] Critics such as Guimon argue that the "objectification" of psychiatric diagnoses undertaken within the "biological" paradigm, which "holds that psychiatric disturbances are real illnesses", is invalid in that "psychopathological phenomena" are not susceptible to measurement,[Guimon, 1989 p33] but, as already discussed, such critics are sympathetic to the imputation of a correlational relationship between physical changes and objective quantitative measures for those mental disorders such as "dementia" which are defined as "organic" (see Chapter 3). However, it was the "method of correlation" which was formulated in support of the science of psychometrics by its proponents in the persons of Francis Galton and Karl Pearson (1857-1936),[Fancher, 1985 p38-9] and employed in constructing the concept of "intelligence" as an objective reality.

Arthur Newsholme, in a 1923 discussion of the "measurement of progress in public health" in which he made specific reference to the U.S. army tests, (see Footnote 10) argued that "every step from a qualitative to a quantitative expression of facts, if accurate, is of high value" but there were "difficulties" in accepting "some of the inferences from social data advanced by the biometrical school of statistics". He regarded the "Galton-Pearson methods as a great advance in statistical science",[13] but argued that there is the "necessity of judging .... whether the 'significance' of the co-efficient of correlation is causal or merely accidental". He also proposed that "however interesting the mathematical manipulation of data may be as mathematics, data in the raw mass ...... are not generally susceptible of any considerable amount of mathematical refinement".[Newsholme, 1923 p197-8]

Spearman, in 1904, reviewing research which had failed to find a positive association between academic and physical performance, the latter determined by measures of "natural innate faculties" such as visual and auditory acuity, concluded that such failure was attributable to "attenuation", "constriction", "dilation", and "distortion", all of which represented some form of "measurement error".[Spearman, 1904a p89-95] However, by formulating a "correction factor" for results derived from Pearson's "method of correlation" he arrived at "the proof and measurement of association between two things", i.e. the relationship between "natural ability" and "intelligence".[Spearman, 1904a p72] This same formula, which resulted in at least one correlation coefficient greater than 1.0, was applied to his own research on "general intelligence: objectively determined and measured".[Spearman, 1904b p272] It may therefore be posited that the biological, hereditary concept of intelligence was "reified" by the mathematical manipulation of data and by the

---

[13] Newsholme also took note of "Mr J.M. Keynes destructive criticism" of the Galton-Pearson methods".[Newsholme, 1923 p197]

imputation of "causation" to the correlation coefficients derived therefrom. Furthermore the potential for operationalising the "objectively determined" concept and for making inferences based upon its quantification were also expanded by statistical refinement.

Binet's observation that age was a major predictor of intellectual "level" as determined by test performance in children became translated by William Stern into "mental age" a construct that was subsequently entered into the positivist formula for the "intelligence quotient": I.Q.= Mental Age/Chronological Age) x 100.[Terman, 1916 p17] However, Robert Yerkes, instigator of the U.S. army tests, found the concept of "mental age", which ceased to increase beyond the teenage years, and "the type of measurement which the Binet Scale yields (to be) amenable to statistical treatment only in a restricted way". By converting such test scores into a cumulative "point scale" Yerkes developed a method for establishing "norms or expected results", for age, sex, race, educational or social status. Thereafter, "in order to use the point scale profitably for a new race or social group it is necessary only to make a sufficient number of examinations to yield reliable norms. These immediately become standards of judgment".[Yerkes, 1917 p114-5]

In his 1923 lecture, Newsholme argued that "statistics fulfil an indispensable part in social and medical studies" including serving as "a valuable means of investigation, especially in regard to the incidence and causation of controllable disease". On the other hand, his concerns lay not merely with inferences which may be drawn from spurious correlations and from data that have been "refined" but also with the "accurate definition of each unit" to be counted and the complexity of phenomena subject to statistical investigation and it may be argued that Newsholme's concerns are as valid today as in

1923, but the measuring techniques developed by the intelligence men in the early part of the 20th century, further refined by the quantitative sciences, have been re-employed in establishing "standards of judgment" in a new "social group", i.e. the older population. On the other hand, the "unit" to be counted or quantified in their "mental tests" has been defined within the medical profession, i.e. by psychiatry, which has, arguably disregarded the later work of psychologists who elaborated on the complexity of intellectual functioning and of factors intervening in its "quantification".

"Intelligence" redefined as "cognition" and "neuro-psychology"

Writing in 1923 of the current successes of the mental testing enterprise Spearman proposed that "in the course of such a tremendous movement, the question could not but arise with great urgency as to what ..... is the nature of this 'intelligence'". He noted that the "whole concept (of intelligence) - although among the most familiar in ordinary life - would seem to have been in systematic psychological treatises deliberately ignored".[Spearman, 1923 p2] His own "readily intelligible hypothesis was ..... derivable from physiology" and comprised a "general or quantitative factor", "that was taken, pending further information, to consist in something of the nature of an 'energy' or 'power' which serves in common the whole cortex, or possibly even the whole nervous system". This factor, i.e. "brain power",[Fancher, 1985 p95] "being evoked to explain the correlations that exist between even the most diverse sorts of cognitive performance", is "commonly written in inverted commas or replaced by the simple letter g".[Spearman, 1923 p5]

Wechsler in 1939 described Spearman's "g" as the "basic prop of psychometrics", and he himself defined "intelligence" in this work as a "global capacity" composed of elements which are "qualitatively differentiable and quantitatively measurable".[Wechsler, 1939 p3] However, addressing the American Psychological Association in 1949 he proposed that Spearman was "in error" in "equating 'g' with general intelligence", and in arguing that this was "the only independent factor". He posited that "intelligence tests, such as they are, contain not one but several independent factors which are all cognitive". There were, however, other independent factors or "residual components" in test scores including "conation", i.e. desire or motivation, factors which he cumulatively termed "personality components".[Wechsler, 1950 p80-83] Similarly, in an article published in the same year, a British psychiatrist and psychologist argued that total intelligence test scores were "at best a crude and often misleading measure of global intellectual performance" and "separate aspects of cognitive functioning are now reported on independently". Furthermore, "there is a growing recognition that what is tested, even with so-called tests of intelligence is not the subject's intellectual functions isolated from the rest of his personality".[Brody & Williams, 1950 p163]

It may therefore be proposed that the "cognitive" model of mental testing was constructed upon the empirical observations of "intelligence psychologists" in the "clinical setting" as opposed to evidence produced in "apersonal" group psychometrics.[Wechsler, 1939 p48] In the new model various "cognitive functions" were measured by individual items within a battery of tests, the results of which were interpreted with caution in view of the contribution of independent unquantifiable factors such as "motivation" and "personality". On the other hand, a factor not discussed in either of the papers cited above, which is arguably intrinsic to "cognitive functions" and not independent of them, is learning experience and the knowledge derived therefrom.

Wechsler, while still subscribing to the "global capacity" theory of intelligence in 1939 divided his WAIS battery between items he defined as either "performance" or "verbal". Wechsler, 1939 p11 In a study conducted by psychologists in an army hospital during the 2nd World War, in which a version of the WAIS was employed as the assessment instrument,[14] it was demonstrated that there were significant differences in test scores for some of the "performance" or puzzle-solving items between patients who had undergone brain surgery and a control group who had head injuries treated without surgical intervention. On the other hand it was found that some "verbal" items including general information retrieval and vocabulary were relatively unaffected. Aita et al, 1947 p25-44 [15] Similarly, D.O. Hebb, writing on the "organisation of behaviour", making a "sedulous attempt to find some community of neurological and psychological conceptions", observed that the effects of removing large areas of brain tissue were "often astonishingly small" in terms of its effect on intelligence test scores. He concluded that the "level of intelligence test performance is a function of the concepts a patient has already developed ...... The patient with brain injury at maturity may continue to think and solve problems

---

[14] In 1941 Robert Yerkes, instigator of the mass intelligence testing of U.S. army recruits in 1917 argued that had the military continued with his methods it would "have been much better prepared to meet the efficient threat of the German military machine". Kevles, 1968 p580 The programme was not, however, reinstated but Wechsler was asked to develop a version of the WAIS, the Mental Ability Scale for use by the army which was employed in the study by Aita et al.

[15] The WAIS was administered three to six months after surgery to 333 patients with an age range from 18 to 35 years, and also to a control group of 288 neurology patients which included non-surgical cases recovering from head trauma. Aita et al, 1947 p25-44

normally (in familiar fields) although his intelligence would have been far from normal if a similar injury had happened at birth".[Hebb, 1961 p1-2] Hebb therefore proposed a "major effect of experience", defined as exposure to a "stimulating environment" conducive to "conceptual development", on mental test scores.[Hebb, 1949 p299-302]

Lezak in a textbook on "neuropsychological assessment" cites the 1940s work of Aita and Hebb as seminal in the generation of a discipline, i.e. neuropsychology, which evolved with the mapping of specific brain functions. Lezak argues that the discipline "contributed significantly to the redefinition of the nature of intelligence", replacing the concept of a "global" or "unitary variable" with a multi-dimensional cognitive model of intellectual functioning,[Lezak, 1983 p20-1] although it is unclear whether the concept of "cognition" which replaced that of "intelligence" relates to knowledge that has been acquired or to the ability to learn. Nevertheless it may be argued that neuropsychology has undermined the potential for imputing a simple correlational relationship between quantitative measures *in vivo* and either "natural ability" or physical changes in the brain and this has been consistently confirmed by pathologists and other neuroscientific investigators.[Hebb, 1949 p1-2, Blessed, 1968 p808, McLoughlin et al, 1994 p439] It may be proposed, however, that the psychiatric "cognitive paradigm" for dementia does impute a relationship between mental testing and what Guimon described as "hard physical indices"[Guimon, 1989 p35] and that it has legitimated the imputation of incipient or early organic disease to test scores even in the absence of "dementia behaviour", discounting the knowledge which has been elaborated with regard to the complexity of mental functions, the interpretation of test results, and the deliberations which have been given to the assessment of older people.

Testing for "mental decline"

Wechsler consistently observed that the ability to achieve high scores on "intelligence" tests declines with age.[Wechsler, 1939 p63] On the other hand he was diffident with regard to the assessment of "mental decline" since conclusions were based upon "cross-sectional and not longitudinal studies", i.e. "studies derived from the comparison of test scores of successive age groups and not from scores derived from the successive examination of the same individuals at different ages".[Wechsler, 1958 p135] Furthermore, in addressing a 1954 gerontology congress he spoke of both "intrinsic" and "extrinsic" problems with regard to such appraisal, the latter including the potential for recruitment of a representative sample of the older population, and the former relating to such factors as physical impairment and loss of visual acuity making for difficulty in "interpreting the data".[Wechsler, 1954 p275] It may be argued, however, that the medical profession has become progressively less inhibited than psychologists in the development of its own mental tests for the elderly, in expanding their use, and in ascribing meaning to the results obtained.

In 1953, Martin Roth, in collaboration with a lay psychologist published the results of a study of "psychological test performance in patients over 60". Several assessments, including selected items from the WAIS, were administered to subjects admitted to a psychiatric hospital in order to assist in the diagnosis of mental disorders in old age. The tests were found to distinguish between those with "senile psychosis" and others with "affective disorders" who manifested behavioural signs of dementing disease but without impairment in "intellectual functioning", i.e. suffering from a type of depression now defined as "pseudodementia" (see Chapter 3). The primary tests which distinguished between the two groups were "orientation" and "knowledge of public affairs", but the ability to perform well on some mental tests was relatively well preserved even in "senile psychotics".[Roth & Hopkins, 1953 p439-450] Also in England in 1953, Lovett Doust et al. conducted a "psychometric investigation" of "hospitalised senile patients" in a study to measure "cerebral blood flow", one which identified a spurious correlation between "physiologic and psychometric indices" in men but not in women.[16] The WAIS was administered to all subjects in the study but the investigators also tested a "simple 22-item screening questionnaire", i.e. the Tooting Bec Questionnaire. The instrument included a number of questions on "public affairs" which required frequent updating,[17] but the researchers concluded that it was a "reliable, shorter substitute" for the WAIS.[Lovett Doust et al, 1953 p383-98] Similarly, in the 1970s two instruments were developed by American psychiatrists as shorter alternatives to the WAIS. The Mini-Mental State Examination (MMSE) was published by Folstein et al. as a "quantified assessment of cognitive state" and was found, in a study similar to that conducted in 1953 by Roth and Hopkins, to distinguish between patients admitted to a psychiatric hospital with dementia, defined by Folstein as a "global deterioration of intellect", and those diagnosed with "affective disorders" and other conditions.[Folstein, 1975 p189-92] In the same year the Short Portable Mental Status

---

[16] A later study by Hachinski et al. which is widely cited showed "no correlation between the degree of dementia and cerebral blood flow" in primary degenerative dementia, i.e. that which has been termed Alzheimer's disease. Individuals who experience multiple small strokes and develop "multi-infarct dementia" do show a decrease in cerebral blood flow to specific areas of the brain,[Hachinski, 1975 p632-7] but it is accepted that there is no support for the general concept of "hardening of the arteries" in the aetiology of dementia.

[17] For example the 1953 Tooting Bec Questionnaire asked: "Where is Korea? What is happening there?".[Lovett Doust et al, 1953 p398]

Questionnaire (SPSMQ) was published by Pfeiffer,[Pfeiffer, 1975 p433-441] but whereas the MMSE was first promoted to assist in the diagnostic process among older psychiatric patients Pfeiffer advocated widespread use of his instrument "for the assessment of organic brain deficit" by "clinicians", not necessarily psychiatrists, "whose practice includes any significant number of elderly patients". On the other hand, both the SPSMQ and the MMSE were validated on a sample of non-psychiatric patients to yield "norms" which Yerkes defined as "standards of judgment".[Yerkes, 1917 p114] However, it may be argued that norms do not address Wechsler's concerns with regard to baseline scores against which to measure "mental decline" nor the effect of intervening variables such as visual and hearing impairment. Furthermore, for "non-patients" they do not indicate how the quantification of "cognitive state" is reflected in behaviours indicative of dementia or "organic brain deficit". In spite of such questions regarding the validity and reliability of scores derived from mental testing, Pfeiffer's SPSMQ was adopted for quantifying "cognitive functioning" in the multi-site Established Populations for Epidemiologic Studies of the Elderly (EPESE) research conducted by the National Institute on Aging,[Cornoni-Huntley et al, 1986 p7] (see Chapter 2) and Folstein's MMSE was legitimated by the National Institute of Mental Health to measure "current cognitive state", the "closest approximation to organic mental disorders" in operationalising the DSM-III criteria in the Epidemiologic Catchment Area (ECA) Study,[George et al, 1991 p291] (see Chapter 3).

Panoptic surveillance of the elderly

In 1980, the year in which the DSM-III dementia model was published, Morton Kramer, a member of the task force on nomenclature and statistics for the DSM-III and for many years chief of the "biometrics branch" of the National Institute of Mental Health, argued that "the world is in the midst of a rising pandemic of mental disorders and associated chronic diseases and disabilities" based on increasing life expectancy among those at "high risk" of developing these conditions. He used "currently available morbidity and mortality rates and population projections" to arrive at prevalence estimates for mental health problems, including "senile brain disease" that "may face the world by the year 2000". He also argued that his projections "merely provide a basis for speculating on changes likely to occur" and that there was a "lack of epidemiological data on the incidence, duration and prevalence of specific mental disorders in most countries in the world".[Kramer, 1980 p382-397] However, in a 1987 review of studies of the "prevalence of dementia" Jorm et al. identified twenty-three investigations conducted in the five years between 1980 and 1985 in countries ranging from Germany to South Africa, and from Britain to the U.S.S.R., Japan, Taiwan and the United States.[Jorm et al, 1987 p465-479] Rocca et al. in a similar comparative study identified twenty-two "current prevalence studies of dementia" undertaken in European countries alone including Italy, Spain, Sweden and Finland between 1980 and 1990.[Rocca et al, 1991 p381-90] The case finding methods were not described for all the studies identified by the authors of these papers, of which less than half employed comparable methodology, but "objective cognitive testing" was one of the covariates in the re-analyses conducted by Jorm et al.[Jorm et al, 1987 p473] In a review of specific investigations conducted for this discussion, including four of the six which qualified for the reanalysis by Rocca et al. to determine the "frequency and distribution of Alzheimer's disease in Europe", "mental testing" of the elderly population was conducted in surveys in Cambridge,[O'Connor et al, 1989 p190-8] and in Cambridgeshire, England,[Brayne et al, 1989 p214-9] in Bordeaux, France,[Dartigues et al, 1992 p981-8] in Appignano, Italy,[Rocca et al, 1990 p626-31] in Zaragoza, Spain,[Lobo et al, 1990] in Shanghai Province, China,[Zhang, et al, 1990 p428-37] and in Tasmania, Australia.[Kay et al, 1985 p771-87] In each of these studies, the "screening instrument" used was the MMSE.

Guimon proposed that "changes in psychiatric fashion in the United States have international impact",[Guimon, 1989 p33] and it may be proposed that this is supported by the apparent legitimacy achieved for the instrument which was first employed by the National Institute of Mental Health (NIMH) in the ECA study for the detection and quantification of organic mental disorders within the DSM-III model formulated by American psychiatry, and its use in promoting Kramer's speculations with regard to the pandemic of mental disorders as a reality in "panoptic surveillance" of older populations. On the other hand, the method of operationalising the "cognitive paradigm" for dementia has also been employed in a range of clinical and other studies in the United States allocated to the biomedical community in the search for a cure for senile dementia, redefined as Alzheimer's disease, the aetiology of which has not yet been fully distinguished from "normal ageing" (see Chapter 2). It was from one of these studies, sponsored by the Mental Disorders of Ageing Research Branch of the NIMH to evaluate surgery as a "risk factor" for the onset of dementia indicative of the disease that data were derived to evaluate the validity of the concerns expressed by Wechsler with regard to mental testing in older people and the interpretations given to the results of mental testing.

Chapter 5

EMPIRICAL EVIDENCE: A CASE STUDY FROM THE ALZHEIMER'S ENTERPRISE

The study and the role of the author

In November 1988 a grant was awarded by the Mental Disorders of Aging Research Branch of the National Institute of Mental Health for a three-year study entitled "Mental status change after surgery in an elderly population". The proposal was submitted by, and the award made to, Nathan Billig, M.D., Professor in the Department of Psychiatry, and Director of the Center on Aging, Georgetown University Medical Center, Washington, D.C..

In January 1989 the author of this thesis, a Research Associate at Georgetown University, was appointed Project Director of the study. The role of the Project Director was to supervise and participate in the recruitment of subjects and the conduct of interviews as well as having responsibility for data management, i.e. questionnaire coding, the development of a computerised database and the statistical analyses. While carrying out these activities in accordance with the pre-approved methodology of the study, the author began to question the validity of the instruments being employed in this and other investigations for the assessment of "cognitive impairment". She therefore sought, and was given permission by the grant holder, Dr. Billig, to conduct separate and independent analyses of the data collected. Study methods, data management, response rates, demographic and other baseline characteristics of the sample are described in Chapter 6, together with the rationale for selection of variables used in the independent analyses as well as the statistical techniques employed. The results are presented in

Chapters 7 and 8.

The Center on Aging at Georgetown Medical School and the NIH

The Georgetown University Center on Aging was established by the Dean of the Medical School in 1987 "to organize and promote a multidisciplinary approach to the clinical and scientific problems of aging".[Dean's Annual report for the School of Medicine, 1986-87] A psychiatrist was appointed Director of the Center, in the face of lobbying by other medical departments, including Community and Family Medicine, for the directorship to be assigned to one of their faculty members.[1] It may be argued that the appointment was an acknowledgement by the Dean of Georgetown Medical School of the "reality" that in the 1980s the ageing brain and mind had become an appropriate and potentially rewarding focus for "ageing" research. Furthermore, the designation of a psychiatrist as director legitimated the submission of proposals seeking research funding for the activities of the Center on Aging to the National Institute of Mental Health and the National Institute on Aging, as well as to voluntary trusts and foundations such as the Alzheimer's Association.[Fox, 1989 p80]

Stevens argued that grants from the National Institutes of Health (NIH) for extra-mural research came to represent a major "back door" subsidy for "nominally independent medical schools" in the 1950s.[Stevens, 1971 p359] The continuing importance of such grants may be gauged by the fact that in 1989, the first full year of funding for the

---

[1] Prior to being appointed Project Director for the study: Mental status change after surgery in an elderly population, the author was a Research Associate in the Department of Community and Family Medicine and therefore privy to negotiations being conducted with regard to the appropriate credentials of a director for the proposed Center on Aging.

Center on Aging study under discussion, research support for the Georgetown School of Medicine from the NIH was $22.8 million. A further $11.75 million was received in "service contracts" with clinical departments including Anaesthesia, Radiology and Radiation Medicine.Georgetown Medical Bulletin, 43:1:22, 1991 Not only does such public funding provide direct support for salaries and other research resources for the successful applicants, including the purchase of capital equipment, they also contribute additional substantial "overhead" funding to the parent institution. NIH grants are therefore a more significant source of support for academic medical establishments than those from private foundations, which tend to award lesser sums, and are seldom supplemented with overhead and capital equipment financing. Government grants thus engender intense competition both internally and externally between academic departments and institutions.[2]

The NIH funding cycle, i.e. the time between submission of a grant proposal, its passage through the review process, and the allocation of funds if the proposal is approved is approximately one year. Some 14 months after the establishment of the Center on Aging, its first major award was received, i.e. from the National Institute of Mental Health (NIMH), the objectives and background of which are discussed below. It may be proposed that receipt of this grant legitimated the decision of the Dean of Georgetown Medical School to endorse "psychiatry" as the discipline for attracting funding in the field of ageing research in the 1980s.

---

[2] The Georgetown School of Medicine ranked 45th among the 124 American medical schools in 1991 in terms of the total funding received from the various NIH institutes.

The study objectives, its merit, and the review process

The proposal entitled "Mental status change after surgery in an elderly population" was approved by the Mental Disorders of Aging Research Branch of the NIMH with the following objectives:

i. To assess the prevalence of perioperative cognitive impairment and depression in a sample of patients aged 60 and over undergoing elective surgery.

ii. To assess the effects of hospitalisation and surgery as precipitating factors for the onset of cognitive impairment and depression in the elderly.

iv. To estimate the extent to which elective surgery affects the recovery of pre-morbid functioning in the year following surgery.

The proposal was submitted with the written consent of the Professors of Ophthalmology, Urology and Orthopaedics at Georgetown Medical Center to allow the investigators access to patients undergoing elective cataract, prostate, or joint replacement surgery. It may therefore be proposed that the submission fulfilled the "multidisciplinary" remit of the Center on Aging and that the collaboration may have commended the investigation to those by whom it was approved. On the other hand it may be posited that the selection of the proposal for funding by the NIMH may be perceived as evidence of the fragmentation of "knowledge" which may occur with a high degree of occupational specialisation as in American medicine (see Chapter 1).

Stevens described the approval of NIH research grant applications in the 1960s as a "Byzantine system of national peer review groups and advisory committees".[Stevens, 1971 p360]

A guide to "preparing a grant application for the National Institutes of Health" states that review groups are composed "primarily (of) non-government scientists", and that "the goal of peer review is to achieve an objective evaluation of the merit and significance of the application in the context of the current state of knowledge in a given field, and the probability that the applicant can succeed in carrying out the proposed project".[Murphy & Dean, 1989 p4] In the case of the NIMH the "non-government scientists" are primarily psychiatrists with academic appointments who, in common with specialist committees reviewing proposals to the other institutes of the NIH, combine clinical practice, teaching and research, and subjectively appraise the applications for grant funding in accordance with their own definitions of reality and their own awareness and perceptions of the "current state of knowledge".

Applications to the National Institutes of Health have a clearly stated requirement for inclusion of a review of relevant "published and unpublished work" in order to demonstrate "that you are thoroughly familiar with your field and have a balanced knowledge of it".[Eaves, 1989 p153] A literature search conducted for the present discussion, which does not claim to be exhaustive, identified numerous investigations in which pre- and post-operative assessments, primarily of cognition but also of depressive or affective symptoms, were conducted in older people undergoing surgery. The Georgetown proposal, however, stated specifically that "in our search over the past ten years we could find no study which compared pre and post-operative mental status (relating to) elective surgery ..... in the elderly" and only four published reports, all by psychiatrists, addressing conditions and using methods which did not conflict with the rationale for the proposed investigation were cited in the bibliography (see below).

Intuitively it might be assumed that had the committee members been aware of earlier studies they would have sought a more extensive review of the "current state of knowledge", clarification of how the proposed study differed in its methodology and objectives from those previously conducted, and input from the specialists who had produced the greatest volume of work in the field and were not consulted for the proposal, i.e anaesthetists. The Georgetown study was, however, funded without requesting a further review of the literature and a reasonable assumption may be made that previous research into the effects of sugery in older people (see below), was unknown to the review committee of psychiatrists by whom it was approved.[3]

Are there adverse effects from surgery? Other related studies

For more than a century it has been proposed that those undergoing surgery are at risk of experiencing and exhibiting a range of psychiatric symptoms and behaviours, of both short and long-term duration, which may be attributable to the specific effects of anaesthetic agents or to other variables associated with the procedure. It has further been proposed that there are predisposing factors for these phenomena, both emotional and

---

[3] All applicants for grants from the National Institutes of Health receive an evaluation of the proposals submitted, whether or not they are approved for funding. Any deficiencies detected by the reviewers are noted and the applicant may resubmit the proposal after revisions have been made to address these concerns.

physical.[4] Savage, 1887 Age is the most frequently posited risk factor, and it has been suggested repeatedly, both in the medical literature and from reports by relatives and friends, that elderly people show evidence of permanent and progressive impairment in cognitive functioning following surgical procedures, i.e "He's never been the same since his operation".Bedford, 1955 p259 [5]

When individuals are rendered unconscious by anaesthetic agents they are dependent on "life-support" by the anaesthetist for the maintenance of breathing, oxygenation and blood pressure. Potential anaesthetic catastrophes may occur from physiological changes brought about by the surgery due to blood loss etc.; adverse reactions to any of the various inhaled or intravenous drugs administered; technical failure of equipment, and lack of vigilance by the anaesthetist in recognising and acting to correct these problems. The question that has been explored many times since 1887 is whether mental status changes occur in the absence of any overt and unpredicted intra-operative events.

A literature review and case study report by Bedford, published in 1955 concluded

---

[4] An article published in 1887 in the British Medical Journal, entitled "Insanity following the use of Anaesthetics in Operations", described several cases of individuals, both young and old, exhibiting significant and often florid mental changes after surgery. In some cases there was remission and recovery, but in others the symptoms persisted. The author proposed that the phenomena were attributable variously to the administration of chloroform, ether and nitrous oxide. He concluded by questioning whether "neurotic inheritance" or previous "attacks of insanity" should "interfere with operations of convenience not essential for the prolonging or saving of life".Savage, 1887 p2000

[5] ".....careful, there are bones in the fish. I worry because of the Queen Mother. They gave her a full anaesthetic, it's wicked. When you have an anaesthetic after fifty, half your brain dies." Feebly we protest that surely it comes back? "No. Never."Interview with Barbara Cartland reported in The Times, October 6, 1993

by arguing that older people were susceptible to both post-operative "extreme dementia",[6] and to subtle but potentially progressive mental changes or "lesser degrees of dementia" related to a number of factors some of which were modifiable, while others were unavoidable, such as the use of multiple drugs for anaesthesia and pain relief. It was therefore suggested that "operations on elderly people should be confined to unequivocally necessary cases".[Bedford, 1955 p263]

In response to this latter article, a major study to determine whether there was statistical evidence to support the retrospective and anecdotal reports linking the onset of dementing disease to surgery, was conducted by the Nuffield Department of Anaesthetics, Oxford University, with reported negative findings.[Simpson et al, 1961] All patients aged 65 and over, admitted for elective surgical procedures during a one-year period (n=681), were evaluated prior to surgery using standardised questionnaires for the assessment of physical and social activity and mental status, and again three months after discharge.

---

[6] The early literature does not distinguish between what are now regarded as two separate conditions: acute confusion or "delirium", and "dementia". Delirium is defined as an organic brain syndrome of acute onset, characterised by diffuse cognitive dysfunction, and change in level of consciousness, i.e. the patient may be either somnolent, or hyperactive and agitated. Hallucinations or delusions also frequently occur.[DSM-111R, 1987 p103]

The aetiology of delirium has been associated with most diseases with major systemic effects, and with many classes of drugs, some used in the administration of general anaesthesia.[Tune, 1981 p651] Although delirium may be indicative of a life-threatening medical emergency and there is high associated mortality, it is often transient and of short duration, particularly where an underlying cause can be identified and treated, e.g. reduction of a high fever or identification of adverse drug reactions.[Lipowski, 1983]

Delirium is distinguished from dementia by its sudden onset, by change in level of consciousness, and by its potential reversibility, e.g. alcoholic intoxication is a form of delirium. Most studies indicate a higher susceptibility with advancing age, and pre-existing dementia has been defined as a predisposing factor.[Liston, 1982] However, the extent to which an episode of delirium precipitates the onset of dementing disease is unknown.

The evaluations were conducted during home visits to subjects from the anaesthetist investigator.

Simpson, in his discussion of the results, acknowledged that anaesthetic catastrophes can occur, particularly in emergency surgery, and that extra vigilance may need to be exercised in the care of older people, but the study concluded that "neither the drugs used during anaesthesia, nor anaesthesia in its widest sense did any harm to the patients studied in this series".[Simpson et al, 1961 p893] On the other hand mortality rates were high (8 percent of study cases), and there were some apparent inconsistencies in the methods for pre- and post-operative evaluation. For example, a subset of the total, i.e. subjects undergoing major surgery and those of advanced age, was defined as "most susceptible to deterioration", and subjected to more intensive testing by a psychologist.[Simpson et al, 1961 p887] However, only the pre-operative test results from one hundred of the group of two hundred and fifty were reported,[Williams, 1960 p215] and there was no record of post-operative follow-up.

Since that time the topic of post-operative change in cognitive functioning has not ceased to attract investigators. Numerous longitudinal or prospective studies to detect mental status change among surgical patients, predominantly but not exclusively in older age groups, have been conducted in Europe and North America using a variety of different assessment techniques, number of follow-up evaluations and total investigative time periods. In order to achieve comparability across cases most studies have focussed on specific procedures, which have included: cardiac surgery[Aberg et al, 1974 & 1977, Smith et al, 1986]; hip replacement[Hole et al, 1980 Riis et al, 1983]; varicose vein stripping[Ryhanen et al, 1978];

prostatectomy,[Chung et al, 1987, Chung et al, 1989(a), Haan et al, 1991] and cataract surgery.[Karhunen et al, 1982, Applegate et al, 1987, Chung et al, 1989 (b)] Some researchers have sought to detect differences in mental recovery time in the immediate post-operative period from the administration of different specific anaesthetic agents.[Simpson et al, 1976, Denis et al, 1984] More frequently the studies have evaluated the use of different techniques, e.g. spinal as opposed to general anaesthesia for procedures such as hip surgery or prostatectomy,[Riis et al, 1983 Bigler et al, 1985 Chung et al, 1987 Haan et al, 1991] and general anaesthesia compared with local anaesthesia with mild sedation for surgery such as cataract extraction.[Karhunen et al, 1982 Chung et al, 1989 (b)] The follow-up times have varied from twenty-four hours or less,[Simpson et al, 1976 Cohen et al, 1982] to one month,[Chung et al, 1989(a)] to three months,[Haan et al, 1991] to one-year.[Applegate et al, 1987 Elam et al, 1988]

With the exception of investigations relating to open heart surgery,[7] only one of the studies cited above detected prolonged differences in cognitive functioning before and after surgery. The one study of patients aged 56 to 84 scheduled for hip replacement surgery reported "significant mental changes" postoperatively in 23 percent of patients (7/31, one of whom died) randomly assigned to the use of general anaesthesia, changes which continued to affect the "quality of life" of 5 of these cases several months later.[Hole et al, 1980 p279] In contrast, only 1 of the 29 patients receiving spinal anaesthesia for the same procedure demonstrated postoperative psychological change, and this was of brief duration.[Hole et al, p284]

---

[7] Open heart surgery is a complex and highly technical surgical and anaesthetic procedure conducted on individuals with life-threatening conditions. Several studies have shown that it carries with it a risk for individuals of all ages of damage to the central nervous system from factors such as cerebral emboli or poor perfusion of the brain during cardiopulmonary bypass.[Aberg et al, 1974, Henrikson, 1984, Shaw et al, 1985, Smith et al, 1986]

Although the study by Hole et al. described all the patients as "mentally normal for their age" prior to surgery, no standardised pre- or post-operative testing was carried out. By comparison, all the other investigations since 1961 have employed various cognitive screening tests used singly or in combination. The assessment instruments employed have included the Tooting Bec Questionnaire (Lovett Doust et. al, 1953),[Simpson et al, 1961] subtests from the WAIS (Wechsler 1939) and other instruments, either established or developed *ad hoc* by psychologists,[Simpson et al, 1976, Karhunen et al, 1982, Cohen et al, 1982, Riis et al, 1983, Bigler et al, 1985] and the Mini-Mental State Examination (Folstein et al, 1975).[Chung et al, 1987, Chung et al, 1989, Haan et al, 1991]

As might be predicted, since the majority of patients are given peri-operative sedation and narcotics for pain relief, those studies which included testing in the immediate post-operative period, i.e. within twenty-four hours, detected some decline in test performance compared with preoperative values, a decline which was greater and of more prolonged duration in patients undergoing general anaesthesia than in those where other anaesthetic techniques were employed.[Chung et al, 1987 Chung et al, 1989(a)] On the other hand, recovery to pre-operative levels was reported to occur in all groups within days, and studies which did not institute post-operative testing until four or more days later found insignificant differences between groups, both during the period of hospital admission and in follow-up assessments up to several months later.[Riis et al, 1983, Bigler et al, 1985, Haan et al, 1991]

Changes over time in the management of surgical patients

During the past 35 years there has been an increase in pre-operative testing to detect previously undiagnosed comorbidity such as anaemia or cardiac problems, and there have been changes in surgical techniques, in understanding pharmacological effects, and in physiological monitoring of patients during anaesthesia. It may therefore be argued that outcomes, either overt or subtle, which might be attributable to factors such as poor oxygenation or precipitate falls in blood pressure, discussed by Bedford in his 1955 review, would become progressively less likely to occur during the administration of general anaesthesia, and that this would be particularly applicable to surgical cases in academic medical centres where the majority of research is conducted.

All the investigations cited above which have included standardised assessments, reported insignificant findings with regard to persistent adverse cognitive effects associated with any frequently-performed elective procedure or anaesthetic method, in either younger or older samples. On the other hand, there is consistent evidence from these and other investigations of more rapid physical recovery and improved subjective well-being with the use of regional or spinal as opposed to general anaesthesia, even where a spinal block is supplemented with sedation.[Jones, 1988 p172 Chung, 1989 p382] Davis et al. have proposed that for procedures such as hip replacement or prostate surgery, when spinal anaesthesia is employed, there is less change in body temperature and blood pressure and less intra-operative bleeding, indicating reduced physiological stress on the body.[Davis et al, 1987 p569] Furthermore it may be argued that there is a lower potential risk of undetected adverse events during surgery when normal breathing and a degree of consciousness is maintained. At the same time, with less debilitating effects from spinal anaesthesia, individuals tolerate more rapid mobilisation,[Bigler, 1985 Jones, 1988 p173] thereby reducing the risk of post-operative pulmonary and circulatory problems, as well as affording a rationale for reductions in length of hospital stay.

Consequently there has been a progressive decrease in the use of general anaesthesia for individuals of all ages, except where technically necessary, e.g. for heart and lung surgery. This it may be argued has further reduced the overall risk of adverse effects from anaesthetic agents or their administration. It is therefore reasonable to assume that the probability of detecting significant or prolonged changes in cognitive functioning related to elective procedures commonly performed in older people, such as cataract and prostate surgery, has also declined.

Organic versus emotional factors and the unanswered questions

It may be argued that it is significant that the investigators in all the studies cited thus far were anaesthetists, often in collaboration with the relevant surgical specialists. Freidson has posited that since the "foremost claim of a profession is to special expertise it follows that the first question to ask about various forms of professional practice concerns technical performance, or the quality of service provided".[Freidson, 1973 p99] Anaesthesia has been defined by Freidson as a "service specialty" with less autonomy and lower status than surgery from which it claims legitimacy as a professional entity.[Freidson, 1973 p120] [8] The "technical performance" and "quality of service provided" by the anaesthetist who is in the "second-line of authority", with a remit to ensure the patient's comfort and survival during surgery, is therefore subject to evaluation by the surgeon

---

[8] Stevens has described how anaesthesia developed as an "adjunct of surgery", doctors competing initially with "nurse anaesthetists" in the early part of the century. The Board of Anaesthesiology was approved in 1938 as an "affiliate" of the American Board of Surgery by the Advisory Board for Medical Specialties and the AMA Council on Medical Education. The anaesthestists were granted independent status in 1941.[Stevens, 1971 p241-2]

who is ultimately responsible for the "success" of any procedure. It may therefore be argued that the special situation of the anaesthetist in the medical division of labour and a requirement to demonstrate professional competence to a higher authority, has prompted the numerous studies which have been conducted as a response to anecdotal and/or implied critiques of their practice. At the same time, the focus of the investigations by anaesthetists has been on the evaluation of "technique", based upon the assumption of a purely organic relationship between factors associated with their professional *modus operandi* and any adverse mental status effects. A small number of studies evaluating the relationship between emotional or affective symptoms and recovery from surgical procedures have, however, been published by psychiatrists. Millar evaluated 100 Scottish patients aged 65 and over, before and one week after surgery, both for the prevalence and the peri-operative onset of "psychiatric morbidity", including depression and "intellectual impairment". He found a close association between the course of the physical illness and depression, as well as between post-operative complications and the onset of acute confusion, i.e. delirium.[Millar, 1981 p17] The literature reviewed for the present discussion indicated that psychiatric studies of the effects of surgery have derived primarily from consultations requested for delirious or otherwise troublesome patients,[9] and the citations in the proposal submitted to the NIMH by the Georgetown University Center on Aging were a report of two cases of delirium following cataract extraction,[Summers et al, 1979] and three studies relating to "post-cardiotomy

---

[9] The author of the 1887 paper who proposed "neurotic inheritance" and previous "attacks of insanity" as predisposing to post-anaesthetic mental disorder[Savage,1887 p2000] was a psychiatrist or Lecturer in Mental Diseases at Guy's Hospital and the Royal Bethlem Hospital.

delirium".[Kornfeld et al, 1974, Heller et al, 1974, Tune et al, 1981]

It has been argued that psychiatry has occupied a relatively marginal status within the profession (see Chapter 3) and the primary area where access to surgical patients in a consultative capacity has been reported by psychiatrists in the United States has been for assistance in the prevention and management of delirium, a complex syndrome with a high incidence after open-heart surgery.[10] As a result, a small number of studies have assessed such factors as personality and pre-operative levels of anxiety as predictors of a disorder, the aetiology of which is poorly understood, but which presents problems for the nursing and surgical staff and may put the patient at risk of self-injury. One or two studies in which pre-operative interviews were conducted by psychiatrists, in order to evaluate the level of stress being experienced by patients and to listen to their fears, reported a lower incidence of the condition following this intervention.[Smith et al, 1989 p455] It may therefore be argued that there is a rationale for collaborative studies between medical specialties in addressing issues such as the interaction between anxiety, depression and the events associated with surgery; in identifying those at risk; and in developing interventions to ameliorate emotional distress.

---

[10] In addition to the risk of permanent impairment, (see Footnote 7)[Aberg, 1974, Henriksen, 1984 Smith et al, 1986] there is a large body of literature relating to the phenomenon of "post-cardiotomy delirium". This has been found to affect more than 30 percent of individuals undergoing open-heart surgery, a prevalence rate that remained constant, as reported in a review of 44 studies that were published between 1963 and 1987.[Smith et al, 1989]

A wide range of demographic, psycho-social, pharmacological, intra- and post-operative variables, including sleep deprivation and time in intensive care, were evaluated as risk factors in the plethora of studies, but with inconsistent findings. The extent to which the phenomenon is "organic" rather than "psychological" is not known,[Layne et al, 1971] nor is the extent to which the syndrome is totally reversible or may leave permanent if marginal impairment.

It is not, however, proposed that anecdotal evidence of adverse cognitive effects in older people associated with surgery has been totally refuted by the negative results of the studies previously discussed. For example, Bigler et al. did conduct pre- and post-operative assessments on 40 elderly patients randomly assigned to receive general or spinal anaesthesia for hip fracture surgery within forty-eight hours of admission,[Bigler et al, 1985] but there are significant problems in establishing a pre-operative baseline for cognitive or any other level of functioning in those admitted for emergency procedures.[11] Furthermore, other than the 1961 Oxford series,[Simpson et al, 1961] all the anaesthesia studies of older people cited above have selected very specific common procedures such as cataract and hip replacement surgery, thereby excluding operations such as resection of malignant tumours, where the patient may be more debilitated and the pre-, intra-, and post-operative course is less predictable and more stressful. At the same time, (see Footnote 7) studies of cardiac surgery in patients of all ages, undergoing procedures with a range of intervening variables which cannot be predicted and controlled, have consistently shown a high level of both short and long-term adverse neurological and cognitive effects.[Aberg et al, 1974, Shaw et al, 1985 Smith et al, 1986]

---

[11] It should be noted that the other group of surgical patients for whom both a high level of post-operative delirium, and long-term cognitive decline has been reported are those undergoing fixation of hip fracture.[Williams et al, 1985, Berggren et al, 1987, Gustafson et al, 1991] This is clearly not an "elective" procedure and factors such as pre-existing dementia, malnutrition, dehydration and stress associated with the events surrounding the fracture and hospital admission have all been identified as predictors of post-operative confusion among these predominantly frail older individuals.

Anaesthetists and nurses have been the primary investigators in the published work. It therefore appears either that these patients have not merited psychiatric consultation requests or that their care has not generated a specific sphere of interest for studies by psychiatrists.

Replicating and operationalising studies within the new paradigm

It may be proposed that the Georgetown study, i.e. "Mental status change after surgery in an elderly population" (Grant Number RO1 MH42103) was innovative in that it sought to monitor mental status, both in terms of depression as well as cognition, thereby adding an emotional component to the "organic" focus of most previous studies cited in this review. Furthermore it combined three different surgical procedures, i.e. cataract extraction, prostatectomy and joint replacement. On the other hand it may be argued that the "merit and significance" of the proposal derived in part from an apparent lack of awareness of previous observational studies of cognitive change in older people undergoing elective surgical procedures, both by the psychiatrist proposing the study, and among members of the peer review group by whom it was approved.

Since the mid-1980's, however, other investigations of the mental status effects of anaesthesia and surgery in older people have been undertaken in the United States and other countries in which earlier work was cited and which also had a specific focus on the assessment of cognition before and after surgical procedures. While the Georgetown study was in progress, a research proposal entitled "Neuropsychiatric change after surgery in the elderly" (Grant Number RO1 MH45048) was also funded by the Mental Disorders of Aging Research Branch of the NIMH. This study, proposed by a psychiatrist at the State University of New York, Buffalo School of Medicine, was designed to investigate the "relationship between geriatric surgery and neuropsychiatric decline". Circular from the National Institute of Mental Health, 1990 The research design was almost identical to that undertaken at Georgetown and included three common surgical procedures, i.e.

hernia repair, gall-bladder removal and again joint replacement. The studies were, however, defined within different demographic parameters: whereas it was proposed to assess the prevalence of "cognitive impairment" and "depression" in all surgical subjects aged 60 and over in the Georgetown study, in Buffalo the effects of "geriatric surgery" were to be evaluated in those aged "55 to 75 years old".[Circular from the National Institute of Mental Health, 1990] Furthermore, the investigator, to whom the Buffalo grant was awarded, was co-author of a "meta-analytic review" of the "effects of surgery upon the mental status of older persons", designed "to clarify the current status of scientific knowledge". The review, which omitted depression from the analysis, included several investigations cited in the present discussion, and concluded that previous cognitive studies were "generally lacking in methodological rigour".[Cryns et al, 1990 p184-91]

The Georgetown and Buffalo studies were initiated by psychiatrists but the same research topic has recently attracted a range of medical disciplines including: anaesthesia together with psychiatry in Canada;[Chung et al, 1987 Chung et al, 1989 (a) Chung et al, 1989 (b)] neurology and anaesthesia in The Netherlands;[Haan et al, 1991]; medicine and community medicine in the U.S.A.[Applegate et al, 1987 Elam et al, 1988] Rather than addressing the more complex issues associated with high technology and emergency interventions, however, these later investigations, even by anaesthetists, have tended to replicate the simple pre- and post-operative observational studies of the elderly from which there had previously been no statistical evidence of adverse cognitive effects associated with frequently performed elective procedures. It may therefore be proposed that just as the Buffalo group concluded that earlier studies were "generally lacking in methodological rigour",[Cryns et al, [1990 p191] psychiatrists and members of other medical specialties, not merely in the United States, appear to have determined previous research, where reviewed, to be in some way inadequate and the results therefrom suspect or invalid.

It is acknowledged that there is a considerable time lag between conceptualising, initiating and conducting a longitudinal study and publishing its results, and that this may result in overlap and duplication in research efforts particularly between groups in different countries. On the other hand, the present review indicates that there has also been a change in the rationale for more recent investigations when compared with those published prior to the mid-1980s. When Bedford, a consultant physician, expressed his concerns in 1955 with regard to the "adverse cerebral effects of anaesthesia on old people",[Bedford, 1955 p259] the study conducted in response to his retrospective case reviews was designed to investigate the evidence regarding iatrogenic assaults on mental status in order that older people should not be denied the potential benefits of elective surgery when indicated,[Simpson, 1961 p893] (see above). Furthermore, the legitimation for subsequent investigations which evaluated the use of different anaesthetic agents and techniques related to the minimisation of risks to the patient and a reduction in the debilitating effects of surgical interventions. The authors of recent reports have, however, tended to justify their research with observations regarding "the increasing prevalence of geriatric patients in our population",[Chung et al, 1989] and the "life expectancy of the western population" resulting in "more and more surgery .... in geriatric patients".[Haan, 1991 p596] It may therefore be proposed that there has been a shift from evaluating practice in the interests of the well-being of patients, to negative perceptions of the older population as both a demographic and medical management problem, the exacerbation of which should be avoided. It may be further proposed that more recent studies have been conceived not merely in the context of the "problem of ageing" but also explicitly or implicitly within the "cognitive paradigm" of dementia,[Berrios, 1989 p11] as demonstrated by the use of the instruments developed by American psychiatry for its operationalisation.

Earlier investigations reviewed for this discussion either acknowledged the assistance of psychologists in the selection and development of a range of different test instruments and/or co-opted them in their design and conduct.[Simpson et al, 1961, Aberg et al, 1974 Simpson et al, 1976, Ryhanen et al, 1978 Cohen et al, 1982 Riis et al, 1983] In 1987, however, the results of a study of cataract surgery were published which evaluated outcome in terms of cognitive status as well as visual change and physical functioning over a period of a year after the procedure.[Applegate, 1987 p1064 Elam, et al, 1988] No psychologists were identified in the research team, and mental status before and after surgery was assessed with the Short Portable Mental Status Questionnaire developed by a psychiatrist.[Pfeiffer, 1975 p433-441] Other recent surgical studies reviewed, have adopted the psychiatric Mini-Mental State Examination,[Folstein et al, 1975] as the primary, but not necessarily the only, assessment instrument and indicator of change.[Chung et al, 1987, Chung et al, 1989(a), Chung et al, 1989(b), Haan et al, 1991][12] It was the latter that was approved by the NIMH for use in both the Georgetown University and the New York, Buffalo study.

---

[12] As previously noted, the 10-item Short Portable Mental Status Questionniare developed by Pfeiffer[Pfeiffer, 1975 p433-41] was adopted by the National Institute on Aging for assessment of mental status in its multi-site epidemiological study, whereas the Folstein Mini-Mental State Examination,[Folstein, 1975 p189-98] was favoured by the National Institute of Mental Health in the 1980 ECA study of adult mental disorders to determine "cognitive impairment" (see Chapter 4).

It was earlier proposed that the cognitive paradigm for dementia was constructed by American psychiatry and operationalised as a case-finding method for the condition, ignoring the body of knowledge that had previously been elaborated by psychologists with regard to "mental testing" and the behavioural component of dementing disease (see Chapter 4). It may therefore be proposed that evidence of the international dominance of the cognitive paradigm among unrelated medical specialties, as demonstrated by the replication of earlier studies of the mental status effects of surgery on older people employing the new instruments validated for its detection, is a tribute to the success of the enterprise on behalf of which the paradigm was constructed and of its proponents in defining studies conceived outside the paradigm as inadequate or "lacking in methodological rigour". Furthermore, although positive benefits may derive from the "standardisation" of international indicators in the field of medical care and from the reporting of research findings which may be generalisable to other populations, the adoption of unreliable or invalid investigative indicators and the dissemination of the results of data based upon them may serve merely to reinforce perceived "problems", with negative effects upon those who are the focus of scrutiny. It was from this perspective that data from the study conducted by the Center on Aging at Georgetown University were analysed, a study funded by the Mental Disorders of Aging Research Branch of the NIMH with public financing made available on behalf of the Alzheimer's enterprise.

Chapter 6

STUDY METHODS, RESPONSE RATES AND SAMPLE CHARACTERISTICS

6.1 The study site and population

Georgetown University, founded in 1789 the year before a site was selected for the federal capital of the United States of America, remains on its original site on which the medical school was also established in 1849.[1] Georgetown now lies in the Northwest quadrant of Washington, D.C. and is one of 7 universities within the 63 square miles of the District of Columbia, three of which have medical schools and teaching hospitals.

Washington, D.C. is a city with a population of approximately 600,000 residents of whom 400,000 (66 percent) are African-Americans and 180,000 (30 percent) are white, with 4 percent Hispanic and other recent immigrants. Metropolitan Washington, comprising the city together with the suburbs of Maryland and Virginia has a combined population of 3.9 million. In suburban Maryland the population of 1.8 million is 65 percent white, as is 81 percent of the 1.5 million suburban Virginia population.[U.S. Bureau of the Census: 1990 Summary Tape File 3A] Because of its location and lack of public transport the main patient catchment areas for Georgetown Medical Center are the affluent and very predominantly white neighbourhoods of Washington, D.C., and the demographically similar Maryland suburbs north of the city and the nearby Virginia areas.

---

[1] "At George Town there is a Jesuit College; delightfully situated and as far as I had an opportunity of seeing, well managed. Many persons who are not members of the Romish Church avail themselves, I believe, of these institutions and the advantageous opportunities they afford for the education of their children. The heights in this neighbourhood above the Potomac River are very picturesque, and are free I should conceive, from the insalubrities of Washington".[Charles Dickens, 1843 p85]

6.2 Study Methods

The study entitled "Mental status change after surgery in an elderly population", consisted of a series of assessment interviews conducted in the subjects' homes. A consecutive, i.e. non-randomised, sample of patients, scheduled for surgery by participating surgeons in the Departments of Ophthalmology, Urology and Orthopaedics was enrolled in the study between January, 1989 and September 1990. Criteria for inclusion were: age 60 years or over; resident in Metropolitan Washington, DC; scheduled for elective surgery for a non-malignant condition, and not undergoing concurrent treatment for a major systemic disease such as renal failure or cancer.

A weekly list was compiled from the surgical schedules for all those fitting the age and residence criteria, and a letter was sent to each potential participant providing information about the study. A follow-up telephone call was then made, asking individuals to take part in the investigation, with interviews to be arranged at the following times: 1 week prior to surgery; 1 week after surgery (conducted in the hospital if the subject had not been discharged), and 6 weeks, 6 months and 1 year post-surgery. Written consent for the interviews and for access by the study team to medical records was given by all who agreed to participate.

Similarly, a consecutive sample of patients to act as "control cases" for the study was enrolled from a list generated by the out-patient billing department of the Department of Ophthalmology. Criteria for enrollment of controls were the same as for surgical cases, except that individuals were ineligible if they were anticipating having to undergo any surgical procedures during the study year. Control subjects were asked to consent to three in-home interviews at time 1 (baseline), and 6 months and 1 year later.

All interviews were conducted by two interviewers (the author, who was also the Project Officer, and a research nurse).

6.3 Changes to the protocol

Approximately 3 months after the data collection had begun a meeting was arranged between the Center on Aging research team from Georgetown University and members of the Mental Disorders of Aging Research Branch of the National Institute of Mental Health (NIMH) to request two inter-related changes to the study protocol:

1) **Eligibility for the study:**

   The original study protocol proposed the exclusion of anyone taking more than two concomitant medicines of any kind, either prescribed or "over-the counter" purchases, in order to "limit the variables with regard to medication effects". Numerous studies in the United States have found mean values for concomitant drug use by those aged 65 and over living in the community to be 3-4 per day.[Stockton & Jones, 1993 p2] Lower than average medication use as an eligibility criterion for the study resulted in the exclusion of all but a small initial sample of healthy, affluent and well-educated enrollees. In particular, elderly African-Americans had higher levels of morbidity and consequently higher medication use. Therefore most of the relatively small percentage of the majority Washington, D.C. population receiving treatment at Georgetown University Hospital, who were willing to participate in study, were deemed ineligible after the first interview. It was therefore proposed that medication use should be discontinued as an exclusionary criterion, both to increase the sample size and to introduce greater variability between subjects.

2) **Cognitive testing**

The original protocol approved the use of the Mini-Mental State Examination (MMSE)[Folstein et al, 1975] as the sole instrument for the evaluation of cognitive functioning at all stages of the study. Little or no variation from near-perfect initial MMSE test scores was found, either before or after surgery, among the early study participants. It was therefore proposed by a clinical psychologist, who had joined the Georgetown University Center on Aging as Research Director, that the "psychiatric" MMSE might be insensitive to the detection of change and that two additional, potentially more challenging measures of cognitive functioning be introduced.

The NIMH representatives approved the change in eligibility criteria and the addition of two "psychology" instruments to the mental status assessment battery (see below).

6.4 Study instruments and the assessment schedule

Final study instruments consisted of the following:

i. Demographic, medical and social activity questionnaire

The questionnaire (see Appendix A) sought baseline information on age, sex, race,[2] marital status, education and occupation. Respondents were also asked to

---

[2] Race is institutionalised in the collection and reporting of all data collected by the Bureau of the Census, National Center for Health Statistics and other agencies in the United States. See discussion below.

list the number of chronic medical conditions for which they had received treatment, and the number of medicines they were taking. Information on the latter was sought at the time of each interview.

The interview also asked participants to report level of "social integration" or "social activity" as determined by indicators ranging from the number of current activities in which subjects were engaged, including employment, voluntary work and hobbies, and the hours per week spent on them. Questions were also asked relating to the amount of contact and interaction the respondent had with other people. All "social activity" data were collected at each assessment interview with the exception of the one conducted one week after surgery.

ii. Activities of Daily Living (ADL) Scale

The Activities of Daily Living (ADL) Scale, based on the Index of ADL developed some thirty years ago,[Katz et al, 1963] assesses physical status as determined by the ability to maintain 7 basic activities including bathing, dressing, feeding and walking. This instrument was administered at each interview.

iii. Geriatric Depression Scale (GDS)

The Geriatric Depression Scale (GDS)[Yesavage,1983] is a 30-item instrument designed to elicit Yes/No answers to questions relating to emotions and behaviour symptomatic of depression. It can be self-administered or the questions may be asked by the interviewer for those with difficulty reading (see Appendix A). The questionnaire was administered at each assessment.

iv. Cognitive screening tests:

Mini-Mental State Examination (MMSE)[Folstein et al, 1975]

b. The Controlled Oral Word Association (Verbal Fluency Test).[Benton et al, 1978]

c. The Similarities Test from the Wechsler Adult Intelligence Scale.[Wechsler, 1939]

Each of the cognitive assessment instruments is described in detail in the subsequent relevant chapters (see Chapter 7 for the MMSE, and Chapter 8 for the Verbal Fluency and Similarities Test).

Inter-rater reliability testing was conducted by the two interviewers for all the mental status assessment instruments. Tests were conducted by each interviewer with the same 10 patients within 24 hours of each other and showed total score correlations ranging from .86 to .90 ($p<.001$) for the GDS, the MMSE, Verbal Fluency and the Similarities Test.

Figure 6.1 shows the schedule for administration of the various assessment instruments relating to cognitive functioning, as well as depressive symptoms, social activity, and ADL functioning for both surgical subjects and control cases.

Figure 6.1

SCHEDULE OF ASSESSMENTS FOR CASES AND CONTROLS DURING STUDY YEAR

	TIME 1[1]	TIME 2[2]	TIME 3[3]	TIME 4[4]	TIME 5[5]
CASES	Cognition	Cognition	Cognition	Cognition	Cognition
	Depression	Depression	Depression	Depression	Depression
	ADL	ADL	ADL	ADL	ADL
	Social Activity		Social Activity	Social Activity	Social Activity
CONTROLS	Cognition			Cognition	Cognition
	Depression			Depression	Depression
	ADL			ADL	ADL
	Social Activity			Social Activity	Social Activity
	1 Baseline	2 1 Week Post Surgery	3 6 Weeks Post Surgery	4 6 Mths Post Baseline	5 1 Year Post Baseline

6.5 Data management

Questionnaires were checked for recording errors immediately after completion of the interviews and again prior to coding. Data were coded according to a standardised format developed by the Project Officer for the study. Coded data from the questionnaires were then entered into computer files developed in dBASE 111 Plus, Version 1.1 (Ashton Tate Inc., 1986). A separate file was created for each interview, i.e time 1 (baseline), time 2 (one week post surgery) etc. Range and consistency checks were run frequently to monitor accuracy of data entry, so that any typographical errors might be corrected. When all interviews were completed files were converted into SAS 6.04, and these files were later merged into one combined dataset.[3]

An additional file containing data on all potential study subjects was kept throughout the study to track response rates for all those asked to participate in the investigation. Information on sex, age and race for those who refused to take part was obtained from computer files in the Medical Records Department. From these data, descriptive statistics, i.e. chi-square and Student's t-tests were computed to compare the characteristics of those consenting to participate with those who refused.

6.6 Sample pool and response rates

Within the District of Columbia, 17 percent of the population is aged 60 and over, almost identical with the national average of 16.8 percent.[U.S. Bureau of the Census: 1990 Summary Tape File 3A] Black women, the poorest of all demographic groups constitute more than 40

---

[3] SAS (Statistical Analysis System) for Personal Computers, SAS Institute Inc., Cary, North Carolina, 27511-8000, U.S.A.

percent of all older age groups in the city of Washington. In contrast, blacks aged 60 and over (both men and women combined) represent no more than 14 percent of the older population in the Maryland suburbs and 6 percent in Virginia.[D.C. Office on Aging, 1991]

During the nearly 2-year study enrollment period, 486 subjects fitting the age and residence criteria were scheduled for surgery or asked to enter the study as control cases. Three hundred and eighty-five agreed to participate, a response rate of 79 percent (see Table 6.1), and were interviewed at least once. There was a slightly higher proportion of men than women in the sample pool (249/486 or 51 percent men), accounted for by the fact that prostate surgery was one of the surgical procedures on which the study was based. However, differences by gender in the total sample size were not significant. On the other hand, it was immediately apparent that the sample pool was not representative of the racial composition of Washington D.C. since only 21 percent of those in the sample pool were African-Americans (103/486 or 21 percent). Furthermore, chi-square tests showed that there were significantly fewer black men than women in the potential sample (40/103 or 39 percent African American men, $p=.007$), making them the least represented demographic group in the study.

Although there were major differences in the demographic structure of the older patient population attending Georgetown Medical Center, as represented by the study sample, there were no significant differences in total or within-group participation rates between men and women, white or African-American. White women and black men had the highest refusal rates, (25 percent) and white men the lowest (17 percent) (see Tables 6.1 and 6.2).

Table 6.1

TOTAL SAMPLE: STUDY PARTICIPATION RATES BY SEX AND BY RACE

	TOTAL SAMPLE		PARTICIPANTS		REFUSALS		(Chi–square) p value
	Number	(%)	Number	(%)	Number	(%)	
TOTAL	**486**	**(100.0)**	**385**	**(79.2)**	**101**	**(20.8)**	
SEX							
Men	249	(100.0)	203	(81.5)	46	(18.5)	0.199
Women	237	(100.0)	182	(76.8)	55	(23.2)	
RACE							
White	383	(100.0)	304	(79.4)	79	(20.6)	0.871
Black	103	(100.0)	81	(78.6)	22	(21.4)	

Table 6.2

COMPARISON OF MALE AND FEMALE PARTICIPATION RATES BY RACE

	TOTAL		PARTICIPANTS		REFUSALS		(Chi–square) p value
	Number	(%)	Number	(%)	Number	(%)	
MEN							
White	209	(100.0)	173	(82.8)	36	(17.2)	0.246
Black	40	(100.0)	30	(75.0)	10	(25.0)	
WOMEN							
White	174	(100.0)	131	(75.3)	43	(24.7)	0.361
Black	63	(100.0)	51	(81.0)	12	(19.1)	
WHITE							
Men	209	(100.0)	173	(82.8)	36	(17.2)	0.071
Women	174	(100.0)	131	(75.3)	43	(24.7)	
BLACK							
Men	40	(100.0)	30	(75.0)	10	(25.0)	0.473
Women	63	(100.0)	51	(81.0)	12	(19.1)	

Table 6.3

COMPARISON OF MEAN AGE AND AGE RANGE FOR PARTICIPANTS AND NON-PARTICIPANTS BY SEX AND BY RACE

	PARTICIPANTS				REFUSALS				Student's t-test
	Number	Mean Age	(S.D.)	Range	Number	Mean Age	(S.D.)	Range	p value
TOTAL SAMPLE	385	72.7	(7.2)	60–93	101	72.8	(7.7)	60–88	.895
Men	203	71.5	(6.6)	60–89	46	72.8	(7.9)	60–86	.310
Women	182	74.1	(7.5)	60–93	55	72.9	(7.6)	60–88	.303
White	304	72.5	(7.0)	60–93	79	72.7	(8.0)	60–88	.867
Black	81	73.5	(7.6)	60–93	22	73.4	(6.4)	60–83	.942

Student's t-tests showed no overall differences in mean age between those who agreed to participate and those who refused, [72.7 years (s.d 7.2) age range 60-93 for participants, and 72.8 (s.d.7.7) age range 60-88 for refusals] (see Table 6.3). Mean ages of women were higher than those of men, and blacks on average were older than whites, but differences were not significant between those who agreed to be enrolled in the study and those who refused. The data did, however, indicate higher rates of refusal among "younger" elderly women, mean age 74.1. (s.d. 7.5) for female participants and 72.9 (s.d. 7.6) for non-participants.

6.7 Data analysis

Data derived from this study were analysed to fulfill the approved objectives of the original protocol, as described in Chapter 5, on behalf of the research team at the Georgetown Center on Aging. Based on the concerns that resulted in changes to the protocol (see above), and the unusual characteristics of the study sample in terms of the distribution and range of years of formal education, the author sought permission to conduct independent analyses with a view to evaluating the validity of the assessment instruments employed in this and similar studies. The subsequent discussion of the data analysis and it results is therefore presented from this independent perspective.

Data were analysed to evaluate the validity of test scores for three cognitive assessment instruments: the MMSE, Verbal Fluency and Similarities Test, the specific content of which are discussed in Chapters 7 and 8. Except where otherwise specified, these constituted the outcome or "dependent" variables in the various analyses. All other variables, including the GDS (depression) score, were treated as intervening or "independent" and therefore potentially confounding factors.

In common with comparisons made between respondents and non-respondents in the total sample pool (described above), descriptive statistics, i.e frequencies, mean values and standard deviations, were computed for within and between-group baseline characteristics of study participants. Chi-square tests were used to determine the significance of differences between categories for nominal scale variables. Student's t-test was used to evaluate the significance of the difference between mean values of interval scale variables including age. Correlation coefficients (Pearson's r) were also calculated to measure the strength and direction of relationships between all interval scale variables, including age and education, and between interval scale variables and dichotomised nominal scale variables such as sex, as a basis for further analyses of data.

The primary test statistics used were:

i) Analysis of Variance (ANOVAs).[4] Two-way ANOVAs were computed to compare values for dependent variables by age group, educational level etc. Within subject, and between group changes over time, i.e. differences in mean values of dependent variables, such as Mini-Mental State Examination, by age group, at different assessment interviews, were evaluated using "repeated measures" ANOVAs. In all the ANOVAs, Tukey's studentised range tests were used to test for the significance of the difference between mean values of each dependent variable by class category.

---

[4] The ANOVA F-statistic ratio reflects the variability of cases in the combined sample relative to the variability within groups. The use of t-tests for multiple comparisons increases the probability of rejecting the null hypothesis of no significant difference between means, since it is predicted that one test in twenty will yield significance at the .05 level even if all population means are identical. ANOVA is therefore a more appropriate statistical test.[Blalock,1979, p253]

ii) Standardised multiple linear regression models were constructed to determine the relative strength, direction and predictive value of selected "independent" demographic, morbidity and social activity variables used in the ANOVAs. Beta weights were computed to indicate the level of change in the dependent variable produced by a standardised change in each independent variable when controlling for others in the model.[5] Blalock, 1979 p480

6.8 Selection of independent variables for the analyses

The study produced a large volume of data from which a number of specific variables were selected as potential predictors of test scores in the present analyses. The rationale for selecting the following intervening variables was based upon reviews of published reports of studies employing cognitive testing in old age and/or their potential association with the results of such assessments:

Age: Prevalence estimates vary widely, but advancing age has been the consistent and primary predictor of dementing disease in all studies reviewed. Bliwise et al, 1987 p7-11 In this investigation, standard demographic categories were used in evaluating test scores by age, i.e. 60-64, 65-74, 75-84 and 85 years and over.

Years of Education: A relationship between education and cognitive assessment test scores has been observed in several studies which found low levels of education to be a confounding factor in screening tests. Anthony et al, 1982 Scherr et al, 1988, Murden et al, 1991 It has also been proposed that there is a positive association between higher education and cognitive

---

[5] Beta weights, together with mean values of the dependent variable and the multiple correlation coefficient, R^2 are presented in tables in Chapters 7 and 8.

assessment scores.[Weissman et al, 1985, O'Connor et al, 1989(a)] On the other hand there is a marked inverse relationship between age and years of education in the United States. The mean years of education for the total adult population in 1983 was 12.5, but it fell to 11.0 for those aged 65-69 and to 8.8 for those aged 75 and over. Furthermore nearly 55 percent of the older age group finished school before the age of 14, and 7.5 percent attended college, compared with 28 percent with limited schooling and more than 10 percent with college education in the age group 60-64.[U.S. Senate, 1983 p35] In view of this relationship earlier studies included few, if any older individuals with very high levels of education, in contrast with the distribution in the present study (see below).

For purposes of the analysis, standard United States classifications were used for years of formal education: 8 years or less (elementary school), 9-12 years (high school), 13-16 years (college), 17 or more years (postgraduate education).

Medical Diagnoses All medical conditions with systemic effects, from Parkinson's disease to diabetes, may result in impairment in physical and in some cases mental status. Since older people may have significant comorbidity and the level of disability produced by any specific disease or disease combination should be reflected in the ADL Score, (see below) no attempt was made to rank or classify medical conditions and the total number reported was the variable used in the present analyses.

Medication Use Numerous studies have shown adverse effects on mental status from both individual and multiple concomitant drug use, or "polypharmacy", and interaction effects may occur not merely with the use of drugs designed to have psychotropic effects such as antidepressants and sleeping pills.[Jones, 1989, Blazer et al, 1983] The multiple pharmaceutical agents reported by study subjects, the many different drug categories, and the relatively small sample size all precluded evaluation of the effects of any specific drugs. Therefore, although it is a relatively crude indicator of the complexity of medication use the total number of medicines reported was the only indicator available for the analyses. A record of the number of medicines in current use was made at each interview.

Disability or ADL Score Many studies have found an association between physical disability and mental illness, a relationship identified particularly with depression.[Lurie, 1987] Advanced dementia is also associated with impairment in physical functioning, i.e. the development of incontinence, difficulty dressing, etc. Some thirty years ago an association was reported between mental test scores and physical status among older people in institutions,[Kahn et al, 1960] and intuitively it might be assumed that physical disability, related to factors such as arthritis or poor vision, would result in lower cognitive assessment scores for instruments requiring writing or observation, and although largely overlooked "hearing and visual deficits" were identified as confounding factors in cognitive testing in one study reviewed.[Blessed et al, 1991 p193]

Assessment of physical status was made on a 3-point scale of ability to maintain the following 7 Activities of Daily Living (ADL): bathing, dressing, using the toilet, continence, feeding, transfer from bed to chair, and walking.[Katz, 1963] Scores in the present study were scaled by ability to perform different activities either independently (score=0), with assistance (score=1), or incapable (score=2). A cumulative score of 0 denotes total independence and a score of 14, total disability.

Depression or GDS Score Major depression is widely accepted as a condition which may present as "pseudodementia" in older people.[Roth, 1976 p77, DSM III-R, 1987 p106] An association has also been reported between "number of depressive symptoms" and cognitive function test scores.[Scherr et al, 1989 p1090] Therefore although "depression" was proposed in the original protocol as a mental status "outcome" variable related to perioperative events, and was used as such in other analyses of the study data,[6] the number of depressive symptoms, as measured by the Geriatric Depression Scale (GDS), was defined as a potential predictor variable for cognitive assessment scores in the present analyses. For the 30-item GDS a score of 0-10 has been reported from testing to represent the "normal range". A score of 11 or more has been reported to indicate possible clinically significant depression.[Yesavage et al, 1983 p46]

Social activity

In the Oxford pre- and post-operative assessment study conducted in 1961,[Simpson et al, 1961] a measure of "social integration" was included as a potential indicator of mental status change. In a more recent investigation, findings were inconsistent with regard to an association between "social isolation" and cognition test results.[Scherr, 1988 p1090] In the present study, several questions were asked which might be defined as yielding indicators of "social integration" or "social activity" and 3 variables were constructed:

Number of Activities: Information was sought on the number and types of activities

---

[6] Billig N., Stockton P., Cohen-Mansfield J. Cognitive and affective changes after cataract surgery in an elderly population. The American Journal of Geriatric Psychiatry 4:1, 1996 (in press).

engaged in by participants including: employment, voluntary work, social activities and hobbies (excluding television watching). The total number reported was used in the analyses.[7]

Hours of Activity: This was a cumulative total of the estimated number of hours per week spent on the above activities.

Company of Others: The average number of waking hours spent in the company of other people was recorded and used in the analyses. A five level variable was developed ranging from 0=less than 5 hours to 4=more than 50 hours per week.[8]

Additional variables:

As noted above, race and gender are the two defining characteristics for the collection, classification and reporting of all statistics in the United States, to the exclusion of data on the relationship between class, as determined by income and occupation, and health status etc. However, it may be argued that the institutionalisation of "race" and "gender" both afford a genetic or biological rationale for differences in mortality, morbidity, income distribution etc.[9]

---

[7] Thirty percent of the sample was employed, but a number of others engaged in a range of voluntary work, daycare for infant grandchildren etc. which involved as great a commitment of time and effort as some of the participants in paid employment.

[8] By an oversight the questionnaire did not include a question related to whether participants were "living alone". Therefore this could not be used as an indicator of "social support" or "social isolation", and "marital status" did not appear to be an appropriate proxy.

Fifty-five percent of the sample was married. A number of individuals, however, had been widowed, divorced or had never married but had informal living arrangements with others, and a number lived with children or grand-children.

[9] There has been some limited academic debate outside the mainstream, regarding the extent to which "race" should be considered as independent of "class" in its effects on the lives of African Americans.[Wallace, 1991 p257, Navarro, 1991]

Sex A number of epidemiological studies, particularly among those conducted in America have reported higher rates of dementia in women.[Holzer et al, 1984, Weissman et al, 1985, Jorm et al, 1987] In contrast, some European studies reviewed reported either no differences between sexes or marginally higher rates in men.[Rorsman et al, 1986, O'Connor et al, 1989]

Race Several studies have reported higher levels of dementia, among the older black compared with the white population, as determined by the use of screening instruments.[Weissman et al, 1985, Schoenberg et al, 1985 Klein et al, 1985] One recent study, however, reported no significant differences among inner city blacks and whites of the same educational level.[Murden et al, 1991]

Preliminary descriptive statistics were computed for the sample categorised by age and sex, as described below. In addition to showing significantly lower levels of education for women compared with men, and African-Americans compared with whites, as well as correlations, either positive or negative, between education and all other selected variables, the data also showed considerable heterogeneity within all sample subgroups. Therefore, beyond the preliminary descriptive statistics, the data were analysed for the sample as a whole, with the exception of the regression analyses, described in Chapters 7 and 8, computed to discriminate between the different levels of formal education of men and women.

6.9 Sample characteristics and preliminary analysis of data

i) Distribution by age, sex, race and education:

As previously shown, (see Table 6.3) the initial total study sample, all of whom completed at least one interview, consisted of 385 individuals aged 60-93 with a mean age of 72.73 (s.d. 7.16). Fifty-three percent of the participants (n=203) were men and 47 percent women (n=182). As might be predicted, in view of the longer life expectancy of women, there were significant differences in the age group distribution by sex [chi-square=14.77 (3 d.f)., p<.005]. More than 49 percent of the women were aged 75 and over (n=90) compared with 33.5 percent of men (n=68) (see Table 6.4). African-Americans represented only 21 percent of the sample (n=81) but there was no significant difference in the age group distribution by race.

There were significant differences in the education group distribution of the sample both by sex and by race (see Table 6.5). Only 21 percent of men in the study had not attended college, i.e. had received 12 years or less of education (42/203) compared with 46 percent of women (84/182), and 45 percent of men (n=92) had postgraduate education compared with 15 percent of women (n=28) [chi-square=47.44 (3 d.f), p=.000]. African-Americans had significantly less education than whites with 32 percent (n=26) having only elementary education (8 years or less) compared with less than 2 percent of whites, 78 percent of whom had received either undergraduate or postgraduate education [chi-square=113.50 (3 d.f.) p=.000]. It is, however, of note that 26 percent of the African-Americans in the sample had attended college, and of these 7 percent (6/81) had postgraduate education.

ii. Baseline mean values of select variables by gender and by race

Women had a significantly higher mean age than men [74.10 (s.d. 7.5) years compared with 71.5 (s.d. 6.6), p<.001]. Women also had significantly less education on average than men [13.07 (s.d. 3.6) mean years for women compared with 16.10 (s.d.

Table 6.4

SAMPLE DISTRIBUTION BY AGE GROUP, SEX AND RACE

			Sex*				Race**			
Age	Total		Men		Women		White		Black	
	N=385	(100.0)	N=203	(100)	N=182	(100.0)	N=304	(100.0)	N=81	(100.0)
60–64	58	(15.06)	38	(18.72)	20	(10.99)	48	(15.79)	10	(12.35)
65–74	169	(43.90)	97	(47.78)	72	(39.56)	134	(44.08)	35	(43.21)
75–84	140	(36.36)	64	(31.53)	76	(41.76)	109	(35.86)	31	(38.27)
85 & over	18	(4.68)	4	(1.97)	14	(7.69)	13	(4.28)	5	(6.17)

* Chi–square value for age distribution by sex, (3 d.f.)=14.767, p=.002.
**Chi–square value for age distribution by race, (3 d.f.)=1.11, p=.775

Table 6.5

SAMPLE DISTRIBUTION BY EDUCATION, SEX AND RACE

			Sex*				Race**			
Education	Total		Men		Women		White		Black	
	N=385	(100.0)	N=203	(100.0)	N=182	(100.0)	N=304	(100.0)	N=81	(100.0)
0–8 years	31	(8.05)	9	(4.43)	22	(12.09)	5	(1.64)	26	(32.10)
9–12 years	95	(24.68)	33	(16.26)	62	(34.07)	61	(20.07)	34	(41.98)
13–16 years	139	(36.10)	69	(33.99)	70	(38.46)	124	(40.79)	15	(18.52)
17 & over	120	(31.17)	92	(45.32)	28	(15.38)	114	(37.50)	6	(7.41)

* Chi–square value for years of education by sex, (3 d.f)=47.44, p=.000.
**Chi–square value for years of education by race, (3 d.f)=113.48, p=.000

3.8), for men $p < .001$] (see Table 6.6).

When gender group differences in mean values of selected morbidity indicators were compared, i.e. medical diagnoses, number of medicines used, disability and depressive symptom score, women had higher mean values than men for all these variables. On the other hand, the only value that was significantly different for women was the ADL or disability score which was 0.66 (s.d. 1.6) for compared with 0.28 (s.d. 0.9) for men, ($p < .05$) (see Table 6.6).

Intuitively it might be predicted, that an inverse relationship would be found between morbidity and social activity indicators and this was confirmed. Men had significantly higher mean values for all the various "social integration" variables, although the range was great for both sexes. On average, men reported significantly more hours of activity per week [35.20 (s.d. 19.3) compared with 24.00 (s.d. 16.0) for women $p < .001$]. They also reported a higher mean number of activities [5.19 (s.d. 2.2) compared with 4.68 (s.d. 2.5) $p < .05$]. Women were also, on average, more socially isolated than men, reporting a lower mean number of waking hours spent in the company of others, [3.58 (s.d. 0.8) reported by men compared with 2.99 (s.d. 1.3) by women, $p < .001$] [10] (see Table 6.6).

When the mean values of these same selected variables were compared for white and African-Americans in the sample the differences were significant for most of the variables tested. Blacks were older on average, but not significantly so [mean age 73.48

---

[10] As described above, the variable for "waking hours spent in the company of others" was scaled from 0 to 4, with "0=less than 5 hours per week", to "4=more than 50 hours per week".

Table 6.6

COMPARISON OF BASELINE MEAN VALUES FOR SELECTED VARIABLES, BY SEX

	MEN (n=203)			WOMEN (n=182)			Student's t-tes
	Mean	S.D.	Range	Mean	(S.D.)	Range	p value
Age	71.52	(6.6)	60-89	74.10	(7.5)	60-93	0.0005**
Education (years)	16.10	(3.8)	4-22	13.07	(3.6)	0-20	0.0001**
Medical Diagnoses	3.12	(1.4)	0-8	3.42	(1.7)	0-8	0.0635
Medicines Used	2.91	(2.4)	0-13	3.34	(2.3)	0-11	0.0733
ADL Disability Score	0.28	(0.9)	0-7	0.66	(1.6)	0-11	0.0047*
GDS Depression Score	5.19	(4.2)	0-25	6.00	(4.6)	0-25	0.0800
Weekly Hours of Activity	38.99	(21.1)	0-96	24.00	(16.0)	0-78	0.0001**
Number of Activities	5.19	(2.2)	0-13	4.68	(2.5)	0-17	0.0364*
Company of Others	3.58	(0.8)	0-4	2.99	(1.3)	0-4	0.0001**

Table 6.7

COMPARISON OF BASELINE MEAN VALUES FOR SELECTED VARIABLES, BY RACE

	WHITE (n=304)			BLACK (n=81)			Student's t-tes
	Mean	S.D.	Range	Mean	(S.D.)	Range	p value
Age	72.53	(7.0)	60-93	73.48	(7.6)	60-93	0.3132
Education (years)	15.72	(3.2)	6-22	10.58	(4.0)	0-19	0.0001**
Medical Diagnoses	3.05	(1.4)	0-8	4.06	(1.8)	1-8	0.0001*
Medicines Used	3.05	(2.5)	0-13	3.34	(2.0)	0-10	0.2697
ADL Disability Score	0.33	(1.1)	0-11	0.94	(1.9)	0-11	0.0064*
GDS Depression Score	5.27	(4.3)	0-25	6.65	(4.8)	0-18	0.0190*
Weekly Hours of Activity	35.20	(19.3)	0-96	19.47	(19.0)	0-72	0.0001**
Number of Activities	5.32	(2.3)	0-17	3.54	(2.1)	0-10	0.0001**
Company of Others	3.39	(1.0)	0-4	2.98	(1.4)	0-4	0.0119*

* Significant p<.05
** Significant p<.001

(s.d. 7.6) compared with 72.53, p=.313] (see Table 6.7). The differences in the education group distribution shown in Table 6.5 were, however, clearly reflected in the mean years of education of African Americans compared with whites [10.58 (s.d. 4.0) years as opposed to 15.72 (s.d. 3.2), p<.001].

African Americans reported higher mean levels of morbidity and the differences were significant for all variables with the exception of medication use. Blacks in the sample had a higher mean number of medical diagnoses [4.06 (s.d. 1.8) compared with 3.05 (s.d. 1.4), p<.001]. They had significantly higher mean disability scores [0.94 (s.d. 1.9) compared with 0.33 (s.d. 1.1), p<.01], and a slightly higher mean number of depressive symptoms than whites [6.65 (s.d. 4.8) as opposed to 5.27 (s.d. 4.3), p<.05], (see Table 6.7).

In common with the results of the analyses comparing men and women, there was an inverse relationship between the higher mean values of morbidity indicators for African Americans and "social integration" variables. Blacks, on average, reported significantly fewer hours of activity per week, [19.47 (s.d. 19.0) as opposed to 35.20 (s.d. 19.3.), p<.001]. They also reported participation in fewer activities [3.54 (s.d. 2.1) compared with 5.32 (s.d. 2.3), p<.001], and less time spent in the company of others [2.98 (s.d. 1.4) compared with 3.39 (s.d. 1.0), p<.05]. A high proportion of the difference between white and African-Americans and between men and women was attributable to the characteristics of black women who represented 63 percent of the African-Americans in the study (51/81) and 28 percent of the female sample (51/182). They were significantly more "disadvantaged" in terms of education, morbidity and social indicators than all other groups.[11] Beyond the statistical differences in mean values of the various selected indicators, for men and women and for racial categories, there was also considerable heterogeneity within all demographic groups as demonstrated by the very similar range of values shown for all variables in Tables 6.6 and 6.7.[12]

Since women were on average significantly older than men in the sample and African-Americans had a higher mean age than whites, it was hypothesised that a high proportion of the difference between groups for the various indicators might be attributable to age. Mean values for each of the selected variables were therefore computed for age groups categorised by gender.

iii. Age group comparisons of select variables by sex

When the sample characteristics were compared by age group and by sex the most interesting finding was with regard to mean years of education, both within and between

---

[11] African-American women had significantly fewer mean years of education than white women, [10.16 (s.d. 3.8) compared with 14.20 (s.d. 2.8), $p<.0001$]. They had higher values for all morbidity indicators than other demographic groups: mean number of medical diagnoses was 4.41 (s.d. 1.8) compared with 3.04 (s.d. 1.5) for white women, $p<.0001$, mean disability score=1.16 (s.d. 2.1) compared with 0.47 (s.d. 1.3) for white women) $p<.05$. Although not achieving statistical significance, levels of depression and medication use were also higher. They had significantly lower mean values for social activity variables and also for time spent in the company of others.

In contrast, the small sample of African-American men (n=30) differed from white men only in mean years of education [11.30 (s.d. 4.3) compared with 16.88 (s.d. 3.0) $p<.0001$), and in social activity levels.

[12] There was the same heterogeneity when groups were categorised by gender and by race, for all variables tested, i.e. the education range for black women was 0-18 years, 6-20 for white, 4-19 years for black men and 6-22 for white. The range of medical diagnoses was 0-8 for white men and women and 1-8 for black. Disability scores were 0-11 for women of both races, 0-7 for white men and 0-5 for African American men. Weekly hours of activity ranged from 1-78 for white women and 0-59 for black, and from 0-96 for white men and 0-72 for black.

age categories. For all age groups (60-64, 65-74, 75-84, and 85 and over) men had significantly more mean years of education than women (see Table 6.8). In contrast with other studies, however, which have found an inverse relationship between advancing age and years of education, there was no such correlation for either men or women. Although the numbers were small, those aged 85 and over had more mean years of education in both gender groups, [18.25 (s.d. 2.4) for men (n=4), and 13.57 (s.d. 4.5) for women] (n=14). The fact that education was found to be independent of age among study subjects afforded a focus for all further analyses.

Women tended to report more medical diagnoses, higher medicine use and higher levels of physical disability than men, but there were no significant differences between men and women by age group (see Tables 6.9 to 6.11). The mean number of medical conditions for the total sample was 3.26 (s.d. 1.6, range 0-8). Women reported a higher mean number of medical conditions with advancing age [3.30, (s.d. 1.4) at ages 60-64, rising to 3.86 (s.d. 2.0) at ages 85 and over]. Men, in contrast, reported a mean of 2.66 (s.d. 1.3) at ages 60-64, increasing to a mean of 3.58 (s.d. 1.5) at ages 75-84 (see Table 6.9). Those few men aged 85 and over (n=4) reported a slightly lower number of medical conditions, [3.50 (s.d. 0.6)] indicating possible differential survival rates for those with lower incidence of morbidity.

This observation was supported by the data relating to medicines in daily use. The mean number of medications for the total sample was 3.11 (s.d. 2.4, range 0-13).[13]

---

[13] As noted earlier (see 6.3 Changes to the protocol) one of the original eligibility criteria for the study was use of 2 or less medications. Had this criterion not been changed, only 176/385 (46 percent) of those who agreed to participate would have been eligible for the study.

Table 6.8
YEARS OF EDUCATION

	Men				Women				p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t–test)
60–64	38	15.45	(3.8)	(4–20)	20	13.50	(2.1)	(9–17)	0.016*
65–74	97	16.55	(3.1)	(8–22)	72	13.33	(3.3)	(5–20)	0.000*
75–84	64	15.53	(4.6)	(4–22)	76	12.61	(4.0)	(0–20)	0.000*
85 & over	4	18.25	(2.4)	(15–20)	14	13.57	(4.5)	(6–19)	0.019*

Table 6.9
MEDICAL DIAGNOSES

	Men				Women				p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t–test)
60–64	38	2.66	(1.3)	(1–6)	20	3.30	(1.4)	(1–6)	0.098
65–74	97	2.99	(1.4)	(0–8)	72	3.25	(1.8)	(0–8)	0.325
75–84	64	3.58	(1.5)	(1–7)	76	3.54	(1.5)	(1–8)	0.879
85 & over	4	3.50	(0.6)	(3–4)	14	3.86	(2.0)	(2–8)	0.558

Table 6.10
MEDICINES IN DAILY USE

	Men				Women				p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t–test)
60–64	38	2.32	(2.1)	(0–6)	20	3.15	(2.7)	(0–11)	0.240
65–74	97	2.67	(2.3)	(0–10)	72	3.31	(2.3)	(0–11)	0.080
75–84	64	3.64	(2.6)	(0–13)	76	3.50	(2.2)	(0–10)	0.735
85 & over	4	2.50	(0.6)	(2–3)	14	2.93	(2.7)	(0–8)	0.589

Table 6.11
DISABILITY SCORE

	Men				Women				p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t–test)
60–64	38	0.26	(0.9)	(0–5)	20	0.20	(0.4)	(0–1)	0.707
65–74	97	0.18	(0.5)	(0–3)	72	0.31	(0.8)	(0–4)	0.209
75–84	64	0.39	(1.3)	(0–7)	76	0.89	(2.0)	(0–11)	0.077
85 & over	4	1.00	(0.0)	(1–1)	14	1.93	(2.7)	(0–9)	0.226

*** Significant, p<.05**

Men reported lower usage rates than women in all age groups except those aged 75-84 years. Women aged 85 and over, however, had lower mean rates of medication use than other female age groups, [2.93 (s.d. 2.7) compared with a low of 3.15 (s.d. 2.7) for those aged 60-64 years]. Similarly men aged 85 and over reported rates of medicine use [2.50 (s.d. 0.6)] comparable with men aged 60-64 years (see Table 6.10).

In contrast with values indicative of lower morbidity in terms of medical diagnoses and medication use amongst those of advanced age, there was a correlation between age and physical disability or frailty. The overall mean disability score was 0.50 (s.d. 1.3, range 0-11). Of the total sample, 299 (78 percent) reported independence in all activities of daily living (ADL) and only 29 cases (7.5 percent) had a disability score of 2 or more out of a possible 14. For both men and women, however, disability increased with age, although it should be noted that there were individuals with no reported impairment in all age groups (see Table 6.11).

There was a trend to higher mean numbers of depressive symptoms, as measured by the Geriatric Depression Scale (GDS), with advancing age. The overall mean was 5.56 (s.d. 4.4, range 0-25). Fifty-three subjects (14 percent of the total) had baseline GDS scores of 11 or higher, defined as indicative of possible clinical depression.[Yesavage, 1983 p46] Although the mean number of symptoms was higher for women than men, and increased with advancing age, ranging from 4.37 (s.d. 4.0) for men aged 60-64 to 6.75 (s.d 2.1) for those aged 85 and over, and from 4.80 (s.d. 4.3) for women aged 60-64 to 7.07 (s.d. 5.4) for the oldest age group, the differences between men and women were not significant (see Table 6.12).

Table 6.12
DEPRESSION SCORE

		Men				Women			p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t-test)
60-64	38	4.37	(4.0)	(0-17)	20	4.80	(4.3)	(0-14)	0.712
65-74	97	5.15	(4.7)	(0-25)	72	6.24	(4.7)	(0-16)	0.139
75-84	64	5.63	(3.4)	(0-15)	76	5.84	(4.5)	(0-25)	0.749
85 & over	4	6.75	(2.1)	(5-9)	14	7.07	(5.4)	(0-18)	0.859

Table 6.13
WEEKLY HOURS OF ACTIVITY

		Men				Women			p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t-test)
60-64	38	49.13	(22.4)	(0-85)	20	33.10	(20.2)	(7-78)	0.008*
65-74	97	40.15	(20.2)	(0-96)	72	25.25	(15.5)	(1-57)	0.000*
75-84	64	32.38	(19.4)	(0-95)	76	21.72	(14.2)	(0-57)	0.000*
85 & over	4	20.25	(12.3)	(2-28)	14	16.64	(16.2)	(0-47)	0.647

Table 6.14
NUMBER OF ACTIVITIES

		Men				Women			p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t-test)
60-64	38	4.92	(2.1)	(1-10)	20	5.70	(2.7)	(1-10)	0.269
65-74	97	5.67	(2.1)	(0-13)	72	5.07	(2.6)	(1-17)	0.106
75-84	64	4.75	(2.1)	(0-10)	76	4.28	(2.3)	(0-10)	0.212
85 & over	4	3.00	(2.3)	(1-5)	14	3.43	(2.4)	(0-9)	0.758

Table 6.15
COMPANY OF OTHERS

		Men				Women			p value
Age Group	N=203	Mean	(S.D.)	(Range)	N=182	Mean	(S.D.)	(Range)	(t-test)
60-64	38	3.66	(0.8)	(1-4)	20	3.55	(0.9)	(1-4)	0.650
65-74	97	3.55	(0.8)	(1-4)	72	2.92	(1.3)	(0-4)	0.0007*
75-84	64	3.58	(0.9)	(0-4)	76	2.96	(1.3)	(0-4)	0.001*
85 & over	4	3.50	(0.6)	(3-4)	14	2.79	(1.5)	(0-4)	0.175

Significant, p<.05

For social activity, however, there were wide variations and some significant differences by gender. The overall mean reported "hours of activity" per week was 31.90 (s.d. 20.3, range 0-96). The number of hours declined with each advancing age group for men and women, from a mean of 49.13 (s.d. 22.4) for men aged 60-64, to 16.64 (s.d. 16.2) for women aged 85 and over. Men reported more hours of activity at all ages than women, and for all but the oldest age group, the differences were significant (see Table 6.13). Similarly, the cumulative number of different activities reported tended to decline with advancing age, for both men and women, and there were no significant age-group differences between the sexes (see Table 6.14).[14] In terms of time spent in the company of others, however, women were more "socially isolated" than men after the age of 65, and the differences were significant for all but the oldest age group (see Table 6.15).

In summarising the characteristics of the sample by age group and sex, there were no significant differences in any of the morbidity indicators, i.e. medical diagnoses, medicine use, disability or depression scores, between men and women, although physical disability did increase with advancing age, and was higher for women than men. Age effects for the other indicators were inconsistent, and the range of values was great for all but the small group of men aged 85 and over who appeared to be the most homogeneous group as determined by all the variables tested. Women reported

---

[14] Thirty percent of study participants (predominantly men) were in paid employment, either full or part-time. This probably accounts for the high number of "hours of activity" but relatively lower number of different activities reported by men aged 60-64 years.

significantly less social activity than men and spent less time in the company of others after the age of 65. The greatest difference between men and women, both overall and within age-groups was, however, with regard to mean years of education, which were significantly higher for men than for women. Of particular interest, however, was the lack of association between age and educational level for either gender.

Summary and presentation of data

The study conducted by the Center on Aging at Georgetown University Medical Center was designed to assess the effects of elective surgery on the mental status of people aged 60 and over, defined both in terms of cognitive functioning and number of depressive symptoms. Changes to the protocol (see 6.3 above), resulted both in the enrollment of a more varied sample of study participants than would otherwise have occurred, and in the addition of two additional cognitive assessment instruments to the Mini-Mental State Examination (MMSE),[Folstein et al, 1975] the only test of cognition in the original proposal. All study subjects consented to be interviewed at least once, and complete demographic information and data on morbidity, depressive symptoms and social indicators, as well as MMSE scores were recorded.[15]

It is apparent from the frequency distributions shown in Tables 6.1 to 6.2 that relatively few of the majority African-American older population of Washington, D.C. receive medical treatment at Georgetown Medical Center and it is not proposed that the

---

[15] The Verbal Fluency and Similarities Tests were administered to a 10 percent smaller sample (n=347) since their use was initiated after the study had begun. Any differences from the characteristics of the total sample are discussed in Chapter 8.

study sample is representative of the older population either of the United States in general or of its capital city. On the other hand, the sample was exceptional both in terms of the very high mean educational levels of both men and women, not merely compared with national statistics for the elderly, but also for the total adult population.[U.S. Senate, 1983 p35] It was also heterogeneous as demonstrated by the range of values for all the variables selected for testing among all demographic groups, male and female, white and African-American. Futhermore, a possibly unique characteristic demonstrated in the preliminary analyses was the independence of level of education from age. The predictive value of education versus age, morbidity, and other indicators, and the relationship between them, therefore afforded a focus for evaluating the validity of the cognitive assessment instruments administered in the conduct of the study, and to this end the level of depression was redefined as an independent predictor rather than an outcome variable as in the original proposal.

Epidemiological studies of dementing disease typically consist of one interview or assessment session. The Georgetown Center on Aging study was, however, designed as a longitudinal investigation and this afforded an opportunity for evaluating the reliability of cognitive assessment instruments as determined by their sensitivity to "practice" or "learning effects" when the tests are repeated over time. The results of these analyses, combined with the cross-sectional data generated by the first interview of the total sample, are presented in Chapters 7 and 8, as evidence for questioning the overall validity and reliability of the techniques employed under the "cognitive paradigm" of dementia for determining "impairment" in epidemiological and clinical studies.

Chapter 7

THE MINI-MENTAL STATE EXAMINATION

7.1 A method for "grading the cognitive state of patients"[Folstein, 1975] The Mini-Mental State Examination (MMSE), developed in the mid-1970s by a psychiatrist, Marshal Folstein, was described in the original publication as a "quantified assessment of cognitive state, of demonstrable reliability and validity", permitting "comparisons to be drawn between intellectual changes and other aspects of mental functioning".[Folstein et al, 1975 p195] The term "mini" was used, not because the test is relatively brief to administer, but because it concentrates "only on the cognitive aspects of mental functioning, and excludes questions concerning mood, abnormal mental experiences and form of thinking".[Folstein, 1975, p189]

The MMSE was validated on a sample of 69 patients admitted to a psychiatric hospital with "dementia syndromes, affective disorder, pseudodementia, mania, schizophrenia and personality disorders", and also a group of 63 "normal subjects" living in the community. It was also administered to a further 137 patients admitted to the same hospital for "standardisation" of scores by type of mental disorder.[Folstein et al, 1975 p190] Twenty-six subjects were selected from the above samples for concurrent testing of the MMSE with the Wechsler Adult Intelligence Scale (WAIS). These tests showed Pearson correlation coefficients between the WAIS "verbal" test scores and the MMSE of 0.776 and between the WAIS "performance" tests and the MMSE of 0.660.[Folstein et al, 194]

Reliability of the MMSE was tested by administration of the instrument twice within 24 hours of hospital admission to 41 study subjects, a procedure which resulted in correlation coefficients of 0.89 between test scores when the instrument was administered by the same person and 0.83 when two different interviewers were used. In addition, 23 "elderly depressed and demented patients chosen for their clinical stability" were tested and re-tested approximately 28 days apart (Pearson's $r=0.98$). From these studies it was concluded that the MMSE is a valid and reliable instrument for distinguishing "patients with cognitive disturbance from those without such disturbance". It was also reported that 33 of the initial sample were tested prior to as well as after treatment "appropriate to their conditions", and it was noted that scores correlated with changes in cognitive state "when and if patients recover".[Folstein, 1975 p194-6] On the other hand, with the exception of 3 patients for whom individual charts illustrated recovery from "metabolic delirium", "head injury" and "depression" over periods ranging from 20 to 75 days, the specific number showing such recovery and the magnitude of improvement in MMSE score was not reported.[Folstein, 1975, p193] Furthermore, there was no record of repetitive testing of "normal" subjects over time.

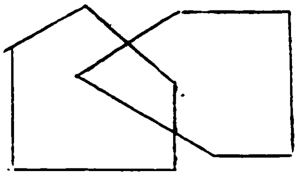
7.2 The content of the Mini-Mental State Examination

The MMSE is an 11-item instrument with a potential total score of 30. Tests included in the formal structure of the MMSE are classified as measures of "orientation", "registration", "attention and calculation", "recall", and "language" (see Figure 7.1):

> **Orientation** has been defined in neuropsychology as "the awareness of self in relation to one's surroundings, (requiring) consistent and reliable integration of attention, perception and memory". It is typically assessed in relation to "person", "place" and "time".[Lezak, 1983 p533] The latter two components are tested

Figure 7.1

MINI MENTAL STATE EXAMINATION QUESTIONNAIRE

Maximum score	Score	ORIENTATION
5	( )	What is the (YEAR) (SEASON) (MONTH) (DATE) (DAY)?
5	( )	Where are we? (STATE) (COUNTY) (TOWN) (HOSPITAL/ FLOOR) or address of interview site.
		REGISTRATION
3	( )	Name 3 objects: 1 second to say each. Then ask the patients all 3 after you have said them. Give 1 point for each correct answer. Then repeat until he learns all three. Count trials and record.
		ATTENTION & CALCULATION
5	( )	Serial 7's, subtracting 7 from 100. 1 point for each correct answer. Stop after 5 answers.
		RECALL
3	( )	Ask for the 3 objects repeated above. Give 1 point for each correct answer.
		LANGUAGE
9	( )	Name a pencil, and a watch (2 points) Repeat the following: "NO IFS ANDS OR BUTS" (1 POINT) Follow a 3-stage command: "TAKE A PAPER IN YOUR RIGHT HAND: FOLD IT IN HALF, AND PUT IT ON THE FLOOR" (3 points) Read and obey the following: CLOSE YOUR EYES (1 POINT) Write a sentence (1 point) Copy a design * (1 point) 
____	____	Total score

*** "On a clean piece of paper, draw intersecting pentagons, each side about 1 inch, and ask him to copy it exactly as it is. All 10 angles must be present and 2 must intersect to score 1 point."** Folstein et al, 1975 p198 in the MMSE. Orientation to "time" is measured by asking the subject the date, month, year, day of week and season of the year, (score=5 points). Orientation to "place" requests details of the address and specific location where the test is taking place (score=5 points).

Registration is the process by which "incoming information is held briefly in sensory store" and is either selected for transfer to short-term memory or ignored.[Lezak, 1983 p25] This is assessed in the MMSE by asking individuals to repeat three unrelated "objects" chosen by the interviewer (score=3 points).

Attention and Calculation The MMSE items classified under this heading relate to three mental processes, "attention, concentration and tracking". Although each may be differentiated theoretically, in practice this is problematic since intact attention is a necessary precondition for both concentration and mental tracking activities.[Lezak, 1983 p548] Two alternative measures are offered to subjects. The first is the "serial sevens test" which requires sequential subtraction of the number 7 from 100 five times to reach the number 65, (score=5 points). Alternatively the person may be asked to spell the word "WORLD" backwards (score=5 points).

Recall Short-term memory is assessed by asking the subject to recall the three nouns repeated above in the "registration" item (score=3 points).

Language The entire second section of the MMSE is described under this heading. The subject is shown a pencil and a watch and asked to name them (score=2 points). The subject is requested to repeat the expression "NO IFS ANDS OR BUTS" (score=1 point); to follow a three stage command relating to folding a sheet of paper (score=3 points); and to obey a written instruction, i.e. CLOSE YOUR EYES (score=1 point). The subject is then asked to write a sentence spontaneously (score=1 point), and to copy a diagram of intersecting pentagons (score=1 point).

The diagram of intersecting pentagons is described by Folstein as a "complex polygon similar to a Bender-Gestalt Figure"[1].Folstein et al, 1975 p190 No rationale, however, was given in the original publication for inclusion of this or any other specific item in the MMSE, or for the classification of all tests in the second section as "language" rather than subdividing them into a "performance" or similar category.

7.3 Normative criteria for scoring the MMSE

When the MMSE was first "validated", the mean score for the control group of "normal elderly persons" whose mean age was 73.9, was reported to be 27.6, s.d. 1.7.Folstein et al, 1975 p191 For patients admitted to the psychiatric unit, a score of 20 out of a possible 30 was reported to distinguish those with "cognitive impairment" from those without.Folstein et al, 1975 p196 Two years later, Folstein, was the co-author of a paper reporting the results of "mental status assessments" of neurological patients admitted to hospital. A mean MMSE score of 22.7 was found among 38 cases determined to have a "cognitive

---

[1] The Bender-Gestalt test, also called the "Visual Motor Gestalt Test", was originated in the 1930s to study the mental development of children. It has been described as assuming "first place position" among the most frequently used psychological tests in the U.S.A. It serves both as a "visuoconstructional task for neuropsychological assessment" and as a "projective technique for studying personality". The test requires the subject to copy nine designs, one of which consists of intersecting hexagons.Lezak, 1983 p385 This latter apparently inspired the pentagon diagram of the MMSE.

defect", and the authors proposed that the instrument had been "revalidated".[DePaulo and Folstein, 1978 p225-7] In subsequent studies reviewed for this discussion the upward revision of the "norm" to an MMSE score of "23 or less" was the standard adopted.

As noted above, it was reported in the original publication that improvement in cognitive state among patients admitted to a psychiatric hospital was reflected by a concomitant increase in MMSE test scores. On the other hand, no significance or normative values were ascribed to any specific improvement in test scores. In 1983, however, Folstein collaborated in an investigation of 20 patients undergoing electroconvulsive therapy (ECT) under general anaesthesia for "major depressive disorder". The mental status of the patients was evaluated with the MMSE, before and after the procedure was carried out. The report of this study proposed that a decrease of two points or more on the MMSE met the "clinical criterion for cognitive decline".[Mondimore et al, 1983 p193] Since that time a two-point "decline" has been used in at least two studies which have employed the MMSE to evaluate mental status change in older people following surgery.[Chung et al, 1989(a) p384, Haan et al, 1991 p599]

7.4 The present study and the focus of the analyses

The administration of the Mini-Mental State Examination[Folstein et al, 1975] to a sample of residents of Metropolitan Washington, D.C. attending Georgetown University Medical Center, afforded the opportunity for evaluating the validity and reliability of the instrument in terms of the association between test scores and a range of demographic, morbidity and social indicators. Since the study from which the data were derived was designed to monitor "mental status change pre- and post-surgery", longitudinal data were also generated for further assessment of the reliability of the instrument in terms of the sensitivity of MMSE scores to repetitive testing. As previously discussed, (see Chapter 6), a number of specific variables were selected for testing as potential independent predictors or confounding factors for MMSE scores. These variables were age, education, medical diagnoses, medicine use, level of disability, depression score, and three "social integration" variables: hours of activity per week, number of activities, and hours per week spent in the company of others. Some significant differences in mean values for these variables were detected between gender subgroups but there was also considerable within-group heterogeneity (see Tables 6.8 to 6.15). In view of this variation, and the great volume of data generated by the study, the analyses presented relate to the total sample, with the specific exceptions described in the tables and narrative.

Analysis of variance modelling was the statistical technique used to assess the significance of the variation in mean values by age, education or other categories. Tukey's test was computed to assess the significance of between-group differences in mean values for categories within the model.

7.5 The relationship between age, MMSE score and other indicators All clinical and epidemiological studies previously reviewed have defined advancing age as the primary predictor of cognitive impairment and/or dementing disease.[Bliwise et al, 1987 pp7-11] It was therefore hypothesised that age would be the primary predictor of MMSE scores in the present study. The results of the analyses demonstrating mean MMSE scores by age category, together with mean values for years of education and all selected morbidity and social indicators is shown in Table 7.1.

As predicted, the analysis of variance (ANOVA) detected a significant overall difference in mean MMSE score by age group [$F(3,384)=11.21, p<.001$]. For the two highest age categories, mean MMSE scores were significantly lower ($p<.05$) than for either of the younger groups, but there was also greater variation in test scores among the older groups ($\bar{X}=27.25$, s.d. 2.9 for those aged 75-84, and $\bar{X}=26.00$, s.d. 3.8 for those aged 85 and over). In addition, scores were on average slightly higher for those aged 65 to 74 ($\bar{X}=28.46$, s.d. 1.7) compared with those aged 60 to 64 ($\bar{X}=28.34$, s.d. 2.1), but the difference was not significant.

When mean years of education were computed by age group, however, relationships were inconsistent and the ANOVA model failed to achieve significance. Those aged 75 to 84 had significantly fewer mean years of education than other age groups ($\bar{X}=13.94$ s.d. 4.5, $p<.05$), and mean years of education were highest for those aged 65 to 74 ($\bar{X}=15.18$, s.d. 3.5), the group with the highest mean MMSE score (see Table 7.1).

The ANOVA models showed significant differences in mean values by age group for all morbidity and disability indicators, with the exception of depression score. There was a consistent increase in the mean number of medical diagnoses with advancing age and the variation was significant [$F(3,384)=4.14, p<.01$]. Those aged 60 to 64 reported significantly fewer medical conditions ($\bar{X}=2.88$, s.d. 1.3, $p<.05$) compared with other age groups. ANOVA indicated a small but significant difference in the mean number of medicines in daily use by age group [$F(3,384)=2.99, p<.05$] but differences in mean values did not correlate with advancing age, the youngest and the oldest age groups

Table 7.1

TOTAL SAMPLE: MEAN BASELINE MMSE SCORES, YEARS OF EDUCATION MORBIDITY AND SOCIAL ACTIVITY LEVELS, BY AGE CATEGORY

		MMSE Score		Years of Education		Medical Diagnoses	
Age	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	58	28.34	(2.1)	14.78	(3.5)	2.88*	(1.3)
65-74	169	28.46	(1.7)	15.18	(3.5)	3.10	(1.6)
75-84	140	27.25*	(2.9)	13.94*	(4.5)	3.56	(1.5)
85 & over	18	26.00*	(3.8)	14.61	(4.6)	3.78	(1.7)

ANOVA of MMSE Scores by Age Category [F(3,384)=11.21, p=.0001]
ANOVA of Age by Years of Education [F(3,384)=2.51, p=.0581]
ANOVA of Age by Number of Medical Diagnoses [F(3,384)=4.14, p=.0066]

		No. of Medicines		Disability Score		Depression Score	
Age	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	58	2.64	(2.4)	0.24	(0.7)	4.51	(4.1)
65-74	169	2.90	(2.3)	0.23	(0.6)	5.62	(4.7)
75-84	140	3.59	(2.5)	0.66	(1.7)	5.74	(4.1)
85 & over	18	2.73	(2.5)	1.72*	(2.4)	7.00	(4.8)

ANOVA of Age by Number of Medicines in Daily Use [F(3,384)=2.99, p=.0311]
ANOVA of Age by Disability Score [F(3,384)=9.53, p=.0001]
ANOVA of Age by Depression Score [F(3,384)=1.83, p=.1411]

		Hrs of Activity		No. of Activities		Company of Others	
Age	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	58	43.60*	(22.8)	5.19	(2.3)	3.62	(0.8)
65-74	169	33.80*	(19.7)	5.41	(2.3)	3.28	(1.2)
75-84	140	26.59	(17.6)	4.49*	(2.2)	3.24	(1.2)
85 & over	18	17.44	(15.2)	3.33*	(2.3)	2.94	(1.4)

ANOVA of Age by Weekly Hours of Activity [F(3,384)=14.57, p=.0001]
ANOVA of Age by Number of Activities [F(3,384)=7.35, p=.0001]
ANOVA of Age by Waking Hours in the Company of Others [F(3,384)=2.38, p=.069]

* $p<.05$

having lower rates of medication use than those aged 65-74 and 75-84.[2]

The most consistent age-associated morbidity indicator was disability score [F(3,384)=9.53, $p<.001$]. The mean value was also significantly higher for the oldest age group than for all others, although the variation was great (1.72, s.d 2.4, $p<.05$). Symptoms of depression tended to increase with advancing age, and to correlate with higher morbidity, but there were no significant overall or between group differences for this variable.

Comparison of social activity indicators by age group and by gender in earlier analyses (see Tables 6.13 to 6.15) showed a marked and consistent decline with advancing age for both sexes, and the findings were therefore the same for the combined sample. The ANOVA model demonstrated a significant difference in mean weekly hours of activity [F(3,384)=14.57, $p<.001$]. Those aged 60-64 ($\overline{X}$=43.60, s.d. 22.8) and those aged 65-74 ($\overline{X}$=33.80, s.d. 19.7) reported significantly more hours of activity than the older groups ($p<.05$). Similarly there was a significant overall difference in the number of activities [F(3,384)=7.35, $p<.001$], those in older age groups reporting a lower mean number than younger subjects ($\overline{X}$=4.49, s.d. 2.2 for those aged 75-84, and 3.33 s.d. 2.3 for those aged 85 and over, $p<.05$). Although there was an inverse relationship between hours spent in the company of others and advancing age, the differences were not significant.

---

[2] As previously noted, lower medicine use among those aged 85 and over may be indicative of differential survival of the healthiest into advanced old age (see Chapter 6, Table 6.10).

7.6 MMSE scores and other indicators by years of education

As described in Chapter 6, a high proportion of the sample, both men and women, had received university education, many undertaking both undergraduate and postgraduate studies. Table 7.2. shows mean values for MMSE scores and other selected variables categorised by years of education. There is a strong correlation between mean MMSE score and educational category, one that is greater than between age groups as reflected in the ANOVA value [$F(3,384)=47.79$, $p<.001$]. Those with 0-8 and 9-12 years of education had significantly lower mean MMSE scores ($\overline{X}=24.10$, s.d. 3.7, and 27.15, s.d.2.4) than groups with education at the college or postgraduate level, ($\overline{X}=28.64$, s.d. 1.5 and $\overline{X}=28.58$, s.d. 1.9 respectively). Differences in mean MMSE scores between the lower two educational groups were also significant ($p<.05$) but not between the two higher education categories.

ANOVA also demonstrated that there was an association between educational category and age, which just achieved significance [$F(3,384)=2.77$, $p=.0415$]. Those with the least education were on average significantly older ($\overline{X}=76.13$, s.d. 7.7 years), compared with a mean of less than 73 years for the other educational groups. Although the difference was small, and did not achieve statistical significance, those with the highest mean MMSE score were in the group with 13-16 years of education, the category which also had the lowest mean age (72.09, s.d. 7.0).

All other variables tested, however, with the exception of number of medicines used, showed a consistent association with educational category and a highly significant overall difference in mean values, as determined by analysis of variance ($p<.001$).

Table 7.2

TOTAL SAMPLE: MEAN BASELINE MMSE SCORES, AGE, MORBIDITY, AND SOCIAL ACTIVITY LEVELS, BY EDUCATIONAL CATEGORY

		MMSE Scores		Age		Medical Diagnoses	
Education	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0–8 years	31	24.10*	(3.7)	76.13*	(7.7)	4.26*	(1.5)
9–12 years	95	27.15*	(2.4)	72.79	(7.3)	3.48	(1.7)
13–16 years	139	28.64	(1.5)	72.09	(7.0)	3.09	(1.6)
17 or more	120	28.58	(1.9)	72.55	(6.9)	3.04	(1.3)

ANOVA of MMSE Scores by Education [F(3,384)=47.79, p=.0001]
ANOVA of Years of Education by Age [F(3,384)=2.77, p=.0415]
ANOVA of Number of Medical Diagnoses by Education [F(3,384)=6.42, p=.0003]

		No. of Medicines		Disability Score		Depression Score	
Education	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0–8 years	31	3.23	(2.1)	1.39*	(2.7)	7.94*	(4.5)
9–12 years	95	3.32	(2.3)	0.54	(1.5)	6.45	(5.3)
13–16 years	139	3.01	(2.3)	0.36	(0.9)	5.45	(4.1)
17 or more	120	3.03	(2.5)	0.28	(0.8)	4.37	(3.5)

ANOVA of Number of Medicines in Daily Use by Education [F(3,384)=0.38, p=.771]
ANOVA of Disability Score by Education [F(3,384)=6.61, p=.0002]
ANOVA of Depression Score by Education [F(3,384)=7.67, p=.0001]

		Hrs of Activity		No. of Activities		Company of Others	
Education	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0–8 years	31	15.74*	(20.1)	2.52*	(1.6)	2.87*	(1.6)
9–12 years	95	25.00*	(16.8)	4.13*	(2.0)	3.06	(1.3)
13–16 years	139	31.97*	(19.6)	5.25	(2.3)	3.40	(1.0)
17 or more	120	41.43	(19.0)	5.90	(2.2)	3.50	(0.9)

ANOVA of Education by Weekly Hours of Activity [F(3,384)=22.21, p=.0001]
ANOVA of Education by Number of Activities [F(3,384)=26.26, p=.0001]
ANOVA of Education by Waking Hours in the Company of Others [F(3,384)=4.61, p=.004]

* p<.05

Furthermore, those in the lowest educational category had significantly higher mean values for all the morbidity and disability indicators and sigificantly lower values for the "social integration" variables, including time spent in the company of others. At the same time, those in the highest educational category (17 years or more) reported the lowest mean values for morbidity and disability, and the highest for social activity, even though, as noted above, the mean age of the post-graduate education group was slightly greater than the age of those with 13-16 years of education (see Table 7.2).

7.7 MMSE score and other indicators by baseline MMSE score level

There was considerable variation in MMSE scores both for those in the oldest age groups and in the lowest educational categories (see Tables 7.1 and 7.2). Therefore, in an attempt to determine the relative effect of age and education, a variable was created to categorise the level of MMSE scores in relation to the norms which have been proposed for assessing the significance of such scores.

As described earlier, the values recorded for MMSE scores when the instrument was first validated were 27.6 for "normal subjects" and 20 for those with "cognitive impairment".[Folstein, 1975 p196] The latter "cut-off" point was raised to 23 in a second publication[De Paulo & Folstein, 1978 p227] and had been largely accepted in the literature reviewed. Based on normative criteria promoted by the developer of the instrument, a "baseline MMSE score" variable was created and categorised as 23 or less= "cognitively impaired", 24-27 = "low normal", and 28-30 = "high normal". Mean values, ANOVAs, and between group significance tests were computed for all predictor variables categorised by the baseline MMSE score level variable.

Twenty-five subjects (6.49 percent of the sample) scored 23 or less on the MMSE, 81 (21 percent) scored in the range 24-27, and 279 (72 percent) scored 28 or more. The mean score for those in the lowest group was 20.96 s.d. 2.2, close to the value first proposed as the indicator of cognitive impairment (see above). Since the lowest category covered a much greater score range (0-23), the within group variation was therefore higher compared with the other groups ($\bar{X}=25.89$, s.d. 1.1, for scores 24-27, and $\bar{X}=29.09$, s.d. 0.8 for scores 28-30). Because the variable was designed to examine the characteristics of those scoring at different levels, both overall variation and between group differences were, predictably, highly significant (see Table 7.3).

When the "baseline MMSE score" variable was evaluated by age and years of education, variation was greater for the latter although both were highly significant, ANOVA for mean age by level of baseline MMSE score [$F(3,384)=14.69$, $p<.001$], compared with [$F(2,384)=44.99$, $p<.001$] for education. Those scoring at the highest level were significantly younger than the other two groups ($\bar{X}=71.56$, s.d. 6.4 for score level 28-30, $\bar{X}=75.60$, s.d. 7.9 for scores in the 24-27 range, and 76.52, s.d. 8.6 for the lowest MMSE score group). There were, however, significant differences in mean years of education between all three groups, ranging from $\bar{X}=9.08$, s.d. 4.1 years for the lowest group to $\bar{X}=15.54$, s.d. 3.4 for those scoring 28-30.

Table 7.2 showed an inverse correlation between years of education and mean values for all morbidity indicators, with the exception of number of medicines in daily use. Similar but more pronounced and consistent relationships were observed with regard to the baseline MMSE score categories (see Table 7.3). For this variable, ANOVA

Table 7.3

TOTAL SAMPLE: MEAN MMSE SCORE, AGE, YEARS OF EDUCATION, MORBIDITY AND SOCIAL ACTIVITY LEVELS BY BASELINE SCORE

Baseline		MMS Score		Age		Years of Education	
MMS Score	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-23	25	20.96*	(2.2)	76.52	(8.6)	9.08*	(4.1)
24-27	81	25.89*	(1.1)	75.60	(7.9)	13.25*	(3.9)
28-30	279	29.09*	(0.8)	71.56*	(6.4)	15.54*	(3.4)

ANOVA of MMS Scores by Baseline Score Level [F(2,384)=940.15, p=.0001]
ANOVA of Age by Baseline Level of MMS Score [F(2,384)=14.69, p=.0001]
ANOVA of Education by Baseline Level of MMS Score [F(2,384)=44.99, p=.0001]

Baseline		Medical Diagnoses		Medicines Used		Disability Score	
MMS Score	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-23	25	4.16	(1.8)	3.64	(2.5)	2.12*	(3.2)
24-27	81	3.64	(1.5)	3.28	(2.2)	0.78*	(1.7)
28-30	279	3.08*	(1.5)	3.01	(2.4)	0.22*	(0.6)

ANOVA of Medical Diagnoses by Baseline Level of MMS Score [F(2,384)=8.77, p=.0002]
ANOVA of Medicines Used by Baseline Level of MMS Score [F(2,384)=1.06, p=.347]
ANOVA of Disability Score by Baseline Level of MMS Score [F(2,384)=31.36, p=.0001]

Baseline		Depression Score		Hours of Activity		Company of Others	
MMS Score	N=385	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-23	25	8.20*	(4.3)	13.96*	(18.0)	2.72*	(1.5)
24-27	81	6.19	(4.4)	27.32*	(19.8)	3.19	(1.2)
28-30	279	5.14*	(4.3)	34.83*	(19.6)	3.39*	(1.0)

ANOVA of Depression Score by Level of Baseline MMS Score [F(2,384)=6.77, p=.0013]
ANOVA of Hours of Activity by Baseline Level of MMS Score [F(2,384)=15.89, p=.0001]
ANOVA of Hours in the Company of Others by Baseline MMS Score [F(2,384)=4.76, p=.0091]

* p<.05

detected significant overall variation for medical diagnoses by MMSE score level, [F(2,384)= 8.77, p<.001]. Those with the highest MMSE scores had a significantly lower mean number of diagnoses than both other groups ($\bar{X}$=3.08, s.d. 1.5), compared with 3.64, s.d. 1.5 for the 24-27 group, and 4.16, s.d 1.8 for those scoring 0-23. A more powerful indicator of the relationship between morbidity and level of MMSE score was the physical disability variable, for which significant overall [F(2,384)=31.36, p<.001] and between group differences were found. Those in the 0-23 MMSE score range had a higher mean disability score ($\bar{X}$=2.12, s.d. 3.2) than the group with baseline score 24-27 ($\bar{X}$=0.78, s.d. 1.7) and this in turn was significantly higher than the mean value for those in the 28-30 group ($\bar{X}$=0.22, s.d. 0.6). Although there was less variation, [F(2,384)=6.77, p<.01] mean depression scores were also significantly different between those in the lowest and the highest MMSE score levels (see Table 7.3).

As would be predicted from earlier analyses, the number of activities and mean hours of activity were inversely correlated with disability scores. Significant differences were found in mean values, both overall [F(2,384)=15.89, p<.001] and between all three groups. There was also significant variation in hours spent in the company of others, [F(2,384)=4.76, p<.01], and between-group differences were significant for those in the 0-23 group compared with the highest MMSE score group (see Table 7.3).

7.8 Summary of the analyses of variance

Reviewing the results presented thus far, there was a significant decline in mean MMSE scores for those aged 75 and over compared with younger age groups (see Table 7.1). On the other hand, there was also a correlation between MMSE scores and years of education, up to the age of 84. An analysis of mean values for MMSE score by educational category (see Table 7.2) showed a greater range than MMSE score by age group. Those subjects in the lowest education category had lower average scores than the oldest age group. Furthermore the mean values for the number of medical diagnoses, disability and depression scores, as well as indicators of social integration, were more consistently significant and more highly correlated (either positively or negatively), with lower education than with advancing age. However any conclusion regarding the relative effect of age and education was confounded by the fact that the mean age of the lowest education group was significantly higher than that of other educational categories.

In attempting to determine the relative influence of age, education and morbidity on test results, a new variable was computed which categorised MMSE scores according to those values which have been proposed as indicating "cognitive impairment", i.e. 0-23, "low normal" scores, ie. 24-27, and "high normal", i.e. 28-30. ANOVAs for all predictor variables demonstrated an absolutely consistent correlation between the three MMSE score levels, and age, education, medicines in daily use, disability and depression scores and indicators of social activity. Those in the highest MMSE score category were significantly younger, had more years of education, less morbidity and medicine use, lower disability and depression scores as well as reporting significantly more hours of activity and time spent in the company of others, and vice versa. Those in the "low normal" range of MMSE scores, 24-27, had intermediate mean values for all indicator variables and those in the "cognitive impairment" category were on average the oldest, with significantly less education, higher morbidity and disability, and lower levels of social activity (see Table 7.3).

There was thus a strong correlation between the level of MMSE score, years of education, rates of morbidity, and levels of social activity or social integration. However, because those in the highest MMSE score category were significantly younger than in the other groups, age served as a confounding factor. Since none of the analyses thus far were able to determine the relative effect of age, education or other specific variables, linear regression analyses were employed to discriminate more precisely between age, education or other independent variables in predicting MMSE scores.

7.9 Multiple regression analyses

Many of the selected indicators were highly correlated and therefore all predictor variables were initially entered into a "stepwise" forward selection multiple regression procedure.[3] A multiple linear regression model was then computed for those variables meeting the significance level selection criterion for the stepwise procedure. In this secondary regression procedure, beta coefficients (standardised beta weights) were computed to indicate the relative effect on the dependent variable, i.e. MMSE score, of a change in any of the predictor variables. The overall significance of the model was determined by the F-statistic, and the multiple correlation coefficient (R^2) defined the overall proportion of the variation in mean MMSE score explained by the independent

---

[3] The maximum multiple correlation coefficient, R^2, occurs when correlations between independent variables are zero. If the intercorrelations are high, the total R^2 will be little greater than the highest individual correlation with the dependent variable, i.e. where there are two highly interrelated independent variables the second will explain essentially the same variation as the first. If they are uncorrelated they will explain a different proportion of the total variation.[Blalock, 1979 p485]

or predictor variables.[4] The stepwise procedure was also used to demonstrate the percentage of variance uniquely explained by each independent variable selected for the model.

For the total sample, 4 variables met the "stepwise" selection criteria for the linear regression model. These variables were, in hierarchical order: education, disability score, age, and number of activities reported. The mean MMSE value for the total sample was 27.89. The multiple correlation coefficient for all independent variables was $R^2=0.36$, and the four-variable model was highly significant [$F(4,380)=54.44$, $p<.001$] (see Table 7.4). Education and disability score were the most significant predictors of MMSE score ($p<.001$), followed by age ($p<.01$) and number of activities reported ($p<.05$). Education explained 21 percent of the variance (R^2 value), disability score 12 percent, and age merely 2 percent. When beta coefficients were examined, education was found to have the highest predictive value for MMSE scores when controlling for other variables in the model (beta=0.332). Only slightly less powerful as a negative predictor of mean MMSE score was disability level (beta=-0.286). Age also was a negative but relatively weak predictor (beta=0.139), approximately equal in its negative effect to the positive power of the number of activities reported (0.130).

Education was the most powerful predictor of test scores for the total sample, and therefore in view of the previously recorded differences in mean years of education between men and women (see Table 6.6), the same "stepwise" variable selection

---

[4] Beta weights indicate how much change in the dependent variable is produced by a standardised change in one of the independent variables when the others are controlled.[Blalock, 1979, p480]

Table 7.4

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Total Sample (N=385)

VARIABLE	Standardised Beta Weights	p value
Age	−0.139	0.0013*
Education	0.332	0.0001**
Disability score	−0.286	0.0001**
Number of Activities	0.130	0.0069*

Total mean score=27.89
Model: R−square 0.36, F(4,380)=54.44, p=.0001

Table 7.5

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Men (N=203)

VARIABLE	Standardised Beta Weights	p value
Age	−0.160	0.0139*
Education	0.400	0.0001**
Disability score	−0.309	0.0001**

Mean score=28.19
Model: R−square 0.30, F(3,199)=28.25, p=.0001

** p<.001
* p<.05

procedure was conducted for each gender group. For both men and women, age, education, and disability score met the significance level for entry into the multiple regression model. For men the mean MMSE score was 28.19, $R^2=0.30$, $F(3,199)=28.25$, $p<.001$ (see Table 7.5). Education explained 17 percent of the variance, disability score 10 percent and age 3 percent. Education was the most powerful predictor of the dependent variable (beta=0.400), followed by the negative effects of disability (beta=-0.309), with age, by comparison, exerting a lesser negative effect (beta=-0.160).

The total mean MMSE score was lower (27.55) for women, but the three-variable multiple correlation coefficient was higher, [$R^2=0.39$, $F(3,178)=38.21$, $p<.001$] (see Table 7.6). The stepwise procedure demonstrated that education explained 23 percent of the variance, disability score 13 percent, and age again explained only 3 percent. Beta coefficient values for the predictive power of education (0.397), disability (-0.312) and age (-0.181) were comparable with those for men.

With increased life-expectancy, old age, as socially constructed by retirement income policy, covers a thirty to forty year age span and in the present study the age range was 33 years, i.e. 60 to 93 (see Table 6.6). Furthermore, as noted previously, the mean age of those with the least education was significantly higher than the mean age of other educational groups even though the range was great. It was therefore decided to dichotomise the sample by age, and conduct separate multiple regression procedures for those aged 60-74 and those aged 75 and over.

For the younger sample, education, disability score, and number of activities reported

Table 7.6

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Women (N=182)

VARIABLE	Standardised Beta Weights	p value
Age	−0.181	0.0037*
Education	0.397	0.0001**
Disability score	−0.312	0.0001**

Mean score=27.55
Model: R−square=0.39, F(3,178)=38.21, p=.0001

Table 7.7

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Subjects Aged 60−74 (N=227)

VARIABLE	Standardised Beta Weights	p value
Education	0.372	0.0001**
Disability score	−0.142	0.0170*
Number of Activities	0.212	0.0009**

Mean score=28.43
Model: R−square=0.31, F(3,223)=32.77, p=.0001

** $p<.001$
* $p<.05$

were all selected by the stepwise procedure for the multiple regression model, but age ceased to have any predictive power as an independent variable (see Table 7.7). The mean MMSE score was 28.43, $R^2=0.31$, $F(3,223)=32.77$, $p<.001$. More than 23 percent of the variance was uniquely explained by education, 6 percent by the number of activities reported, and only 2 percent by disability. Education, as in the previous regression models, achieved the greatest significance and was the most powerful predictor of MMSE score (beta=0.372). As might be predicted intuitively, the negative effect of disability was less of a determining factor for the younger age category than for the total sample or gender subgroups (beta=-0.142), but the number of activities reported was both a highly significant and relatively powerful independent variable (beta=0.212).

When the stepwise regression model was computed for the older age group, i.e. those aged 75 and over, the selected variables were education, disability score, and time spent in the company of others, but not advancing age. Once again the model was highly significant, but in contrast with the regression models for other groups, 20 percent of the variance was explained by disability compared with only 12 percent by education. The multiple correlation coefficient for those aged 75 and over was $R^2=0.35$, $F(3,154)=27.58$, $p<.001$. As predicted by the variance, the negative effect of disability (beta=-0.394) was a more powerful indicator than the positive effects of education (beta=0.316). Time spent in the company of others was also a significant predictor of MMSE score (beta=0.180) (see Table 7.8).

In evaluating the results of the regression analysis for those aged 75 and over the significance of "disability score" was further examined. This variable was positively

Table 7.8

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Subjects Aged 75 and Over (N=158)

VARIABLE	Standardised Beta Weights	p value
Education	0.316	0.0001**
Disability score	−0.394	0.0001**
Company of Others	0.180	0.0069*

Mean score=27.11
Model: R−square=0.35, F(3,154)=27.58, p=.0001

Table 7.9

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Subjects Aged 75 and Over, Excluding the Most Disabled (n=152)

VARIABLE	Standardised Beta Weights	p value
Education	0.336	0.0001**
Disability score	−0.193	0.0102*
Company of Others	0.234	0.0013*

Mean score=27.31
Model: R−square=0.29, F(3,149)=16.44, p=.0001

** p<.001
* p<.05

correlated with age and the number of medical diagnoses, medicine use, and depression score in Tables 7.1 to 7.3. It may therefore be argued that disability score was a cumulative indicator of the debilitating effects of disease and ageing as a predictor of MMSE scores for the total sample, and for all subsets examined. Only for those aged 75 and over, however, was disability a more powerful predictor than education (see Tables 7.4 to 7.7). At the same time, it was known from preliminary analyses that, in general, study subjects reported low levels of physical disability, the overall mean score being 0.66, s.d. 1.6 for women and 0.28, s.d. 0.9 for men (see Table 6.6). A re-examination of the distribution showed that 6 subjects (1.6 percent), all in the older age group, were "outliers" with a mean disability score of 9.0 (range 7-11). Two of these subjects were almost totally disabled by dementing disease.[5] The other 4 were recruited to the study following hospital admission for serious medical problems, and elective surgery for other conditions was recommended during their hospital stay. The high level of disability of this latter group was largely related to the acute physical problems for which they were being treated.[6]

Since a few extreme values may have confounding effects for correlational techniques

---

[5] Dementing disease was not an exclusionary criterion for the study and five individuals had established diagnoses of "organic mental disorders", one following a massive stroke, one with a diagnosis of multi-infarct dementia, and three with diagnoses of Alzheimer's disease. Relatives and caregivers of these five welcomed the inclusion of these subjects hoping that the testing might improve their mental status.

[6] These subjects included a 92-year old woman who had fallen and broken her arm, and cataract surgery was recommended during the recovery period, and two men admitted with congestive heart failure, who underwent elective prostate surgery.

such as linear regression, potentially producing a relatively high correlation where none exists among the remaining cases, it has been recommended that a more accurate model may result from the exclusion of such cases.[Blalock, 1979 p401] When the regression analysis was computed with this "high disability" group excluded, education regained its power as the most significant predictor of MMSE score (beta=0.336). Although the relative effect of disability score declined in the new model, it nevertheless continued to be a strong and significant predictor (beta=-0.193). Of even greater predictive power, however, was the variable reflecting "socialisation" or time spent in the company of others by this older age group (beta=0.234) (see Table 7.9).

Dementing disease would, by definition, be a confounding factor for cognitive assessment score. Therefore, since those with established diagnoses of the condition were all in the category aged 75 and over, a further regression model was computed with this specific group (n=5) excluded, increasing the mean score to 27.41.[7] Again the model was highly significant, $R^2=31$, (F(3,149)=21.70, $p<.001$. Education explained 20 percent of the variance, disability 8 percent, and the company of others 3 percent and was the most significant and powerful predictor of MMSE score (beta=0.389). Although remaining a strong negative predictor of test scores (beta=-0.289) disability achieved less signficance in this model. Time spent in the company of others also had independent predictive power (beta=0.147) for the older age group when the most severely demented were excluded (see Table 7.10).

---

[7] It is not to be assumed that there were no others in the sample with mild to moderate dementing disease, but there were only 5 cases in which such a diagnosis had been formally made.

Table 7.10

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Subjects Aged 75 and Over, Excluding those with a Diagnosis of Dementing Disease (n=153)

VARIABLE	Standardised Beta Weights	p value
Education	0.378	0.0001**
Disability score	−0.276	0.0102*
Company of Others	0.140	0.0446*

Mean score=27.41
Model: R−square=0.31, F(4,149)=21.70, p=.0001

Table 7.11

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR MMSE SCORES

Total Sample Excluding those with Elementary Education (n=354)

VARIABLE	Standardised Beta Weights	p value
Age	−0.179	0.0003**
Education	0.219	0.0001**
Disability score	−0.195	0.0001*
Number of Activities	0.146	0.0057*

Mean score=28.22
Model: R−square=0.23, F(4,349)=23.65, p=.0001

** p<.001
* p<.05

The results of all the regression analyses showed education to be a consistently significant and powerful predictor of MMSE score, the positive effects of which were far greater than the negative effects of age. Furthermore age was superceded by level of disability as a negative indicator of MMSE test score in all but one of the analyses. In a 1982 study using the MMSE as a screening instrument for "delirium and dementia" among adults of all ages admitted to a hospital general medical ward, a high proportion of "false positives" was reported among those with "less than 9 years of education".[Anthony et al, 1982 p397] Guided by this study, a further regression analysis was undertaken, excluding those with the least education. As might be assumed, intuitively, when study subjects with 0-8 years of formal education (n=31) were omitted from the model, education showed less power than in any of the previous analyses (see Table 7.11). Nevertheless, the regression model remained significant [$R^2=0.23$, F(4,349)=23.65, $p<.001$] and education (beta=0.219) continued as a significant independent predictor of MMSE score, explaining 11 percent of the variance and retaining slightly more predictive power than disability level (beta=-0.195) which explained 6 percent of the variance, and age (beta=-0.179) which uniquely explained 3 percent of the variance.

The regression analyses leave little doubt of the strong and positive correlational relationship between education and MMSE scores. This held constant for the total sample, for men, women, those aged 60-74 and those aged 85 and over. Even when those with the least education were omitted from the analysis, education continued as a significant predictor of MMSE score. A further finding, not previously reported, was the significant negative effect of disability as a confounding factor for MMSE score, even when the most disabled and demented were excluded from the regression models. In none of the models tested did age achieve the power and significance of education or disability. A further notable finding was the small but significant independent predictive effect of social activity variables. Specifically, the number of activities reported was a significant independent and positive predictor for the sample as a whole, for men, and for those aged 60-74, while hours spent in the company of others was significant for those aged 75 and over.

7.10 The reliability of the MMSE in repeated testing

In order to evaluate the reliability of the MMSE in terms of the sensitivity of test scores to repeated administration of the instrument, it was decided to focus on those subjects who were tested three times in the peri-operative period, i.e. before surgery, 1-week post surgery and 6-weeks post surgery. By focussing on this group it was also possible to evaluate the significance of any change in test scores during this period against the criteria which had been adopted in other studies involving pre-and post-operative evaluations of mental status using the MMSE as the assessment instrument.[Chung et al, 1989, Haan et al, 1991]

Three-hundred and eighty-five study subjects formed the total sample for whom data have been analysed and described thus far. Of these, 274 (71 percent) underwent surgical procedures approximately 1-week after the baseline interview and then participated in further assessment interviews 1-week and 6-weeks after surgery. Table 7.12 shows a comparison of the mean values of selected variables for the surgical subjects with those of others in the total sample, i.e. those deemed to be "ineligible"

Table 7.12

MEAN VALUES OF SELECTED VARIABLES FOR THOSE COMPLETING PERIOPERATIVE INTERVIEWS COMPARED WITH OTHERS IN TOTAL SAMPLE

	Perioperative Interviews Completed (n=274)		Others (n=111)		Student's t-test
	Mean	S.D.	Mean	(S.D.)	p value
MMSE baseline score	27.94	(2.5)	27.78	(2.4)	0.518
Age	72.41	(7.3)	73.53	(6.7)	0.148
Education (years)	14.85	(3.8)	14.12	(4.4)	0.123
Medical Diagnoses	3.27	(1.5)	3.26	(1.7)	0.978
Medicines Used	3.04	(2.3)	3.28	(2.5)	0.404
ADL Disability Score	0.56	(1.4)	0.22	(1.0)	0.008*
GDS Depression Score	5.55	(4.6)	5.58	(3.9)	0.963
Weekly Hrs of Activity	32.06	(20.6)	31.48	(19.7)	0.795
Number of Activities	5.00	(2.4)	4.86	(2.0)	0.588
Company of Others	3.31	(1.1)	3.29	(1.2)	0.887

* Significant, p<.05

(n=38), "controls" (n=59) who were not interviewed according to the same assessment schedule as those undergoing surgery (see Figure 6.1), and "dropouts" (n=14) who refused more than one interview.[8] Overall the characteristics of those participating in the 3 peri-operative interviews showed little difference from the total sample. Subjects who had surgery, which included hip replacement and cataract extraction, had a significant and predictably higher initial mean disability score ($p < .01$), but there was no significant difference between the two groups in other repects.

The statistical technique adopted to compute within- and between-group changes in mean values for the MMSE and other variables at each of the assessment periods, was Repeated Measures Analysis of Variance. This technique tests for within-group "effects of time", or overall differences in the dependent variable, i.e. in mean MMSE score, from time 1, to time 2, and time 3. It also tests for variation over time in mean values by categories of the independent variable, e.g. age group, baseline MMSE score level etc. This is described as "time by group interaction". Tukey's studentised range tests were again computed to assess the significance of the difference in mean values between categories of the dependent variable, e.g. between MMSE score by age group at each

---

[8] As described earlier (see Chapter 6), initial eligibility criteria included the exclusion of individuals from further interviews if they were using more than 2 medications, and 38/385 or nearly 10 percent were deemed to be "ineligible" for the study for this or other reasons including concurrent treatment for conditions such as renal failure. As might be predicted, a number of individuals were lost to the study at each assessment period. Some refused to continue after the first two or three interviews, or moved from the Washington area. Eight subjects, with mean age 82.63, two of whom were "controls" died during the study year for reasons apparently unrelated to any surgical procedure. A much smaller sample of control cases was enrolled than originally proposed (n=48) and four underwent surgery before the end of the study year. The assessment schedule for control subjects did not include interviews in the post-operative period which are the focus of the data presented in the present discussion (see Figure 6.1).

assessment period.

Since the surgical subjects in the analyses shown in Tables 7.13 to 7.15 represent only 71 percent of the initial total sample, analysis of variance F-statistic values are also shown for the dependent variable at time 1. This is to enable a comparison to be made between the frequency distributions and characteristics of the surgical sample at time 1 (baseline) with those of the total sample shown in Tables 7.1 to 7.3.

7.11 Change in test scores by age group and level of education

The analysis of variance model by age group for the surgical subgroup (n=274) shows that those aged 85 and over and 75-84 had significantly lower mean MMSE scores than the younger groups [F(3,270)=8.88, p<.001] (see Table 7.13), values which are comparable with those of the total sample (see Table 7.1). The ANOVA for the effects of time, i.e. repeating the tests on three occasions, detected a small but significant overall within-group difference [F(2,540)=4.24, p<.05]. On the other hand, there was no significant interaction between time and age categories, i.e. the extent to which any group experienced change did not differ from change which occurred in the other groups. There was a tendency for test scores to increase between time 1 and time 3, but those aged 85 and over showed a slight decline in mean MMSE scores at time 2 (1-week post surgery) but recovered to improve on initial values by time 3.

As shown in Table 7.14 there was considerable variation in mean MMSE scores by educational group at time 1 among those who underwent surgery [F(3,270)=34.83, p<.001], the differences being comparable with those shown for the total sample in Table 7.2. There were significant differences between those in the lower two categories,

Table 7.13

MEAN MMSE SCORES IN THE PERI-OPERATIVE PERIOD BY AGE

		Time 1		Time 2		Time 3	
Age	N=274	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	46	28.63	(2.0)	28.57	(2.1)	28.93	(1.5)
65-74	120	28.53	(1.7)	28.63	(1.6)	28.72	(1.4)
75-84	94	27.12*	(3.0)	27.12*	(3.3)	27.36*	(3.3)
85 & over	14	26.64*	(3.6)	25.57*	(3.4)	26.71*	(2.7)

ANOVA of MMSE Score at Time 1 by Age Group[F(3,270)=8.88, p=.0001]
ANOVA of within-group change over Time [F(2,540)=4.24, p=.0158]
ANOVA of interaction between Time and Age Group [F(6,540)=1.53, p=.166]

Table 7.14

MEAN MMSE SCORES IN THE PERI-OPERATIVE PERIOD BY EDUCATION

		Time 1		Time 2		Time 3	
Education	N=274	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-8 years	20	23.80*	(3.8)	23.90*	(3.9)	24.65*	(3.7)
9-12 years	63	27.22*	(2.4)	27.13*	(2.6)	27.27*	(2.9)
13-16 years	101	28.76	(1.5)	28.58	(1.9)	28.81	(1.7)
17 or more	90	28.50	(2.1)	28.69	(2.0)	28.92	(1.5)

ANOVA of MMSE Score at Time 1 by Education Group [F(3,270)=34.83, p=.0001]
ANOVA of within-group change over Time [F(2,540)=4.23, p=.0151]
ANOVA of interaction between Time and Education Group [F(6,540)=1.38, p=.2199]

Table 7.15

MEAN MMSE SCORES IN THE PERI-OPERATIVE PERIOD BY BASELINE SCORE

		Time 1		Time 2		Time 3	
Baseline MMSE Score	N=274	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-23	16	20.44*	(2.4)	20.81*	(3.6)	21.88*	(3.6)
24-27	55	25.85*	(1.1)	26.16*	(1.9)	26.91*	(2.1)
28-30	203	29.12*	(0.8)	28.99*	(1.1)	29.03*	(1.1)

ANOVA of MMSE Score at Time 1 by Baseline Score [F(2,271)=723.42, p=.0001]
ANOVA of within-group change over Time [F(2,542)=4.78, p=.0088]
ANOVA of interaction between Time and Baseline Score [F(4,542)=9.05, p=.0001]

* p<.05

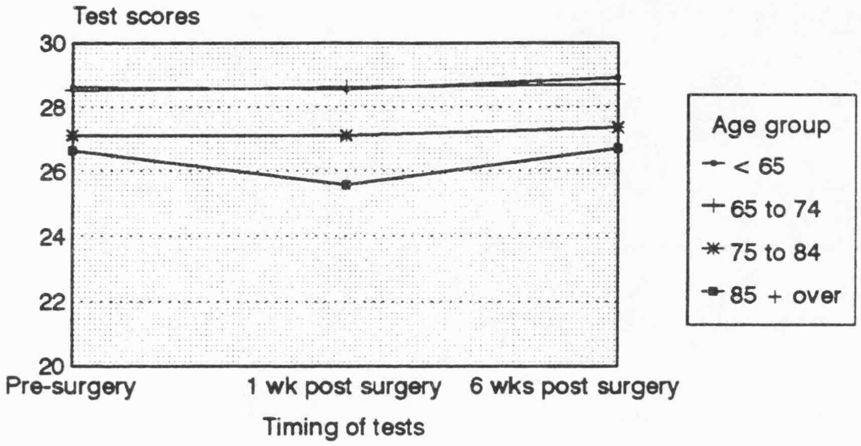
and also between each of them and the higher education groups. Repeated measures ANOVA demonstrated that although these relationships did not change, there were within-group differences between time 1 and time 3 [F(2,540)=4.23, $p<.05$]. Over the same period, the interaction between time and education category was not significant. When the MMSE values for each education group were examined over the 6 week period, a small increase in mean scores was shown by all groups between time 1 and time 3, but the increase was marginally higher for those in the lowest education category.

7.12 Repeated measures by baseline MMSE score level

Table 7.3, which compared mean values for independent variables by categories of baseline MMSE score demonstrated that baseline score was strongly correlated with years of education. Again the mean score distribution for those undergoing surgery was almost the same as observed in the total sample, with very high F-statistic values [F(2,271)=723.42, $P<.001$] reflecting the design of the variable (see Table 7.15). The significant differences in mean values between groups were found at each of the assessment times. There was, however, a significant within-group change between time 1 and time 3 [F(2,542)=4.78, $p<.01$]. Furthermore, and in contrast with the repeated measures analyses by age and education group, the interaction between time and baseline MMSE score category was also highly significant [F(4,542)=9.05, $p<.001$]. When the mean values by group over time are examined, it is apparent that those in the lowest band had the greatest improvement in mean scores, from 20.44 to 21.88 at time 3, the middle group improved from 25.85 to 26.91, and the highest scoring group experienced a small decline in mean MMSE score between time 1 and time 3.

Figure 7.2

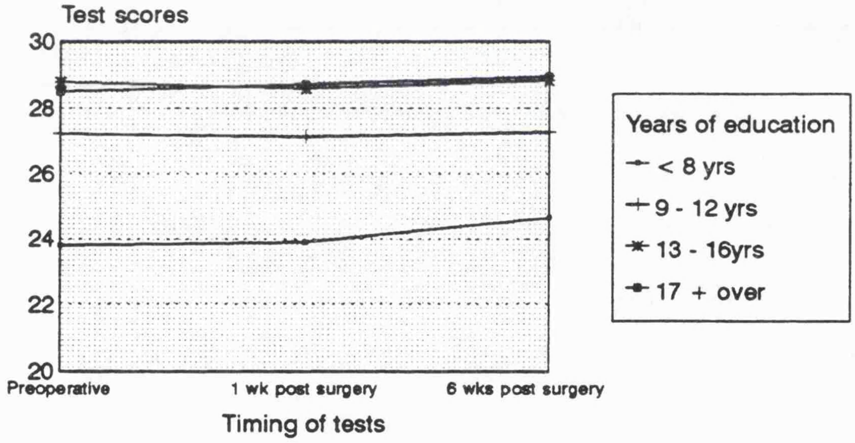
Change in Perioperative Mean MMSE Scores
By Age Group



ANOVA of effects of time, $p < .05$
ANOVA of interaction betwen time and age group, $p = .166$

Figure 7.3

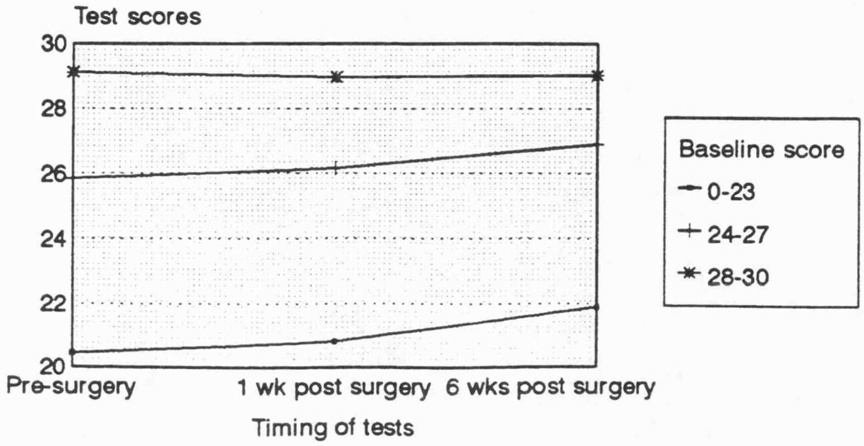
Change in Perioperative Mean MMSE Scores
By Level of Education



ANOVA of effects of time, $p < .05$
ANOVA of interaction between time and group, $p = .220$

Figure 7.4

Change in Perioperative Mean MMSE Scores
By Baseline MMSE Level



ANOVA of effects of time, $p < .01$
ANOVA of interaction between time and group, $p < .001$

The results of the repeated measures analysis of variance by age, educational group, and baseline MMSE score level, are illustrated in Figures 7.2 to 7.4, showing the much greater variation in mean MMSE scores by educational category than by age group. Figure 7.2 demonstrates the relative decline in MMSE scores at time 2 (1-week after surgery) for those aged 85 and over, possibly reflecting the debilitating effect of surgery in the oldest age group. On the other hand mean scores returned to baseline values by time 3 (6-weeks after surgery), indicating that the response was transient, and there was relative stability in test scores for other age groups. Figure 7.3 shows the relatively greater increase in mean MMSE scores for those in the lowest education group, and Figure 7.4 illustrates the gradient of improvement, particularly between times 2 and 3 for those in the two lower baseline MMSE score groups. It may be argued that the latter differences are in part artefactual since there are "ceiling effects" with regard to improvement for those with perfect or near perfect scores in the 28-30 range, and the potential for change is greater in those with an initial MMSE score in the ranges 0-23, and 24-27. Nevertheless, there was an observable and statistically significant improvement among those with lower initial MMSE scores.

In order to evaluate the change in mean MMSE values, categorised by baseline score, against change in morbidity between time 1 and time 3, repeated measures analysis of variance was also conducted for medication use, disability score and depressive symptoms (see Table 7.16). There was an increase in medication use for all three groups between time 1 and time 3 [$F(2,271)=6.46$, $p<.05$], and the interaction between time and group was not significant. There was also a small but significant decrease in

Table 7.16

PERIOPERATIVE CHANGE IN MORBIDITY VARIABLES BY INITIAL MMSE SCORE LEVEL

Baseline MMSE Score	N=274	Medicines 1 Mean	(S.D.)	Medicines 3 Mean	(S.D.)
0–23	16	3.94	(2.7)	4.38	(3.0)
24–27	55	3.29	(2.2)	3.69	(2.7)
28–30	203	2.91	(2.3)	3.08	(2.3)

ANOVA of Medicine Use at Time 1 by Baseline MMSE Score [F(2,271)=1.91, p=.1502]
ANOVA of within–group change over Time [F(2,271)=6.46, p=.0116]
ANOVA of interaction between Time and Baseline Score Group [F(2,271)=0.59, p=.5538]

Baseline MMSE Score	N=274	Disability Score 1 Mean	(S.D.)	Disability Score 3 Mean	(S.D.)
0–23	16	2.69*	(3.2)	1.94*	(2.4)
24–27	55	1.00*	(2.0)	0.89*	(1.5)
28–30	203	0.27*	(0.6)	0.33*	(0.8)

ANOVA of Disability Score at time 1 by Baseline MMSE Score [F(2,271)=30.70, p=.0001]
ANOVA of within–group change over Time [F(2,271)=0.19, p=.6631]
ANOVA of interaction between Time and Baseline Score Group [F(2,271)=5.36, p=.0052]

Baseline MMSE Score	N=274	Depression Score 1 Mean	(S.D.)	Depression Score 3 Mean	(S.D.)
0–23	16	8.50*	(3.9)	8.25*	(4.1)
24–27	55	6.40	(4.6)	5.67	(4.1)
28–30	203	5.09*	(4.5)	4.63*	(4.7)

ANOVA of Depression Score at time 1 by Baseline MMSE Score [F(2,271)=5.45, p=.0048]
ANOVA of within–group change over time [F(2,271)=8.50, p=.0039]
ANOVA of interaction between Time and Baseline Score Group [F(2,271)=0.25, p=.7778]

depression score for all groups between time 1 and time 3 [F(2,271)=8.50, p<.01], and again the interaction of group by time did not achieve significance.

Mean disability values correlated inversely with baseline MMSE score level at time 1, [F(2,271)=30.70, p<.001] a relationship that remained constant over time. However, in contrast with medicine use and symptoms of depression, disability score showed a marked group by time interaction [F(2,271)=5.36, p<.01]. Between time 1 and time 3 there was a decline in level of disability for the two groups scoring at lower levels, (in whom an improvement in MMSE scores was detected, as illustrated in Figure 7.4), and a small increase in disability for those in the highest initial MMSE category, who also experienced a small decline in MMSE scores.

7.13 Review of the repeated measures analyses of variance

Summarising the repeated measures analyses by baseline MMSE category, it is apparent that there were changes in MMSE scores between the pre- and post-operative assessments, with slightly greater improvement among those with the lowest initial scores. An inverse relationship was also found between change in level of disability and MMSE scores. On the other hand, although the mean values for the lower two MMSE score groups improved at each assessment time, the standard deviation increased, indicating that there was greater within-group heterogeneity (see Table 7.15). Increased heterogeneity was true also of the top scoring group, for whom there was a small decline in mean values. The data from the repeated measures analyses therefore indicate that some MMSE scores improved while others declined, and that there was a possible interaction between change in test scores and level of disability.

Further analyses were therefore undertaken in an attempt to determine the extent to which change in MMSE score was predicted by change in physical status or other demographic or morbidity factors, or whether improvement was primarily attributable to a test-retest or learning effect.

7.14 The significance of a two-point change in test scores

As previously noted, it has been proposed that a "decrease of two or more points" meets the "clinical criterion for cognitive decline".[Mondimore et al, 1983 p931 Chung et al, 1989 p384 Haan et al, 1991 p599] Therefore in order to evaluate the relevance of such a change in MMSE score, either positive or negative, its relationship to age and education, and any clinical correlates, i.e. change in disability score, medication use or level of depression, a variable was created to compare those in whom the score remained stable, those with a two-point "decline", and those with a two-point "increase". Repeated measures analysis of variance was again the statistical technique used to evaluate the effects of time, as well as time by group interaction, for each category of the two-point change variable.

There was significant variation in mean MMSE scores at time 1 between the three variable categories, those who improved having a significantly lower initial mean score than the other two groups [F(2,271)=26.69, p=<.001] (see Table 7.17). There was no change in MMSE scores for 73 percent of cases (201/274), who maintained a mean score of 28.50 between the baseline assessment prior to surgery, and the third assessment, 6 weeks after operation. Eleven percent (29/274) declined, from a mean of 27.97 (s.d. 3.4) to a mean of 25.52 (s.d. 3.8), and 16 percent (44/274) showed an improvement in mean MMSE score, from 25.68 (s.d. 2.8) to 28.52 (s.d. 2.0), which was

Table 7.17

COMPARISON OF GROUPS WITH AND WITHOUT A TWO-POINT or GREATER CHANGE IN MMSE SCORES IN THE PERIOPERATIVE PERIOD

Level of		MMS Score 1		MMS Score 3	
Change	**N=274**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**
No change	201	28.46	(1.9)	28.50	(2.0)
Decline	29	27.97	(3.4)	25.52*	(3.8)
Improvement	44	25.68*	(2.8)	28.52	(2.0)

ANOVA of MMSE Scores at Time 1 by Level of Change [F(2,271)=26.69, p=.0001]
ANOVA of within-group change over time [F(1,271)=17.38, p=.0001]
ANOVA of interaction between Time and Level of Change [F(2,271)=319.16, p=.0001]

Level of		Age		Years of Education	
Change	**N=274**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**
No change	201	71.95	(6.8)	14.97	(3.7)
Decline	29	73.28	(8.6)	14.38	(3.7)
Improvement	44	73.93	(8.4)	14.64	(4.4)

ANOVA of Mean Age by Level of Change [F(2,271)=1.57, p=.210]
ANOVA of Mean Years of Education by Level of Change [F(2,271)=0.39, p=.675]

Level of		Medicines 1		Medicines 3	
Change	**N=274**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**
No change	201	3.00	(2.4)	3.23	(2.3)
Decline	29	3.45	(2.2)	3.48	(2.9)
Improvement	44	2.95	(2.0)	3.36	(2.7)

ANOVA of Medicine Use at Time 1 by Level of Change [F(2,271)=0.51, p=.6006]
ANOVA of within-group change over time [F(1,271)=6.46, p=.012]
ANOVA of interaction between Time and Level of Change [F(2,271)=0.52, p=.593]

reflected in significant within group variation [F(2,271)=17.38, $p<.001$] and also in a highly significant group by time interaction, [F(2,271)=319.16, $p<.001$].

When the demographic and morbidity characteristics of those in the three categories of the two-point change variable were evaluated, there was little association with variation in MMSE test scores and there were no significant time by group interactions. The mean age of those in whom no change was observed was lower than the other two groups, but not significantly so, and there was no difference in mean years of education. There were no differences between groups in medicine use at time 1, and there was a small but significant increase at time 3 for all groups [F(2,271)=6.46, $p<.05$].[9] Those with stable MMSE scores had significantly lower mean disability scores at time 1 than those who improved, [F(2,271)=3.40, $p<.05$] (see Table 7.17 contd.). There was an increase in mean level of disability for those in whom MMSE score declined, and a marginal decrease in those with improved mean scores, but there was no significant within group change and no group by time interaction effect. There was also a small but significant within group decline in depressive symptoms between time 1 and time 3 for all groups [F(2,271)=8.55, $p<.01$] but again no significant group by time interaction.

Twenty-seven percent of the sample experienced a two-point or greater change in MMSE scores between time 1 and time 3, the majority (16 percent) being positive, a factor which has merited little discussion in studies which have attached significance to "decline". At the same time the repeated measures ANOVAs indicated that this level of

---

[9] The increase in medicine use may be explained by the fact that all groups were recovering from surgical procedures which typically results in increased use of analgesics, antibiotics and other drugs which may be specific to the procedure.

Table 7.17 (contd.)

COMPARISON OF GROUPS WITH AND WITHOUT A TWO-POINT or GREATER CHANGE IN MMSE SCORES IN THE PERIOPERATIVE PERIOD

Level of Change	N=274	Disability Score 1 Mean	(S.D.)	Disability Score 3 Mean	(S.D.)
No change	201	0.43*	(1.3)	0.41*	(1.1)
Decline	29	0.83	(1.5)	1.10*	(1.5)
Improvement	44	0.98*	(1.6)	0.73	(1.3)

ANOVA of Disability Score at Time 1 by Level of Change [F(2,271)=3.40, p=.0349]
ANOVA of within-group change over time [F(1,271)=0.19, p=.6663]
ANOVA of interaction between Time and Level of Change [F(2,271)=2.53, p=.081]

Level of Change	N=274	Depression Score 1 Mean	(S.D.)	Depression Score 3 Mean	(S.D.)
No change	201	4.82*	(4.3)	4.41	(4.3)
Decline	29	8.62	(6.0)	8.34*	(5.6)
Improvement	44	6.89	(3.5)	5.80	(4.0)

ANOVA of Depression Score at Time 1 by Level of Change [F(2,271)=11.82, p=.0001]
ANOVA of within-group change over time [F(1,271)=8.55, p=.0037]
ANOVA of interaction between Time and Level of Change [F(2,271)=1.14, p=.3215]

change in MMSE scores occurred independently of age, education and morbidity, with the exception of small variations in disability score, for which the lack of statistical significance may indicate that differences in disability levels over time were artefacts brought about by major changes in a few individuals which did not reflect the experience of the majority.

On the evidence thus far it may be argued that as far as could be determined by the data, a two-point improvement in MMSE scores was attributable to an independent learning effect, and that the two-point decrease shown by 11 percent of the sample indicated "cognitive decline" as an outcome of surgery which was again independent of demographic factors or any significant change in morbidity. In a detailed examination of the demographic characteristics of those in whom test scores improved by two points between time 1 and time 3 it was observed that the age range was from 60 to 93, the education range from 5 to 20 years and the gender distribution was approximately equal between men and women. The demographic distribution was almost identical for those in whom MMSE scores declined, i.e. age range 61 to 91, 6 to 20 years of education, with 52 percent women compared with 48 percent men. It was further observed that of the latter group, seventy-nine percent (23/29) had scored in the top level of 28-30 at time 1, and the score of only one of this group changed by more than two points.

Therefore, in view of the high proportion of the sample in whom both positive and negative change was observed, the lack of association between change in morbidity indicators and MMSE score, and the fact that a two point "decline" was observed predominantly among subjects from the top scoring group, it seemed reasonable to question the "clinical" significance of the indicator itself, and the "criterion" was further evaluated by testing for differences in demographic and morbidity indicators among those with a ten percent (3/30) or greater change in MMSE scores over time.

7.15 Evaluating a three-point change in MMSE test scores

In accordance with the previous analysis, a three-level variable was constructed to evaluate differences between those with a less than 3-point change in MMSE values, defined as "no change", a 3-point "decline" and a 3-point "improvement". Based upon these categories, 91 percent of the sample remained stable between time 1 and time 3 (250/274), 3 percent "declined" (9/274), and 5 percent (15/274) showed "cognitive improvement" with MMSE scores increasing on average by more than four points (see Table 7.18). Those without change were younger with a mean age of 72.00 compared with more than 74 years for the other groups, but the difference was not significant. There was, however, significant overall variation in mean years of education [$F(2,271)=4.18, p<.05$]. Those in whom no change was observed had a mean of 15.06 (s.d. 3.6) years compared with less than 13 years for the other groups. For all the morbidity indicators, those who "declined" had significantly higher medication use, disability scores, and higher mean depression scores than the other groups at time 1, all of which increased at time 3. Furthermore, the interaction between time and level of change in disability was significant [$F(2,271)=3.26, p<.05$]. Since there was greater variation in mean values and in statistical significance associated with a three-point change, it was decided to investigate further the individual characteristics of those in whom change was observed.

Table 7.18

COMPARISON OF GROUPS WITH AND WITHOUT A THREE–POINT or GREATER CHANGE IN MMSE SCORES IN THE PERIOPERATIVE PERIOD

Level of Change	N=274	MMS Score 1 Mean	(S.D.)	MMS Score 3 Mean	(S.D.)
No change	250	28.30*	(2.0)	28.40	(2.0)
Decline	9	25.78	(4.8)	22.33*	(5.1)
Improvement	15	23.60	(3.4)	28.10	(2.4)

ANOVA of MMS Scores at time 1 by Level of Change [F(2,271)=36.26, p=.0001]
ANOVA of within–group change over time [F(1,271)=11.21, p=.0009]
ANOVA of interaction between Time and Level of Change [F(2,271)=157.65, p=.0001]

Level of Change	N=274	Age Mean	(S.D.)	Years of Education Mean	(S.D.)
No change	250	72.20	(7.2)	15.06	(3.6)
Decline	9	74.67	(8.8)	12.56	(3.9)
Improvement	15	74.47	(8.2)	12.87	(5.2)

ANOVA of Mean Age by Level of Change [F(2,271)=1.13, p=.326]
ANOVA of Mean Years of Education by Level of Change [F(2,271)=4.18, p=.016]

Level of Change	N=274	Medicines 1 Mean	(S.D.)	Medicines 3 Mean	(S.D.)
No change	250	2.99	(2.3)	3.17	(2.3)
Decline	9	5.33*	(2.4)	6.11*	(3.4)
Improvement	15	2.53	(2.1)	3.40	(3.0)

ANOVA of Medicine Use at time 1 by Level of Change [F(2,271)=5.06, p=.007]
ANOVA of within–group change over Time [F(1,271)=6.53, p=.0112]
ANOVA of interaction between Time and Level of Change [F(2,271)=1.99, p=.139]

The range of initial MMSE scores for those who improved by three points was from 16 to 27, and of those who declined from 18 to 30. Of those with a decrease of 3 or more points, two subjects experienced greater visual impairment following eye surgery and two others had major complications (one renal failure, and one systemic infection) following hip replacement. Two other participants had high depression scores (16 and 22) at each assessment, and for one case there was no apparent reason for a transient decline which increased to baseline at subsequent testing. Also among the subjects who declined by 3 points between time 1 and time 3 were two with an established diagnosis of dementing disease. On the other hand two of this small group were also among the 15 subjects who "improved" by 3 points. Since impairment in short-term memory, defined as "inability to learn new information", is the primary diagnostic criteria for dementia in the DSM-IIIR,[DSM-IIIR, 1987 p107] no "learning effect" can be assumed and the results of this study indicate that cognitive test scores tend to be erratic over the short term among those with dementia. On the other hand, the number of subjects was small, and all scored in the "impaired range" of the MMSE, i.e. 23 or less at each assessment.

Of the others for whom MMSE test scores improved by 3 to 7 points between time 1 and time 3, two subjects experienced dramatic improvement in visual acuity as a result of cataract surgery. These, together with three others, i.e. one third of the total in whom a three-point improvement was observed, also represented twenty-five percent of those in the lowest education category (5/20) (see Table 7.14). At the same time, four of these same subjects had scored in the "impaired range" of the MMSE at time 1 but in the "normal" range at time 3. A significant improvement was thereby demonstrated for a

Table 7.18 (contd.)

COMPARISON OF GROUPS WITH AND WITHOUT A THREE-POINT or GREATER CHANGE IN MMSE SCORES IN THE PERIOPERATIVE PERIOD

Level of Change	N=274	Disability Score 1 Mean	(S.D.)	Disability Score 3 Mean	(S.D.)
No change	250	0.49*	(1.3)	0.49	(1.1)
Decline	9	1.78*	(2.2)	2.11*	(2.0)
Improvement	15	0.93	(1.7)	0.33	(0.9)

ANOVA of Disability Score at time 1 by Level of Change [F(2,271)=4.28, p=.0148]
ANOVA of within-group change over time [F(1,271)=0.19, p=.665]
ANOVA of interaction between Time and Level of Change [F(2,271)=3.26, p=.04]

Level of Change	N=274	Depression Score 1 Mean	(S.D.)	Depression Score 3 Mean	(S.D.)
No change	250	5.36*	(4.5)	4.80*	(4.5)
Decline	9	9.00*	(6.8)	9.44*	(6.6)
Improvement	15	6.67	(2.9)	6.53	(4.1)

ANOVA of Depression Score at time 1 by Level of Change [F(2,271)=3.26, p=.0398]
ANOVA of within-group change over Time [F(1,271)=8.52, p=.0038]
ANOVA of interaction between Time and Level of Change [F(2,271)=0.67, p=.5121]

quarter (4/16) of those determined by the generally accepted criteria to be cognitively impaired, i.e. with a MMSE score less than 23 (see Table 7.15). It may therefore be proposed that for a number of subjects, significant positive change in MMSE performance was attributable either to improvement in physical status and well-being, or to an independent learning effect, or both, and that the improvement was observed disproportionately in those in the lowest education category.

7.16 Summary of results and discussion

The MMSE was developed by Folstein and colleagues in 1975 and promoted as a "quantified assessment of cognitive state" designed to distinguish "patients with cognitive disturbance from those without such disturbance".[Folstein et al, 1975 p195] As such it was incorporated into the Diagnostic Interview Schedule (DIS), the instrument used in the 1980 Epidemiological Catchment Area study, funded by the National Institute of Mental Health (NIMH), from which was derived "the most comprehensive report on psychiatric disorders in America ever assembled".[Regier, 1991 p1] In the DIS the MMSE served as a measure of "current cognitive state".[George et al, 1991 p291] The MMSE was also approved by the NIMH as the instrument to be employed in the study entitled "Mental status change in an elderly population" conducted at Georgetown Medical Center, designed to assess the "prevalence of perioperative cognitive impairment", and the "onset of cognitive impairment in response to elective surgery". It may be posited, however, that the results of the investigation call into question both the validity and reliability of the instrument as either a measure of "impairment" or its onset.

The sample of individuals recruited to the study conducted at Georgetown Medical

Center was unusual, if not unique, in having an age range from 60 to 93, with a high proportion of individuals educated to graduate and post-graduate levels in all age groups, and among both sexes. There was also a smaller sample subset, primarily elderly women, with very limited formal education. It may be argued that education is a proxy for lifetime income, and there was an inverse correlation between education and morbidity, including the number of medical diagnoses and level of physical disability, a finding that was congruent with the results of a national survey reporting an association between income and health status among the older population.[U.S. Senate, 1983 p375] At the same time there was a positive correlation between education and social integration in terms of the number of activities pursued, hours of activity reported and time spent in the company of other people.

There was an association between age and MMSE performance, mean test scores declining with advancing age, subjects in older age groups, i.e. 75 to 84, and 85 and over, having significantly lower mean test scores than those in the younger age categories. On the other hand, there was greater and more significant variation between years of education and MMSE score. Furthermore, in the regression analyses education was selected as the most significant and powerful predictor of MMSE test scores, uniquely explaining the greatest percentage of the variance for the total sample, as well as for men, and for women. As noted previously, a 1982 study in which Folstein participated, which was designed to detect "dementia and delirium", found that the MMSE yielded a high rate of "false positives" among those with low levels of education.[Anthony et al, 1982] In the Georgetown sample, even when those with the 8 years or less of education were excluded from the regression model, the variable continued to be the most significant predictor of MMSE test scores.

In order to respond appropriately to the MMSE, hearing and visual acuity are required, as well as motor skills in terms of the ability to hold and manipulate a pencil in order to follow commands, write, read and copy a diagram. It is therefore of note that rather than age, it was disability score, based on restricted ability to perform Activities of Daily Living,[Katz, 1963] that was shown to be the second most powerful and significant independent, predictor, with negative effects on MMSE score for the total sample, and for both men and women. Furthermore, when the group was dichotomised by age, level of disability exerted the strongest predictive power for MMSE performance among those aged 75 and over.

It may be argued that disability score is a relatively crude indicator, the significance of which is difficult to evaluate, since it may include impairments ranging from moderate difficulty walking, to arthritic hands and poor visual acuity. The score may also be determined by loss of physical control from end-stage dementing disease. On the other hand, when the most disabled and those with an established diagnosis of dementia were excluded from the analyses, disability continued to exercise a significant independent effect on MMSE scores, although education again assumed primacy in predicting test performance. In spite of its significance and power in the present study, there was little discussion in the literature reviewed, of physical disability as a potentially confounding factor in MMSE performance.

It may therefore be argued that based on the strong evidence of education as the most powerful predictor of MMSE score, with physical disability exerting an independent negative effect on performance, the validity of the MMSE as a screening instrument for "cognitive impairment", dementia, Alzheimer's disease or other constructs ascribed to the ageing brain and mind is subject to challenge. At the same time, it may be proposed that further doubt may be cast upon the reliability of any imputations and extrapolations based upon scores from MMSE screening by the results of repetitive testing from the study under discussion.

The rationale for the study was to detect potential deterioration in test scores following surgery and any associated correlates, as well as the proportion of subjects with a 2-point decrease in MMSE scores, proposed as a "clinical criterion for cognitive decline".[Mondimore et al, p931] The data indicated that MMSE scores varied either positively or negatively within a two-point range for nearly one third of participants, the change being primarily positive. There were no significant differences in demographic or morbidity indicators associated with movement in either direction, and nearly eighty percent of those in whom MMSE test scores "declined" scored in the top range at the initial interview. On the other hand it should be noted that a two-point change in MMSE scores in either direction may result simply from failing to state the date and day of the week correctly, i.e. Tuesday the 16th, rather than Wednesday the 17th at one of the assessment interviews (see Figure 7.1). It may therefore be argued that overall, little "clinical" significance could be ascribed to a 2-point decline in MMSE scores. On the other hand, the repeated measures analysis of variance showed that overall there was a tendency for MMSE scores to increase over time, the greatest improvement being among those with the lowest initial scores, i.e. with the least education (see Figures 7.3 and 7.4).

When a 3-point or ten percent change (3/30) was adopted as the indicator of decline, however, there were statistically significant correlates, including very high rates of medication use, depressive symptoms, and an increase in disability score among the three percent (9/274) in whom such a fall in MMSE scores was observed (see Table 7.18). At the same time, five percent of subjects (15/274) showed a 3-point or greater improvement in test performance, including twenty-five percent of those with 0-8 years of education. Furthermore, twenty-five percent of those who scored in the impaired range, i.e. MMSE score 23 or less at time 1 (4/16) improved to the non-impaired range at time 3.

Examination of the characteristics of those in whom decline was observed tended to confirm its association with the onset of physical complications following surgery. On the other hand, with the exception of one or two subjects who experienced major improvement in visual acuity following surgery, there was little association between an increase in test scores and change in morbidity. It may therefore be proposed that the analyses indicated that improvement in test performance was in part attributable to improved physical status but also to an independent learning effect.

The number in whom a "significant" change was observed, either positive or negative was small, and it may be argued that the results of the MMSE testing, were in some way artefactual. It was therefore of value to be able to compare the results of assessments employing the "psychiatric" MMSE with two instruments devised by psychologists, in order either to refute or confirm the findings which call into question the validity and reliability of the MMSE as an instrument in screening for "cognitive impairment" and "decline" in the older population in either clinical or population studies.

Chapter 8

VERBAL FLUENCY AND SIMILARITIES TESTING AND COMPARISON WITH THE MINI-MENTAL STATE EXAMINATION

8.1 The rationale for testing "language skills" and "abstraction"

When the Mini-Mental State Examination (MMSE) was first published it was noted that the instrument excluded "questions concerning mood, abnormal mental experiences and the form of thinking", but it was proposed by the author that "within the cognitive realm it is thorough".[Folstein et al, 1975 p189] On the other hand, the validity of the claim to "thoroughness" in assessment of cognitive functioning has been challenged by both psychiatrists and psychologists.

A Task Force on Alzheimer's Disease, of which Folstein was a member, established under the auspices of the U.S. Department of Health and Human Services to develop criteria for clinical diagnosis of the condition, endorsed the MMSE for documenting "dementia". It also proposed the simultaneous evaluation of other "cognitive processes" in conjunction with the MMSE, such as "language skills tested by examination of verbal fluency of the semantic or category type".[McKhann et al, 1984 p942] Roth, in discussing the Cambridge Mental Disorders of the Elderly Examination (CAMDEX), argued that the rationale for inclusion of the MMSE in the CAMDEX assessment battery was its widespread use in the United States. On the other hand he noted that the instrument failed to assess "perceptual ability and abstract thinking".[Roth, 1986 p700] Similarly, Lezak in her work on "neuropsychological assessment" included the MMSE in a "compendium of tests and assessment techniques". She also posited that the MMSE failed to include "abstraction" and proposed its use in conjunction with the "Similarities" test from the

Wechsler Adult Intelligence Scale (WAIS).[Lezak, 1983 p583]

As noted previously (see Chapter 6, 6.3) when the present study was first approved, the MMSE was the only instrument included in the protocol to pursue the primary study objective, i.e. the detection of change in cognition after elective surgery. When some thirty interviews had been completed, however, preliminary data analyses revealed the high educational level of the majority of study participants and the close association between years of education and MMSE test results. The possibility was therefore raised that the MMSE might be insensitive to detecting change among those with near-perfect scores. In the ensuing discussion of this methodological concern, and in the light of critiques of the MMSE as a comprehensive measure of cognitive functioning (see above), it was suggested that it should be supplemented with other instruments in an attempt to reduce potential doubts with regard to the reliability of the study results. It was therefore proposed by a clinical psychologist on the research team, and approved by members of the Mental Disorders of Aging Branch of the National Institute of Mental Health, sponsors of the study, that two widely used and recommended neuropsychological tests be added to the assessment battery. These were the "Controlled Oral Word Association" or "Verbal Fluency" test[Benton et al, 1978 p11] and the "Similarities" test.[Wechsler, 1939 p87-9]

8.2 The Verbal Fluency Test

The Controlled Oral Word Association (Verbal Fluency) test, was published in 1978 as a component of the Multilingual Aphasia Examination, developed at the University of Iowa. The eleven-test battery was designed to "evaluate the presence, severity, and qualitative aspects of aphasic disorder", i.e. the ability to use and comprehend language.[Benton & Hamsher, 1978 p1] In reviewing neuropsychological assessment techniques, Lezak proposed that word fluency tests provide "an excellent means of finding out whether, and how well, the subject organises his thinking". She argued that "fluency tests requiring the subject to generate words according to an initial letter give the greatest scope to the subject seeking a strategy for guiding his search for words and are most difficult for subjects who cannot develop strategies of their own".[Lezak, 1983 p330]

The Verbal Fluency test consists of a word naming trial of three different letters of the alphabet, each trial lasting for one minute. The examiner asks the subject to say as many words as he can think of beginning with a given letter, excluding proper nouns, numbers, and the same word with a different suffix (see Figure 8.1). Three letters of progressively increasing difficulty are presented successively. The difficulty level of each letter was determined by the relative frequency of words beginning with that letter found in standard dictionaries of the English language. Two comparable sets of letters were defined by this method: **C F L** and **P R W**.[Benton & Hamsher, 1978 p11-14]

The test was "standardised" on a sample of 360 subjects ranging in age from 16 to 69 years, "whose native language was English and who showed no evidence or history of hemispheric brain disease".[Benton & Hamsher, 1978 p1] In validating the test, it was found that "both educational background and age are important determinants of performance level" and therefore norms were provided for the adjustment of scores for these two variables.[Benton & Hamsher, 1978 p12] On the other hand, since the upper limit of the age range of the subjects on whom the sample was standardised was 69, no adjustments were offered for those beyond that age.

Figure 8.1

The Verbal Fluency Test

I am going to say a letter of the alphabet. Then I want you to say as many words as you can think of beginning with that letter, as quickly as you can.

If I give you the letter B, you might say BAD, BATTLE, BED, but I do not want you to use words that are the names of people or places such as BETTY or BOSTON. Also I do not want you to use the same word with a different ending such as BITE and BITING.

Can you think of any words beginning with B?

(If the subject responds appropriately, proceed with the test, otherwise repeat the instructions.)

Group 1

The first letter is **C**. Tell me in one minute all the words you can think of beginning with the letter **C.** Go ahead.

____

____

Total:_________

The next letter is **F**. Now tell me all the words you can think of beginning with the letter **F**. Go ahead.

____

____

Total:_________

The last letter is **L**. Tell me all the words you can think of beginning with the letter **L.** Go ahead.

____

____

Total:_________

Cumulative total =________

Group 2: At alternate interviews the letters **P R W** were given but the administration of the test was exactly as above.

In 1984 the Verbal Fluency test was among eight tests selected by the same University of Iowa research team, and published as the Iowa Screening Battery for Mental Decline. This was described as a "brief neuro-psychological test for examining older persons for whom a question of abnormal mental decline has been raised". The battery was standardised on 53 "normal volunteers" and 53 patients aged 60 to 88 years referred to a Department of Neurology for the "evaluation of dementia".[Eslinger et al, 1984] In reporting the results of the 106 subject standardisation procedure it was proposed that a combination of three of the tests, including Controlled Oral Word Association (Verbal Fluency), "correctly classified 89 percent of cases with a high degree of probability", i.e. either as demented, or without evidence of abnormal mental decline.[Eslinger et al, 1984 p670]

8.3 Verbal fluency and other indicators by age group

Tables 8.1 and 8.2 show the frequency distributions by age group and level of education of the sample to whom the Verbal Fluency test was administered (n=346), as well as the distribution by gender and race. The Verbal Fluency and Similarities tests were introduced some three months after the study had begun, and the number of subjects was thirty-nine less than the total assessed using the MMSE (n=385). On the other hand, a comparison of Tables 8.1. and 8.2 with Tables 6.4 and 6.5, indicates that there was little difference in the distribution of the two samples by age, education, sex, or race. Forty-one percent were aged 75 and over, 67 percent had received formal education at the college or postgraduate level, and 79 percent were white, proportions that were common to both samples. Analysis of variance modelling was the statistical technique employed to assess the significance of the variation in mean values of

Table 8.1

VERBAL FLUENCY TESTING:
SAMPLE DISTRIBUTION BY AGE GROUP, SEX AND RACE

			Sexual Distribution*				Racial Distribution**			
Age			Men		Women		White		Black	
	N=346	(100.0)	N=186	(100.0)	N=160	(100.0)	N=275	(100.0)	N=71	(100.0)
60-64	49	(14.16)	31	(16.67)	18	(11.25)	42	(15.27)	8	(9.86)
65-74	154	(44.51)	92	(49.46)	62	(38.75)	120	(43.64)	34	(47.89)
75-84	128	(36.99)	59	(31.72)	69	(43.13)	102	(37.09)	26	(36.62)
85 & over	15	(4.34)	4	(2.15)	11	(6.87)	11	(4.00)	4	(5.63)

* Chi-square value for age distribution by sex, (3 d.f.)=11.45, p=.01
**Chi-square value for age distribution by race, (3 d.f.)=1.75, p=.626

Table 8.2

VERBAL FLUENCY TESTING:
SAMPLE DISTRIBUTION BY EDUCATION, SEX AND RACE

			Sexual Distribution*				Racial Distribution**			
Education			Men		Women		White		Black	
	N=346	(%)	N=186	(%)	N=160	(%)	N=275	(%)	N=71	(%)
0-8 years	25	(7.23)	7	(3.76)	18	(11.25)	5	(1.82)	20	(28.17)
9-12 years	87	(25.14)	31	(16.67)	56	(35.00)	56	(20.36)	31	(43.66)
13-16 years	128	(36.99)	65	(34.95)	63	(39.38)	114	(41.45)	14	(19.72)
17 & over	106	(30.64)	83	(44.62)	23	(14.38)	100	(36.36)	6	(8.45)

* Chi-square value for age distribution by sex (3 d.f.)=44.31, p=.000
**Chi-square value for age by race (3 d.f.)=87.97, p=.000

demographic, morbidity and social activity variables by age group, and Tukey's test was computed to assess the significance of between-group differences.

Analysis of variance showed a significant overall difference in mean Verbal Fluency values by age group [F(3,345)=6.13, p<.001]. On the other hand, only those aged 75-84 scored significantly lower ($\overline{X}$=30.95, s.d. 14.1) than the other three age categories. At the same time it should be noted that there was considerable within-group heterogeneity in performing this test as reflected in the high standard deviation of mean scores, ranging from 12.8 to 15.1, for all groups (see Table 8.3). The ANOVA model for variation in mean years of education by age group was not significant, but those with significantly lower mean Verbal Fluency scores also averaged fewer years of education, (13.98, s.d. 4.4) compared a mean with 15.00 years for all other groups.

ANOVA demonstrated a small but significant overall difference in the mean number of medical diagnoses, [F(3,345)=3.25, p<.05]. The average number increased with advancing age but there were no significant between-group differences. For medication use by age group, the ANOVA model just achieved significance [F(3,345)=2.70, p=.0459] but the pattern was different. The age group with the lowest mean Verbal Fluency scores and the lowest mean years of education (aged 75-84) also had higher rates of daily medication use, ($\overline{X}$=3.59 compared with less than 3.00 for other groups), but again the between-group differences were not significant.

Mean values for all other selected morbidity and social activity variables, were directly correlated, either positively or negatively with advancing age. There was significant variation in disability level [F(3,345)=7.88, p=.0001], and the mean

Table 8.3

BASELINE MEAN VERBAL FLUENCY SCORES, YEARS OF EDUCATION MORBIDITY AND SOCIAL ACTIVITY LEVELS, BY AGE CATEGORY

		Verbal Score		Years of Education		Medical Diagnoses	
Age	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	49	38.43	(12.8)	15.20	(2.9)	2.98	(1.4)
65-74	154	37.72	(15.1)	15.07	(3.5)	3.14	(1.7)
75-84	128	30.95*	(14.1)	13.98	(4.4)	3.58	(1.5)
85 & over	15	36.60	(15.1)	15.00	(4.4)	3.93	(1.8)

ANOVA of Verbal Fluency Scores by Age Group [F(3,345)=6.13, p=.0005]
ANOVA of Years of Education by Age Group [F(3,345)=2.30, p=.0772]
ANOVA of Number of Medical Diagnoses by Age Group [F(3,345)=3.25, p=.0221]

		No. of Medicines		Disability Score		Depression Score	
Age	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	49	2.69	(2.4)	0.27	(0.8)	4.63*	(4.3)
65-74	154	2.90	(2.3)	0.25	(0.7)	5.60	(4.7)
75-84	128	3.59	(2.5)	0.61	(1.6)	5.72	(3.8)
85 & over	15	2.73	(2.5)	1.67*	(2.3)	7.93*	(4.5)

ANOVA of Number of Medicines in Daily Use by Age Group [F(3,345)=2.70, p=.0459]
ANOVA of Disability Score by Age Group [F(3,345)=7.88, p=.0001]
ANOVA of Depression Score by Age Group [F(3,345)=2.31, p=.0761]

		Hrs of Activity		No. of Activities		Company of Others	
Age	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	49	44.00*	(21.8)	5.24*	(2.2)	3.65	(0.8)
65-74	154	33.88*	(19.7)	5.41*	(2.3)	3.33	(1.1)
75-84	128	27.54	(17.6)	4.65	(2.2)	3.23	(1.2)
85 & over	15	16.00	(14.3)	3.13*	(1.9)	3.13	(1.2)

ANOVA of Weekly Hours of Activity by Age Group [F(3,345)=12.84, p=.0001]
ANOVA of Number of Activities by Age Group [F(3,345)=6.52, p=.0003]
ANOVA of Hours in the Company of Others by Age Group [F(3,345)=2.04, p=.1085]

* p<.05

disability score of those aged 85 and over was significantly higher than other age groups, but the standard deviation indicated that the group was heterogeneous with regard to level of disability ($\bar{X}$=1.67, s.d. 2.3). The mean number of depressive symptoms increased with advancing age and those aged 85 and over reported a significantly higher mean number ($\bar{X}$=7.93, s.d. 4.5) than those aged 60-64 ($\bar{X}$=4.63, s.d. 4.3), but the ANOVA model was not significant.

As shown in previous analyses (see Tables 6.13, 6.14 and 7.1) there was an inverse relationship between age and social activity variables. ANOVA detected a highly significant variation in weekly hours of activity [F(3,345)=12.84, p<.001]. Those aged 60-64 ($\bar{X}$=44.00, s.d. 21.8) and those aged 65-74 ($\bar{X}$=33.88, s.d. 19.7) reported significantly more hours of activity per week than the older age groups. Similarly, there was significant overall variation in the number of activities reported [F(3,345)=6.52, p<.001], those in the two younger age groups reporting significantly more activities than those aged 85 and over. Time spent in the company of others declined with advancing age but there was no significant overall variation or between group difference (see Table 8.3).

8.4 Verbal fluency and other indicators by educational group

Verbal Fluency scores were highly correlated with years of education, and analysis of variance modelling showed both significant variation [F(3,345)=33.93, p<.001], as well as differences between groups. Those with 0-8 years of education had a mean score of 18.96, s.d. 11.7 in contrast with a mean of 42.19, s.d. 13.5 for those with 17 or more years of education. Again, the high standard deviations indicated considerable within-group heterogeneity. When values were computed for mean age by education category, those with the least education and correspondingly low Verbal Fluency scores were also significantly older, with a mean of nearly 77 years, compared with less than 73 years for other educational groups, and the variation was significant overall [F(3,345)=3.13, $p<.05$] (see Table 8.4).

The ANOVA models showed significant variation in mean values by education category for all morbidity indicators with the exception of medicines in daily use, the variation being determined primarily by the difference in mean values for those with the least education. The mean number of medical diagnoses for those with 0-8 years of education (4.26, s.d. 1.5) was significantly higher than for all other education categories, ANOVA value [F(3,345)=7.90, $p<.001$]. Disability score (1.24, s.d. 2.2) was also significantly higher for this same group [F(3,345)=4.32, $p<.01$]. The mean number of depressive symptoms differed between education groups, the two groups with the least education having significantly more symptoms than the higher education categories [F(3,345)=6.99, $p<.001$].

Similarly, comparison of social activity indicators showed considerable variation in mean values, those in the lowest education category consistently reporting less social interaction. For weekly hours of activity there was a significant difference between all groups, those with 0-8 years of education reporting a mean of 17.12 (s.d 20.8) compared with 41.44 (s.d. 18.6) for those with postgraduate education [F(3,345)=17.97, $p<.001$]. The number of activities reported was significantly different between all groups with the exception of those in the two higher education categories [F(3,345)=22.05, $p<.001$].

Table 8.4

MEAN BASELINE VERBAL FLUENCY SCORES, AGE, MORBIDITY, AND SOCIAL ACTIVITY LEVELS, BY EDUCATIONAL CATEGORY

		Verbal Scores		Age		Medical Diagnoses	
Education	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-8 years	25	18.96*	(11.7)	76.88*	(6.3)	4.26*	(1.5)
9-12 years	87	27.90*	(13.3)	72.48	(7.0)	3.54	(1.8)
13-16 years	128	37.73*	(12.6)	72.37	(7.0)	3.14	(1.6)
17 or more	106	42.19*	(13.5)	72.59	(7.0)	3.04	(1.3)

ANOVA of Verbal Fluency Scores by Education Level [F(3,345)=33.93, p=.0001]
ANOVA of Age by Education Level [F(3,384)=3.13, p=.0257]
ANOVA of Number of Medical Diagnoses by Education Level [F(3,345)=7.90, p=.0001]

		No. of Medicines		Disability Score		Depression Score	
Education	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-8 years	25	3.32	(2.0)	1.24*	(2.2)	8.36*	(4.5)
9-12 years	87	3.31	(2.4)	0.49	(1.4)	6.37	(5.0)
13-16 years	128	3.10	(2.4)	0.38	(1.0)	5.48*	(4.1)
17 or more	106	2.93	(2.6)	0.30	(0.9)	4.48*	(3.6)

ANOVA of Number of Medicines in Daily Use by Education Level [F(3,345)=0.45, p=.7190]
ANOVA of Disability Score by Education Level [F(3,345)=4.32, p=.0052]
ANOVA of Depression Score by Education Level [F(3,345)=6.99, p=.0001]

		Hrs of Activity		No. of Activities		Company of Others	
Education	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-8 years	25	17.12*	(20.8)	2.60*	(1.6)	2.88*	(1.5)
9-12 years	87	25.37*	(17.0)	4.23*	(1.9)	3.10	(1.3)
13-16 years	128	32.12*	(19.4)	5.27	(2.2)	3.45	(0.9)
17 or more	106	41.44*	(18.6)	5.90	(2.2)	3.48	(0.9)

ANOVA of Weekly Hours of Activity by Education Level [F(3,345)=17.97, p=.0001]
ANOVA of Number of Activities by Years of Education Level [F(3,345)=22.05, p=.0001]
ANOVA of Hours in the Company of Others by Education Level [F(3,345)=4.00, p=.0080]

* p<.05

Finally, those with the least education reported significantly less time spent in the company of others [F(3,345)=4.00, $p<.05$].

8.5 Verbal Fluency and other indicators by baseline score level

The range of Verbal Fluency scores for the total sample was 5 to 72, and as described above, there was great heterogeneity both within and between age and education groups. A variable was therefore created to categorise Verbal Fluency by baseline score level in order to define more precisely the relationship between age, education, morbidity and social activity indicators, and test performance.

When the Controlled Oral Word Association (Verbal Fluency) test was published by Benton and Hamsher in 1978, the following table was produced for interpretation of results: Benton & Hamsher, 1978 p14

Score	Interpretation
53+	Superior
45-52	High Average
31-44	Average
26-30	Low Average
24-25	Borderline
17-23	Defective
Less than 17	Very Defective

As previously reported, the Verbal Fluency test was standardised on subjects with an age range from 16 to 69 years. In the present study the age range was 60 to 93 years, but in view of the wide differences in test performance, and lacking other published criteria in the literature reviewed, the 1978 standards were adopted and modified for classification of scores and for further analyses.

In order to achieve consistency with the methods adopted for analysis of the results of the MMSE (see Table 7.3) a 3-level variable was constructed with the variable levels defined as follows: score 0-25, combining the lowest three categories; score 26-44, combining "low average" and "average"; score 45 and over, combining "high average" and "superior". Based upon these categories, twenty-seven percent of the sample (95/346) was classified as "borderline or defective", forty-six percent (159/346) as "average", and twenty-seven percent (92/346) as "superior or high average" (see Table 8.5).

Predictably, there was highly significant variation in mean Verbal Fluency score by level of baseline performance, but the groups were more homogeneous as demonstrated by the standard deviations. Mean scores were 16.92, s.d. 5.5, 35.58, s.d. 5.3, and 53.67, s.d. 7.1 for the three variable levels [F(2,345=908.82, $p<.001$]. Those with the highest mean Verbal Fluency score were significantly younger than those in the lowest performance group (71.32, s.d. 7.0, compared with 74.12, s.d. 6.7), but neither differed significantly by age from the middle range group, although variation within the model did achieve significance [F(2,345)=3.80, $p<.05$]. There was much greater variation in mean years of education by baseline Verbal Fluency level with significant differences between all three groups [F(2,345)=40.13, $p<.001$] (see Table 8.5). The lowest scoring group had a mean of 12 years of education compared with a mean of nearly 17 years for those with the best performance.

Table 8.5

MEAN VERBAL FLUENCY SCORES, AGE, YEARS OF EDUCATION, MORBIDITY AND SOCIAL ACTIVITY LEVELS BY BASELINE SCORE

Baseline Verbal		Verbal Score		Age		Years of Education	
Fluency Score	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-25	95	16.92*	(5.5)	74.12*	(6.7)	12.12*	(4.2)
26-44	159	35.58*	(5.3)	72.86	(7.1)	15.14*	(3.3)
45 and over	92	53.67*	(7.1)	71.32*	(7.0)	16.54*	(2.9)

ANOVA of Verbal Fluency Scores by Baseline Score Level [F(2,345)=908.82, p=.0001]
ANOVA of Age by Baseline Verbal Fluency Score Level [F(2,345)=3.80, p=.0232]
ANOVA of Education by Baseline Verbal Fluency Score Level [F(2,345)=40.13, p=.0001]

Baseline Verbal		Medical Diagnoses		Medicines Used		Disability Score	
Fluency Score	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-25	95	3.64	(1.5)	3.42	(2.1)	0.87*	(1.9)
26-44	159	3.23	(1.6)	3.12	(2.6)	0.26	(0.7)
45 and over	92	3.13	(1.6)	2.79	(2.3)	0.34	(0.7)

ANOVA of Medical Diagnoses by Baseline Verbal Fluency Score Level [F(2,345)=2.85, p=.0594]
ANOVA of Medicines in Daily Use by Baseline Verbal Fluency Score Level [F(2,345)=1.58, p=.2079]
ANOVA of Disability Score by Baseline Verbal Fluency Score Level [F(2,345)=8.48, p=.0003]

Baseline Verbal		Depression Score		Hours of Activity		Company of Others	
Fluency Score	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-25	95	5.83	(3.9)	27.36*	(21.4)	3.11*	(1.2)
26-44	159	5.72	(4.8)	32.79	(18.6)	3.38	(1.0)
45 and over	92	5.17	(3.9)	36.16*	(20.0)	3.48*	(1.0)

ANOVA of Depression Score by Baseline Verbal Fluency Score Level [F(2,345)=0.64, p=.5280]
ANOVA of Hours of Activity by Baseline Verbal Fluency Score Level [F(2,345)=4.76, p=.0092]
ANOVA of Hours in the Company of Others by Baseline Verbal Score Level [F(2,345)=3.11, p=.0457

* p<.05

There was no significant overall variation or between-group difference in the number of medical diagnoses, medicines in daily use, or in mean depression score by Verbal Fluency category. The ANOVA model for level of disability was significant [F(2,345)=8.48, p<.001], those with the lowest performance having a significantly higher mean score (0.87, s.d. 1.9) than the other two groups, but the range was great. Level of disability was thus the only morbidity indicator showing significant variation, but the social activity variables demonstrated both overall and between-group differences in mean scores. Those with the highest verbal performance reported significantly more weekly hours of activity than those with the poorest performance (36.16, s.d. 20.0, compared with 27.36, s.d. 21.4), [F(2,345)=4.76, p<.01]. The mean number of activities (not shown in Table 8.5) was significantly lower for those in the lowest Verbal Fluency group compared with the other two categories, and there was significant overall variation for time spent in the company of others [F(2,345)=3.11, p<.05].

The analyses demonstrated that there was a significant correlation between age, years of education, social activity variables and Verbal Fluency, but the between group variation for mean values of these variables was less than for the three levels of the baseline MMSE score variable (see Table 7.3). On the other hand, in common with the MMSE analyses, the effects of age and education were confounded by the fact that those scoring in the highest performance category were significantly younger than those with the lowest mean scores, and had the highest mean years of education. Therefore multiple linear regression modelling was the technique adopted in order to discriminate between the effects of age, education or other independent variables on Verbal Fluency performance, and the same stepwise forward selection technique used in the analyses of the MMSE data (see Chapter 7, 7.9) was employed.

8.6 Multiple regression analyses by demographic group

The mean Verbal Fluency score for the total sample was 35.27, s.d. 14.75. The stepwise forward selection procedure defined 3 variables for inclusion in the multiple linear regression model. In hierarchical order, these were education, number of activities reported, and disability score, but age did not meet the selection criteria. The model was highly significant, $R^2=0.27$ [$F(4,342)=30.76$, $p<.001$] (see Table 8.6). When beta coefficients were computed, education had the highest predictive value when controlling for others in the model, (beta=0.389 $p<.001$) followed by number of activities (beta=0.182, $p<.001$), while disability score proved to be a weaker negative predictor of higher Verbal Fluency score (beta=0.118, $p<.05$). In the stepwise procedure, education uniquely explained 22 percent of the variance, number of activities 3 percent and disability score 2 percent.

In view of the gender differences in mean years of education (see Tables 6.6 and 6.8) separate models were again constructed for men and women. For the male sample ($n=186$) the mean Verbal Fluency score was 37.48, s.d. 14.23, and the stepwise procedure selected 3 variables, which were in rank order: education, number of activities, and age. These variables again produced a highly significant multiple regression model, $R^2=0.22$, [$F(3,182)=16.40$, $p<.001$] (see Table 8.7). Beta coefficients showed that education had the highest predictive power (beta=0.29,

Table 8.6

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE VERBAL FLUENCY SCORES
Total Sample (N=346)

VARIABLE	Standardised Beta Weights	p value
Education	0.389	0.0001**
Disability score	−0.118	0.0181*
Number of Activities	0.182	0.0009*

Mean score=35.27
Model: R−square=0.27, F(3,342)=30.76, p=.0001

Table 8.7

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE VERBAL FLUENCY SCORES
Men (N=186)

VARIABLE	Standardised Beta Weights	p value
Age	−0.154	0.0224*
Education	0.290	0.0001**
Number of Activities	0.195	0.0099*

Mean score=37.48
Model: R−square=0.22, F(3,182)=16.40, p=.0001

**p<.001
* p<.05

$p < .001$), followed by the number of activities reported (beta=0.195, $p < .01$), while age was a less powerful negative predictor of test scores (beta=-0.154, $p < .05$). The model had a lower explanatory power for the variance of Verbal Fluency scores among men than for the total sample, but education uniquely accounted for 15 percent of the variance, the number of activities 4 percent, and age 2 percent.

The mean Verbal Fluency score for women (n=160) was lower, 32.69, s.d. 14.98, and the variables selected by the stepwise procedure for inclusion in the multiple regression model were education, disability score, and number of activities, but age did not meet the selection criteria. The model was again highly significant, $R^2=0.31$, [$F(3,156)=21.36$, $p < .001$]. Beta coefficients indicated that education had the highest predictive power when controlling for other variables in the model (beta=0.417, $p < .001$), disability level was a weaker negative predictor of performance (beta=-0.144, $p < .05$) and number of activities was a positive predictor (beta=0.142) although this variable just failed to achieve significance (see Table 8.8). Education explained 24 percent of the variance in female performance, disability 3 percent, and number of activities 2 percent.

In view of the 33 year age range of the sample (60-93) it was again dichotomised and separate analyses were undertaken for the sample aged 60 to 74 (n=203) and those aged 75 and over (n=143). The mean Verbal Fluency score for the younger sample was 37.89, s.d. 14.57. A significant 3-variable model was again selected by the stepwise regression procedure, $R^2=0.23$, [$F(3,199)=20.01$, $p < .001$] (see Table 8.9). Education was the most significant and powerful predictor of Verbal Fluency score, (beta=0.361,

Table 8.8

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE VERBAL FLUENCY SCORES

Women (N=160)

VARIABLE	Standardised Beta Weights	p value
Education	0.417	0.0001**
Disability score	−0.144	0.0392*
Number of Activities	0.142	0.0605

Mean score=32.69
Model: R−square=0.31, F(3,156)=21.36, p=.0001

Table 8.9

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE VERBAL FLUENCY SCORES

Ages 60−74 (N=203)

VARIABLE	Standardised Beta Weights	p value
Education	0.361	0.0001**
Number of Activities	0.218	0.0017**
Company of Others	0.135	0.0350

Mean score=37.89
Model: R−square=0.23, F(3,199)=20.01, p=.0001

**p<.001
* p<.05

$p < .001$), followed by two "social integration" variables, i.e. number of activities reported, (beta=0.218, $p < .01$) and time spent in the company of others, (beta=0.135, $p < .05$). Education explained 18 percent of the variance, the number of activities 3 percent, and the company of others only 2 percent.

For those aged 75 and over, the mean Verbal Fluency score was 31.54, and the significant independent variables selected by the stepwise procedure were education, disability score, and time spent in the company of others (see Table 8.10). Once again the model was highly significant, $R^2=0.33$, [$F(3,139)=22.71$, $p < .001$] and beta coefficients showed education to be the most powerful independent predictor of higher test performance (beta=0.418, $p < .001$). Disability score was a negative predictor (beta=0.246, $p < .001$) and time spent with others was a less powerful but significant positive predictor (beta=0.210, $p < .01$). Education accounted for 23 percent of the variance, disability score 6 percent, and time spent in the company of others 4 percent.

8.7 Adjustment of scores to control for the effects of education

When the Verbal Fluency test was first standardised it was found that "both educational background and age are important determinants of performance level", but as noted above, criteria for adjustment of scores to control for these effects did not extend beyond age 69.$^{\text{Benton \& Hamsher, 1978 p12-14}}$ The Iowa Screening Battery for Mental Decline, published by the same research team, and standardised on both "normal" and "dementia" subjects aged 61 to 86 years, provided the following "correction factors".$^{\text{Eslinger et al, 1984 p15}}$

Table 8.10

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE VERBAL FLUENCY SCORES

Ages 75 and Over (N=143)

VARIABLE	Standardised Beta Weights	p value
Education	0.418	0.0001**
Disability score	−0.246	0.0006**
Company of Others	0.210	0.0037*

Mean score=31.54
Model: R−square=0.33, F(3,139)=22.71, p=.0001

Table 8.11

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE VERBAL FLUENCY SCORES

Excluding those with Limited Education (N=321)

VARIABLE	Standardised Beta Weights	p value
Education	0.303	0.0001**
Disability score	−0.110	0.0387*
Number of Activities	0.173	0.0020*

Mean score=36.53
Model: R−square=0.20, F(4,316)=18.91, p=.0001

**p<.001
* p<.05

Years of Education	Adjustment Points Men	Women
0-8 years	11	9
9-11 years	6	6
12-15 years	4	3
16 or over	0	0

The adjusted mean test scores, or "mean performance levels" reported from the Eslinger assessment procedures were 37.9, s.d. 10.9 for the "normal sample", and 20.1, s.d. 11.6 for the "dementia" group.[Eslinger et al, 1984 p11, Eslinger et al, 1985 p672] It is of note that according to these criteria, men receive a slightly higher education adjustment score than women, but no correction factor was developed for either sex for advancing age. In the present study, however, men had significantly higher levels of education and also higher mean Verbal Fluency performance scores. Therefore, in the light of the characteristics of the sample under discussion, two separate analyses were undertaken to adjust for the effects of education.

In the first analysis, a multiple regression model was constructed excluding all those with 0-8 years of education, a group consisting predominantly of older women. In the second analysis, all test scores were adjusted, using the correction factors shown above, and analysis of variance and multiple regression techniques were used to assess the extent to which the effect of education was reduced as a determinant of test scores.

When those with little formal education were excluded from the sample (n=25) a significant 3-variable model was again selected by the stepwise procedure, $R^2=0.19$ [$F(3,317)=23.91$, $p<.001$] (see Table 8.11). Education continued to be the most significant and powerful predictor of Verbal Fluency score, ($beta=0.303$), followed by number of activities ($beta=0.173$, $p<.01$) and disability score ($beta=-0.110$, $p<.05$). As would be predicted, education explained less of the variance than in the previous models, but it uniquely accounted for 14 percent, compared with 3 percent for number of activities and 2 percent for disability score.

In the second analysis, Verbal Fluency scores were adjusted according to the criteria published in the Iowa Screening Battery for Mental Decline. This raised the mean score for the total sample from 35.27, s.d. 14.75 to 37.68, s.d 13.75. ANOVAs were then computed for adjusted mean Verbal Fluency score, age, and education by the three level baseline score variable.

A comparison of results from analysis of the unadjusted scores shown in Table 8.5 with those in Table 8.12 shows that the mean Verbal Fluency score for each variable level was raised by addition of the education correction factors. As might be predicted, the increase was greatest for those in the lowest baseline category, (unadjusted score=16.92, s.d. 5.5, adjusted score=19.46, s.d. 4.2), raising the mean to approximately the level of the standardised adjusted score for "dementia cases", i.e. 20.[Eslinger et al, 1984 p11] At the same time there was a small change in the frequency distribution of the sample between the three variable levels, the proportion in the lowest scoring group being reduced from 27 to 23 percent (78/346), and the proportion in the highest score category increasing to 30 percent (103/346). There was little change in the mean age of the three variable levels, those in the highest performance group continuing to be significantly younger than those in the lowest [$F(2,345)=3.23$, $p<.05$].

Table 8.12

MEAN VERBAL FLUENCY SCORES ADJUSTED FOR YEARS OF EDUCATION,** BY AGE, EDUCATION AND BASELINE SCORE LEVEL

Baseline Verbal		Verbal Score		Age		Years of Education	
Fluency Score	N=346	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0–25	78	19.46*	(4.2)	74.17*	(6.8)	12.97*	(3.9)
26–44	165	36.08*	(5.5)	72.93	(7.1)	14.59*	(4.0)
45 and over	103	54.03*	(7.1)	71.53*	(6.9)	16.12*	(3.1)

ANOVA of Adjusted Verbal Fluency Scores by Baseline Score Level [F(2,345)=809.43, p=.0001]
ANOVA of Age by Baseline Verbal Fluency Score Level [F(2,345)=3.23, p=.0407]
ANOVA of Education by Baseline Verbal Fluency Score Level [F(2,345)=16.03, p=.0001]

Table 8.13

REGRESSION MODEL FOR BASELINE VERBAL FLUENCY SCORE ADJUSTED FOR YEARS OF EDUCATION**

Total Sample (N=346)

VARIABLE	Standardised Beta Weights	p value
Education	0.2349	0.0001**
Disability score	–0.1180	0.0278*
Depression score	–0.1191	0.0238*
Number of Activities	0.1885	0.0014*

Total adjusted mean score=37.68
Model: R–square=0.15, F(4,341)=15.09, p=.0001

**p<.001
* p<.05

** Adjustment according to Eslinger et al., 1984

Furthermore, highly significant overall and within-group differences in mean years of education remained between the three variable categories [$F(2,345)=16.03$, $p<.001$].

The stepwise forward selection procedure for adjusted Verbal Fluency scores, resulted in a 4-variable multiple regression model, $R^2=0.15$, [$F(4,341)=15.09$, $p<.001$]. In hierarchical order the significant variables were: education, number of activities, depressive symptoms and disability score. Education continued to be the most powerful and significant predictor of adjusted Verbal Fluency score (beta=0.235, $p<.001$), followed by number of activities (beta=0.189, $p<.01$). Both disability (beta=-0.118, $p<.05$) and depression (beta=-0.119, $p<.05$) were weaker negative indicators (see Table 8.13). Education uniquely explained 10 percent of the variance, number of activities 3 percent, and depression and disability score just 1 percent each. Thus although the predictive power of education was reduced by the adjustments developed by Eslinger et al., its effect was not eliminated.

8.8 The sensitivity of Verbal Fluency scores to repetitive testing

In reviewing the publications by the group who developed and standardised the Verbal Fluency test, and later incorporated it into the Iowa Screening Battery for Mental Decline, the reliability of performance scores and their sensitivity to change was not addressed.[Benton & Hamsher, 1978, Eslinger et al, 1984] However, as previously described, the longitudinal design of the present investigation made possible an assessment of the sensitivity of all study instrument scores to the effects of repetitive testing.

The reliability of Verbal Fluency test scores as determined by a single assessment was evaluated by analysing the performance of those tested three times in the perioperative period, i.e. before surgery, 1-week post-surgery, and 6-weeks after surgery. Table 8.14 compares the characteristics of the subjects who underwent surgery and participated in all three interviews (n=241) compared with "others" in the sample (n=105), i.e. those who were deemed "ineligible", "dropouts" and "controls" (see Chapter 7, Footnote 8). Those who participated in the sequence of three assessments tended to be younger, but not significantly so. The only significant difference between the two groups was a slightly higher mean disability score ($p<.05$) for those who underwent surgery and were therefore in the repetitive testing group.

Repeated Measures Analysis of Variance was the statistical technique used to compute within and between-group changes in mean values for Verbal Fluency scores at each of assessment interviews. Tukey's studentised range tests were again computed to test for the significance of the difference in mean values between categories of the dependent variable.

8.9 Change in Verbal Fluency scores with repetitive testing

The analysis of variance model for Verbal Fluency values by age group for subjects undergoing surgery (n=241) showed that at time 1 those aged 75 to 84 had significantly lower mean scores than other age categories [$F(3,237)=5.17$, $p<.01$] (see Table 8.15). This was consistent with the distribution of mean values for the total sample (n=346) shown in Table 8.3. The ANOVA for the effects of time, i.e. repeated testing on three occasions, showed a highly significant within-group difference [$F(2,474)=8.71$, $p<.001$]. Those aged 85 and over had a decline in mean Verbal Fluency scores at time 2, but at time 3 there was a 3-point gain to a level higher than at time 1, a pattern that

Table 8.14

MEAN VALUES OF SELECTED VARIABLES FOR THOSE COMPLETING PERIOPERATIVE INTERVIEWS COMPARED WITH OTHERS IN TOTAL SAMPLE

	Perioperative Interviews Completed (n=241)		Others (n=105)		Student's t-test
	Mean	S.D.	Mean	(S.D.)	p value
Baseline Verbal Fluency	36.05	(15.0)	34.50	(14.2)	0.3291
Age	72.31	(7.1)	73.90	(6.6)	0.0511
Education (years)	14.90	(3.7)	14.17	(4.1)	0.1203
Medical Diagnoses	3.32	(1.5)	3.30	(1.8)	0.9408
Medicines Used	3.04	(2.3)	3.29	(2.6)	0.4106
ADL Disability Score	0.55	(1.3)	0.22	(1.0)	0.0120*
GDS Depression Score	5.63	(4.5)	5.54	(4.0)	0.8504
Weekly Hrs of Activity	32.33	(20.5)	31.89	(19.0)	0.8462
Number of Activities	5.01	(2.4)	4.99	(1.9)	0.9280
Company of Others	3.34	(1.1)	3.30	(1.1)	0.7045

* Significant, p<.05

was consistent with MMSE performance for this age group (see Figure 7.2). For all other age groups, mean test scores increased consistently at each assessment period, but the greatest increase (4.8 points) was for those with a significantly lower mean score at time 1, i.e. those aged 75-84. By time 3, however, there was no significant difference in mean values between the four age groups. ANOVA did not demonstrate any significant interaction effect between time and age categories, i.e. the extent to which any one group experienced change did not differ from that of any other (see Table 8.15).

As would be predicted from previous analyses, ANOVA of Verbal Fluency score by education group at time 1, showed significant variation [F(3,237)=22.34, p<.001] (see Table 8.16). Those with the least education, in the two categories 0-8 and 9-12 years, had significantly lower mean performance scores than the higher groups. However, ANOVA for within-group changes over time were also highly significant [F(2,474)=12.54, p<.001]. Test scores improved for all groups at each assessment and there was no significant time by education group interaction.

Repeated measures ANOVAs were also computed for the 3-level baseline Verbal Fluency score variable. Once again, reflecting its structure, there was highly significant variation in mean scores at time 1, [F(2,237)=642.36, p<.001], and significant differences in mean values continued between all three variable categories at each assessment interview (see Table 8.17). On the other hand, ANOVA also demonstrated significant within-group changes over time [F(2,476)=19.88, p<.001]. At time 3, all groups had higher test scores than at time 1. However, the change was greatest (4.17 points) for those in the lowest performance group, compared with less than 2 points for

Table 8.15

MEAN VERBAL FLUENCY SCORES IN THE PERI–OPERATIVE PERIOD BY AGE

		Time 1		Time 2		Time 3	
Age	**N=241**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**
60–64	40	39.88	(12.5)	41.10	(12.6)	43.25	(13.1)
65–74	107	38.31	(15.0)	39.21	(14.8)	40.68	(15.6)
75–84	83	31.00*	(15.2)	33.92	(15.9)	35.82	(17.4)
85 & over	11	38.18	(13.0)	36.36	(17.5)	39.36	(14.1)

ANOVA of Verbal Fluency Score at Time 1 by Age Group [F(3,237)=5.17, p=.0018]
ANOVA of within–group change over Time [F(2,474)=8.71, p=.0002]
ANOVA of interaction between Time and Age Group [F(6,474)=1.25, p=.2785]

Table 8.16

MEAN VERBAL FLUENCY SCORES IN THE PERI–OPERATIVE PERIOD BY EDUCATION

		Time 1		Time 2		Time 3	
Education	**N=241**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**
0–8 years	16	21.75*	(12.8)	24.00*	(12.4)	25.00*	(12.0)
9–12 years	56	27.05*	(13.3)	27.23*	(13.7)	29.04*	(13.8)
13–16 years	90	38.26	(13.1)	40.44	(13.2)	42.46	(13.8)
17 or more	79	42.80	(13.7)	44.38	(13.3)	46.10	(15.1)

ANOVA of Verbal Fluency Score at Time 1 by Education Group [F(3,237)=22.34, p=.0001]
ANOVA of within–group change over Time [F(2,474)=12.54, p=.0001]
ANOVA of interaction between Time and Education Group [F(6,474)=0.60, p=.7317]

Table 8.17

MEAN VERBAL FLUENCY SCORES IN THE PERI–OPERATIVE PERIOD BY BASELINE SCORE

		Time 1		Time 2		Time 3	
Baseline Verbal Fluency Score	**N=241**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**	**Mean**	**(S.D.)**
0–25	63	17.16*	(5.4)	20.00*	(7.4)	21.33*	(7.8)
26–44	107	35.28*	(5.0)	37.51*	(8.9)	39.24*	(8.6)
45 and over	71	53.96*	(7.5)	53.25*	(9.5)	55.58*	(12.2)

ANOVA of Verbal Fluency Score at Time 1 by Baseline Score [F(2,238)=642.36, p=.0001]
ANOVA of within–group change over Time [F(2,476)=19.88, p=.0001]
ANOVA of interaction between Time and Baseline Score [F(4,476)=2.30, p=.0578]

those in the highest scoring group, a change that just failed to achieve significance for time by group interaction [F(4,476)=2.30, p=.06].

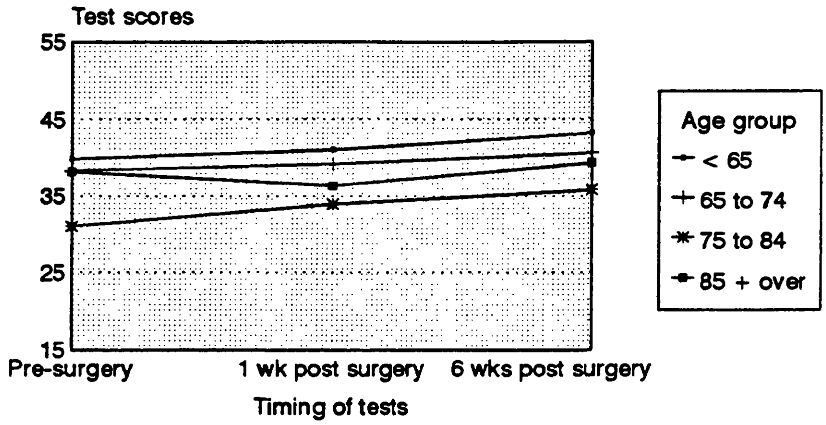
The pattern of change in mean Verbal Fluency values by age group, educational category and baseline score level is shown graphically in Figures 8.2, 8.3 and 8.4. The figures illustrate the greater variation in mean test scores by education than by age. They also show that at time 3 all groups had a slightly higher mean score than at time 1. Figure 8.2 demonstrates that those in the age group 75-84, who also had the lowest mean value for years of education, and the lowest score at time 1 (see Table 8.3), had a slightly higher and more consistent increase over time. Only those aged 85 and over showed a small decline at time 2, which rebounded at time 3. In evaluating the pattern of change in Verbal Fluency performance it should be noted both that the assessment at time 2 was within a week of surgery, and also that the test was administered with a different set of letters at the second assessment (P R W), than at times 1 and 3 (C F L) (see Figure 8.1).

8.10 Evaluating the significance and correlates of change

In view of the significant changes in mean Verbal Fluency scores identified by the repeated measures ANOVAs, further analyses were undertaken in an attempt to determine if improvement or decline in performance was associated with demographic factors and/or change in morbidity indicators. Frequency distributions were computed for Verbal Fluency scores in which any change was observed, and these demonstrated that no less than 68 percent of study subjects (163/241) showed some improvement in test scores between time 1 and time 3, 27 percent (66/241) had lower scores, and

Figure 8.2

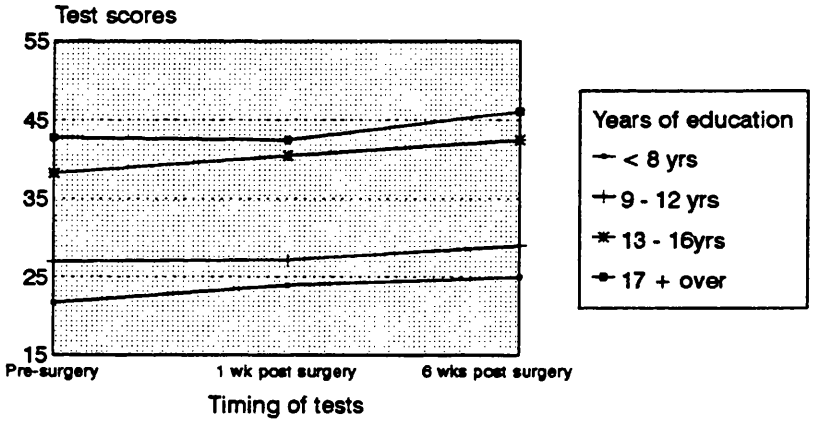
Change in Perioperative Mean Verbal Fluency Scores by Age Group



ANOVA of effects of time, p<.001
ANOVA of interaction between time and age group, p=.278

Figure 8.3

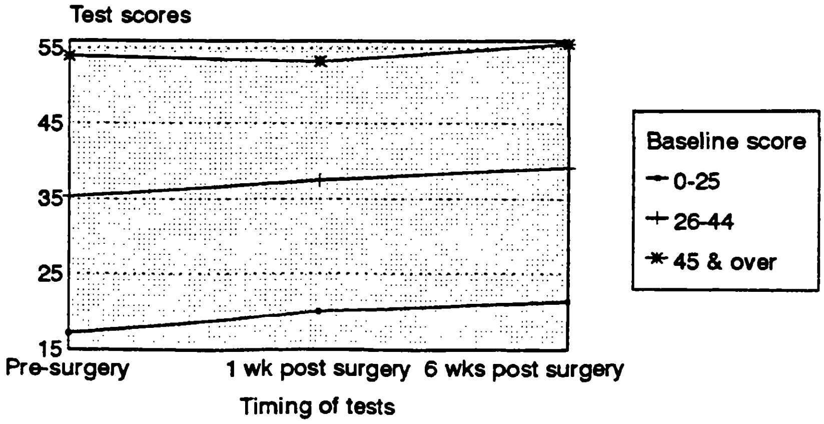
Change in Perioperative Mean Verbal Fluency Scores, by Level of Education



ANOVA of effects of time, p=.0001
ANOVA of interaction between time and group, p=0.7317

Figure 8.4

Change in Perioperative Mean Verbal Fluency Score By Baseline Level



ANOVA of effects of time, p=.0001
ANOVA of interaction between time and group, p=.0576

performance was stable for only 5 percent (12/241) (not shown in Tables). At the same time, repeated measures ANOVAs failed to identify any significant difference by age, education, medications, disability, or depressive symptoms between those with improvement, decline, or stability in test scores between time 1 and time 3.

As discussed in Chapter 7, a 2-point decrease in MMSE scores was defined as meeting the "clinical criterion for cognitive decline".[Mondimore et al, 1983 p931] No similar norms were found in the literature reviewed for determining the significance of any magnitude of change in Verbal Fluency. However, in view of the fact that rather than a 2-point decline, a 3/30, i.e. ten percent or greater change from maximum potential MMSE score was associated with significant variations in morbidity indicators between time 1 and time 3 (see Table 7.18), the same criterion was adopted to compare the characteristics of subjects with change in Verbal Fluency scores. As noted previously (see section 8.5), the maximum score achieved at time 1 by any subject in this open-ended test was 72 points. A three-level variable was therefore constructed, based on a 7-point or 10 percent change, either positive or negative, from the highest baseline score.

Thirty-two percent of subjects (77/241) improved by 7 or more points, compared with 10 percent (23/241) who declined, and 58 percent (141/241) whose scores remained within a 6-point range over time (see Table 8.18). There was significant overall variation in mean test scores at time 1 between subjects with negative or positive change and those with stable performance F(238)=7.91, p<.001]. The mean Verbal Fluency score was significantly greater 47.30 (s.d. 14.1), at time 1 for subjects whose scores decreased compared with the other groups (see Table 8.18), and it is of note that the

Table 8.18

COMPARISON OF GROUPS WITH AND WITHOUT A SEVEN–POINT or GREATER CHANGE IN VERBAL FLUENCY SCORES IN THE PERIOPERATIVE PERIOD

Level of Change	N=241	Verbal Score 1 Mean	(S.D.)	Verbal Score 3 Mean	(S.D.)
No change	141	34.30	(14.8)	35.50	(14.7)
Decline	23	47.30*	(14.1)	36.91	(13.9)
Improvement	77	35.88	(14.3)	47.60*	(15.7)

ANOVA of Verbal Fluency at Time 1 by Level of Change [F(2,238)=7.91, p=.0005]
ANOVA of within–group change over time [F(1,238)=190.02, p=.0001]
ANOVA of interaction between Time and Level of Change [F(2,238)=342.38, p=.0001]

Level of Change	N=241	Age Mean	(S.D.)	Years of Education Mean	(S.D.)
No change	141	72.19	(6.8)	14.62	(3.8)
Decline	23	72.30	(7.5)	15.09	(3.5)
Improvement	77	72.52	(7.6)	15.38	(3.7)

ANOVA of Mean Age by Level of Change [F(2,238)=0.05, p=.9493]
ANOVA of Mean Years of Education by Level of Change [F(2,238)=1.06, p=.3466]

Level of Change	N=241	Medicine Use 1 Mean	(S.D.)	Medicine Use 3 Mean	(S.D.)
No change	141	3.15	(2.3)	3.30	(2.4)
Decline	23	2.70	(2.6)	3.00	(2.3)
Improvement	77	2.95	(2.3)	2.94	(2.2)

ANOVA of Medication Use at Time 1 by Level of Change [F(2,238)=0.46, p=.6300]
ANOVA of within–group change over time [F(1,238)=1.50, p=.2220]
ANOVA of interaction between Time and Level of Change [F(2,238)=0.54, p=.5862]

mean score at time 1 for those with a 7-point fall was in the "high average" range, declining to "average", and vice versa for the group with positive change between the first and the third assessments.

The mean age of all three groups was nearly identical, mean years of education ranged only from 14.62 to 15.38 years for the three groups, and there was no variation or significant between-group difference over time in medications used or in level of disability. There was a marginal increase in level of depression compared with a small decrease for other groups, and the variation was reflected in a significant within group change [$F(2,238)=8.21$, $p<.005$] but no between-group interaction (see Table 8.18 contd.).

The data analyses indicate that the 7-point or greater change in mean Verbal Fluency scores, observed in more than forty percent of the sample, occurred independently of age, education, or change in morbidity indicators, with the exception of a small increase in depressive symptoms for subjects with reduced scores. It may therefore be argued that improvement was attributable to a "learning" or "practice" effect, with lower scores indicative of "cognitive decline". In examining the demographic characteristics of those in different categories of change, however, it was noted that the age range for those with a 7-point or greater improvement was from 60 to 91, and of decline from 60 to 93 years, while the education range for those with increased scores at time 3 was from 4 to 22 years, and for those with lower performance, from 8 to 21 years. At the same time, sixty-five percent (15/23) of those with a 7-point decrease were in the highest performance category (score 45 and over) at time 1, and the scores of more than 50

Table 8.18 (contd.)

COMPARISON OF GROUPS WITH AND WITHOUT A SEVEN–POINT or GREATER CHANGE IN VERBAL FLUENCY SCORES IN THE PERIOPERATIVE PERIOD

Level of Change	N=241	Disability Score 1 Mean	(S.D.)	Disability Score 3 Mean	(S.D.)
No change	141	0.67	(1.5)	0.63	(1.2)
Decline	23	0.35	(0.8)	0.61	(1.5)
Improvement	77	0.39	(0.9)	0.31	(0.7)

ANOVA of Disability Score at Time 1 by Level of Change [F(2,238=1.48, p=.2295]
ANOVA of within–group change over time [F(1,238)=0.10, p=.7521]
ANOVA of interaction between Time and Level of Change [F(2,238)=1.02, p=.3637]

Level of Change	N=241	Depression Score 1 Mean	(S.D.)	Depression Score 3 Mean	(S.D.)
No change	141	5.52	(4.3)	4.87	(4.5)
Decline	23	6.87	(5.5)	7.04	(6.1)
Improvement	77	5.48	(4.5)	4.94	(4.2)

ANOVA of Depression Score 1 at Time by Level of Change [F(2,238)=0.96, p=.3827]
ANOVA of within–group change over time [F(1,238)=8.21, p=.0045]
ANOVA of interaction between Time and Level of Change [F(2,238)=0.79, p=.4553]

percent remained in the same category after the "decline".

It may therefore be proposed that it is difficult to interpret the significance of the decline in test scores for the majority of those in whom it was observed. At the same time, more than thirty percent of the sample showed a 7-point or greater improvement in Verbal Fluency, and the change occurred in approximately one third of individuals in each baseline score category. On the other hand, a 7-point or greater increase represents, at minimum, a twenty-eight percent (7/25) gain for those with the lowest baseline score of 0-25 at time 1. The magnitude of improvement was therefore far greater for the twenty-nine percent of subjects (18/63) in the "defective and borderline" category (see Table 8.17), a few of whom raised their scores by seventy to more than one hundred percent between time 1 and 3, i.e. from 22 to 38 and from 20 to 44 points.

8.11 Summarising the results of Verbal Fluency testing

There was a closer association between education and Verbal Fluency scores than between age and test performance (see Tables 8.3 and 8.4), a finding that was congruent with the results of the MMSE analyses (Tables 7.1 and 7.2). Similarly, there was a negative correlation in the word testing data between level of education and morbidity, a positive correlation between morbidity indicators and advancing age, and vice versa for social activity variables. At the same time, there was considerable heterogeneity in Verbal Fluency scores in all age and education categories.

In the interests of consistency in the analysis of data from the different study assessment instruments, a three-level variable was constructed to categorise Verbal Fluency performance by baseline score level, based upon values standardised by developers of the instrument for the interpretation of results.[Benton & Hamsher, 1978 p14] Frequency distributions showed that twenty-seven percent of the sample (95/346) scored in the lowest category which combined "defective and borderline" performance, and another twenty-seven percent (92/346) in the highest category combining "high average and superior". Analyses of variance were computed for demographic, morbidity, and social activity variables by baseline score level and there were again significant differences by age, education, disability, social activity, and time spent in the company of others, for different categories of the dependent variable (see Table 8.5). On the other hand, the variation among mean values of independent variables was less than for the different levels of the similarly constructed MMSE variable (see Table 7.3). For example education ranged from a mean of 12 to 16.5, a spread of 4.5 years for the three levels of baseline Verbal Fluency, compared with 9 to 15.5, or 6.5 years for the MMSE baseline score variable.

It may be argued that intuitively, a strong association would not necessarily be assumed between level of education and "verbosity", "loquaciousness" or even "eloquence", and that intuition was confirmed by the heterogeneity of Verbal Fluency scores within education groups, and also by statistical evidence that the lowest baseline score category was not restricted to individuals with limited formal education. On the other hand, when multiple regression models were constructed, education was selected as the most significant predictor of performance for the total sample, men, women, those aged 60 to 74 and those aged 75 and over. Furthermore even when the least educated, i.e. with 0-8 years of education, were excluded from the analyses, education continued to be the primary predictor of test scores (see Tables 8.6 to 8.11).

When Benton and Hamsher first standardised the Verbal Fluency test, it was observed that "educational background" was an important determinant of test performance, and adjustment scores were proposed.[Benton & Hamsher p12] When the same team of neuro-psychologists selected a battery of tests for the detection of "abnormal mental decline" in the elderly and for discriminating between a "normal" and a "dementia" sample, education adjustments were again standardised.[Eslinger et al, 1984 p11] When Verbal Fluency performance was adjusted according to these criteria, the mean score of subjects in the lowest category of the present sample was raised only to the norm for "dementia", i.e. 20 points, with a small number, four percent moving into the "low average" level. At the same time, education continued to be the most powerful predictor of test results (see Table 8.12) and its effects were therefore reduced but not eliminated by the Eslinger adjustments.

In the regression analyses of both the Verbal Fluency test and the MMSE, education was the most significant predictor of test scores and uniquely explained a similar proportion of the variance in performance for the total sample and for subsets within it (see 7.9 and 8.6). In the MMSE test data, however, physical disability was the second most powerful independent predictor of instrument scores (see Tables 7.4 to 7.11). In contrast, the most powerful predictor of Verbal Fluency performance after education, for the total sample and most subgroups, was the number of activities reported, with time spent in the company of others also exerting an independent, positive effect, notably among those aged 75 and over, both being variables which reflect the level of social interaction.

The Verbal Fluency test is not dependent on visual acuity or motor skills for reading, writing or drawing, and it may be argued that the very limited power of physical disability in predicting scores for the verbal assessment instrument tends to confirm its independent negative effect on MMSE performance.[1] On the other hand, the total R^2 correlation coefficients were mostly lower for the Verbal Fluency models than for the MMSE, indicating that the "social interaction" variables selected for inclusion in the regression models had weaker explanatory power for the variance in test scores. The test is timed, however, and requires that individuals generate as many words as possible in one minute. It may therefore be proposed that concepts such as personality, mood,[2] motivation and self-confidence, may account for a high proportion of the variance in a test, the score for which depends upon speed of performance, and that the social activity variables selected for inclusion in the regression models, with their relatively weak independent predictive power, were indicators of more complex interaction effects not captured in the data.

It may further be proposed that this contention is supported by the highly significant improvement in Verbal Fluency scores for a high proportion of the sample when the test

---

[1] The only group for whom physical disability explained more than 5 percent of the variance in Verbal Fluency scores was subjects aged 75 and over. As noted in discussing Table 7.10 this group included five individuals with physical disability related to confirmed diagnoses of dementing disease.

[2] Mood as reflected in number of depressive symptoms was a relatively weak negative predictor of Verbal Fluency when scores were adjusted for education using the standards developed by Eslinger et al. (see Table 8.13).

was repeated over time. Repeated measures analyses of variance showed a significant increase between time 1 and time 3 for all age groups, for every education category, and for all baseline Verbal Fluency performance levels, the greatest improvement being for those with the lowest score at time 1 (see Tables 8.15 to 8.17). The only group for which there was a small decline in test performance at time 2 was the age group 85 and over, one which recovered to a slightly higher level by time 3 (see Figure 8.2). This finding was consistent with the analyses of the MMSE data (see Figure 7.2) and may indicate that performance in the tests was negatively, if transiently, affected by surgery in the very elderly.

In order to determine the demographic and/or morbidity correlates of improvement or decline in Verbal Fluency scores, a variable was constructed to compare groups experiencing a 7-point or greater change in performance over time, based upon ten percent of maximum achieved score, an indicator associated with significant variation in demographic and morbidity factors in the MMSE analyses (see Tables 7.18). Some thirty percent of subjects in all baseline score categories improved by 7 or more points, the magnitude of change being greater for those with the lowest initial scores in the range 0-25. At the same time, ten percent of subjects, predominantly in the highest scoring category showed a 7-point "decline" between time 1 and time 3.

In contrast with the MMSE, the analyses indicated that change in Verbal Fluency scores of this degree, i.e. ten percent or more, occurred independently of age, education, or variation in morbidity or physical disability. It may therefore be argued that a marked "practice" or "learning" effect was observed for testing with this instrument. It may further be proposed that the lack of association between independent variables and change in test scores, either positive or negative, among individuals of all educational levels, and the relatively weak explanatory power for variation in Verbal Fluency performance of variables other than education in the regression models, indicate that differences were determined in part by factors, unquantified in the data, such as personality, and motivation. At the same time, the opportunity was presented by the study to compare the results of the MMSE and the Verbal Fluency testing, against those of another widely employed neuropsychology instrument.

8.12 The WAIS Similarities test and its use in the present study

The Similarities test is one of the Verbal Scale Subtests from the eleven item Wechsler Adult Intelligence Scale (WAIS). Wechsler proposed that the Similarities test was found to be "among the best of our entire battery" being "recognised by all investigators as containing a great amount of 'g'". This he defined as a "psycho-mathematical quantity" comparable with "some of the constants met with in physical or atomic measurements".[Wechsler, 1939p8] He further noted that "correlations for the Similarities test .... are uniformly among the highest obtained either with the individual tests or with total (WAIS) score".[Wechsler, 1939 p87-9]

More recently, Lezak, in her compendium of tests and assessment techniques, described the Similarities test as one of "verbal concept formation" and "an excellent test of general intellectual ability". She further proposed that of all the WAIS verbal subtests "it is least affected by the subject's background and experiences" and reiterated that it is "relatively independent of social or educational background". Lezak, 1983 p265

The MMSE and the Verbal Fluency test were administered exactly as defined by those who developed them. However, in view of the fact that subjects were requested to respond to a morbidity and social activity questionnaire, and submit to mental status assessment with each of the test instruments at every interview, the Similarities test was modified and abbreviated for the study under discussion in order to maintain the interest and cooperation of participants. The 12 items of the test were separated into two comparable groups of 6 items (see Figure 8.5), and the groups of items were asked at alternate interviews.[3]

Scoring criteria for the Similarities test were defined as:

2 credits (points) "for the most fundamental likeness".
1 credit "for any essential likeness".
0 credit "for failure to respond, giving a difference, or mentioning some inconsequential similarity." Wechsler, 1939 p174

In contrast with the MMSE and the Verbal Fluency test, it may be argued that scoring for the Similarities test is potentially subject to variation according to the reasoning of the person administering the test. For example, samples of responses and appropriate scores given by Wechsler are as follows:

---

[3] Two different sets of letters were administered at alternate interviews for the Verbal Fluency test (see above).

Figure 8.5

Similarities Test

Tell me in what way the following things are similar:

Group 1*

		Response	Score
a)	Orange and Banana	________________	______
b)	Axe and Saw	________________	______
c)	North and West	________________	______
d)	Air and Water	________________	______
e)	Egg and Seed	________________	______
f)	Wood and Alcohol	________________	______

Total Score: ________

Group 2*

a)	Coat and Dress	________________	______
b)	Dog and Lion	________________	______
c)	Eye and Ear	________________	______
d)	Table and Chair	________________	______
e)	Poem and Statue	________________	______
f)	Praise and Punishment	________________	______

Total Score: ________

* Group 1 items were asked at Time 1 and Time 3.
Group 2 items were asked at Time 2.

EGG - SEED

Score 2	Both embryonic substances Both beginnings of life Both reproduce what they come from Young come from both
Score 1	Both give life (incompletely explained) Means of reproduction (not further explained) Both can create Both have shells Something grows from both
Score 0	You can eat both Both have a yolk Both are round Wechsler, 1939 p177

8.13 Similarities test scores by age, education and baseline score

The Similarities test was administered to the same individuals (n-1) who participated in Verbal Fluency testing. Similarities test scores ranged from 0 to the maximum possible score of 12, and in common with the results of the MMSE and Verbal Fluency tests, showed much greater variation by level of education than by age group. Scores ranged from 8.56, s.d. 2.0, for those aged 60-64, to 6.13, s.d. 2.4, for the age group 85 years and over. The mean score for the youngest group was significantly higher than for the two oldest and there was significant overall variation [$F(3,341)=6.08$, $p<.001$]. On the other hand, mean scores by educational group ranged from 4.24, s.d. 2.9, for those with 0-8 years of education to 8.92, s.d. 1.8, for those with 17 or more years, and the differences were significant between all groups [$F(3,341)=54.94$, $p<.001$] (not shown in Tables). The mean values for morbidity and social activity indicators by age group and by educational category, and the significance of any overall or between-group differences were as shown in Tables 8.3 and 8.4 and previously discussed.

In the interests of consistency in the data analysis, a 3-level mean baseline Similarities score variable was constructed. In contrast with the MMSE and the Verbal Fluency test, however, no norms were available as a guide for classifying Similarities scores in the literature reviewed, and as noted above, the test was modified for the present study. An *ad hoc* rationale was therefore developed in order to compare demographic and other independent variables by level of performance. Since a maximum potential score of two points was available for each of six items, and the mean score for the total sample was 7.60, the variable was categorised as follows: mean score 0-5 (less than 1 point per item); 6-8 in the middle range; and 9-12 (high score).

Based upon this classification, eighteen percent of cases (n=63) scored in the lowest range ($\bar{X}$=3.24, s.d. 1.4), thirty-five percent (n=122) in the middle range ($\bar{X}$=7.20, s.d. 0.8), and forty-six percent (n=160) in the high range ($\bar{X}$=9.62, s.d. 0.8), with highly significant overall variation and between-group differences, F(2,342) =1048.12, p<.001] (see Table 8.19). There was significant variation in mean age for the three variable levels, those in the highest scoring group, ($\bar{X}$=70.91, s.d. 6.3), being significantly younger than the other two categories [F(2,342)=11.16, p<.001]. Mean years of education were closely correlated with baseline score level for all three groups with significant between-group differences. Subjects in the lowest scoring category had a mean of 10.51, s.d. 3.7, compared with 14.43, s.d. 3.1, for the middle level and 16.51, s.d. 3.1, for those with the highest initial scores [F(2,342)=79.65, p<.001].

As shown in the previous analyses for the MMSE and Verbal Fluency test, low

Table 8.19

MEAN SIMILARITIES TEST SCORES, AGE, YEARS OF EDUCATION, MORBIDITY AND SOCIAL ACTIVITY LEVELS BY BASELINE SCORE

Baseline Score	N=345	Similarities Score Mean	(S.D.)	Age Mean	(S.D.)	Years of Education Mean	(S.D.)
0-5	63	3.24*	(1.4)	74.95	(7.0)	10.51*	(3.7)
6-8	122	7.20*	(0.8)	74.02	(7.4)	14.43*	(3.1)
9-12	160	9.62*	(0.8)	70.91*	(6.3)	16.51*	(3.1)

ANOVA of Similarities Scores by Baseline Score Level [F(2,344)=1048.12, p=.0001]
ANOVA of Age by Baseline Similarities Score Level [F(2,344)=11.16, p=.0001]
ANOVA of Education by Baseline Similarities Score Level [F(2,344)=79.65, p=.0001]

Baseline Score	N=345	Medical Diagnoses Mean	(S.D.)	Medicines Used Mean	(S.D.)	Disability Score Mean	(S.D.)
0-5	63	4.14*	(1.7)	3.60	(2.4)	1.02*	(2.0)
6-8	122	3.21	(1.7)	3.15	(2.5)	0.44	(1.2)
9-12	160	3.06	(1.5)	2.88	(2.4)	0.24	(0.6)

ANOVA of Medical Diagnoses by Baseline Similarities Score Level [F(2,344)=11.27, p=.0001]
ANOVA of Medicines in Daily Use by Baseline Similarities Score Level [F(2,344)=2.08, p=.1266]
ANOVA of Disability Score by Baseline Similarities Score Level [F(2,344)=9.68, p=.0001]

Baseline Score	N=345	Depression Score Mean	(S.D.)	Hours of Activity Mean	(S.D.)	Company of Others Mean	(S.D.)
0-5	63	6.33	(4.1)	21.14	(19.2)	2.89*	(1.4)
6-8	122	6.33	(4.8)	27.84	(16.7)	3.28	(1.1)
9-12	160	4.82*	(3.9)	39.62*	(19.7)	3.56	(0.8)

ANOVA of Depression Score by Baseline Similarities Score Level [F(2,344)=5.31, p=.0053]
ANOVA of Hours of Activity by Similarities Score Level [F(2,344)=27.15, p=.0001]
ANOVA of Hours in the Company of Others by Baseline Score Level [F(2,344)=9.43, p=.0001]

* $p<.05$

levels of education were associated with significantly higher levels of morbidity and physical disability, and lower levels of social activity. There was significant variation in the ANOVA models, those with the lowest baseline score having a significantly higher mean number of medical diagnoses [F(2,344=11.27, p=.0001], disability score [F(2,344)=9.68, p=.0001] and number of depressive symptoms [F(2,344)=5.31, p=.0001]. At the same time these same subjects reported significantly less hours of activity and time spent in the company of others (see Table 8.19).

8.14 Regression analyses of Similarities scores

Once again, using the techniques described in the analyses of data generated by the two other study instruments, multiple regression models were constructed in order to determine the relative effects of age, education or other independent variables on Similarities test scores. For the total sample, the mean Similarities score was 7.60, s.d. 2.5. Four variables were selected by the stepwise procedure for inclusion in a highly significant multiple linear regression model. These were, in hierarchical order, education, number of activities, age, and disability score, $R^2=0.41$, [F(4,340)=59.58, $p<.001$]. Beta coefficients showed that education was an extremely powerful predictor of Similarities performance (beta=0.475, $p<.001$), followed by number of activities (beta=0.162, $p<.001$). Both age (beta=-0.143, $p<.01$) and disability (beta=-0.116, $p<.05$) were weaker negative indicators of test score. Education accounted uniquely for no less than 33 percent of the variance, the number of activities reported 5 percent, age 2 percent, and disability score just 1 percent (see Table 8.20).

Because of the known gender differences in mean years of education, separate models

Table 8.20

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE SIMILARITIES TEST SCORES

Total Sample (N=345)

VARIABLE	Standardised Beta Weights	p value
Age	−0.143	0.0010*
Education	0.475	0.0001**
Disability score	−0.116	0.0107*
Number of Activities	0.162	0.0009*

Mean score = 7.60.
Model: R−square=0.41, F(4,340)=59.58, p=.0001

Table 8.21

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE SIMILARITIES TEST SCORES

Men (N=186)

VARIABLE	Standardised Beta Weights	p value
Age	−0.210	0.0003*
Education	0.562	0.0001**
Disability score	−0.152	0.0078*
Company of Others	0.141	0.0158*

Mean score = 8.28.
Model: R−square=0.43, F(4,181)=34.31, p=.0001

**p<.001
* p<.05

were again constructed for men and for women. The mean Similarities score for men was 8.28. The stepwise forward selection procedure produced a highly significant 4-variable model, comprising education, age, disability score and time spent in the company of others $R^2=0.43$, [$F(4,181)=34.31$, $p<.001$]. Education was the most powerful predictor (beta=0.562, $p<.001$), followed by age (beta=-0.210, $p<.001$), and disability score (beta=-0.152, $p<.01$) both of which were negative predictors, while time spent in the company of others was a positive indicator (beta=0.140, $p<.05$) (see Table 8.21). Education accounted for 34 percent of the variance in the male sample, age for 5 percent, disability for 2 percent and the company of others for 1 percent.

As with the results of other study instruments, women had a lower mean Similarities test score, 6.80. Three variables were selected by the stepwise procedure, producing a highly significant but less powerful model for women than for men, $R^2=0.31$ [$F(3,155)=22.77$, $p<.001$] (see Table 8.22). Education was the strongest independent predictor (beta=0.347, $p<.001$), followed by number of activities (beta=0.231, $p<.01$). Disability score was a negative indicator (-0.168, $p<.05$), but age was not selected. Education uniquely explained 21 percent of the variance, number of activities 7 percent, and disability score accounted for 3 percent, values that were very different from those of the male subjects.

In keeping with the analyses of other study instruments, the sample was dichotomised for the construction of regression models for subjects aged 60-74 years, and 75 and over. For the younger group, the mean Similarities score was 7.99 and a significant 3-variable model was once again determined by the stepwise forward selection procedure, $R^2=0.35$,

Table 8.22

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE SIMILARITIES TEST SCORES

Women (N=159)

VARIABLE	Standardised Beta Weights	p value
Education	0.347	0.0001**
Disability score	−0.168	0.0210*
Number of Activities	0.231	0.0022*

Mean score = 6.80.
Model: R−square=0.31, F(3,155)=22.77, p=.0001

Table 8.23

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE SIMILARITIES TEST SCORES

Age 60−74 (N=203)

VARIABLE	Standardised Beta Weights	p value
Age	−0.185	0.0015
Education	0.467	0.0001**
Number of Activities	0.179	0.0053**

Mean score = 7.99
Model: R−square=0.35, F(3,199)=35.10, p=.0001

**p<.001
* p<.05

[F(3,199)=35.10, $p<.001$]. It is of note that age was one of the independent variables selected as a negative predictor of Similarities scores for this sample subset (beta=-0.185, $p<.01$). Education, however, was the most powerful predictor (beta=0.467, $p<.001$), with number of activities having a less significant effect (beta=0.179, $p<.01$) (see Table 8.23). Education explained 29 percent of the variance, with age and number of activities each explaining 3 percent.

The mean score for those aged 75 and over was 7.04. A very powerful and significant 3-variable model was once more selected by the stepwise procedure for those aged 75 and over, $R^2=0.44$, [F(3,138)=36.65, $p<.001$]. Education (beta=0.489, $p<.001$), followed by number of activities (beta=0.221, $p<.01$) were positive predictors of higher scores, and disability level (beta=-0.159, $p<.05$) was a weaker negative predictor (see Table 8.24). Education uniquely explained 36 percent of the variance in Similarities performance (a higher proportion than for any other group), number of activities 6 percent, and disability 2 percent.

In an attempt to control for the effects of education, a final regression model was constructed, excluding subjects with 8 or fewer years of formal schooling. Four independent variables were selected for what was once more a significant and relatively powerful model, $R^2=0.35$, [F(4,315)=43.00, $p<.001$]. Education continued as the most powerful indicator (beta=0.431, ($p<.001$), followed by age as a negative predictor (beta=-0.174, $p<.001$), number of activities (beta=0.166, $p<.01$) and disability score (beta=-0.138) (see Table 8.25). Education uniquely accounted for 25 percent of the variance, even when the least educated were excluded from the analysis, number of

Table 8.24

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE SIMILARITIES TEST SCORES
Age 75 and Over (N=142)

VARIABLE	Standardised Beta Weights	p value
Education	0.489	0.0001**
Disability score	−0.159	0.0214*
Number of Activities	0.221	0.0030*

Mean score = 7.04
Model: R−square=0.44, F(3,138)=36.65, p=.0001

Table 8.25

MULTIPLE REGRESSION MODEL OF PREDICTOR VARIABLES FOR BASELINE SIMILARITIES TEST SCORES
Excluding those with Limited Education (N=320)

VARIABLE	Standardised Beta Weights	p value
Age	−0.174	0.0002**
Education	0.431	0.0001**
Disability score	−0.138	0.0043*
Number of Activities	0.166	0.0010*

Mean score = 7.86
Model: R−square=0.35, F(4,315)=43.00, p=.0001

**p<.001
* p<.05

activities for 5 percent, age 4 percent and disability score 1 percent.

8.15 Change in Similarities scores with repetitive testing

In view of the power of education as a determinant of Similarities performance in the present sample intuitively it might be predicted that test scores would remain stable when the instrument was administered more than once to the same subjects. Tables 8.26 to 8.28 show the repeated measures ANOVAs for mean scores by age group, educational category, and baseline score level for those who completed the three assessment interviews pre-surgery, and 1-week and 6-weeks after surgery. At time 1, there was significant overall variation by age group, those aged 75-84, and 85 and over, having significantly lower mean scores than those aged 60-64 and 65-74 [F(3,239)=6.60, $p<.001$]. By time 3, however, scores had improved for all age groups as reflected in the ANOVA for within-group change [F(2,427=9.56, $p<.001$], with a marginally higher improvement for those aged 85 and over, after a decline at time 2, but there was no time by age group interaction.

There was significant variation in mean test scores by level of education at time 1, [F(3,239)=30.76, $p<.001$]. Those in the higher two educational categories had significantly higher mean scores than the two lower groups and the between-group differences persisted at each assessment. On the other hand, all groups showed a within-group improvement at time 3, [F(2,472)=10.98, $p<.001$]. There was also a small but significant time by group interaction [F(6,472)=2.23, $p<.05$], possibly reflecting the consistent increase in test scores for those with higher levels of education and the decline at time 2 for the lower two categories (see Table 8.27).

Table 8.26

MEAN SIMILARITIES SCORES IN THE PERI-OPERATIVE PERIOD BY AGE

		Time 1		Time 2		Time 3	
Age	N=240	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
60-64	40	8.68*	(2.1)	8.80*	(2.3)	9.05*	(1.7)
65-74	106	7.99	(2.2)	8.10	(2.7)	8.27	(2.2)
75-84	83	6.88*	(2.8)	6.72*	(3.2)	7.55*	(2.9)
85 & over	11	6.64*	(2.2)	6.00*	(2.3)	7.64	(1.3)

ANOVA of Similarities Score at Time 1 by Age Group [F(3,236)=6.60, p=.0003]
ANOVA of within-group change over Time [F(2,472)=9.56, p=.0001]
ANOVA of interaction between Time and Age Group [F(6,472)=1.75, p=.1070]

Table 8.27

MEAN SIMILARITIES SCORES IN THE PERI-OPERATIVE PERIOD BY EDUCATION

		Time 1		Time 2		Time 3	
Education	N=241	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-8 years	16	4.56*	(3.3)	3.94*	(3.3)	5.25*	(3.1)
9-12 years	56	6.11*	(2.7)	5.61*	(3.1)	6.61*	(2.9)
13-16 years	89	8.16	(1.7)	8.26	(2.3)	8.60	(1.7)
17 or more	79	8.85	(1.9)	9.18	(1.8)	9.27	(1.4)

ANOVA of Similarities Score at Time 1 by Education Group [F(3,236)=30.76, p=.0001]
ANOVA of within-group change over Time [F(2,472)=10.98, p=.0001]
ANOVA of interaction between Time and Education Group [F(6,472)=2.23, p=.0395]

Table 8.28

MEAN SIMILARITIES SCORES IN THE PERI-OPERATIVE PERIOD BY BASELINE SCORE

		Time 1		Time 2		Time 3	
Baseline Verbal Fluency Score	N=241	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
0-5	41	3.15*	(1.4)	3.56*	(2.9)	4.63*	(2.6)
6-8	84	7.21*	(0.9)	7.36*	(2.4)	8.06*	(1.7)
9-12	115	9.61*	(0.8)	9.33*	(1.5)	9.43*	(1.3)

ANOVA of Similarities Score at Time 1 by Baseline Score [F(2,237)=718.10, p=.0001]
ANOVA of within-group change over Time [F(2,474)=17.86, p=.0001]
ANOVA of interaction between Time and Baseline Score [F(4,474)=7.66, p=.0001]

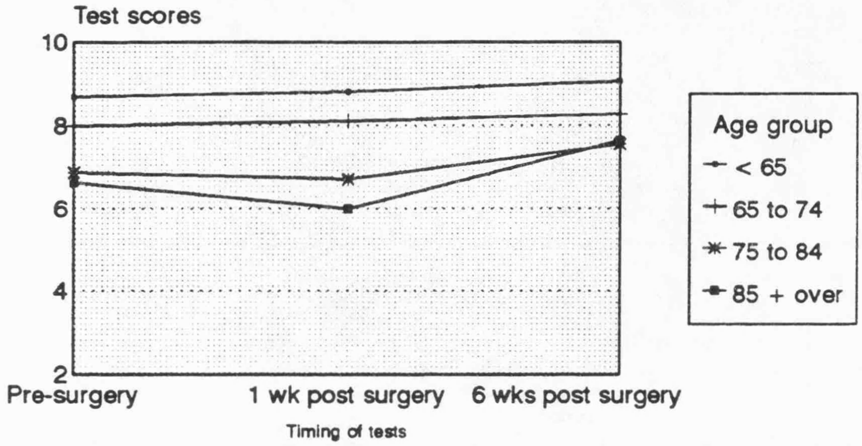
When test scores, categorised by baseline score level, were analysed, the significant overall and between-group variation at time 1 [$F(2,237)=718$, $p<.001$] persisted at each assessment period. However, there were significant within-group changes over time [$F(2,474)=17.86$, $p<.001$]. Subjects in the lower two groups showed improvement in mean test score at each assessment, compared with small loss by those in the highest scoring group. These differences were reflected in a significant interaction between time and baseline score category [$F(4,474)=7.66$, $p<.001$].

The variation in test performance by age group, level of education and baseline score is illustrated in Figures 8.6, 8.7 and 8.8. As described above, there was a small decline in mean scores for those aged 85 and over at time 2 which rebounded to a higher level at time 3 (see Figure 8.6). This was consistent with the pattern for the MMSE (see Figure 7.2) and the Verbal Fluency test (see Figure 8.2), and as previously suggested, may be attributable to transient debilitating effects of surgery in the oldest age group (see Figure 8.6). There was also a fall in mean scores at time 2 for both lower education categories, with recovery to a higher mean score at time 3, a pattern not observed for those with higher levels of education (see Figure 8.7). At the same time, (see Figure 8.8) there was an improvement in mean scores for subjects in the two lower baseline categories at each assessment.

In the 1958 edition of the WAIS a table of "difficulty values" for individual items in the Similarities test was published. This demonstrated that less than 40 percent of subjects achieved any "credit" for three items, i.e. "wood - alcohol", "poem - statue", "praise -punishment".[Wechsler, 1958 p247] In the present study, one of these items was in the

Figure 8.6

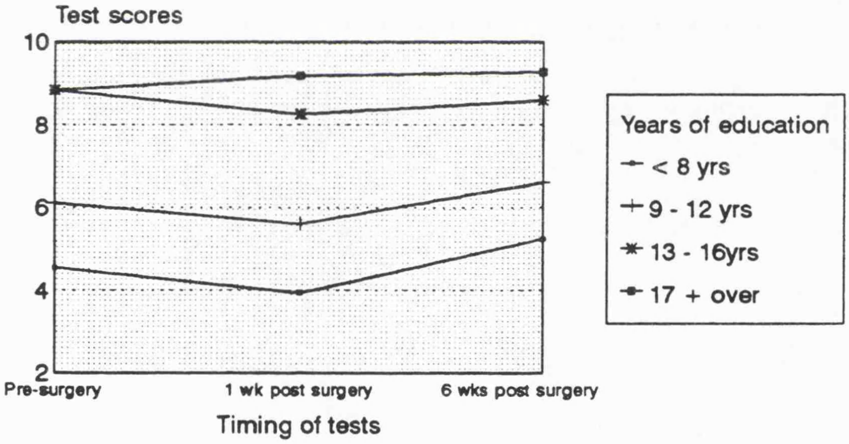
Change in Perioperative Similarities Scores
By Age Group



ANOVA of effects of time, p=0.0003
ANOVA of interaction between time and group, p=.1070

Figure 8.7

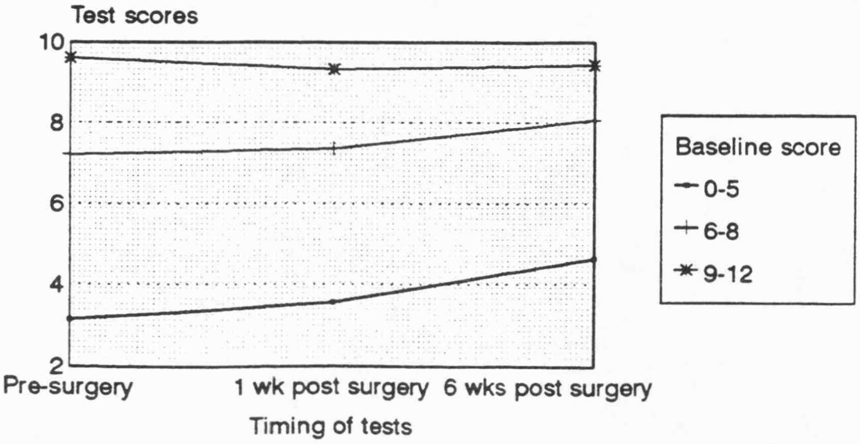
Change in Perioperative Similarities Scores
By Level of Education



ANOVA of effects of time, p=0.0001
ANOVA of interaction between time and group, p=0.0395

Figure 8.8

Change in Perioperative Mean Similarities
Score By Baseline Level



ANOVA of effects of time, p=.0001
ANOVA of interaction between time and group, p=.0001

group administered at times 1 and 3, and two in the second group administered only at time 2 (see Figure 8.2). It may therefore be proposed that the plotting of Similarities scores illustrates an "education" effect, based upon the different levels of difficulty for the two sets of test items, one that was apparent in the reaction of subjects with the least education to the more difficult test presented at time 2. At the same time it may be argued that the graphs indicate a probable "learning" effect from repetitive testing, as depicted in the upward gradient for all lower score groups between times 2 and time 3.

8.16 Significance and correlates of change in Similarities scores

To be consistent in evaluating change in test performance between the three study instruments, and in order to determine whether differences over time occurred independently of age, education or variation in morbidity and disability, further analyses were undertaken to compare these variables for subjects in whom scores remained stable, those with improvement and those with a decline in mean test values. The score range of the Similarities test precluded measurement of a precise ten-percent change, the criterion for evaluating mean values of independent variables against the MMSE and Verbal Fluency performances over time, and a 2-point (2/12 or 17 percent) difference was therefore adopted as the measure of change in Similarities score between times 1 and 3.

The frequency distribution for groups with this level of variation showed that scores remained stable within a two-point range for sixty-nine percent of subjects (166/240), decreased for eight percent (20/240) and improved by two points or more for twenty-three percent of subjects (54/240) (see Table 8.29). The mean score for those with a

Table 8.29

COMPARISON OF GROUPS WITH AND WITHOUT A TWO-POINT or GREATER CHANGE IN SIMILARITIES SCORES IN THE PERIOPERATIVE PERIOD

Level of Change	N=240	Similarities Score 1 Mean	(S.D.)	Similarities Score 3 Mean	(S.D.)
No change	166	8.22*	(2.2)	8.25	(2.2)
Decline	20	7.85*	(2.7)	5.30*	(2.5)
Improvement	54	5.89	(2.5)	8.81	(2.2)

ANOVA of Scores at time 1 by Level of Change [F(2,237)=20.64, p=.0001]
ANOVA of within-group change over time [F(1,237)=73.94, p=.0001]
ANOVA of interaction between Time and Level of Change [F(2,237)=389.66, p=.0001]

Level of Change	N=240	Age Mean	(S.D.)	Years of Education Mean	(S.D.)
No change	166	71.74	(7.1)	15.37	(3.5)
Decline	20	72.50	(7.4)	13.40	(4.2)
Improvement	54	73.67	(7.4)	14.02	(4.1)

ANOVA of Mean Age by Level of Change [F(2,237)=1.48, p=0.2289]
ANOVA of Mean Years of Education by Level of Change [F(2,237)=4.55, p=.0116]

Level of Change	N=240	Medicines 1 Mean	(S.D.)	Medicines 3 Mean	(S.D.)
No change	166	2.92	(2.2)	3.08	(2.2)
Decline	20	2.55	(1.9)	2.90	(2.1)
Improvement	54	3.54	(2.7)	3.44	(2.7)

ANOVA of Medicine Use at time 1 by Level of Change [F(2,237)=1.90, p=.1516]
ANOVA of within-group change over Time [F(1,237)=1.62, p=.2048]
ANOVA of interaction between Time and Level of Change [F(2,237)=0.88, p=.4175]

minimum 2-point increase at time 3, 5.89, s.d. 2.5, was significantly lower than the score of the other two groups, which were 8.22, s.d. 2.2, for subjects with stable scores, and 7.85, s.d. 2.7 for those with decline [F(2,237)=20.64, p<.001]. There was a significant within-group change over time as mean values improved to 8.81, s.d. 2.2, for the "improved" subjects, and fell to 5.30, s.d. 2.5 for the "decline" group [F(1,237)=73.94, p<.001]. These changes were, reflected in a very significant time by group interaction [F(2,237)=389.66, p<.001].

ANOVA showed no significant differences in mean age between the three groups although those who improved were older than subjects in whom scores decreased (X=73.67, s.d. 7.4 compared with 72.50, s.d. 7.4). There were no significant between-group differences in mean years of education, which ranged from 13.40 for those with lower scores at time 3 to 15.37 for those with stable scores, but the variation did achieve significance [F(2,237)=4.55, p<.05]. At the same time, there were no significant differences in mean values for medication use, disability score, or symptoms of depression, and a small decrease in the latter was experienced equally by all groups (see Table 8.29 contd.).

In common with the results of the Verbal Fluency assessments, the age and education range for subjects with both positive and negative change in mean Similarities scores was great. The age range was from 60 to 90 years for those with a minimum 2-point improvement and from 61 to 87 for those with a 2-point or greater decline, with years of education ranging from 5 to 21, and from 4 to 20 respectively. Thirty-five percent (19/54) of subjects with a 2-point or greater increase in test scores were in the lowest

Table 8.29 (contd.)

COMPARISON OF GROUPS WITH AND WITHOUT A TWO-POINT or GREATER CHANGE IN SIMILARITIES SCORES IN THE PERIOPERATIVE PERIOD

Level of Change	N=240	Disability Score 1 Mean	(S.D.)	Disability Score 3 Mean	(S.D.)
No change	166	0.49	(1.3)	0.49	(1.1)
Decline	20	0.50	(0.9)	0.60	(1.1)
Improvement	54	0.76	(1.3)	0.69	(1.1)

ANOVA of Disability Score at time 1 by Level of Change [F(2,237)=0.89, p=.4121]
ANOVA of within-group change over time [F(1,237)=0.04, p=.8513]
ANOVA of interaction between Time and Level of Change [F(2,237)=0.226, p=.8042]

Level of Change	N=240	Depression Score 1 Mean	(S.D.)	Depression Score 3 Mean	(S.D.)
No change	166	5.63	(4.5)	5.02	(4.7)
Decline	20	5.50	(4.2)	4.80	(4.3)
Improvement	54	5.83	(4.7)	5.54	(4.7)

ANOVA of Depression Score at time 1 by Level of Change [F(2,237)=0.06, p=.9462]
ANOVA of within-group change over Time [F(1,237)=8.45, p=.004]
ANOVA of interaction between Time and Level of Change [F(2,237)=0.27, p=.7672]

scoring group at the first assessment, representing forty-six percent of subjects (19/41) in this category (see Table 8.28). At the same time, fifty percent of those with a decrease of two or more points (10/20) were in the highest scoring category at time 1. It may therefore be proposed that the analyses indicated that there was a small association between education and a 2-point or greater change in Similarities performance, but variation was independent of age, morbidity or physical disability, and improvement was shown disproportionately by those with the lowest initial test scores.

8.17 Summarising the results of the Similarities testing

The Similarities and Verbal Fluency tests were both conducted with the same subjects, and the demographic and morbidity characteristics by age and education were therefore as shown in Tables 8.4 and 8.5, and previously discussed. In order to be consistent in the analyses of all three study instruments, the sample was categorised by baseline Similarities scores, using an *ad hoc* rationale rather than published norms which were available for the MMSE and Verbal Fluency tests. Subjects in the highest scoring category were significantly younger, and there were very significant differences in mean years of education between the three variable levels, ranging from 10.5 for the lowest to 16.5 years for the highest (see Table 8.19). Those with the lowest mean scores had significantly more medical diagnoses, higher disability and lower values for social activity variables, all of which were comparable with subjects in the MMSE range 0-23, (see Table 7.3).

Multiple regression models constructed to evaluate the relative effects of age, education or other variables as predictors of Similarities performance tended to be more powerful than those developed for the other two study instruments, with multiple correlation coefficients, i.e. R^2 values greater than 40 for the total sample, for men, and for those aged 75 and over. As noted above, the Similarities test was described in a textbook of neuro-psychological assessment as "relatively independent of social or educational background". Lezak, 1983 p265 In the present study, however, education uniquely accounted for thirty-three percent of the overall variance in test scores, thirty-four percent for men, and thirty-six percent for subjects aged 75 and over, thereby being a stronger predictor of Similarities scores than of MMSE or Verbal Fluency test performance. On the other hand, the model for women had less predictive power ($R^2=.31$) and education explained only twenty-one percent of the variance.

The next most powerful predictor, explaining from five to seven percent of the variance for the sample as a whole, for women, and those in the older age group, was the number of activities reported a variable reflecting the level of "social interaction" of the participant. Age also uniquely accounted for five percent of the variance in scores for male participants but a lesser proportion in three other models, while physical disability explained no more than one to three percent of the variance in any of the models tested (see Tables 8.20 to 8.25).

As previously noted, intuitively it might be assumed that the strength of formal years of education in predicting Similarities scores would increase the reliability of initial test scores and reduce the potential for improvement with repetitive testing, and that any change would therefore tend to be negative and associated with increased morbidity or disability. It may be argued that the education effect was evident in response to the administration of a more difficult set of test items at time 2 compared with times 1 and 3. As illustrated in Figures 8.7, there was a small decrease in test performance at the second assessment for the lower education categories. On the other hand, scores improved to a higher level at time 3, a pattern also observed for the two older age groups. Furthermore, there was a small improvement in Similarities performance by subjects in the lower baseline score categories at each assessment.

An increase of 2 or more points was observed in twenty-three percent of participants (54/240) in whom the tests were repeated on three occasions before and after surgery, and a decline of the same magnitude occurred in only eight percent of subjects, both positive and negative change being independent of age, education or variation in morbidity or physical disability. At the same time, nearly half of those in the lowest baseline score category (19/41) improved by 2 or more points. It may therefore be argued that in addition to the effects of education as a predictor of Similarities scores, there was a marked learning or "practice" effect occurring disproportionately in those with the poorest initial performance, and the lowest mean years of formal education.

8.18 Comparing the MMSE, Verbal Fluency and Similarities results

The MMSE, promoted as a "quantified assessment of cognitive state, of demonstrable reliability and validity",[Folstein, 1975 p189] was the primary instrument adopted in the present study for determining the "prevalence" of cognitive impairment, and for detecting its "onset" following surgery. Among the validity tests conducted by Folstein in developing the MMSE a small number of subjects were assessed against the full battery of the WAIS resulting in correlation coefficients of 0.776 and 0.66 respectively between the MMSE and the WAIS "verbal" and "performance" subscales. The Similarities test from the WAIS verbal battery, and the Verbal Fluency test, both recommended for use in conjunction with the MMSE to assess processes such as "abstract thinking",[Lezak, 1983 p583, McKhann, 1984 p942 Roth, 1986 p700] were also introduced into the study under discussion.

The prevalence of dementing disease, in common with all chronic degenerative disorders increases with advancing age, and it may be argued that the validity of any instrument employed in screening for "cognitive impairment" as an indicator of dementia in elderly populations, for which age is not the primary predictor of positive findings is therefore subject to challenge. At the same time, an "inability to learn new information" is one of the primary psychiatric diagnostic criteria for dementia,[DSM-III-R, 1987 p107] and it may be argued that improvement in assessment scores, particularly among those identified as "cases", as determined by classifications according to normative values renders the assessments unreliable and further invalidates the instrument upon which they are based.

In the present study, which recruited a sample of individuals aged 60 to 93 with unusually high mean levels of education among men and women of all ages, together with a group of subjects with little formal schooling, education was identified as the primary predictor of MMSE and Verbal Fluency scores, with even greater predictive power for Similarities performance. At the same time, there were further differences between the instruments, as determined by the multiple regression models constructed from a range of variables selected as potentially significant intervening or confounding factors for test scores. These indicated that different "processes" were indeed evaluated by the three assessment instruments.

The second most powerful predictor of MMSE scores was the level of physical disability which uniquely accounted for as much as 12 percent of the variance in initial performance (see Chapter 7, 7.9). In contrast the variable explained no more than 1 to 3 percent of the variance in Verbal Fluency and Similarities testing. As previously discussed, successful completion of all items of the MMSE requires both visual acuity and motor skills, and it may be argued that the lack of power in predicting Verbal Fluency and Similarities test scores, tends to confirm the independent negative effect of physical disability on MMSE performance.

For the Verbal Fluency and Similarities tests the number of activities reported by subjects, and time spent in the company of others, variables which may be construed as indicators of "social interaction", were the secondary predictors of test performance, but accounted for a lower proportion of the variance in Verbal Fluency scores. The Similarities and Verbal Fluency tests elicit only verbal responses, but the latter test is timed and there was considerable heterogeneity of Verbal Fluency scores within both age and education categories, and evidence that low performance level was not necessarily associated with limited education. It was therefore proposed that the relative lack of predictive power of "social activity" variables in the Verbal Fluency regression models, the dependent variable value for which depends upon speed of response, was indicative of interaction between predictor variables and factors such as personality, motivation and self-confidence, not captured in the data.

It may be further argued that this proposition is supported by the repeated measures analyses which demonstrated that thirty-two percent of subjects, approximately one-third in each initial performance category, showed a 10 percent, i.e. 7-point or greater increase in Verbal Fluency when the tests were repeated over time, the change being independent of age, education, and variation in morbidity indicators. A similar independent "learning" or "practice effect" was detected in relation to a 2-point improvement in Similarities scores which was identified in twenty-three percent of study subjects, and was observed disproportionately in nearly half of those in the lowest baseline score category. It may therefore be posited that responses to this test, the initial scores for which were so strongly predicted by education, were also determined in part by level of social activity, serving as a proxy for other interaction effects, of which one of some importance may have been increasing comfort or self-confidence with the assessment process, the lack of which may well have accounted for initial poor performance, particularly among women.

It may also be argued that the 2-point or greater improvement in MMSE scores shown by sixteen percent of study participants, which was largely independent of age, education or morbidity, was also attributable to a similar learning or practice effect (see Table 7.17). A 2-point decrease in MMSE scores has been adopted as meeting the "clinical criterion for cognitive decline" in a number of studies, but as previously observed, a 2-point change may result from simply mis-stating the date by one day, and it was therefore proposed that little significance could be attributed to either a positive or negative change in MMSE performance of this magnitude. This argument was supported by comparing the mean values for Verbal Fluency and Similarities testing for all those with a 2-point or greater change in MMSE scores (see Table 8.30). This demonstrates that there was significant within-group improvement in Verbal Fluency [F(1,237)=48.19, $p < .001$] and Similarities mean test scores [F(1,237)=17.34, $p < .001$] between time 1 and time 3 for all groups, whether showing stability, decline or improvement in MMSE scores of 2 or more points.

On the other hand, a 3-point, i.e. ten percent or greater change in MMSE scores (3/30) was associated with some highly significant variation in morbidity indicators, particularly in the mean level of physical disability, a statistical result that was confirmed by investigation of the individual characteristics of those in whom change was observed. Most subjects with this level of decline suffered major complications following surgery, while benefits in terms of improved visual acuity were associated with higher MMSE scores in other study participants. At the same time, for a number of individuals, including several in the lowest education category who scored in the "impaired" range at time 1, improvement was independent of change in disability and may therefore be attributed once more to a learning or practice effect.

The discussion thus far has not addressed the significance of a decrease in Verbal Fluency and Similarities test scores between time 1 and time 3 which was observed in ten and eight percent of subjects respectively. Although there were no significant between-group differences in age and level of change for any of the study instruments, those without change in test scores tended to be younger. It may therefore be proposed that instability in test performance was attributable to the effect of age, which was an independent predictor of test scores in a number of the regression models constructed for

Table 8.30

COMPARISON OF TEST PERFORMANCE FOR THREE STUDY INSTRUMENTS BASED ON A TWO-POINT CHANGE IN MMSE SCORES

Level of Change	N=240	MMS Score 1 Mean	MMS Score 1 (S.D.)	MMS Score 3 Mean	MMS Score 3 (S.D.)
No change	175	28.49	(1.7)	28.55	(1.8)
Decline	25	28.32	(3.4)	25.88*	(3.9)
Improvement	40	25.75*	(2.9)	28.55	(2.0)

ANOVA of MMSE Scores at Time 1 by Level of Change [F(2,237)=27.15, p=.0001]
ANOVA of within-group change over time [F(1,237)=18.19, p=.0001]
ANOVA of interaction between Time and Level of Change [F(2,237)=263.68, p=.0001]

Level of Change	N=240	Verbal Fluency 1 Mean	Verbal Fluency 1 (S.D.)	Verbal Fluency 3 Mean	Verbal Fluency 3 (S.D.)
No change	175	37.07	(14.8)	40.90	(15.9)
Decline	25	36.04	(12.2)	39.24	(15.5)
Improvement	40	32.75	(16.1)	34.63	(15.4)

ANOVA of Verbal Fluency Scores at Time 1 by Level of Change [F(2,237)=1.38, p=.2572]
ANOVA of within-group change over time [F(1,237)=48.19, p=.0001]
ANOVA of interaction between Time and Level of Change [F(2,237)=1.07, p=.3438]

Level of Change	N=240	Similarities 1 Mean	Similarities 1 (S.D.)	Similarities 3 Mean	Similarities 3 (S.D.)
No change	175	7.87*	(2.4)	8.40*	(2.3)
Decline	25	7.60	(2.3)	7.68	(2.4)
Improvement	40	6.80*	(2.8)	7.30*	(2.8)

ANOVA of Similarities Score at Time 1 by Level of Change [F(2,237)=3.05, p=.0491]
ANOVA of within-group change over time [F(2,237)=17.34, p=.0001]
ANOVA of interaction between Time and Level of Change [F(2,237)=0.69, p=.5026]

the different instruments, was observed in the transient decline in test performance by the oldest age group after surgery, and was also noted in MMSE assessments of subjects with an established diagnosis of dementing disease (see Chapter 7). At the same time a major proportion of those with a decline in Verbal Fluency and Similarities performance, across the age spectrum, were among those with initial scores in the highest category at time 1. It may therefore be posited that regardless of age, decline may also have been attributable to factors such as variation in subjective well-being, and what may be construed as a "reverse" learning effect in which some subjects with high initial scores became less comfortable and tolerant of the assessment procedure with repetition. All of which begs the question of whether initial test performance or the change demonstrated by study participants was indicative of cognitive impairment or its onset.

Chapter 9

THE COGNITIVE PARADIGM AND THE ALZHEIMER'S ENTERPRISE

"Cognitive impairment has many meanings"[Folstein, 1985 p231]

Ten years after publishing the Mini-Mental State Examination (MMSE) as a "quantitative" and "valid test of cognitive function", with scores which "correlate with a standard test of cognition, the Wechsler Adult Intelligence Scale" (WAIS),[Folstein et al, 1975 p195] Marshal Folstein discussed "the meaning of cognitive impairment".[Folstein, et al 1985 p228] In support of this discussion data were presented from interviews conducted in East Baltimore, one of the sites of the Epidemiologic Catchment Area (ECA) study of psychiatric disorders in the adult population of the United States funded by the National Institute of Mental Health in 1980. The diagnostic criteria for prevalence estimates of "mental disorders in both treated and untreated populations" in the investigation were those which "had evolved into the official nomenclature of the American Psychiatric Association, published as the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III, 1980),[Regier & Robins, 1991 p1-9] (see Chapter 3). The ECA study instrument, the Diagnostic Interview Schedule (DIS), incorporated the MMSE as a measure of "current cognitive status", determined by the investigators to be the "closest approximation to organic mental disorders" using DSM-III criteria.[George et al, 1991 p291-3]

Accompanying the article by Folstein in the same edition of the journal was another by Kramer et al. describing the methods of the ECA study and the prevalence rates by age group for all mental disorders detected in the East Baltimore interviews.[Kramer et al, 1985 p236-45] As previously observed, Morton Kramer, formerly Chief of the Biometrics Branch of the National Institute of Mental Health,[DSM, 1952 px DSM-II, 1968 pxi] and a member of the Committee on Nomenclature and Statistics for the DSM-III, had argued in 1980 that "the world is in the midst of a rising pandemic of mental disorders and associated chronic diseases and disabilities" which he attributed to growth in the "number of persons in age groups at high risk for developing mental disorders and chronic diseases", together with medical interventions "prolonging the lives of affected individuals" including those with "brain syndromes associated with senile brain disease". He used 1980 morbidity and mortality data in projections "to demonstrate the potential size of mental health problems that may face the World by the year 2000".[Kramer, 1980 p382-5] At the same time he proposed that "there was a lack of epidemiological data on the incidence, duration and prevalence of specific mental disorders in most countries in the world", including the United States,[Kramer, 1980 p395] (see Chapter 4).

Kramer's thesis on the "alarming rate" of increase in mental disorders associated with longer life expectancy was repeated in another paper on the "worldwide pandemic",[Kramer, 1983 p14] but there was little evidence in its support in "point prevalence rates", i.e. "the state of subjects at the date of interview" in the ECA study. These showed lower overall rates of mental disorders in the age group 65 and over (17.3 percent) compared with subjects aged 18-64 (24.8 percent). Moreover, the most prevalent DSM-III diagnosis in all adults, accounting for 14 percent of total "mental illness" in the younger and 11.5 percent in the older population was "phobic disorder",[Kramer et al, 1985 p240] defined in the ECA study instrument as an "unreasonable fear" of "heights", "tunnels", "storms", "spiders, bugs, mice, snakes", etc.[Robins & Regier, 1991 p408-9]

Kramer, however, gave "special attention to the measure of severe cognitive impairment", i.e. an MMSE score of 17 or less out of 30, reported to be 5.1 percent in adults aged 65 and over. This attention was merited by "the increasing public health importance of dementing illnesses" and because the measure "serves among the elderly to index the odds of being diagnosed as a case of dementia and other mental disorders that manifest in cognitive functioning" (*sic*).[Kramer et al, 1985 p236]

Folstein, in his associated 1985 paper, argued that when "cognitive impairment is identified" clinicians search for its "cause and mechanism", i.e. "the way in which abnormal structure or function of the organism lead to the impairment".[Folstein et al, 1985 p228] As previously discussed (see Chapter 7), when Folstein first published the MMSE he reported that a score of 20 or less had been found to distinguish those with cognitive impairment from those without,[Folstein et al, 1975 p196] but two years later, based on "mental status assessments" of neurological patients admitted to hospital, the MMSE was "revalidated" and the standard for "normality" revised upwards to 24 or more.[DePaulo & Folstein, 1978 p225-7] In East Baltimore four percent of those in the age group 18-64 scored 23 or less on the MMSE as did twenty-one percent of those aged 65 and over, but the former group was not further investigated for the "cause and mechanism" of the impairment. However, all elderly individuals "likely to have" a DSM-III diagnosis as determined by the ECA study instrument, i.e. with an MMSE score of 23 or less or some other disorder, together with a proportion of those with no detected anomaly, were subjected to a range of clinical and psychiatric examinations, a difference which may be attributable to the universal medical insurance provided by the Medicare programme to pay for such examinations in the older age group. Among those with an MMSE score in the "impaired" range the clinical investigations identified dementia, delirium, depression, "neurotic disorders", i.e. the "phobias", "alcohol use disorder", and others but thirty-three percent "had no diagnosable DSM-III condition". Furthermore, when the clinical data were analysed for elderly subjects scoring 17 or less, i.e. with "severe cognitive impairment" as defined for purposes of the ECA study, fourteen percent did not meet criteria for any DSM-III diagnosis.[Folstein et al, 1985 p230-2]

Dementia was identified in only twenty-six percent of all those scoring 23 or less on the MMSE while a further thirty percent were diagnosed with some other "treated or untreated" mental disorder. Among subjects scoring 17 or less (n=35) a diagnosis of dementia was ascribed to fifty-one percent with twenty-six percent meeting DSM-III criteria for a different diagnosis but, as noted above, no "disorder" was detected in fourteen percent of these participants identified as having "severe cognitive impairment". Final prevalence estimates for dementia in those aged 65 and over, based upon clinical and psychiatric investigations of the twenty-one percent of subjects scoring 23 or less on the MMSE were 6.1 percent (11.7 percent in those aged 75 and over).[Folstein, 1985 p232]

Folstein, in first promoting the MMSE as a "quantitative assessment of cognitive state of demonstrable reliability and validity" which "makes more objective" a "subjective impression of cognitive disability during an assessment of a (psychiatric) patient" (see Chapter 7) argued that the MMSE alone does not carry "diagnostic responsibility".[Folstein, 1975 p195] In the report of the Baltimore study he again argued that "the MMSE does not make a diagnosis" but that "a low MMSE score indicates a need for further evaluation".

Based upon the data generated by employing the instrument in this community study he proposed that "cognitive impairment", which he defined as a "diminished capacity to know the world",[Folstein et al, 1985 p228] was a "frequently occurring impairment" which "has many meanings and many implications" and that the "meaning of cognitive impairment as an objective syndrome should be distinguished from the neurologic diseases such as Alzheimer's disease that cause the syndrome". In concluding the article Folstein further proposed that "future research should focus on the diversity of the meaning of cognitive impairment in the elderly and the broad range of possible sources of that impairment, including Alzheimer's disease".[Folstein et al, 1985 p232]

Gubrium, in reporting on his participant observation of the families and others involved in the care of those afflicted with senile dementia and the outcome of their collective efforts, coordinated by the Alzheimer's Disease and Related Disorders Association (ADRDA), in establishing a "public culture" of the "disease", argued that these efforts may be understood as a "cognitive, not a social movement" which represents the need for elaboration of those concerned, either as professionals or as a labour of love" in attempting "to do something about the suffering believed to be caused by Alzheimer's",[Gubrium, 1986 p69-71] (see Chapter 2). Based upon earlier discussions and the empirical data that has been presented, it may be argued that Folstein's paper may be analysed both in the context of the federal government's legislative response to the "public culture" in the allocation of resources for biomedical research (see Table 2.1) and also of the psychiatric paradigm within which "cognitive impairment" was conceptualised as an "objective syndrome" of which the source may be Alzheimer's disease.

"Biologism" and the DSM-III diagnostic models

The introduction to the DSM-III observes that the manual includes "such new features as diagnostic criteria ..... much expanded descriptions of the disorders (some with newly coined names)" and that it "reflects an increased commitment in our field to reliance on data as the basis for understanding mental disorders".[DSM-III, 1980 p1] This commitment by American psychiatry was, arguably, expressed in the conduct of clinical "inter-rater reliability" field trials for the new diagnostic criteria prior to publication of the DSM-III,[DSM-III, 1980 p467] and in the ECA investigation in which they were operationalised in the questionnaire instrument. On the one hand the development of defined criteria or "symptoms", subjected to reliability testing, may be interpreted as a reasonable attempt by American psychiatry to introduce greater objectivity and standardisation into the diagnostic process thereby responding to the symbolic interactionist model of "mental illness" as deriving from subjective and arbitrary labelling of those persuaded to become its clients.[Downes & Rock, 1983 p151-3] On the other hand, although the DSM-III claims to be "atheoretical with regard to aetiology", Guimon and other psychiatrists argue that the DSM-III diagnostic models were conceived within a "Neokraepelinian"[1] biological paradigm which proposes that all "psychiatric disturbances are real illnesses" associated with brain pathology, and seeks to impute a correlational relationship between such changes and quantitative measures.[Guimon, 1989 p35] Critics such as Guimon and Lipowski define the inference of biological causation to behavioural and

---

[1] The term Neokraepelinian derives from the name of Emil Kraepelin the 19th century "organic psychiatrist" who was director of the institute in which Alois Alzheimer identified his eponymous disease (see Chapter 3).

emotional phenomena, thereby discounting the interaction between physical and socio-environmental factors and their subjective interpretation as "biologism".

Guimon proposes that "biologism" or the "unwarranted extension of biological explanations to psycho-social issues as 'science'" may conceal the socio-political factors that encourage the use of certain forms of treatment or bias the choice of research topics.[Guimon, 1989 p34] In support of this proposition, Lipowski posits that the adoption of the new paradigm by American psychiatry was motivated by immediate "political and economic concerns". He interprets the model within which all mental illness is amenable to pharmacotherapy to correct "chemical imbalance" or other imputed biological aetiologies as a means by which the specialty sought greater "medical" legitimacy within the profession while restricting competition from lay therapists.[Lipowski, 1989 p252] However, although Lipowski argues that biological psychiatry "confuses the distinction between aetiology and correlation" and although he tends to be dismissive of the benefits of neuropharmacological research to psychiatric practice he concedes the potential for neuroscience to discover a "cure" for the dementias.[Lipowski, 1989 p252-3] Similarly, Guimon proposes that "standards of 'objectivity' in measuring and description which are demanded by statistical analysis have not yet been achieved in psycho-pathology" but he accepts that biological researchers "can afford to look for hard physical indices" when studying "organic brain disorders".[Guimon, 1989 p35] On the other hand, although Guimon and Lipowski focus on the constructs which have been formulated for "functional" or "psychogenic disorders" within the paradigm, it may be argued that similar critiques are applicable to the interpretations and legitimations which have been proffered with regard to the operationalisation of the DSM-III models for "organic mental disorders" as "current cognitive status" quantified by the MMSE or other mental tests.

Based upon the papers by Folstein and Kramer it may be posited that there was a lack of support for the biological paradigm in relating the results derived from administering the MMSE as the "closest approximation" to "dementia" and other organic conditions to follow-up clinical diagnoses. When elderly East Baltimore participants were subject to clinical and psychiatric investigations, it was found that in some seventy-percent of cases an MMSE score below 23 was associated either with no DSM-III diagnosis or with mental disorders such as an "unreasonable fear" of "spiders, bugs, mice, snakes" etc. Furthermore it was only among the 5.1 percent of subjects with MMSE scores of 17 or less that a high prevalence of "organic" disorders was identified, and even among this group there was no necessary correlation between MMSE score and either dementia or other mental illness.[Folstein et al, 1985 p228-35, Kramer et al, 1985 p236-45] (see above). Therefore, intuitively it might be assumed that in response to the clinical findings Folstein would suggest either that the normative or "expected" score for the MMSE merited reappraisal or, since the primary diagnostic criterion for dementia in the 1980 DSM-III model was in fact "a loss of intellectual abilities of sufficient severity to interfere with social or occupational functioning",[DSM-III, 1980 p111] that the operationalisation of the model in "current cognitive status" should be re-evaluated. In the event, however, no downward revision of the MMSE norm was recommended and Folstein formulated "cognitive impairment", i.e. a score below 24, as a new medical condition or "objective syndrome" with "many meanings". Moreover, in what Berrios describes as the "cognitive paradigm", two years after Folstein's paper was published, "functional" disability was displaced as the primary diagnostic symptom for "dementia" and replaced by items detectable by mental testing with the MMSE or other instruments.[DSM-III-R, 1987 p107]

Berrios argued that the 1987 DSM-III-R dementia model does not distinguish reliably in "cross-sectional diagnosis" between "true dementia and it behavioural phenocopies" including depressive pseudodementia. Furthermore, he proposes that having been defined "within a very narrow cognitive compass" the paradigm discounts personality, emotional and perceptual changes which may be more characteristic of early dementia than mild memory impairment,[Berrios, 1989 p11-15] (see Chapter 3). Nor, like Kitwood who suggests that those afflicted by dementing disease become subject to the deleterious effects of a "malignant social psychology",[Kitwood, 1990 p184-5] does Berrios reject the possibility of psychodynamic interaction in the genesis of dementia. He argues, however, that non-cognitive symptoms are difficult to operationalise and that the "cognitive paradigm" was formulated in the interests of developing and refining diagnostic instruments for "research".[Berrios, 1989 p11] On the other hand, Blessed et al. had previously combined mental testing with "functional", "behavioural" and other phenomena in an instrument devised to comparing "quantitative measures" with post-mortem studies of the brain, albeit with many items reported by a family member or other informant.[Blessed et al., 1968 p806-9] It may therefore be proposed that Folstein's construction of "cognitive impairment" in the symbolic language of medicine as a "syndrome" and its formal accreditation by American psychiatry as the predominant indicator for the diagnosis of dementia in the DSM-III-R model may be perceived as retrospective legitimation by American psychiatry for the operationalisation of the DSM-III model in the ECA study solely in "objective" mental testing with the MMSE and for imputing a correlation between the scores achieved and brain pathology. Furthermore, the legitimation of the cognitive paradigm may be construed as endorsing both the Neokraepelinian model for the diagnosis of organic mental disorders and its adoption as a "case-finding" method for Alzheimer's disease.

Biologism and Alzheimer's disease

Robert Katzman is introduced into Gubrium's 1986 publication as a neurologist who "has pointed out ...... Alzheimer's presenile dementia and senile dementia as essentially the same disorder, on both clinical and pathologic bases".Gubrium, 1986 p7 Gubrium argues that the Alzheimer's Disease and Related Disorders Association (ADRDA) became the "focal apparatus for formulating, embellishing and publicly diffusing ..... the concrete personal experiences of its members and spokespersons" as a "public culture" but that "the basic facts of the realities at stake" were sorted out by "the serious scientists and practitioners (who were) hardly propagandistic".Gubrium, 1986 p112-3 However, based upon Katzman's own writings in which he made the as yet unsubstantiated statement that senile dementia, redefined as Alzheimer's disease "may rank as the fourth or fifth most common cause of death in the United States" and his employment of the terms "malignant" and "major killer" in relation to the condition,Katzman, 1976 p217 Gubrium's observation with regard to the dispassionate role of the serious scientists and practitioners may be disputed, particularly in view of evidence cited by Fox of Katzman's part in establishing the ADRDA, in serving as a member of its Executive Committee and co-chairman of the Medical and Advisory Board, and in promoting support for biomedical research rather than service provision as a priority focus for the organisation.[Fox, 1989 p82-5] Therefore whether one interprets his actions as an "elaboration of concern" or as moral or economic "entrepreneurialism", (see Chapter 2) it may be argued that in the 1970s Katzman operated as the primary "expert in legitimation" for public and private funding of Alzheimer's disease research and that subsequently others, notably Folstein, became "officially accredited definers of reality" in support of its continuation.

Berger and Luckmann in their treatise on the sociology of knowledge propose that "theories" are not merely a reflection of "underlying institutional processes" but that "institutions and symbolic universes are legitimated by living individuals who have concrete social locations and concrete social interests". Furthermore, "experts in legitimation may operate as theoretical justifiers of the *status quo* (or) as revolutionary ideologists".[Berger & Luckmann, 1966 p128] From this perspective it may be argued that Katzman's assertions with regard to senile dementia were ideological in that they were vested in "a concrete power interest" comprising those disciplines and specialties within the biomedical community with the potential to benefit from public financing of Alzheimer's disease research. Furthermore, it may also be argued that Folstein, in his article on "the meaning of cognitive impairment" in which he proposed that when the syndrome is detected, i.e. by an MMSE score of 23 or less, "clinicians search for its cause and mechanism" the source of which may be Alzheimer's disease, was acting as a theoretical justifier of the *status quo* on behalf of these same interests and the paradigms and concepts within which investigation of the redefined disease was being pursued in epidemiology and in clinical studies and practice.

In 1980, soon after Alzheimer's disease was adopted as a research focus by the National Institute on Aging (NIA), the report of a task force established by the institution to consider treatment possibilities for "mental impairment" was published. This promoted the concept of "reversible dementias", i.e. potentially treatable conditions manifesting with symptoms of senile dementia, which merited investigation with a range of diagnostic assessments.[National Institute on Aging, 1980 p260-6] Similarly in 1984 the related concept of "differential diagnosis" was expanded upon in the published deliberations of the working group on the "clinical diagnosis of Alzheimer's disease" set up by the National Institute of Neurological and Communicative Disorders and Stroke and the ADRDA under the auspices of the Department of Health and Human Services (DHHS),[McKhann, 1984 p939] (see Chapter 2). The MMSE was cited as an instrument to "confirm and document" symptoms of "mental impairment" in the NIA task force report, and both Katzman and Folstein served on the DHHS working group on the clinical criteria for Alzheimer's disease. This latter promoted the MMSE as a "quantitative aid for cognitive screening" of "possible" and "probable cases" to be determined by excluding other aetiologies with laboratory tests and the application of high technology to the ageing brain,[2] while awaiting "definite" confirmation by "histo-pathologic evidence obtained from a biopsy or autopsy".[McKhann et al, 1984 p940] The combined deliberations of the task force and working group therefore endorsed the diagnostic model operationalised by psychiatry for organic

---

[2] Such technologies include "positron emission tomography",[Chase et al, 1984 p1244, Claus et al, 1994 p454] "computed tomography",[Johnson et al, 1987 p1075] and "magnetic resonance imaging",[Davis et al, 1992 p1676]. These are in addition to the wide range of other laboratory tests, X-rays, electro-encephalograms etc. which are legitimated within the concept of "differential diagnosis".

mental disorders and it may be argued that an acknowledgment by Folstein in 1985 of inconsistencies between MMSE scores and diagnoses of such disorders would formally have undermined the validity the model. Furthermore, any suggestion that normative scores for "impairment" be lowered would have restricted what may be interpreted either as an expression of "concern" that treatable conditions should not go undetected, or as a rationale for coopting an increasing number of older people to the acute medical market based upon evidence of incipient or early dementia. It may be argued that there is little to support the premise that there is a high prevalence of treatable conditions which may only be discriminated from incurable senile dementia/Alzheimer's disease with extensive use of diagnostic technologies, and that the pathology is subject to "definite" confirmation, after death.

The clinical course of senile dementia is characterised by an "absence" of physical signs and symptoms such as pain, paralysis, visual disturbances or lack of coordination, and it is therefore unclear that it is indistinguishable simply by history and physical examination from other conditions which may enter into the differential diagnosis, since as observed in the report of a National Institutes of Health conference on dementing disease the great majority of those subjected to a full range of diagnostic testing will be ascribed a diagnosis of Alzheimer's disease.[National Institutes of Health, 1987 p4] Moreover, many other neurological disorders such as Creutzfeldt-Jakob and Huntington's disease which enter into the differential diagnosis are similarly "irreversible" and with the rare exceptions of "space occupying lesions" such as brain tumours, which commonly become manifest with specific physical signs, changes revealed by technologies such as computed tomography and magnetic resonance imaging are not necessarily distinguishable from "normal ageing".[Jacoby et al, 1980 p256, Buonanno, 1987 p58] Furthermore, this lack of distinction between age-related changes in the brain and abnormal "pathology" applies not merely to macroscopic observations but also to the neuropathological evidence which afforded the primary legitimating rationale for public investment in Alzheimer's disease research.

Alois Alzheimer first linked early onset dementia and *post mortem* pathology in the brain as evidence of a "specific disease process".[Alzheimer, 1907 p1-2] On the other hand, three years later Kraepelin wrote in a treatise on "senile and pre-senile dementias" that it was an open question as to whether Alzheimer's findings were indicative of a specific disease or merely represented one stage of a more common pathological process,[Kraepelin, 1910 p2] (see Chapter 2). Blessed et al. in 1968 reviewed the evidence described by investigators over the years with regard to the "neurofibrillary tangles and senile plaques" characteristic of Alzheimer's disease, and also identified in senile dementia, and observed that "far from being specific, such changes could be demonstrated in the brains of well-preserved old people". He also observed that "senile dementia could occur without the presence of senile plaques".[Blessed et al, 1968 p797] In their own prospective study linking dementia behaviour to brain pathology they found a "general tendency for functional incapacity to be correlated with the number of senile plaques" but there was, in some cases, discrepancy between the two measures.[Blessed et al, 1968 p808] In 1988 Crystal et al. published a paper subtitled "non-demented subjects with pathologically confirmed Alzheimer's disease" which found numerous senile plaques and other characteristic changes in six out of nine non-demented subjects. They concluded that for those aged over 75, "quantitative pathological criteria for Alzheimer's disease" did not distinguish between demented and non-demented individuals.[Crystal et al, 1988 p1682-7]

Recently, McLoughlin and Lovestone proposed that advances in molecular biology may demonstrate a "functional link" to explain the lack of anatomical correlation between plaques and tangles and variations in and between study results. They further propose that biological and genetic research indicates that the "amyloid cascade hypothesis", i.e. genetically determined triggering of neurotoxic protein deposits,[Hardy & Higgins, 1992 p184-5] remains the best model of the pathogenic process. On the other hand the development of treatments is a "distant prospect" and the first clinical benefits predicted by them lie with the detection of a "marker" to improve diagnostic accuracy.[McLoughlin & Lovestone, 1994 p439-40] However, some years previously Wurtman discussed the "amyloid deposition" and genetic models for Alzheimer's disease together with others, based upon supporting evidence, including infectious agents, toxins, and neuro-transmitters.[Wurtman, 1985 p62-74] He argued that just as a range of different hypotheses were advanced and found to have validity in the aetiology of cancer none could be entirely discounted, an argument that leaves open the possibility of psychodynamic factors in the genesis of "dementia", the *in vivo* indicator of Alzheimer's disease, as suggested by Berrios, a psychiatrist,[Berrios, 1989 p14] and Kitwood, a psychologist.[Kitwood, 1989 184]

In 1986 Gubrium observed that "those concerned are reminded repeatedly that Alzheimer's is a disease for which there is no cure ...... (and) medical researchers, service providers, and the ADRDA itself remind patients, caregivers and others that there is only one source of hope for a cure and that will be provided when it comes by medical research".[Gubrium, 1986p105-6] However, he also observed that "the descriptive tension between the distinct facts of normal ageing and Alzheimer's disease is revealed in the common caveat, namely, that all the symptoms of Alzheimer's disease may be found, to some degree, in the normal and healthy elderly (and) this applies to the full range of brain/behaviour observables".[Gubrium, 1986 p78] It may be argued that the validity of this observation is exemplified by the unresolved neuroscientific problem of defining and establishing a functional link between pathological as opposed to "normal" age related changes in the brain, and twenty years after Katzman first expressed his "belief" that "senile as well as pre-senile forms of Alzheimer *(sic)* are a single disease ...... whose aetiology must be determined, whose course must be aborted, and ultimately a disease to be prevented",[Katzman, 1976 p218] none of these objectives have been achieved. Furthermore, although there has been a very substantial allocation of resources to the National Institutes of Health (NIH) for both intramural and extramural neuroscientific research it appears that, beginning with the prospective study by Blessed et al. in the 1960s, major advances in understanding the pathological processes of senile dementia from which potential diagnostic and therapeutic hypotheses have been conceptualised, have derived primarily from research conducted in the United Kingdom rather than in the United States.[Blessed et al, 1968 p797-811, Perry et al. 1977 p189 & 1978 p1457-9, Hardy & Higgins, 1992 p184-5] Furthermore, the clinical criteria for Alzheimer's disease and methods for their determination have been formulated by groups funded by the NIH are based upon the exclusion of other potential disorders, reversible or irreversible,[McKhann et al, 1984 p939-44] employing techniques which largely fail to distinguish between physical indicators of ageing versus pathology (see above). Moreover, the "case-finding" method advocated for clinical research and practice, and adopted in epidemiological studies is one with great potential for overlap between "the distinct facts of normal ageing" and evidence of disease. On the other hand, it may be proposed that the statistical evidence derived from the latter serves to legitimate continued public financing of investigations related to Alzheimer's disease.

Gubrium, while observing that all the symptoms of Alzheimer's disease may be identified to some degree in "normal and healthy" older people, also observes that "formal descriptions of Alzheimer's disease are commonly introduced with alarm" and that "the demographic urgency of the Alzheimer's disease problem" is best captured by the slogan "silent epidemic", one that informs us that the disease is spreading in the population, and that "until now its enormity has gone unnoticed". However, "those concerned aim to make the silent epidemic known in order to marshal resources for its cure and treatment".[Gubrium, 1986 p34-7] Three years after Gubrium published these words the "alarming facts" of the "realities at stake" as derived from one study funded by the National Institute on Aging were made known in newspaper headlines: **"Alzheimer's Toll Heavier: Illness May Affect 4 Million in U.S."**,[The Washington Post, November 10, 1989] and since that time the allocation of resources by the U.S. Congress in the search for a "cure" has been sustained (see Table 2.1). On the other hand, while neuroscientists continue working to "improve diagnostic accuracy" and while therapies based upon microbiology remain a "distant prospect",[McLoughlin & Lovestone, 1994 p440] the "cognitive paradigm" legitimated by American psychiatry has been employed in a range of research financed with the same resources.

Among such research was the study conducted at Georgetown University, funded by the National Institute of Mental Health (see Chapters 5 to 8). This study afforded an opportunity for evaluating the diagnostic method adopted for detection of "cognitive impairment" which, as observed by Kramer, a biostatistician, "serves among the elderly to index the odds of being diagnosed as a case of dementia or other mental disorders that manifest in cognitive functioning".[Kramer, 1985 p236] This method, operationalised in mental testing with the MMSE in the ECA study derives from an innate, biologically determined global model of intellectual functioning which had been largely invalidated by psychology, the social science within which it had been elaborated but which has been legitimated by the diagnostic authority of the medical profession for the detection of cognitive impairment as an "objective syndrome" and as a screening method for dementia and for Alzheimer's disease, not merely in studies undertaken within the United States.

The reliability of mental testing in the elderly

In 1953, Martin Roth, a pioneer and long time investigator of mental disorders of the elderly, both functional and organic,[Roth, 1950, Roth, 1955, Blessed et al, 1968, Roth, 1976] conducted a study with a lay psychologist of patients admitted to a psychiatric hospital using several mental testing instruments including selected items from the Wechsler Adult Intelligence Scale (WAIS) (see Chapter 4). Results showed that various of the tests distinguished between subjects diagnosed with "senile psychosis" and others with "affective disorders" who manifested signs of dementing disease but showed no impairment in intellectual functioning, i.e. suffering from depressive pseudodementia. The primary tests which distinguished between the two groups were "orientation" and "knowledge of public affairs" but the ability to perform well on mental tests was relatively well preserved even in some "senile psychotics".[Roth & Hopkins, 1953 p450] In the early 1960s Roth collaborated with Kay in a community study of "old age mental disorders" based upon interviews carried out by psychiatrists using "a detailed medical and psychiatric inventory" from which diagnoses were determined by a combination of patient responses and subjective interpretation of observable items.[Kay et al, 1964 p147] However, Roth et al., in developing the CAMDEX (Cambridge Mental Disorders of the Elderly Examination), incorporated the MMSE among the diagnostic measures because of its "popularity" and adoption by the National Institutes of Health for community studies of the elderly.[Roth et al, 1986 p700] Furthermore, in the year in which cognitive impairment, as determined by an MMSE score of 23 or less was constructed by Folstein as an objective syndrome, a paper was published by Kay which employed the MMSE in a study of dementia and depression in elderly residents of Hobart, Tasmania.[Kay et al, 1985 p771-87] Since that time, in what may be construed as a continuum of evidence in support of the contention by Guimon that "changes in psychiatric fashion in the U.S.A. have had international impact",[Guimon, 1989 p33] the MMSE has been administered in studies including the following conducted in: Cambridge, England,[O'Connor et al, 1989 (a) O'Connor et al, 1989 (b)] and Cambridgeshire,[Brayne et al, 1989] in Zaragoza, Spain,[Lobo et al, 1990] in Appignano, Macerata Province,[Bonaiuto et al, 1990] and in Ospitaletto, Brescia,[Frisoni et al, 1993] both in Italy, in the Bordeaux region of France,[Dartigues et al, 1992] and in Shanghai, China.[Zhang et al, 1990] These and numerous other investigations have been variously designed to detect mental impairment indicative of dementia, defined by

Folstein as a "global deterioration of intellectual functioning",$^{\text{Folstein, 1985 p228}}$ and its associated variables.

In 1939 Wechsler proposed that intelligence is a "global capacity" composed of elements that are "qualitatively differentiable" and "quantitatively measurable". The global concept was supported by Spearman's correlational "proof" of general intelligence, "a recurrent mathematical quantity or '*g*'", which Wechsler perceived to be the "basic prop" of psychometrics.$^{\text{Wechsler, 1939 p3-8}}$ Some ten years later, based upon the use of the adult intelligence scale (WAIS) in clinical practice and research, as opposed to the "apersonal psychometrics" of group mental testing, Wechsler rejected the simple unitary concept arguing that "intelligence tests such as they are" contain several factors which are all "cognitive". He also proposed that there were other independent variables which he termed "personality" or "residual components", including "motivation", and that these accounted for a high proportion of the unexplained variance in test scores,$^{\text{Wechsler, 1950 p80-82}}$ (see Chapter 4). Similarly, it was posited by British psychologists that total intelligence scores were "at best a crude and misleading measure of global intellectual performance" and that "there is a growing recognition that what is tested, even with so-called tests of intelligence is not the subject's intellectual functions isolated from the rest of his personality".$^{\text{Brody \& Williams, 1950 p163}}$ However, as previously discussed, Berrios, argues that the "cognitive paradigm" of dementia has discounted personality and other non-cognitive factors in the interests of developing a one-dimensional model which may be operationalised more easily in diagnostic research instruments,$^{\text{Berrios, 1989 p11-12}}$ (see above).

The study conducted by the Center on Aging at Georgetown University Medical

Center, in which assessments were made of "mental status" in subjects aged 60 and over undergoing elective surgical procedures, replicated many earlier investigations whose results were ostensibly unknown or discounted when conceived before the cognitive paradigm achieved dominance (see Chapter 5). It also duplicated more recent studies in which the MMSE, internationally legitimated for operationalisation of the paradigm, was used in the Netherlands to document "global cognitive performance",[Haan et al, 1991, p599] and in Canada "because it combines high validity and reliability".[Chung et al, 1989(a) p383] Folstein, in first validating the MMSE against the WAIS noted that the instrument excluded questions concerning abnormal mental experiences and other psychiatric symptoms, but proposed that "within the cognitive realm it is thorough".[Folstein, 1975 p189] On the other hand, although Roth et al. included the MMSE in the CAMDEX because of its widespread use in the United States, they observed that the instrument failed to assess "perceptual ability and abstract thinking and permits only a rudimentary assessment of most other functions".[Roth, 1986 p700] Similarly, the task force on clinical criteria for Alzheimer's disease on which Folstein served and which endorsed his instrument for the detection of "probable" and "possible" cases proposed evaluating other cognitive processes in conjunction with the MMSE such as "language skills".[McKhann et al, 1984 p942] In the Georgetown study the MMSE was used together with the Similarities test from the WAIS as a measure of "abstraction"[Lezak, 1983 p583] and the Controlled Oral Word Association or Verbal Fluency test.[Benton & Hamsher, 1978 p11] Thus the investigation afforded the potential for comparing measures achieved for the three instruments. Furthermore, since the tests were administered once before and twice after surgery, when data were analysed and interpreted from an alternative, i.e. non-medical perspective, it was possible to evaluate the reliability of scores achieved from a single testing and the validity of norms legitimated for the detection of cognitive impairment and for the assessment of significant decline in MMSE test scores.

Wechsler had expressed diffidence with regard to the appraisal of older people in part because of "intrinsic factors" such as physical impairment and loss of visual acuity created difficulty in "interpreting the data", and also because of "extrinsic" problems including the potential for recruiting a representative sample of the older population.[Wechsler, 1954 p275] Furthermore, although he consistently observed that the ability to achieve high scores on "intelligence" tests decreases with age, he was reluctant to make conclusions with regard to "mental decline" with advancing age since such conclusions were largely dependent upon "cross-sectional and not longitudinal studies".[Wechsler, 1958 p135] The sample enrolled in the Georgetown study was not representative of the United States population aged 60 and over since a high proportion in all age groups had graduated from college and there was no evidence of the inverse correlation that prevails both in the United States and in other countries between age and education.[U.S. Senate, 1983 p35] However, the sample was equally distributed between men and women, the age range was 60-93 years, and it did include a proportion of poorly educated subjects, primarily African-American women (see Chapter 6). Moreover, the analyses of the three instruments provided statistical evidence which tended to confirm the limitations of the MMSE as a general measure of cognitive status, as perceived by Roth et al. (see above); the earlier observation by Wechsler and other psychologists that mental test results reflect a range of both cognitive and non-cognitive personality and other "residual" components; and Wechsler's concerns regarding psychometrics in older people and the confounding effects of "intrinsic factors" such as physical impairment.

O'Connor, a member of the Roth team who led a Cambridge investigation of 2302 general practice patients aged 75 and over to whom the MMSE was administered as a screening test for dementia observed that "the test includes a number of items requiring manual dexterity and at least partial vision" and that "only a small proportion of respondents were unable to attempt them".O'Connor et al, 1989(b), p94 However he did not discuss the effects of physical impairment on test scores of those who did "attempt" them. In the Georgetown study it was found that physical disability, as measured by the Activities of Daily Living (ADL) scale was a relatively powerful negative predictor of performance for the MMSE with its tests of reading, writing and drawing (see Tables 7.4 to 7.6). The validity of this finding, and also the proposition that the instrument is restricted in the functions it assesses, was supported by the relative lack of predictive power for physical disability in scores from the Verbal Fluency and Similarities tests, neither of which require sight or manual skills. At the same time there was a slightly greater influence of "social activity" variables on performance for these latter instruments (see Chapter 8), a finding which may indicate an interaction effect between mental test scores and unquantified variables such as personality, motivation, or unease with the interview and assessment process, etc., factors which may also be related to the most powerful predictor of test scores in the Georgetown study, i.e. education. For example, O'Connor observed that "many poorly educated respondents felt threatened when asked to complete serial sevens, spell WORLD backwards and write a sentence,[O'Connor, 1989(b), p94] and in another paper he reported that "education and social class" influenced scores on all sections of the MMSE except one. He therefore proposed that "performance anxiety" may contribute to poor scores in working class populations and that in contrast "well educated, middle-class subjects usually expected to do well and tackled the test with confidence".[O'Connor, 1989(a) p775]

Lezak has highlighted a 1940s study as seminal in relating specific functions to different areas of the brain, in generating the discipline of neuropsychology, and in redefining the nature of "intelligence", replacing the concept of a global or unitary variable with a multi-dimensional model of cognitive and other independent functions.[Lezak, 1983 p20-21] The study, by Aita et al., found that in administering the WAIS to wartime casualties with head injuries, both to a "surgical" and a "control" group some verbal and "general information retrieval" items were relatively unaffected by severe brain trauma.[Aita et al, 1947 p43] Similarly, Hebb observed in other studies that the effects of removing large areas of brain tissue were often "astonishingly small" and that "the level of intelligence test performance is a function of the concepts a patient has already developed", based upon "a major effect of experience".[Hebb, 1949 p299-32] In the Georgetown study it was found that the number of years of formal education was the most significant independent predictor of initial performance for all three assessment instruments but that improvement in test scores of ten percent or more occurred with repeated testing among five percent of subjects for the MMSE, thirty-two percent for the Verbal Fluency test, an instrument of unlimited range, and twenty-three percent for Similarities testing, the initial scores for which were most strongly associated with years of formal education (see Tables 7.18, 8.18, and 8.29). However, the improvements tended to be greatest among those with the least education and/or the lowest baseline scores. The study therefore indicated that results derived from a single screening assessment based upon any one instrument may be unreliable for up to a third of older subjects whose improvement arguably derives from a learning experience which is greatest for those with little previous knowledge or understanding of the mental testing concept, and also from an associated reduction in "performance anxiety". On the other hand there was further evidence from the repeated testing that MMSE scores may be independently related to physical impairment.

In the Georgetown data there was little to support the "officially accredited" indicator for the "onset" of decline, i.e. a decrease of two MMSE points, as adopted in other similar surgical studies,[Mondimore,et al, 1983 p930, Chung et al, 1989(a) p383, Haan et al, 1991 p599] since a change in mean MMSE test scores of this dimension, was accompanied by an increase in mean scores for the other two instruments (see Table 8.30). However, a three-point or ten percent (3/30) fall in MMSE scores was not associated with improvement in Verbal Fluency or Similarities performance, but was correlated with an increase in the measure of disability. Furthermore, a ten percent improvement in MMSE scores following surgery, which included cataract extraction, was associated with a decline in mean ADL scale values. The number of subjects with a three-point change, either positive or negative, was small (see Table 7.18), and impairment in functional capacity is not independent of the effects of dementing disease, but since physical disability was also identified as a predictor of MMSE scores for the total sample it may be argued that it is a factor which merits greater attention in conducting and assessing the results of screening for cognitive impairment. However, Blessed et al. in a study comparing rates of diagnosis of dementia using the computerised GMS/AGECAT system,[Copeland et al, 1986] (see Chapter 3) and the CAMDEX which includes the MMSE, identified hearing and visual deficits as impeding the achievement of cut-off scores for the "cognitive examination" section of the latter, scores by which subjects would be excluded not merely from further investigations based upon screening for cognitive decline but also from an actual diagnosis of dementia employing "objective" methods.[Blessed et al, 1991 p193]

Berrios argues that the cognitive paradigm for dementia has been conceptualised within a narrow compass which discounts personality, volition, emotional experience or other psychological variables which are difficult to operationalise in screening and diagnostic instruments.[Berrios, 1989 p11-12] However, based upon the reviews conducted for this discussion and the analyses of the empirical data it may be proposed that the instruments developed and the results presented also tend to discount not merely psychological but identifiable physical or biological factors as in assessments with the MMSE. On the other hand, in what may be construed as legitimation both for the cognitive mental testing model and for the Neokraepelinian paradigm for mental disorders, both functional and organic, and the correlational diagnostic model formulated by its proponents, a "biological" aetiology has now been imputed for the effects of "education". This variable, which serves as a proxy for lifetime income and is inversely correlated with disability and poor health status,[U.S. Senate, 1983 p375] has more generally been perceived as a "social" factor reflecting differential childhood experiences and opportunities and interpreted as intervening in the reliability of cognitive screening methods, but has now been redefined as a "risk factor" for "dementia",[Mortimer et al, 1993 pS39] and for Alzheimer's disease.[Zhang et al, 1990 p428]

The meaning of education and the cognitive paradigm of dementia

Jorm et al. in a 1987 "quantitative integration of the literature" on the "prevalence of dementia" identified forty-seven studies conducted between 1945 and 1985 of which half were published in the five years after the DSM-III models were disseminated by the American Psychiatric Association in 1980. They found that the actual levels of dementia varied greatly from one study to another since there were differences in methodological variables such as the "definition of dementia", the design of the study, characteristics of the sample and methods of assessment and diagnosis. On the other hand, they found that for twenty-two of the studies which provided comparable data the relationship between prevalence and age was remarkably consistent with rates doubling every five years up to age 95. The mean rates derived from these studies were: 1.4 percent in the age group 65 to 69, 2.8 percent in the age group 70 to 74, 5.6 percent in those 75-79, 10.5 percent in the age group 80-84, and 20.8 percent in the age group 85-89.[Jorm et al, 1987 p472-6] Similarly, in community based screening studies for "cognitive impairment", in which the MMSE was the predominant instrument used, age was the primary predictor of test scores.[Folstein et al, 1985, Weissman et al, 1985, Scherr, 1988, Dartigues et al, 1992,] However, the proportion identified as impaired, i.e. scoring below the normative cut-off level of 23 points, has been substantially higher than any prevalence estimates for dementia, ranging from 9.3

percent among those aged 65 to 74, to 40.6 percent in those aged 85 and over in the New Haven site of the ECA study funded by the National Institute of Mental Health.[Holzer et al, 1984 p26] Furthermore, as previously discussed, Folstein reported that although twenty percent of those aged 65 and over were detected with what may be defined as "MMSE syndrome" in view of the limitations of the instrument, only twenty-six percent were later ascribed a diagnosis of dementia, in the East Baltimore site of the same investigation, [Folstein, 1985 p228] (see above).

The Georgetown University study sample was unusual if not unique in its educational characteristics and education rather than age was identified as the most powerful predictor of mental test scores, a finding which persisted even when those with limited formal schooling were excluded from the analyses. However, a significant correlation between lack of education and low mental test scores in screening for cognitive impairment has been found in most studies including those cited above, and this was originally interpreted as indicative of "selection bias".[Kittner et al, 1986 p164] Therefore the initial debate focussed on rates of "false positives" and "false negatives" and the adjustment of norms to achieve an optimal trade-off between "sensitivity" in detecting all possible "cases" and "specificity" in avoiding the selection of non-cases. On the other hand, recommended adjustment of scores or other changes, some of which have blurred the distinction between the use of the MMSE or other instruments as screening tools or as diagnostic methods, have varied according to the objectives of the research team and the characteristics of the population under investigation.

Folstein himself participated in the first widely reported study to identify what was perceived as the confounding effects of education on MMSE scores when employing the instrument as a screening test for "delirium and dementia in patients on a general medical ward".[Anthony et al, 1982 p397] The investigators in Baltimore found that among those with less than nine years of education thirty-nine percent with a score lower than 24 out of 30 were "diagnosed" as "false positives". They therefore suggested that the scores might be "modified to make it a more specific test for individuals with little education".[Anthony et al, 1982 p406] A study by Murden et al. on the use of the MMSE in two public clinics in New York attended by elderly patients concluded that greater "specificity" was derived from "using a score of 23 or less to define dementia" in better educated groups and "using 17 or less to define dementia" in the "poorly educated".[Murden et al, 1991 p149] However, both the Baltimore and the New York samples had a limited range of formal schooling, and others have sought to address the question of "sensitivity" in detecting "mild to moderate impairments" and more challenging alternatives to the MMSE have been proposed. For example it was reported that the Neurobehavioural Cognitive Status Examination which includes tests of "calculation", "language skills", and "verbal reasoning" can "reduce the forty-three percent MMSE false negative rate to seven in clinical practice".[Schwamm et al, 1987 p486-90] Furthermore, it is not merely in the United State, where a high rate of false negatives among better educated subjects may be interpreted as a loss of market potential for clinicians that there has been concern with the sensitivity of the MMSE, the instrument which has achieved the highest level of international legitimation. For example, Roth, in discussing the "operational criteria for the diagnosis of dementia" in England proposed that the identification of "mild dementia" presents a "problem of central importance in this field".[Roth et al, 1986 p699]

O'Connor, a member of the Roth team who led the community study of more than 2,000 residents aged 75 and over in the university city of Cambridge, in which the MMSE was evaluated as a screening test for dementia, proposed that more elaborate sampling procedures will be needed "to circumvent the tendency for well educated, mildly demented subjects to score above the customary cut-off point of 23/24".[O'Connor, 1989 (b) p96] However, Kay et al. in their study in Hobart, Tasmania reported that although the MMSE identified all cases of "moderate and severe" dementia at a "cutting point of 23/24" when the "criteria were stretched to include 'mild' cases of dementia" it was not until "a cutting point of 26/27" that "the MMSE identified all DSM-III cases" but that this latter score also yielded 41.5 percent "false positives".[Kay et al, 1985 p779] On the other hand, based in part on the interpretation given to the results of a major Sino-American epidemiological investigation, the debate with regard to the sensitivity and specificity of the MMSE and the rates of false positives and false negatives derived from its use as a screening instrument has been undermined. Robert Katzman, arguably the most highly accredited expert in legitimation for Alzheimer's disease research (see Chapter 2), was an investigator in the study which accorded greater validity for mental testing scores by ascribing "biological" causation to the intervening effects of education.[Zhang et al, 1990 p428]

In this survey of "the prevalence of dementia and Alzheimer's disease" in Shanghai, China, funded jointly by the National Institute of Mental Health and Chinese public health authorities, the MMSE was administered as the screening instrument to a sample of 5,055 elderly persons of whom twenty-five percent, primarily women, had received no schooling and none were reported to have advanced beyond high school. In the report it was observed that "subjects without formal education were usually less familiar with the procedures and contents of psychometric tests developed in Western countries, and they exhibited less confidence and motivation during testings".[Zhang et al, 1990 p435-6] However, based upon the administration of several additional mental tests, a psychiatric interview and neurological examination it was proposed that "female gender and low education" are each highly significant and independent risk factors for dementia". This generated the hypothesis that "lack of early education (or education as a surrogate for nutrition or other socioeconomic factors).... lowers 'brain reserve'".[Zhang et al, 1990 p436]

In 1993 Katzman further elaborated on "education and the prevalence of dementia and Alzheimer's disease". He discussed the Shanghai findings and noted that several studies had reported that individuals with a lack of education "were more likely to develop dementia and Alzheimer's disease" and questioned whether "the effect of education on the prevalence of dementia is a surrogate for other childhood deprivation" or whether "education actually increases neocortical synaptic density".[Katzman, 1993 p18-9] More recently, however, it has been proposed from mental testing studies of elderly monozygotic and dizygotic twins using the MMSE that educational correlations with test scores are largely genetically determined. In a study funded by the National Heart Lung and Blood Institute, another of the institutions deriving resources from those allocated for the investigation of Alzheimer's disease (see Table 2.1), Spearman's "*g*" or general ability factor was revived as a determinant of MMSE performance, the final results of the analyses suggesting "genetic effects as the mediating causal pathway for the education-cognition relationship without the presence of intervening environmental variables".[Carmelli et al, 1995 p52] On the other hand, the investigators in a Swedish-American study which also tested fraternal and identical twins whose ages ranged from 50 to 88 proposed that "correlations between MMSE and cognitive ability in both sexes are primarily due to genetic covariation, whereas those between MMSE and education reflect genetic covariation for men, and both genetic and environmental influences for women".[Pedersen, p61, 1996] Such studies provide legitimacy for the revival by Bleecker et al. of the 1900s concept of "mental age", proposing "age-specific norms" for the MMSE beginning at age 40.[Bleecker et al, 1988 p1565] However, these deterministic models do not explain the increase of four years in the median level of education in older people over the past thirty years, associated with government investment and expanded opportunities in the post-War period,[Mortimer et al, 1993 pS39] and evidence from a longitudinal study begun in the United States in 1982 that there is a downward trend in the prevalence of dementia in spite of increased life expectancy.[Manton et al, 1995 pS201] Nor, do they explore the extent to which mental testing scores are associated with "dementia behaviour" or the functional link between such behaviour and its organic basis.

The "biases of psychiatric diagnosis" and the cognitive paradigm for dementia

Folstein, in his 1985 paper in which cognitive impairment was formulated as "an objective syndrome" with many meanings, stressed that "the MMSE does not make a diagnosis".[Folstein, 1985 p233] It may be argued, however, that the development of biological and genetic models for the outcome of mental testing confers ever increasing legitimacy on the MMSE as both a screening and a diagnostic tool, a legitimacy which, when combined with failure to operationalise fundamental behavioural change in a model defined within a narrow cognitive compass,[Berrios, 1989 p11-12] lends support to Guimon's observation that the Neokraepelinian school of biological psychiatry has tended to "elevate symptoms to diseases, and to abuse neuro-biological explanations".[Guimon, 1989 p33-5]

In the investigation conducted in East Boston under the auspices of the National Institute on Aging in the multi-site Established Populations for Epidemiologic Studies of the Elderly (EPESE), the report states that the results were arrived at using diagnostic criteria "for primary degenerative dementia" from the DSM-III. However, although the screening examinations, consisting of "brief performance tests of selected areas of cognitive functioning"[Scherr er al, 1988 p1100] other than the MMSE, were conducted in the subjects' homes, no attempt was made to determine what proportion of those with detected impairment met the primary DMS-III diagnostic criterion for dementia, i.e. "a loss of intellectual abilities sufficient to interfere with social or occupational functioning".[DSM-III, 1980 p111] This omission was justified by the statement that "it is difficult to apply this criterion in a uniform, meaningful way".[Evans et al, 1989 p2553] Similarly in Shanghai, China, where the MMSE was used as the screening instrument, it was observed that most older people lived with their children who perform the household tasks so that it is "difficult for psychiatrists to arrive at a clinical judgment of functional impairment".[Zhang, 1990 p436] On the other hand, a diagnosis not merely of dementia but of Alzheimer's disease was imputed to a percentage of the study participants and major pronouncements were made with regard to its distribution. From the East Boston study, age-specific prevalence estimates, higher than previously found in any community study of dementing disease were derived and extrapolated to the entire elderly population of the United States.[The Washington Post, November 10, 1989] In the Shanghai investigation, "female gender and low education" were determined to be significant and independent risk factors" for "dementia and Alzheimer's disease".[Zhang et al, 1990 p428]

In the absence of consistently operationalised criteria it is unclear that psychiatric diagnosis itself is not associated with the educational and cultural variables which were identified in the various screening procedures such as: "performance anxiety" among the working class in Cambridge;[O'Connor, 1989(a) p775] the translation of the questionnaire instrument in East Boston "for those who mostly spoke Italian";[Evans et al, 1989 p2552] and the observation in Shanghai, China that "subjects without formal education were usually less familiar with the procedures and contents of psychometric tests developed in Western countries, and they exhibited less confidence and motivation during the testings".[Zhang et al, 1990 p435-6] Furthermore, the potential for diagnostic bias was confirmed by O'Connor who observed that "even when diagnoses are made by psychiatrists, less-intelligent people" pose diagnostic problems. For example he cited a study which found that four out of seven community residents with "suspected chronic brain syndrome" showed no evidence of "actually dementing" four years later.[O'Connor et al, 1989(a) p775] Similarly Teri et al., based upon the results of a recent "longitudinal investigation of risk factors for accelerated decline" in Alzheimer's disease, propose that there may be class differentials in assigning a diagnosis. In this study of American subjects from a range of socio-economic backgrounds it was found that following diagnosis those with higher education had more rapid annual deterioration than the less well educated. They acknowledged the proposition by Katzman that lack of education may affect "reserve capacity" in the brain, allowing the symptoms of dementia to appear earlier, but suggest that more rapid decline may simply indicate that the better educated "escape early detection by performing well on cognitive screening instruments".[Teri et al, 1995 M53-4]

O'Connor, in responding to the imputation of biological causation for low mental test results in those with limited education, albeit prior to the advancement of recent genetic models,[Carmelli, et al, 1995, Pedersen et al, 1996] (see above) argues that "lower socio-economic groups have higher rates of many disorders (but) there is no evidence to date that dementia is actually commoner in working class populations".[O'Connor et al, 1989(a) p775] This argument is supported by the results of a "case-control" study of nearly two-hundred and fifty patients identified with dementing illness in the total community Rochester Epidemiology Project, Minnesota between 1975 and 1979 whose investigators found a lack of association between the diagnosis and education or occupation.[Beard et al, 1992 p2063] However, the evidence reviewed for this discussion does indicate that lower socio-economic groups, such as residents of the working class neighbourhood of East Boston, are at higher risk of selection for more testing and investigation in screening studies. Furthermore, although the middle class and the well educated are not immune to dementing disease, within the "cognitive paradigm", with its failure to operationalise consistent, broader diagnostic criteria taking account not merely of personality, motivational and related psychological factors and other psychiatric disorders and physical functioning, intuitively it may be assumed that there will be biases in the diagnostic process. However, as demonstrated by Blessed et al, even more broadly conceived, standardised instruments such as the CAMDEX and the GMS/AGECAT systems yield different diagnostic rates and "false positives" associated with visual and other physical impairment.[Blessed et al, 1991 p193]

Scull observed that psychiatry in the 19th century "abandoned references to real behaviour in favour of surmises concerning future behaviour", and that it was able "to extend the boundaries of mental pathology to encompass marginally deviant affective symptoms and make a plausible case for the reality of partial sanity",[Scull, 1989 p248] It may be argued that in the late 20th century psychiatry and related disciplines have abandoned references to "real dementia behaviour" and have extended the boundaries of mental pathology to encompass marginally deviant "cognitive" symptoms, i.e. failure in mental tests, citing the results as evidence of the "reality of partial sanity", either incipient or "mild" dementia and "possible" or "probable" Alzheimer's disease", conditions in which the neuropathology is not entirely distinguishable from that of "normal ageing" and for which conceptual models have yet to be validated in terms of their functional link with *in vivo* behaviour.

Chapter 10

THE ALZHEIMERISATION OF AGEING

"And one man in his time plays many parts .....
..... last scene of all,
That ends this strange eventful history,
Is second childishness, and mere oblivion,
Sans teeth, sans eyes, sans taste, sans everything."[1]

Gubrium observed in 1986 that "the basic and experiential facts of Alzheimer's disease are not vested fictions (but) the relationship between fact and interest is more complicated. They are in constant dialogue, one being what it is for those concerned because of the other. The Alzheimer's and Related Disorders Association (ADRDA) 'orchestrates' public awareness of the disease because its members believe and have the facts to warrant that a disease and human needs exist. The orchestration, in turn, reveals the realities at stake for what they obviously are taken to be, facts of apparently alarming proportions. Alarm is indicated precisely because of the enormity of the problem, demographically, financially, personally ..... the facts are unintentionally constituted by, yet constitutive of, the related activity of those concerned, something certainly not the focal activity of a medical establishment".[Gubrium, 1986 p207] However, the evidence indicates that the establishment of the ADRDA itself and its activities were indeed orchestrated by members of the medical and biomedical research establishment: Robert Katzman, who formally proposed that "we should drop the term 'senile dementia' and include these cases under the diagnosis of Alzheimer's disease",[Katzman, 1975 p106] describing the redefined

---

[1] William Shakespeare: As You Like It, Act II

disease as "malignant", a "major killer", and the fourth or fifth leading cause of death in the United States; Donald Tower, director of the National Institute of Neurological and Communicative Disorders and Stroke who "agreed that dementia was a public health problem";[Fox, 1989 p76] and Robert Butler, a psychiatrist and first director of the National Institute on Aging, who adopted Alzheimer's disease as a research focus for the instititution and collaborated with Katzman in establishing the ADRDA, perceiving the need to develop "a public constituency to be an advocate for the Alzheimer's disease research cause",[Fox, 1989 p79-82] (see Chapter 2). However, even interpreting all orchestrations as expressions of concern with the realities at stake, twenty years after Katzman's initial proposal, it is not unreasonable to ask what has been achieved to ameliorate the concerns of those afflicted and their families.

Gubrium observed that the demographic urgency of the Alzheimer's disease problem is captured in the promotional literature of "the related movement", in one of its most frequently used slogans: "the silent epidemic". "Those concerned aim to make the silent epidemic known in order to marshal resources for its cure and treatment. Its demography alone warrants focussed attention in medical research and practice. Funding agencies are asked to take cognisance of this."[Gubrium, 1986 p37] However, with supporting evidence of increasing life expectancy and projections of growth in the proportion of elderly people in the population into the next century (see Tables 1.1 to 1.3) the United States Congress did respond to the demographic urgency with substantial public financing of Alzheimer's disease research with resources allocated to the National Institute on Aging (NIA) for distribution to other institutions (see Table 2.1). This funding has supported neuroscientific, clinical and epidemiological investigations, and with the latter the dimensions of "the silent epidemic" were documented, notably in the East Boston study sponsored by the NIA. This produced prevalence estimates of Alzheimer's disease of three percent in those aged 65 to 74, but rising to forty-seven percent in those aged 85 and over the most rapidly growing population group.[Evans et al, 1989] Based upon these figures, demographers at the NIA projected that within fifty years 14 million Americans "could have the disease", and executives of the ADRDA pronounced that the "financial and public health implications could affect us all", observations which were widely disseminated to the public.[The Washington Post, November 10, 1989]

Bond et al. argue that epidemiological studies are necessary for the "planning of services and policy making (but) it is obviously necessary to have sound diagnostic criteria".[Bond et al, 1993 p41] However, it may be argued that it is also necessary to have a representative sample and, as previously discussed, it is unclear that either of these preconditions were met in the East Boston study (see Chapter 9) the results of which may be interpreted as undermining rather than supporting potential policies. For example, in 1987 the research office of the U.S. Congress published a report entitled "Losing a Million Minds: Confronting the Tragedy of Alzheimer's Disease and Other Dementias", produced out of "Congressional concern with the plight of those suffering from dementing disease", in which a number of federal policy options were proposed to expand the eligibility and long-term care provisions of the Medicare programme,[U.S. Congress, 1987 p467] (see Chapter 2, Footnote 7). However, since that time none of these provisions have been enacted to relieve the plight of the sufferers or their families, a reluctance which may be interpreted as deriving from an unwillingness to expand the open-ended commitment of universal medical insurance for the elderly when confronted by current and future projections of the potential increase in its growth. In consequence, the focus has been on the search for a "cure".

Gubrium writes that those concerned are repeatedly reminded that Alzheimer's disease is a disease for which there is no cure and that "medical researchers, service providers, and the ADRDA itself remind patients, caregivers and others that there is only one source of hope for a cure and that will be provided when it comes, by medical research but as Gubrium himself discusses, "the factual status of Alzheimer's as a distinct entity separate from the varied experiences of normal ageing is not obvious".[Gubrium, 1986 p3] However, although there is evidence that neuropathology identified as characteristic of Alzheimer's disease can be found in the brains of "well preserved" old people, and that senile dementia can occur without the identification of such pathology,[Blessed et al, 1968 p797] McLoughlin and Lovestone propose that the "amyloid cascade hypothesis", based upon a theory of genetically determined triggering of neurotoxic protein deposits, is the best aetiological model for a pathogenic process linked to behaviours which in combination are described as Alzheimer's disease. On the other hand, they observe that although genetic research may improve "diagnostic accuracy" developing treatments based upon molecular biology is a "distant prospect".[McLoughlin & Lovestone, 1994 p439-440] In the meantime, however, some 50,000 Americans are currently taking Cognex (tacrine) the first drug to be approved by the U.S. Food and Drug Administration for Alzheimer's disease, the development of which is derived from another model, i.e. the "neurotransmitter" theory of senile dementia.[Perry et al, 1978 p1459] The only alternative to ingestion of this drug which is associated with significant side effects and little therapeutic benefit (see Chapter 2, Footnote 9) is for those afflicted to undergo extensive diagnostic investigations, based upon the legitimating rationale of "differential diagnosis" in the hope of detecting one of the very few "curable" alternative conditions which may present with symptoms of dementing disease. However, the case-finding model developed by American psychiatry implies that there are significant benefits to be derived from early detection of cognitive symptoms indicative of dementing disease and the model has been widely adopted in questioning the mental competence of older people.

In clinical practice, Omer et al. argue that "the importance of routine cognitive measurements" in medical wards cannot be over stressed".[Omer et al, 1983 p266] Litovitz et al. suggest that "clerks and secretaries" might be employed to administer the MMSE to those seeking treatment in emergency departments.[Litovitz et al., 1985 p400-1] Similarly U'Ren urges that "dementia" is not overlooked in primary care and recommends "routine screening".[U'Ren, 1987 p49] In studies funded by the National Institute of Mental Health elderly people have been investigated in Shanghai, China,[Zhung et al, 1990] as well as in Baltimore, Maryland,[Folstein, 1985] and other sites of the Epidemiologic Catchment Area (ECA) study. The medical profession in other countries has set up what Illich described as "dragnets to apprehend those needing care and protection",[Illich, 1976 p91] in Cambridge, and Cambridgeshire, England;[O'Connor et al, 1989, Brayne et al, 1989] in the Bordeaux region of France,[Dartigues, 1992, p981] Zaragoza, Spain,[Lobo et al, 1990 p] Appignano, Macerata Province, Italy,[Bonaiuto, 1990 p346] the town of Ospitaletto, Brescia, also in Italy,[Frisoni, 1993 pS310] and Tasmania, Australia.[Kay et al, 1985 p779] In these studies, the list of which is by no means exhaustive, some have sought to detect "cognitive impairment" others "dementia" or "Alzheimer's disease", but whichever construct has been investigated both their conduct and the results may be perceived as promoting negative images both of specific demographic groups and of old age.

In their discussion of the images of ageing, Featherstone and Hepworth propose that many of the images used to describe ageing and elderly people are "negative stereotypes" and these are related to the traditional image of "senility", a frightening vision of dribbling, drooling, incontinence, an absence of physical coordination and the inability to comport oneself or to communicate coherently. It is the image of one who is less than fully human.[Featherstone & Hepworth, 1993 p313] Similarly, Kitwood, in his discussion entitled the "dialectics of dementia", argues that "there is one form of distress that is faced with particular dread ...., it is that of 'going senile' ..... developing one of those dementing illnesses whose course seems to be little more than an inevitable path to degrading incapacity and, if life continues that long, eventually to a near-vegetable existence".[Kitwood, 1989 p1] These, however, are the images projected onto, and the potential fears aroused, not merely in those diagnosed as "cases", by what may be perceived as "actuarial activities" which "transform people who feel healthy into patients anxious for their verdict".[Illich, 1976 p90-2] In Shanghai, "female gender and low education" were identified as highly significant risk factors for dementia. In the final ECA report "education" and "race" were observed to be strongly related to "cognitive impairment" which "is concentrated among persons with eight or fewer years of education", "whites" being reported to "have much lower prevalence than blacks or Hispanics".[George et al, 1991 p301-2] In Appignano, and Ospitalleto, Italy "farmers", "housewives" and "blue collar workers", and in Bordeaux "farmworkers", "domestic service employees" and "housewives" were identified as having the highest rates of cognitive impairment and/or Alzheimer's disease.[Bonaiuto, 1990 p346, Dartigues, 1992 p981, Frisoni, 1993 S313] In Cambridge, England, however, the concern was with the ability of the well educated to evade detection and, since the "overall cost to the project of finding demented subjects was approximately 400 per head" which was not perceived to be "excessive", there would be "considerable financial benefits" if ways could be found of "increasing the true positive rate without sacrificing sensitivity".[O'Connor, 1989(b) p95] Such investigations are therefore not merely intrusive and demeaning to older populations, but costly, and as demonstrated by the failure of the U.S. Congress to increase service provision on behalf of those afflicted, generate evidence which may prevent the enactment of policies for those facing what Gubrium describes as "individual financial alarm".

Walker in a discussion of the international adoption of the economic "burden of ageing" argument to legitimate "anti-welfare policies" proposed that the theory has generated "an increasingly prevalent assumption that societal ageing should be a source of ...... remorse rather than pride",[Walker, 1990 p378] (see Chapter 1). However, economists do not have what Starr describes as the broad cultural authority, economic power and political influence" of the American medical profession, deriving in part from its bonds with modern science,[Starr, 1982 p4-6] nor do they have what Freidson describes as "a self-deceiving view of the objectivity and reliability of its knowledge and the virtues of its members",[Freidson, 1973 p370] and data presented by economists and policy proposals will therefore be scrutinised and debated by other members of the discipline.[2] On the other hand, the evidence with regard to Alzheimer's disease would seem to support Freidson's contention that at the heart of most questions of social policy bearing on the professions lies the central question of "autonomy" for it bears on "who may determine what the problem is, how

---

[2] Under the heading: "Social Security: It is affordable", a column appeared on the editorial pages of the Washington Post, written by an economist and a former commissioner of the Social Security Administration. This was a response to a previous report which it described as "fomenting unfounded hysteria" with regard to the financing of retirement pensions in the 21st century. They argued that the "true measure of the burden of the dependent population is the ratio of all dependent, young and old, to active workers" and that the "dependency burden" is never projected to be as high as it was in the 1960s.[The Washington Post, February 15, 1994] On the other hand when members of the medical profession use quantitative data from a study conducted in a small working class neighbourhood and extrapolate these figures to the entire U.S. population and to future generations, i.e. "the National Institute on Aging estimates that within 50 years as many as 14 million Americans could have the disease",[The Washington Post, November 10, 1989] the projections are not perceived as "fomenting unfounded hysteria".

the problem is to be dealt with and what price is to be paid for dealing with it".[Freidson, 1973 p369]

Legislation to establish the National Institute on Aging was passed by the U.S. Congress in the Research on Aging Act, 1974 with a mandate to develop "a plan to coordinate and promote research into the biological, medical, psychological, social, educational and economic aspects of ageing", base upon a perception that "biomedical research could help prevent old age disabilities and save taxpayers'money.[Lockett, 1983 p182] However, senile dementia was identified by the medical researchers as the "problem", this was to be dealt with by adopting the condition, redefined as Alzheimer's disease as a focus for the ageing institute, and the price to be paid for dealing with it has been what Adelman, former President of the Gerontological Society of America perceives as a "blatant shift in priority" in conflict with "the language and intent with which Congress established the National Institute on Aging",[Adelman, 1988 pB1-B2] and ultimately to the "Alzheimerisation of Ageing",[Adelman, 1995 p526] the negative benefits of which have not been restricted to the older population of the United States of America.

REFERENCES

Abel-Smith, B. Value For Money in Health Services. London, Heinemann, 1976.

Aberg T, Kihlgren M. Effect of open heart surgery on intellectual function. Scand J Thorac & Cardiovasc Surg Supp 15:4-63, 1974.

Aberg T, Kihlgren M. Cerebral protection during open-heart surgery. Thorax 32:525-533, 1977.

Abram HS, Gill BF. Predictions of postoperative psychiatric complications. N Engl J Med 265:1123-8, 1961.

Abrams P. The Origins of British Sociology: 1834-1914. Chicago and London, University of Chicago Press, 1968.

Adelman RC. The importance of basic biological science to gerontology. Journal of Gerontology 43:1:B1-2, 1988.

Adelman RC. The Alzheimerization of aging. The Gerontologist 35:526-532, 1995.

Aita JA, Armitage SG, Reitan RM et al. The use of certain psychological tests in the evaluation of brain injury. Journal of General Psychology 37:25-44, 1947.

Albert ML. Subcortical dementia. In Katzman R, Terry RD, Bick KL (eds), Alzheimer's Disease: Senile Dementia and Related Disorders. Aging 7. Raven Press, 1978.

Alford RR, Friedland R. Powers of Theory: Capitalism, the state and democracy. Cambridge University Press, 1985.

Alzheimer A. Uber eine eigenartige Erkrankung der Hirnrinde. Allegemeine Zeitschrift fur Psychiatrie und Psychisch-Gerichtliche Medizin 64:146-48, 1907. (About a peculiar disease of the cerebral cortex: Translation by the Western Geriatric Research Institute, 1987).

American Medical Association: Physician characteristics and distribution in the U.S. Chicago, Illinois, 1993.

Anthony JC, LeResche L, Niaz U, et al. Limits of the 'mini-mental state' as a screening test for dementia and delirium among hospital patients. Psychological Med 12:397-408, 1982.

Applegate WB, Miller ST, Elam JT, et al. Impact of cataract surgery with lens implantation on vision and physical function in elderly patients. JAMA 257:1064-6, 1987.

Armstrong D. Political Anatomy of the Body: Medical knowledge in Britain in the twentieth century. Cambridge University Press, 1983.

Arrow, KJ. The welfare economics of medical care. In: Cooper M.H. and Culyer A.J. (eds). Health Economics. Penguin Books, 1973.

Barr N, Whynes D. Introductory Issues. In: Barr N, Whynes D. Current Issues in the Economics of Welfare. New York, St Martin's Press, 1993.

Beard MC, Kokmen E, Offord KP, et al. Lack of association between Alzheimer's disease and education, occupation, marital status, or living arrangement. Neurology 42:2063--2068, 1992.

Bedford PD. Adverse cerebral effects of anaesthesia on old people. Lancet 2:259-63, 1955.

Benton AL, Hamsher KdeS. Multilingual aphasia examination. Iowa City: University of Iowa Press, 1978.

Benton AL, Eslinger PJ, Damasio AR. Normative observations on neuropsychological test performances in old age. J Clin Neuropsychology 3:33-42, 1981.

Berg R., Franzen M., Wedding D. Screening for Brain Impairment: A manual for mental health practice. New York, Springer Publishing Company, 1987.

Berger PL, Luckman T. The Social Construction of Reality: A treatise in the sociology of knowledge. Anchor Books, Doubleday, New York, 1966 (Reprinted 1989).

Berggren D, Gustafson Y, Eriksson B. Postoperative confusion after anesthesia in elderly patients with femoral neck fractures. Anesth Analg 66:497-504, 1987.

Berkman LF. The association between educational attainment and mental status examinations: of etiologic significance for senile dementias or not? J Chron Dis 39:171-4, 1986.

Bernth-Peterson P. Outcome of cataract surgery IV: socio-economic aspects. ACTA Ophthalmol 60:461-8, 1982.

Berrios GE. Non-cognitive symptoms and the diagnosis of dementia: historical and clinical aspects. British Journal of Psychiatry 154 (suppl4):11-16, 1989.

Bigler D, Adelhoj B, Petring OU et al. Mental function and morbidity after acute hip surgery during spinal and general anaesthesia. Anaesthesia 40:672-6, 1985.

Binney E.A., Estes C.L. The retreat of the state and its transfer of responsibility: the intergenerational war. Int. J. Health Services 18:83-96, 1988.

Binney EA, Swan JH. The political economy of mental health care for the elderly. In: Minkler M, Estes CL (eds) Critical Perspectives on Aging. Amityville N.Y.: Baywood Publishing Co., 1991.

Birnbaum, N. Toward a Critical Sociology. New York, Oxford University Press, 1971.

Blalock, H.M. Social Statistics, (Revised Second Edition). St Louis, McGraw-Hill, 1979.

Blazer DG. Depression in Late Life. St Louis, Toronto, London: The C.V. Mosby Company, 1982.

Blazer DG. Depressive Illness In Late Life. In America's Aging: Health in an Older Society. I.O.M. National Academy Press, Washington DC, 1985.

Blazer DG., Federspiel CF, Ray WA, Schaffner W. The risk of anticholinergic toxicity in the elderly: A study of prescribing practices in two populations. J Gerontol 38:1:31-5, 1983.

Blazer DG, George L, Landerman R et al. Psychiatric disorders. Archives of General Psychiatry 42: 651-6, 1985.

Blazer DG, Hughes DC, George LK. The epidemiology of depression in an elderly community population. The Gerontologist 27:3:281-287, 1987.

Bleecker ML, Bolla-Wilson K, Kawas C et al. Age-specific norms for the Mini-Mental State Exam. Neurology 38: 1565-8, 1988.

Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. Brit J Psychiat 114:797-811, 1968.

Blessed G., Black S.E., Butler T, Kay DWK. The diagnosis of dementia in the elderly: A comparison of the CAMCOG (the Cognitive Section of the CAMDEX), the AGECAT Program, DSM-III, the Mini-Mental State Examination and some short rating scales. British Journal of Psychiatry 159:193-8, 1991.

Bliwise N.G., McCall M.E., Swan S.J. The epidemiology of mental illness in late life. In: Lurie E.E. and Swan J.H. Serving the Mentally Ill Elderly. Lexington Books, D.C. Heath and Company, 1987.

Blumenthal HT. The "Alzheimerization of aging": a response. The Gerontologist 35: 721-723, 1995.

Bonaiuto S, Rocca WA, Lippi A, et al. Impact of education and occupation on the prevalence of Alzheimer's disease and multi-infarct dementia in Appignano, Macerata Province, Italy. Neurology 40 (Suppl 1):346, 1990.

Bond J, Briggs R, Coleman P. The study of ageing. In: Bond J, Coleman P, Peace S (eds). Ageing in Society: An introduction to social gerontology (2nd edition). London, Thousand Oaks, New Delhi, SAGE Publications, 1993.

Botwinick J, Birren JE. Differential decline in the Wechsler-Bellevue subtests in the senile psychoses. Journal of Gerontology 6:365-368, 1951.

Brayne C, Calloway P. An epidemiological study of dementia in a rural population of elderly women. British Journal of Psychiatry 155:214-9, 1989.

Bremer J. A social psychiatric investigation of a small community in Northern Norway. Acta Psychiatrica Neurologica Scandinavica Suppl. 62, 1951.

Brielh J. Community medicine under imperialism: a new medical police? In Navarro V. (ed) "Imperialism, health and medicine". New York, Baywood Publishing Co., 1979.

Brink TL, Yesavage JA, Lum O, et al. Screening tests for geriatric depression. Clin Gerontologist 1:37-43, 1982.

Brody MB, Williams M. Intelligence testing. In: Fleming G.W.T.H. (Ed), Recent Progress in Psychiatry, 2nd Edition. London, J.& A. Churchill, 1950.

Broe GA, Akhtar AJ, Andrews GR et al. Neurological disorders in the elderly at home. Journal of Neurology, Neurosurgery and Psychiatry 39: 362-6, 1976.

Buonanno FS. Magnetic resonance and dementia. National Institutes of Health: Consensus Development Conference: Differential Diagnosis of Dementing Disease, p57-59, Bethesda, Maryland, 1987.

Burks TF. Autonomic agents. In Levenson AJ, ed. Neuropsychiatric Side Effects of Drugs in the Elderly (Aging, Vol. 9), Raven Press, 1979.

Burt C. Experimental tests of general intelligence. British Journal of Psychology 3:94-177, 1909

Burt C. The inheritance of mental characteristics. Eugenics Review 4:168-200, 1912.

Burt C. The genetic determination of differences in intelligence: a study of monozygotic twins reared together and apart. British Journal of Psychology 57:137-53, 1966.

Carmelli D, Swan GE, Cardon LR. Genetic mediation in the relationship of education to cognitive function in older people. Psychology and Aging 10:48-53, 1995.

Carrier J, Kendall I. Social policy and social change - explanations of the development of social policy. Journal of Social Policy 2:3:209-24, 1973.

Caselli G. Health transition and cause-specific mortality. In: Schofield R., Reher D., Bideau A. (eds). The Decline of Mortality in Europe. Oxford, Clarendon Press, 1991.

Chase TN, Fedio P, Foster NL, et al. Wechsler Adult Intelligence Scale Performance: cortical localization by Fluorodeoxyglucose F-18-Positron Emission Tomography. Archives of Neurology 41:1244-7, 1984.

Chung F, Meier R., Lautenschlager E., et al. General or spinal anesthesia: Which is better in the elderly? Anesthesiolgy 67:422-7, 1987.

Chung F, Chung A, Meier RH, et al. Comparison of perioperative mental function after general anaesthesia and spinal anaesthesia with intravenous sedation. Can J Anaesth 36:4:382-7, 1989(a).

Chung F., Lavelle P.A., McDonald S., et al. Cognitive impairment after neuroleptanalgesia in cataract surgery. Anesth Analg 68:614-8, 1989(b).

Claus J.J., van Harskamp F., Breteler et al. The diagnostic value of SPECT with Tc 99m HMPAO in Alzheimer's disease: a population-based study. Neurology 44:454-61, 1994.

Cohen RL, MacKenzie AI. Anaesthesia and cognitive functioning: mental function in postoperative day patients at the time of their release. Anaesthesia 37:47-52, 1982.

Cohen S., Scull A. Introduction: Social control in history and sociology. In Cohen S. and Scull A. (eds). Social Control and the State. New York, St. Martin's Press, 1983.

Cooper JE. An overview of the prospective ICD-10 classification of mental disorders. British Journal of Psychiatry 154 (suppl4) 21-3, 1989.

Copeland JRM, Kelleher MJ, Kellett JM and Gourlay AJ. A semi-structured interview for the assessment of diagnosis and mental state in the elderly: the Geriatric Mental State Schedule. Psychological Medicine 6:439-49, 1976.

Copeland JRM, Dewey ME, Griffiths-Jones HM. A computerized psychiatric diagnosis system and case nomenclature for elderly subjects: GMS and AGECAT. Psychological Medicine 16:89-99, 1986

Copeland JRM, Gurland BJ, Dewey ME et al. Is there more dementia, depression and neurosis in New York. A comparative study of the elderly in New York and London using the computer diagnosis AGECAT. British Journal of Psychiatry 151: 466-73, 1987.

Cornoni-Huntley J., Brock DB, Ostfeld AM et al. Established Populations for Epidemiological Studies of the Elderly. Resource Data Book. U.S. Department of Health and Human Services: National Institute on Aging, NIH Publication No. 86-2443, Bethesda, MD, 1986.

Cornoni-Huntley J., Brock DB, Ostfeld AM et al. Established Populations for Epidemiological Studies of the Elderly. Volume II, Resource Data Book. U.S. Department of Health and Human Services: National Institute on Aging, NIH Publication No. 90-495, Bethesda, MD, 1990.

Cosin LZ, Mort M, Post F et al. Persistent senile confusion: study of 50 consecutive cases. International Journal of Social Psychiatry 3:195-202, 1957

Cowdry E. (ed.): The Problems of Ageing: Biological and Medical Aspects. Baltimore, The Williams & Wilkins Company, 1939.

Crook T, Bartus RT, Ferris SH, et al. Age-associated memory impairment: proposed diagnostic criteria and measures of clinical change. Report of a National Institute of Mental Health Working Group. Developmental Neuropsychology 2:261-76, 1986.

Cryns AG, Gorey KM, Goldstein MZ. Effects of surgery upon the mental status of older persons: a meta-analytic review. Journal of Geriatric Psychiatry and Neurology 3:184-191, 1990.

Crystal HA, Dickson D, Fuld P. et al. Clinico-pathologic studies in dementia: Nondemented subjects with pathologically confirmed Alzheimer's disease. Neurology 38:1682-7, 1988.

Culyer AJ. Cost containment in Europe. Health Care Financing Review. Annual Supplement, 1989.

Culyer, AJ. Health care insurance and provision. In: Barr N. and Whynes D. (eds) Current Issues in the Economics of Welfare. New York, St. Martin's Press, 1993.

Cutler NR, Sramek JJ, Murphy MF, Nash RJ. Implications of the study population in early evaluation of anticholinesterase inhibitors for Alzheimer's disease. Annals of Pharmacotherapy 26:1118-22, 1992.

Dalton JE, Pederson SL, Blom BE, Holmes NR. Diagnostic errors using the short portable mental status questionnaire with a mixed clinical population. J Gerontology 42:512-4, 1987.

Dartigues J-F., Gagnon M, Letenneur L. et al. Principal lifetime occupation and cognitive impairment in a French elderly cohort. American Journal of Epidemiology 135:9:981-8, 1992.

Davies B, Andrewes D, Stargatt R, et al. Tacrine in Alzheimer's disease. The Lancet 163-4, July 1989.

Davis FM, McDermott E., Hickton C. et al. Influence of spinal and general anaesthesia on haemostasis during total hip arthroplasty. British Journal of Anaesthesia 59:561-71, 1987.

Davis PC, Gray L, Albert M et al. Reliability of a standardized MRI evaluation of Alzheimer's disease. The Consortium to Establish a Registry for Alzheimer's Disease. Neurology 42:9:1676-80, 1992.

DePaulo JR, Folstein MF. Psychiatric disturbance in neurological patients: detection, recognition and hospital course. Annals of Neurology 4:3:225-8, 1978.

Denis F, Letourneau JE, Londorf D. Reliability and validity of psychomotor tests as measures of recovery from isoflurance or enflurane anesthesia in a day-care surgery unit. Anesthesia and Analgesia 63:653-6, 1984.

Dewey J. Introduction. In: Cowdry EV. (ed.): The Problems of Ageing: Biological and Medical Aspects. Baltimore, The Williams & Wilkins Company, 1939.

Diagnostic and Statistical Manual of Mental Disorders (DSM-I). Washington, D.C., The American Psychiatric Association, 1952.

Diagnostic and Statistical Manual of Mental Disorders (2nd Edition) (DSMII). Washington, D.C., The American Psychiatric Association, 1968.

Diagnostic and Statistical Manual of Mental Disorders (3rd Edition) (DSMIII). Washington, D.C., The American Psychiatric Association, 1980.

Diagnostic and Statistical Manual of Mental Disorders (3rd Edition-Revised) (DSMIII-R). Washington, D.C., The American Psychiatric Association, 1987.

Diagnostic and Statistical Manual of Mental Disorders (4th Edition) (DSM-IV). Washington, D.C., The American Psychiatric Association, 1994.

Dickens, Charles. American Notes for General Circulation, 1843. (2nd Edition). London, Chapman Hall, 1850.

District of Columbia Office on Aging: Older Washingtonians' perceptions of their health needs: A survey of senior citizens living in Washington, D.C., 1986.

District of Columbia Office on Aging: Elderly Population in Washington, D.C. Policy and Planning Unit, 1991.

Downes D, Rock P. Understanding Deviance: A guide to the sociology of crime and rule breaking. Oxford, Clarendon Press, 1982.

Drickamer MA, Lachs MS. Should patients with Alzheimer's disease be told of their diagnosis? New England Journal of Medicine 326:14:947-51, 1992.

Eaves GN. Preparation of the research grant application: opportunities and pitfalls. In: Preparing a Research Grant Application to the National Institutes of Health. Department of Health and Human Services, U.S. Public Health Service, National Institutes of Health, Bethesda, MD, 1989.

Editorial. Psychiatry - a discipline that has lost its way. Lancet i:731-2, 1985.

Ebaugh FG. Psychiatric complications in surgery. Bull Am Coll Surg 22:153-8, 1937.

Ehrenreich B, Ehrenreich J. Health care and social control. Social Policy 26-40, May, 1974.

Elam JT, Graney MJ, Applegate WB, et al. Functional outcome one year following cataract surgery in elderly persons. J Gerontol 43:M122-6, 1988.

Erkinjuntti T, Sulkava R, Wikstrom J, Autio L. Short portable mental status questionnaire as a screening test for dementia and delirium among the elderly. J Am Geriatric Soc 35:412-6, 1987.

Eslinger PJ, Damasio AR, Benton AL. The Iowa Screening Battery for Mental Decline. Department of Neurology, The University of Iowa College of Medicine, Iowa City, Iowa, 1984.

Eslinger PJ, Damasio AR, Benton AL, Van Allen M. Neuropsychologic detection of abnormal mental decline in older persons. JAMA 253:670-4, 1985.

Essen-Moller E. Individual traits and morbidity in a Swedish rural population. Acta Psychiatrica Neurologica Scandinavica 53:283-97, 1956.

Estes CL. The Aging Enterprise. Jossey-Bass, San Francisco, 1979.

Estes CL. Constructions of reality. J. Social Issues 36:117-132, 1980.

Estes CL. Austerity and aging in the United States: 1980 and beyond. International J. Health Services 12:573-85, 1982.

Estes CL. The politics of ageing in America. Ageing and Society 6:121-34, 1986.

Estes CL. The new political economy of aging: introduction and critique. In: Minkler M, Estes CL (eds) Critical Perspectives on Aging. Amityville N.Y.: Baywood Publishing Co., 1991.

Estes CL., Swan JH, Gerard LE. Dominant and competing paradigms in gerontology: towards a political economy of ageing. Ageing and Society 2:2:151-64, 1982.

Estes CL, Binney EA. Toward a transformation of health and aging policy. International J. Health Services 18:69-82, 1988.

Estes CL, Binney EA. The biomedicalization of aging: dangers and dilemmas. In: Minkler M, Estes CL (eds) Critical Perspectives on Aging. Amityville N.Y.: Baywood Publishing Co., 1991.

Estes CL, Swan JH. et al. The Long-Term Care Crisis: Elders trapped in the no-care zone. Sage Publications, 1993.

Evans DA, Funkenstein HH, Albert MS, et al. Prevalence of Alzheimer's disease in a community population of older persons. JAMA 262:2551-6, 1989.

Eyler J.M. Victorian Social Medicine: The ideas and methods of William Farr. Johns Hopkins University Press, Baltimore, 1979.

Fancher RE. The Intelligence Men: Makers of the I.Q. Controversy. New York, WW Norton & Co., 1985.

Farber JF, Schmitt FA, Logue PE. Predicting intellectual level from the Mini-Mental State Examination. Journal of the American Geriatrics Society 36:509-10, 1988.

Farlow, M., Gracon, S.I., Hershey, L.A. et al. A controlled trial of Tacrine in Alzheimer's Disease. Journal of the American Medical Association 268:18:2523-9.

Feller, B.A. Americans needing help to function at home. National Center for Health Statistics: Advance Data, No 92. Hyattsville, MD, September, 1983.

Fennell G, Phillipson C, Evers H. The sociology of old age. Milton Keynes, Philadelphia. Open University Press, 1988.

Fields SD, MacKenzie R, Charlson ME, Perry SW. Reversibility of cognitive impairment in medical inpatients. Arch Intern Med 146:1593-6, 1986.

Fillenbaum G, Heyman A, Williams K et al. Sensitivity and specificity of standardized screens of cognitive impairment and dementia among elderly black and white community residents. J Clin Epidemiol 43: 7: 651-60, 1990.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state" A practical method for grading the cognitive state of patients for the clinician. J Psychiat Res 12:189-98, 1975.

Folstein MF, Anthony JC, Parhad I, et al. The meaning of cognitive impairment in the elderly. J Am Geriatric Soc 33:228-35, 1985.

Foreman MD. Reliability and validity of mental status questionnaires in elderly hospitalized patients. Nursing Research 36:216-20, 1987.

Foucault M. Madness and civilization: a history of insanity in the age of reason. New York: Pantheon Books, 1965.

Foucault M. Discipline and Punish: the birth of the prison. Vintage Books, New York, 1979.

Fox P. From senility to Alzheimer's disease: the rise of the Alzheimer's disease movement. Milbank Quarterly 67:1:58-102, 1989.

Frank LK. Foreword. In: Cowdry EV. (ed.): The Problems of Ageing: Biological and Medical Aspects. Baltimore, The Williams & Wilkins Company, 1939.

Freidson E. Professional Dominance: The social structure of medical care. New York, Atherton, 1970.

Freidson E. Profession of Medicine: A study in the sociology of applied knowledge. New York, Dodd Mead, 1973.

Freidson E. Professional dominance and the ordering of health services: some consequences. In Conrad P. and Kern R. (eds): The Sociology of Health and Illness: Critical Perspectives. New York, St. Martin's Press, 1981.

Friedman M. Capitalism and Freedom. Chicago, The University of Chicago Press, 1972.

Fries J.F. Aging, natural death and the compression of morbidity. New England Journal of Medicine 303: 130-5, 1980.

Fries J.F. The compression of morbidity. Milbank Memorial Fund Quarterly 61:3:397-419, 1983.

Fries J.F. Strategies for reduction of morbidity. American Journal of Clinical Nutrition 2:55:1257S-62S, 1992.

Frisoni G.B., Rozzini R, Bianchetti A. et al. Principal lifetime occupation and MMSE score in elderly persons. Journal of Gerontology 48:6:S310-S314, 1993.

Galton, F. The possible improvement of the human breed under the existing conditions of law and sentiment. Man, 1901.

George LK, Landerman R., Blazer DG et al. Cognitive Impairment. In: Robins LN, Regier DA. (Eds.) Psychiatric disorders in America: The Epidemiological Catchment Area Study. New York, The Free Press, 1991.

Gerety MB, Mulrow CD, Tuley MR et al. Development and validation of a physicl performance instrument for the functionally impaired elderly: the physical disability index. Journal of Gerentology 48:2:M33-M38, 1993.

Giddens, A. Central Problems in Social Theory: Action, structure and contradiction in social analysis. Berkeley and Los Angeles, University of California Press, 1979.

Gilman S, Foster NL. The "Alzheimerization of aging": a response. The Gerontologist 36: 1, 1996.

Goffman E. Asylums: Essays on the social situation of mental patients and other inmates. Chicago, Aldine Publishing Company, 1963.

Gonzalez M. (ed). Socioeconomics of Medical Practice, 1992. Chicago, Illinois: American Medical Association, 1994.

Granick S, Friedman AS. The effect of education on the decline of psychomentric test performance with age. Journal of Gerontology 22:191-5, 1967.

Greenberg, D.S. Anxiety over the future of U.S. psychiatry. The Lancet (United States edition) 345:781, 1995.

Gruenberg E.F. The failures of success. Millbank Memorial Fund Quarterly 55: 3-24, 1977.

Gubrium, J.F. Oldtimers and Alzheimer's: the descriptive organization of senility. Greenwich, Conn. and London, JAI Press Inc., 1986.

Gubrium, J.F. Structuring and destructuring the course of illness: the Alzheimer's disease experience. Sociology of Health and Illness 9:1:1-24, 1987.

Gubrium J.F. Local cultures and service policy. In: Gubrium J.F., Silverman D. (eds): The Politics of Field Research: Sociology beyond Enlightenment. London, Newbury Park, New Delhi, SAGE Publications, 1989.

Guimon J. The biases of psychiatric diagnosis. British Journal of Psychiatry 154 (suppl 4):33-37, 1989.

Gurland BJ, Cross PS. Epidemiology of psychopathology in old age: some implications for clinical services. The Psychiatric Clinics of North America 5:1, 11-26, 1982.

Gurwitz J.H., Avorn J. The ambiguous relation between aging and adverse drug reactions. Annals of Internal Medicine 114:956-66, 1991.

Gustafson Y, Brannstrom B, Berggren D, et al. A geriatric-anesthesiologic program to reduce acute confusional states in elderly patients treated for femoral neck fractures. JAGS 39:7:655-62, 1991.

Haan J, van Kleef JW, Bloem B, et al. Cognitive function after spinal or general anaesthesia for transurethral prostatectomy in elderly men. JAGS 39:6:596-600, 1991.

Hachinski VV, Iliff L., Zihlka E. et al. Cerebral blood flow in dementia. Archives of Neurology 32: 632-7, 1975.

Hagnell O, Lanke J, Rorsman B et al. Does the incidence of age psychosis decrease? A prospective longitudinal study of a complete population investigated during the 25 year period 1947-1972: the Lundby study. Neuropsychobiology 7:201-211, 1981.

Halstead H. A psychometric study of senility. J Mental Science 89: 364-73, 1943.

Hardy J.A., Higgins G.A. Alzheimer's disease: the amyloid cascade hypothesis. Science 256:184-5, 1992.

Hart B, Spearman C. General ability, its existence and nature. British Journal of Psychology 5:51-84, 1912.

Hartmaier SL, Sloane PD, Guess HA et al. Validation of the minimum data set cognitive performance scale: Agreement with the Mini-Mental State Examination 50A:2:M128-M133, 1995.

Hayman M. The use of serial sevens in psychiatric examination. American Journal of Orthopsychiatry 11:341-55, 1941.

Hayman M. Two minute test for measurement of intellectual impairment in psychiatric disorders. Archives of Neurology and Psychiatry 47:454-64, 1942

Hebb DO. The Organization of Behavior. New York, John Wiley and Sons Inc., 1949 (1961 printing).

Heller SS, Frank KA, Kornfeld DS et al. Psychological outcome following heart surgery. Archives of Internal Medicine 134:908-14, 1974.

Hennekens CH, Buring JE. Epidemiology in Medicine. Boston, Little, Brown and Company, 1987.

Henrikson, L. Evidence suggestive of diffuse brain damage following cardiac operations. The Lancet 816-20, 1984.

Hill M. Social Security Policy in Britain. Edward Elgar Publishing Co., 1990.

Hillman AL, Eisenberg JM, Pauly MV et al. Avoiding bias in the conduct and reporting of cost-effectiveness research sponsored by pharmaceutical companies. New England Journal of Medicine 324:1362-65, 1991.

Hole A, Terjesen T, Breivik H. Epidual versus general anaesthesia for total hip arthroplasty in elderly patients. Acta Anaesth Scand 24:279-87, 1980.

Holzer CE, Tischler GL, Leaf PJ et al. An epidemiological assessment of cognitive impairment in a community population. Research in Community and Mental Health 4: 3-32, 1984.

Homer A.C., Honavar M., Lantos P.L. et al. Diagnosing dementia: Do we get it right? British Medical Journal 297:894-6, 1988.

Hosking MP, Lobdell CM, Warner MA, et al. Anaesthesia for patients over 90 years of age. Outcomes after regional and general anaesthetic techniques for two common surgical procedures. Anaesthesia 44:142-7, 1989.

Hughes EC (ed). Education for the Professions of Medicine, Law, Theology and Social Welfare. New York: McGraw-Hill, 1973.

Ingleby D. Mental health and social order. In Cohen S. and Scull A. (eds). Social Control and the State. New York, St. Martin's Press, 1983.

Illich I. Medical Nemesis: the expropriation of health. New York, Pantheon Books, 1976.

Ineichen B. Measuring the rising tide: how many dementia cases will there be by 2001? British Journal of Psychiatry 150:193-200, 1987.

International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). United States DHHS, (PHS)91-260, Public Health Service, 1991.

Irving G, Robinson RA, McAdam W. The validity of some cognitive tests in the diagnosis of dementia. Brit J Psychiat 117:149-56, 1970.

Jacobs JW, Bernhard BA, Delgado A, Strain JJ. Screening for organic mental syndromes in the medically ill. Ann Int Med 86:40-6, 1977.

Jacoby RJ, Levy R, Dawson JM. Computed tomography in the elderly: 1. The normal population. British Journal of Psychiatry 136:249-55, 1980.

Jacoby R.J., Levy R. Computed tomography in the elderly: 2. Senile dementia: Diagnosis and functional impairment. British Journal of Psychiatry 136:256-69, 1980.

Johnson KA, Davis KR, Buonanno FS et al. Comparison of magnetic resonance and roentgen ray computed tomography in dementia. Archives of Neurology 44:1075-80, 1987.

Johnson TJ. Professions and Power. London, Macmillan, 1972.

Jones JK. Drugs and the elderly. In: Reichel W.(Ed), Clinical Aspects of Aging, ed 3. Williams and Wilkins, Baltimore, 1989.

Jorm AF, Korten AE, Henderson AS. The prevalence of dementia: a quantitative integration of the literature. Acta Psychiatr Scand 76: 465-79, 1987.

Kahn RL, Goldfarb AI, Pollack M et al. The relationship of mental and physical status in institutionalized aged persons. Am J Psychiatry 117:120-4, 1960.

Kahn RL, Goldfarb AI, Pollack M et al. Brief objective measures for the determination of mental status in the aged. Am J Psychiatry 117:326-8, 1960.

Kahn RL, Zarit SH, Hilbert NM, Niederehe MA. Memory complaint and impairment in the aged. Arch Gen Psychiat 32:1569-73, 1975.

Karhunen U, Jonn G. A comparison of memory function following local and general anaesthesia for extraction of senile cataract. Acta Anaesth Scand 26:291-6, 1982.

Katz S, Ford AB, Moskowitz RW et al. Studies of illness in the aged. The Index of ADL: a standardized measure of biological and psychosocial function. Journal of the American Medical Association 185:12:914-919, 1963.

Katz S, Akpom A. A measure of primary sociobiological functions. International Journal of Health Services 6:3:493-507, 1976.

Katzman R. The prevalence and malignancy of Alzheimer Disease: A major killer. Arch Neurol 33: 217-8, 1976

Katzman R. Alzheimer's disease. NEJM 314:15:964-73, 1986.

Katzman R, Karasu T. Differential Diagnosis of Dementia. In: W. Fields (ed), Neurological and Sensory Disorders in the Elderly. New York: Stratton Intercontinental Medical Book Corp, 1975.

Katzman R. Senile dementia of the Alzheimer type - defining a disease. Presented to the Duke University Council on Aging and Human Development, 1978. In: Maddox G.L., Auld E. Seminar Proceedings 1976-1980. Durham, N.C., Duke University, 1981.

Katzman R, Brown T, Fuld P et al. Validation of a short orientation-memory-concentration test of cognitive impairment. American Journal of Psychiatry 140:6:734-9, 1983.

Katzman F, Jackson JE. Alzheimer's disease: basic and clinical advances. Journal of the American Geriatrics Society 39:516-25, 1991.

Katzman R. Education and the prevalence of dementia and Alzheimer's disease. Neurology 43:13-20, 1993.

Kay DWK, Beamish P, Roth m. Old age mental disorders in Newcastle upon Tyne. British Journal of Psychiatry 110: 146-158, 1964.

Kay DWK, Bergmann, Foster EM, McKechnie AA, Roth M. Mental illness and hospital usage in the elderly: a random sample followed up. Comprehensive Psychiatry 11: 1: 26-35, 1970.

Kay DWK, Henderson AS, Scott R et al. Dementia and depression among the elderly living in the Hobart community: the effect of the diagnostic criteria on the prevalence rates. Psychological Medicine 15:771-88, 1985.

Keller MB, Manschreck TC. The bedside mental status examination --reliability and validity. Comprehensive Psychiatry 22:500-11, 1981.

Kendell RE. The concept of disease and its implications for psychiatry. British Journal of Psychiatry 127:305-15, 1975.

Kevles DJ. Testing the army's intelligence: psychologists and the military in World War 1. Journal of American History 55:565-81, 1968.

Kidd M. Alzheimer's disease - an electron microscopical study. Brain 87:307, 1964.

Kiernan FJ, Mueller J, Langston JW, Van Dyke C. The neurobehavioral cognitive status examination: a brief but differentiated approach to cognitive assessment. Ann Int Med 107:481-5, 1987.

Kimball CP. Psychological responses to the experience of open heart surgery: I. Amer J Psychiat 126:3:96-107, 1969.

Kimball CP. Psychological responses to the experience of open heart surgery: II. Am J Psychiat 126:348-59, 1969.

Kinsella K.G. Changes in life expectancy 1900-1990. American Journal of Clinical Nutrition 55:1196S-1202S, 1992.

Kinsella K.G., Gist Y.J. Older Workers, Retirement and Pensions: A comparative international chartbook. U.S. Department of Commerce, International Programs Center, Bureau of the Census, Washington, D.C. 1995.

Kittner SJ, White LR, Farmer ME, et al. Methodological issues in screening for dementia: the problem of education adjustment. J Chron Dis 39:163-70, 1986.

Kitwood T. Brain, mind and dementia: with particular reference to Alzheimer's disease. Ageing and Society 9:1-15, 1989.

Kitwood T. The dialectics of dementia: with particular reference to Alzheimer's disease. Ageing and Society 10:177-196, 1990.

Kitwood T, Bredin K. Towards a theory of dementia care: personhood and well-being. Ageing and Society 12:269-287, 1992.

Klein LE, Roca RP, McArthur J, et al. Diagnosing dementia: univariate and multivariate analyses of the mental status examination. J Am Geriatr Soc 33:483-8, 1985.

Kohli M. Retirement and the moral economy: an historical interpretation of the German case. In: Minkler M, Estes CL (Eds) Critical Perspectives on Aging. Amityville N.Y.: Baywood Publishing Co., 1991.

Kokmen E, Naessens JM, Offord KP. A short test of mental status: description and preliminary results. Mayo Clin Proc 62:281-8, 1987.

Kornfeld DS, Zimberg S, Malm JR. Psychiatric complications of open-heart surgery. New England Journal of Medicine 273:287-92, 1965.

Kornfeld DS, Heller SS, Frank KA, Moskowitz R. Personality and psychological factors in postcardiotomy delirium. Archives of General Psychiatry 31:249-53, 1974.

Kraepelin, E. Lectures on Clinical Psychiatry. New York, William Wood and Company, 1906.

Kraepelin E. Senile and pre-senile dementias. In: Psychiatrie: Ein Lehrbuch fur Studierende und Arzte. Leipzig, Verlag von JA Barth, 1910.

Kramer M. The rising pandemic of mental disorders and associated chronic diseases and disabilities. In: Epidemiological Research as Basis for the Organization of Extamural Psychiatry. Acta Psychiat Scand Suppl 285: 62:382-97, 1980.

Kramer M. The continuing challenge: the rising prevalence of mental disorders, associated chronic diseases and disabling conditions. American Journal of Social Psychiatry 3: 4: 13-24, 1983.

Kramer M, German PS, Anthony JC, et al. Patterns of mental disorders among the elderly residents of eastern Baltimore. J Amer Geriatr Society 33:236-45, 1985.

Kuhn, T.S. The Structure of Scientific Revolutions. Chicago, University of Chicago Press, 1962.

Landsborough Thomson, A. Origin of the British legislative provision for medical research. Journal of Social Policy 2:1:41-54, 1973.

Larson E.B. Alzheimer's disease in the community. (Editorial comment). Journal of the American Medical Association 262:18:2591-2, 1989.

Larson EB, Reifler BV, Featherstone HJ, English DR. Dementia in elderly outpatients: a prospective study. Ann Int Med 100:417-23, 1984.

Larsson T, Sjogren T, Jacobson G. Senile dementia. Acta. Psychiatr. Scand. 39 (suppl. 167):3-259, 1963.

Launer LJ, Dinkgreve MAH, Jonker C. et al. Are age and education independent correlates of the Mini-Mental State Exam performance of community-dwelling elderly? Journal of Gerontology 48:6:P271-277, 1993.

Layne O.L. et al. Postoperative psychosis in cardiotomy patients: The role of organic and psychiatric factors. New England Journal of Medicine 284:518-20, 1971.

Lazarus LW, Newton N, et al. Frequency and presentation of depressive symptoms in patients with primary degenerative dementia. Am J Psychiatry 144:1:41-45, 1987.

Leaf PJ, Myers JK, Mc Evoy LT. Procedures used in the Epidemiologic Catchment Area Study. In: Robins LN, Regier DA (Eds.) Psychiatric disorders in America: The Epidemiological Catchment Area Study. New York, The Free Press, 1991.

Lee M, Ganzini L. The effect of recovery from depression on preferences for life-sustaining therapy in older patients. Journal of Gerontology 49:1:M15-M21, 1994.

Levit KR, Lazenby HC, Cowan CA et al. National Health Expenditures, 1990. Health Care Financing Review 13:1:29-40, 1991.

Lezak MD. Neuropsychological Assessment (2nd ed). Oxford University Press, 1983.

Lipowski ZJ. Transient cognitive disorders (delirium, acute confusional states) in the elderly. Am J Psychiatry 140:11:1426-36, 1983.

Lipowski ZJ. Psychiatry: mindless or brainless, both or neither? Canadian J Psychiatry 34:3:249-54, 1989.

Liston EH. Delirium in the aged. Psychiatric Clinics of North America 5:1:49-66, 1982.

Litovitz GL, Hedberg M, Wise TN, et al. Recognition of psychological and cognitive impairments in the emergency department. Am J Emergency Med 3:400-2, 1985.

Lobo A, Saz P, Dia JL et al. The epidemiological study of dementia in Zaragoza, Spain. In Stefaniss CN, Soldators CR, Rabavilas AD (eds). Psychiatry: a world perspective. Proceedings of the VIII World Congress of Psychiatry, Athens, Oct 13-19, 1989. Amsterdam, Elsevier, 1990 p133-7.

Lockett BA. Aging, Politics and Research: Setting the Federal Agenda for Research on Aging. New York, Springer Publishing Company, 1983.

Lovett Doust JW, Schneider RA, Talland GA, et al. Studies on the physiology of awareness: the correlation between intelligence and anoxemia in senile dementia. J Nerv Ment Dis 117:5:383-98, 1953.

Lubitz J, Prihoda R. Use and costs of Medicare services in the last years of life. Health - United States, 1983. DHHS Pub. No. (PHS)84-1232, Washington, D.C. U.S. Government Printing Office, 1984.

Lurie EE. The interrelationship of physical and mental illness in the elderly. In: Lurie EE, Swan JH. Serving the Mentally Ill Elderly. Lexington Books, D.C. Heath and Company, 1987.

Lyman KA. Bringing the social back in: A critique of the biomedicalization of dementia. The Gerontologist 29:5:597-605, 1989.

Mace N. Using mental status tests. J Gerontological Nursing 13:33.

Magaziner J, Bassett SP, Hebel JR. Predicting performance on the Mini-Mental State Examination. Use of age- and education-specific equations. Journal of the American Geriatrics Society 35: 996-1000, 1987.

Manton KG, Patrick CH, Johnson KW. Health differentials between blacks and whites: recent trends in mortality and morbidity. The Milbank Quarterly 65, Suppl 1, 129-199, 1987.

Manton KG, Stallard E., Corder L. Changes in morbidity and chronic disability in the U.S. elderly population: evidence from the 1982, 1984 and 1989 National Long Term Care Surveys. Journal of Gerontology 50B:4:S194-204, 1995.

Marx J. Major setback for Alzheimer's models. Science 255:1200-2, 1992.

McKhann G, Drachman D, Folstein M et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34: 939-44, 1984.

McKinlay JB. On the medical-industrial complex. Monthly Review, 38-42, October, 1978.

McLoughlin DM, Lovestone S. Alzheimer's disease: recent advances in molecular pathology and genetics. International Journal of Geriatric Psychiatry 9:431-444, 1994.

Meakin CJ. Screening for depression in the medically ill. The future of paper and pencil tests. British Journal of Psychiatry 160: 212-216, 1992.

Medawar C. Power and Dependence: social audit on the safety of medicines. London, Social Audit Ltd., 1992.

Meehl PE. Clinical versus Statistical Prediction: a theoretical analysis and review of the evidence. Minneapolis, University of Minnesota Press, 1954.

Meyer Adolf. The Collected Papers. Baltimore: Johns Hopkins Press, 1950.

Miles CC. Influence of speed and age on intelligence scores of adults. Journal of General Psychology 10:208-10, 1934.

Miller NE, Cohen GD. Clinical aspects of Alzheimer's Disease and senile dementia: synopsis and future perspectives in assessment, treatment, and service delivery. In Miller NE, Cohen GD (eds) Clinical Aspects of Alzheimer's Disease and Senile Dementia (Aging, Vol. 15), Raven Press, 1981.

Milstein V, Small JG, Small IF. The subtraction of serial sevens test in psychiatric patients. Arch Gen Psychiat 26: 439-41, 1972.

Minkler M. Chapter1: Overview. In: Minkler M. and Estes C.L. (eds) Critical Perspectives on Aging: the political and moral economy of growing old. Baywood Publishing Co. Amityville, New York, 1991.

Minkler M. Gold in gray: reflections on business' discovery of the elderly market. In: Minkler M. and Estes C.L. (eds) Critical Perspectives on Aging: the political and moral economy of growing old. Baywood Publishing Co. Amityville, New York, 1991.

Mirra S.S., Heyman A., McKeel D. et al. Standardization of the neuropathologic assessment of Alzheimer's disease. Part II. The Consortium to Establish a Registry for Alzheimer's Disease. Neurology 41:4:479-86, 1991.

Mondimore FM, Damlouji N., Folstein MF et al. Post-ECT confusional states associated with elevated serum anticholinergic levels. American Journal of Psychiatry 140:930-1, 1983.

Moore JT, Bobula JA, Short TB, Mischel M. A functional dementia scale. J Family Practice 16:499-503, 1983.

Morris, J.N. Uses of Epidemiology. London, Longmans, 1964.

Morris, J.N. Uses of Epidemiology. London, Longmans, 1975.

Morris J.N. Social inequalities undiminished. The Lancet 87-90, 1979.

Mortimer JA, Graves AB. Education and other socioeconomic determinants of dementia and Alzheimer's disease. Neurology 43(suppl 4):S39-S44, 1993.

Murden RA, McRae TD, Kaner S, Bucknam ME. Mini-mental state exam scores vary with education in blacks and whites. JAGS 39:149-155, 1991.

Murphy DG, Dean DJ. Application and review procedures for the National Institutes of Health. In: Preparing a Reseach Grant Application to the National Institutes of Health. Department of Health and Human Services, Public Health Service, 1989.

Murphy DG, Dean DJ. Biomedical research and research training support by the National Institutes of Health. In: Preparing a Reseach Grant Application to the National Institutes of Health. Department of Health and Human Services, Public Health Service, 1989.

Myles J. Old Age in the Welfare State: The Political Economy of Public Pensions. Boston, Little, Brown and Company, 1984.

Myles J. Old Age in the Welfare State: The Political Economy of Public Pensions. (Revised edition). Boston, Little, Brown and Company, 1989.

Myles J. Postwar capitalism and the extension of social security into a retirement wage. In: Minkler M, Estes CL (Eds) Critical Perspectives on Aging. Amityville N.Y.: Baywood Publishing Co., 1991.

Nagy Z, Esiri MM, Jobst KA et al. Relative roles of plaques and tangles in the dementia of Alzheimer's disease: correlations using three sets of neuropathological criteria. Dementia 6:21-31, 1995.

Natelson BH, Haupt EJ, Fleischer EJ, Grey L. Temporal orientation and education: a direct relationship in normal people. Arch Neurol 36:444-6, 1979.

(NIA) National Institute on Aging Task Force: Senility reconsidered: treatment possibilities for mental impairment in the elderly. Journal of the American Medical Association, 244:259-63, 1980.

(NIH) National Institutes of Health: Consensus Development Conference: Differential Diagnosis of Dementing Disease: Conference Statement: 1-9, July 1987.

National Center for Health Statistics. Health, United States, 1994. Hyattsville, Maryland: Public Health Service, 1995.

Navarro V. Medicine under capitalism. Prodist, New York 1976.

Navarro V. Medical history as justification rather than explanation: a critique of Starr's "The Social Transformation of American Medicine". Int. J. Health Services 14:4:511-28, 1984.

Navarro V. U.S. Marxist scholarship in the analysis of health and medicine. Int. J. Health Services 15:525-45, 1985.

Navarro V. Race or class or race and class: growing mortality differentials in the United States. International J. of Health Services 21:2:229-235, 1991.

Navarro V. Production and the welfare state: the political context of reforms. Int. J. Health Services 21: 585-614, 1991.

Navarro V. An analysis of the American Medical Association's recommendations for change in the medical care sector of the United States. Int. J. Health Services 21: 685-96, 1991.

Newsholme A. The measurement of progress in public health with special reference to the life and work of William Farr. Economica 3:196-202, 1923.

Offe C, Ronge V. Theses on the theory of the state. In: Giddens A and Held D (eds). Classes, Power and Conflict: Classical and Contemporary Debates. Berkeley, University of California Press, 1982.

O'Connor DW, Pollitt PA, Treasure FP, et al. The influence of education, social class and sex on Mini-Mental State test scores. Psychological Medicine 19: 771-76, 1989 (a).

O'Connor DW, Pollitt PA, Hyde JB et al. The reliability and validity of the Mini-Mental State in a British community survey. J Psychiat Res 23: 1: 87-96, 1989 (b).

O'Connor DW, Pollitt PA, Hyde JB. The prevalence of dementia as measured by the Cambridge Mental Disorders of the Elderly Examination. Acta Psychiatr Scand 79:190-8, 1989 (c).

Omer H, Foldes J, Toby M, Menczel J. Screening for cognitive deficits in a sample of hospitalized geriatric patients: a re-evaluation of a brief mental status questionnaire. J Am Geriatric Soc 31:266-8, 1983.

Parsons T. The Social System. Free Press, New York, 1951.

Parsons T. The Social Structure and Personality. The Free Press of Glencoe, 1964.

Pedersen NL, Reynolds CA, Gatz M. Sources of covariation among Mini-Mental State Examination Scores, education, and cognitive abilities. Journal of Gerontology 51B:P55-P63, 1996.

Pepin EP, Eslinger PJ. Verbal memory decline in Alzheimer's disease: A multiple-processes deficit. Neurology 39:1477-82, 1989.

Perry E.K., Perry R.H., Blessed G. et al. Necropsy evidence of central cholinergic deficits in senile dementia. The Lancet, 189, 1977.

Perry E.K., Tomlinson B.E., Blessed G. et al. Correlation of cholinergic abnormalities with senile plaques and mental test scores in senile dementia. British Medical Journal 2:1457-9, 1978.

Perry RH, Irving D, Blessed G et al. Clinically and neuropathologically distinct form of dementia in the elderly. The Lancet 1:166, 1989.

Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. J Am Geriatr Soc 23:433-41, 1975.

Phillipson C. Capitalism and the Construction of Old Age. Macmillan, London, 1982.

Phillipson C. Inter-generational relations: conflict or consensus in the 21st century. Policy and Politics 19:27-36, 1991.

Pickens D.K. Eugenics and the Progressives. Vanderbilt University Press, 1968.

Popper, KR. The Poverty of Historicism. New York, Harper & Row, 1964.

Prior L. Evaluation research and quality assurance. In: Gubrium J.F., Silverman D. (eds): The Politics of Field Research: Sociology beyond Enlightenment. London, Newbury Park, New Delhi, SAGE Publications, 1989.

Rai GS, Scott L, Beston B. Assessment of the objective value of mini-mental state examination. J Clin Experimental Gerontol 11:41-6, 1989.

Regier DA, Myers JK, Kramer M, et al. The NIMH epidemiologic catchment area program. Arch Gen Psychiatry 41:934-41, 1984.

Regier DA, Robins LN. Introduction. In: Robins LN, Regier DA. (eds). Psychiatric disorders in America: The Epidemiologic Catchment Area Study. New YorK, The Free Press, 1991.

Reifler BV, Lardon E, Hanley R. Coexistence of cognitive impairment and depression in geriatric outpatients. Am J Psychiatry 139:5:623-6, 1982.

Reisberg B, Schneck MK, Ferris SH, et al. The brief cognitive rating scale (BCRS): findings in primary degenerative dementia (PDD). Psychopharmacology Bulletin 19:1:47-51, 1983.

Relman AS. The new medical-industrial complex. New England Journal of Medicine 303:17:963-70, 1980.

Relman AS. What market values are doing to medicine. The Atlantic Monthly 99-106, March, 1992.

Rettig RA. Cancer Crusade: The Story of the National Cancer Act of 1971. Princeton, New Jersey, Princeton University Press, 1977.

Riis J, Lomholt B, Haxholdt O et al. Immediate and long-term mental recovery from general versus epidural anesthesia in elderly patients. Acta Anaesthesiol. Scand. 27:44-49, 1983.

Robertson A. The politics of Alzheimer's disease: a case study in apocalyptic demography. In: Minkler M, Estes CL (eds) Critical Perspectives on Aging. Amityville N.Y.: Baywood Publishing Co., 1991.

Robins L.N., Helzer J.E., Croughan J., et al. National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. Archives of general Psychiatry 38:381-389, 1981.

Robins LN, Helzer JE, Weissman MM, et al. Lifetime prevalence of specific psychiatric disorders in three sites. Arch Gen Psychiatry 41:949-58, 1984.

Robins L.N., Regier D.A. Psychiatric disorders in America: The Epidemiological Catchment Area Study. New York, The Free Press, 1991.

Robinson DJ, Merskey H, Blume WT et al. Electroencephalography as an aid in the exclusion of Alzheimer's disease. Archives of Neurology, 51:280-4, 1994.

Rocca WA, Bonaiuto S, Lippi A et. al. Prevalence of clinically diagnosed Alzheimer's disease and other dementing disorders in a door to door survey in Appignano, Macerato Province, Italy. Neurology 40:626-31, 1990.

Rocca WA, Hofman A, Brayne C. Frequency and distribution of Alzheimer's disease in Europe: a collaborative study of 1980-1990 prevalence findings. Annals of Neurology 30: 381-90, 1991.

Rorsman B, Hagnell O., Lanke J. The prevalence of age psychosis and mortality among age psychotics in the Lundby study: changes over time during a 25-year observation period. Neuropsychobiology 13: 167-72, 1985.

Rorsman B, Hagnell O., Lanke J. Prevalence and incidence of senile and multi-infarct dementia in the Lundby study: a comparison between the time periods 1947-1957 and 1957-1972. Neuropsychobiology 15: 122-9, 1986.

Rosen WG. Verbal fluency in aging and dementia. J Clin Neuropsycology 2:135-46, 1980.

Roslaniec A, Fitzpatrick JJ. Changes in mental status in older adults with four days of hospitalization. Research in Nursing and Health 2:177-87, 1979.

Roth M. Problems of old age and the senile and arteriosclerotic psychoses. In: Fleming GWTH (Ed), Recent Progress in Psychiatry, 2nd Edition. London, J.& A. Churchill, 1950.

Roth M. The natural history of mental disorder in old age. Journal of Mental Science, 101:281-301, 1955.

Roth M. The psychiatric disorders of later life. Psychiatric Annals 6:9:417-445, 1976.

Roth M, Morrissey JD. Problems in the diagnosis and classification of mental disorder in old age; with a study of case material. Journal of Mental Science 98:66-80, 1952.

Roth M, Hopkins B. Psychological test performance in patients over sixty. I. Senile psychosis and the affective disorders of old age. Journal of Mental Science 99: 439-50, 1953.

Roth M, Hopkins B. Psychological test performance in patients over sixty. II. Paraphrenia, arteriosclerotic psychosis and acute confusion. Journal of Mental Science 99: 451-63, 1953.

Roth M, Tym E, Mountjoy CQ, et al. CAMDEX: A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. British Journal of Psychiatry 149:698-709, 1986.

Ryhanen P, Helkala EL, Ihalainen O et al. Effects of anaesthesia on the psychological function of patients. Annals of Clinical Research 10:318-322, 1978.

SAS/STAT User's Guide, Release 6.03 Edition, 1992. SAS Institute Inc. Cary, North Carolina, Savage GH. Insanity following the use of anaesthetics in operations. British Medical Journal, ii:1199-2000, 1887.

Scherr PA, Albert MS, Funkenstein HH, et al. Correlates of cognitive function in an elderly community population. Am J Epidemiology 128:1084-101, 1988.

Schieber GJ, Pouillier J-P, Greenwald LM. U.S. health expenditure performance: An international comparison and data update. Health Care Financing Review 13:4:1-72, 1992.

Schneider EL, Brody JA. Aging, natural death and the compression of morbidity: another view. New England Journal of Medicine 854-6, 1983.

Schoenberg BS, Anderson DW, Haerer AF. Severe dementia: prevalence and clinical features in a biracial US population. Arch Neurol 42:740-3, 1985.

Schwamm LH, Van Dyke C, Kiernan RJ, et al. The neurobehavioral cognitive status examination: comparison with the cognitive capacity screening examination and the mini-mental state examination in a neurosurgical population. Ann Int Med 107:486-91, 1987.

Scull A. The decarceration of the mentally ill: a critical view. Politics and Society 6:173-211, 1976

Scull A. Social Order/Mental Disorder: Anglo-American psychiatry in historical perspective. University of California Press, 1989.

Shanas E, Townsend P, Wedderburn D. et al. Old People in Three Industrial Societies. London, Routledge and Kegan Paul, 1968.

Shapiro S, Skinner EA, Kessler LG, et al. Utilization of health and mental health services: Three epidemiologic catchment area sites. Arch Gen Psychiatry 41:971-8, 1984.

Shaw GB. The Doctor's Dilemma, 1906. In Shaw G.B. The Collected Works. New York, Wm. H. Wise & Company, 1930.

Shaw PJ, Bates D, Cartlidge NEF, et al. Early neurological complications of coronary artery bypass surgery. Br Med J 291:1384-7, 1985.

Silverman D. The Theory of Organisations. London, Heinemann, 1974.

Silverman D, Gubrium JF. Introduction. In: Gubrium J.F., Silverman D. (eds): The Politics of Field Research: Sociology beyond Enlightenment. London, Newbury Park, New Delhi, SAGE Publications, 1989.

Simpson BR, Williams M, Scott JF. The effects of anaesthesia and elective surgery on old people. The Lancet 887-93, 1961.

Simpson CJ, Kellett JM. The relationship between preoperative anxiety and post-operative delirium. Journal of Psychosomatic Research 31:491-7, 1987.

Simpson JEP, Glynn CJ, Cox AG et al. Comparative study of short-term recovery of mental efficiency after anaesthesia. British Medical Journal 1:1560-62, 1976.

Sluss TK, Gruenberg EM, Kramer M. The use of longitudinal studies in the investigation of risk factors for Senile Dementia-Alzheimer Type. In Mortimer JA, Schuman LM (eds), The Epidemiology of Dementia, Oxford University Press, 1981.

Small, G. Tacrine for treating Alzheimer's Disease. Journal of the American Medical Association 268:18:2564-5.

Smith A. The serial sevens subtraction test. Arch Neurol 17: 78-80, 1967.

Smith LW, Dimsdale JE. Postcardiotomy delirium: conclusions after 25 years? American Journal of Psychiatry 146:4:452-8, 1989.

Smith, PLC. Cerebral consequences of cardiopulmonary bypass. The Lancet 823-5, 1986.

Snowdon DA, Kemper SJ, Mortimer JA et al. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the nun study. Journal of the American Medical Association 275:528-532, 1996.

Spearman C. The proof and measurement of association between two things. American Journal of Psychology 15:72-101, 1904(a).

Spearman C. "General Intelligence" objectively determined and measured. American Journal of Psychology 15:201-93, 1904(b).

Spearman C. The Nature of Intelligence and the Principles of Cognition. Edinburgh, MacMillan, 1923.

Spearman C. The abilities of man. New York, Macmillan, 1927.

Spitzer RL, Endicott J, Fleiss JL et al. The psychiatric status schedule: A technique for evaluating psychopathology and impairment in role functioning. Archives of General Psychiatry 23:41-55, 1970.

Spitzer RL, Fleiss JL. A re-analysis of the reliability of psychiatric diagnosis. British Journal of Psychiatry 125:341-7, 1974

Starr P. The Social Transformation of American Medicine. Neew York, Basic Books, 1982.

Stern W. The Psychological Methods of Testing Intelligence. Baltimore, Warwick and York, 1914.

Stern Y, Gurland B, Tatemichi et al. Influence of education and occupation on the incidence of Alzheimer's disease. Journal of the American Medical Association 271:13:1004-10, 1994.

Stevens R. American Medicine and the Public Interest. New Haven and London, Yale University Press, 1971.

Stockton P. The Alzheimer's disease epidemic: mental status assessment of elderly people in the United States. Critical Public Health 2:15-20, 1991.

Stockton P, Jones JK. Medication use by the elderly: a review. Aging Clinical and Experimental Research 5:5:337-47, 1993.

Strain JJ, Fulop G. Screening devices for cognitive capacity. Ann Int Med 107:583-5, 1987.

Strickland S.P. Politics, Science, and Dread Disease: A short history of United States medical research policy. Cambridge, Massachusetts, Harvard University Press, 1972.

Summers, W.K., Reich T.C. Delirium after cataract surgery: review and two cases. American Journal of Psychiatry 136:386-91, 1979.

Szasz, T.S. The Manufacture of Madness. New York, Harper & Row, 1970.

Szasz T. The Myth of Mental Illness. New York, Harper and Row, 1972.

Szasz T. Schizophrenia: the sacred symbol of psychiatry. Oxford, New York, Oxford University Press, 1979.

Szasz T. Psychiatry: rhetoric and reality. The Lancet 2:711-712, 1985.

Taylor-Gooby P. Postmodernism and social policy: a great leap backwards? Journal of Social Policy 23:3:385-404, 1994.

Terman LM, Childs HG. A tentative revision and extension of the Binet-Simon scale of intelligence. Journal of Educational Psychology. 3:61-74, 133-143, 198-208, 277-289, 1912.

Terman LM, Lyman G, Ordahl G et al. The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence. Baltimore, Warwick and York, 1917.

Teri L. McCurry SM, Edland SD et al. Cognitive decline in Alzheimer's disease: a longitudinal investigation of risk factors for accelerated decline. Journal of Gerontology 50A:1:M49-M55, 1995.

Terry RD, Gonatas NK, Weiss M. Ultrastructural studies in Alzheimer's presenile dementia. American Journal of Pathology 44:269, 1964.

Terry RD, Wisniewski HM. Pathology and Pathogenesis of Dementia. In: W. Fields (ed), Neurological and Sensory Disorders in the Elderly. New York: Stratton Intercontinental Medical Book Corp, 1975.

Thompson JH. Sites and mechanisms of drug interactions. In Levenson AJ, ed. Neuropsychiatric Side Effects of Drugs in the Elderly (Aging, Vol. 9). Raven Press, 1979.

Thorne, B. Professional Education in Medicine. In: Hughes E.C. (ed). Education for the Professions of Medicine, Law, Theology and Social Welfare. New York: McGraw-Hill, 1973.

Tinklenberg J, Brooks JO, Tanke ED, et al. Factor analysis and preliminary validation of the mini-mental state examination from a longitudinal perspective. International Psychogeriatrics 2:2:123-34, 1990.

Townsend P. The structured dependency of the elderly: a creation of social policy in the Twentieth Century. Ageing and Society 1:5-28, 1981.

Tune LE, Damlouji NF, Holland A et al. Association of postoperative delirium with raised serum levels of anticholinergic drugs. The Lancet 651-653, 1981.

Turner TB. The medical schools twenty years afterwards: Impact of the extramural research support programs of the National Institutes of Health. Journal of Medical Education 42:109-118, 1967.

Turner, TH. The past of psychiatry: why build asylums? The Lancet 709-10, 1985.

U'ren RC. Testing older patients' mental status: practical office-based approach. Geriatrics 42:49-56, 1987.

Uhlmann RF, Larson EB, Buchner DM. Correlations of mini-mental state and modified dementia rating scale to measures of transitional health status in dementia. J Gerontology 42:33-6, 1987.

U.S. Congress, Office of Technology Assessment. Losing a Million Minds: Confronting the Tragedy of Alzheimer's Disease and Other Dementias. OTA-BA-323. Washington, D.C.: U.S. Government Printing Office, April, 1987.

U.S. Department of Commerce: Survey of Current Business 73: April, 1993.

U.S. Department of Health and Human Services: Healthy People 2000: National health promotion and disease prevention objectives. DHHS Publication No. (PHS)91-50213, 1991.

U.S. Senate: Developments in Aging: Volume 1. A report of the Special Committee on Aging, United States Senate, Washington, D.C. U.S. Government Printing Office, 1983.

U.S. Senate: Special Committee on Aging. Long-term care in Western Europe and Canada: Implications for the United States. U.S. Government Printing Office, 1984.

Walker A. The social creation of poverty and dependency in old age. Journal of Social Policy 9:1:49-75, 1980.

Walker A. Towards a political economy of old age. Ageing and Society 1:1:73-94, 1981.

Walker A. The economic 'burden' of ageing and the prospect of intergenerational conflict. Ageing and Society 10:377-96, 1990.

Wallace S. The political economy of health care for elderly blacks. In: Minkler M. and Estes C.L. (eds) Critical Perspectives on Aging: the political and moral economy of growing old. Baywood Publishing Co. Amityville, New York, 1991.

Webber PA, Fox P, Burnette D. Living alone with Alzheimer's disease: effects on health and social service utilisation patterns. The Gerontologist 34:1:8-14, 1994.

Weber M. The Theory of Social and Economic Organisation. Translated by Henderson A.M. and Parsons T. New York, The Free Press, 1968.

Weber M. Legitimate authority and bureaucracy. (1947) In: Pugh, D.S. (Ed). Organization Theory. Penguin Books Ltd., 1973.

Wechsler D. On the influence of education on intelligence as measured by the Binet-Simon tests. Journal of Educational Psychology 17:248-57, 1926.

Wechsler D. Analytic use of the Army Alpha examination. Journal of Applied Psychology 16:254-6, 1932.

Wechsler D. The Measurement of Adult Intelligence. Baltimore, The Williams and Wilkins Company, 1939.

Wechsler D. The Measurement of Adult Intelligence. 2nd edition. Baltimore, The Williams and Wilkins Company, 1941.

Wechsler D. The psychologist in the psychiatric hospital. Journal of Consulting Psychology 8:281-5, 1944.

Wechsler D. Cognitive, conative and non-intellective intelligence. American Psychologist 5:78-83, 1950.

Wechsler D. The measurement and evaluation of intelligence of older persons. In: Old Age in the Modern World. Report of the 3rd Congress of the Association of Gerontology. Edinburgh, Livingstone, 1954.

Wechsler D. The Measurement and Appraisal of Adult Intelligence. Baltimore, The Williams and Wilkins Company, 1958.

Weisenberg T, Roe A, McBride K. Adult Intelligence: A psychological study of test performances. New York, The Commonwealth Fund, 1936.

Weissman MM, Myers JK, Tischler GL et al. Psychiatric disorders (DSM-III) and cognitive impairment among the elderly in a U.S. urban community. Acta Psychiatr Scand 71: 366-79, 1985.

Welsh, KA, Butters N, Hughes JP et al. Detection and staging of dementia in Alzheimer's disease: Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's Disease. Archives of Neurology 49:448-452, 1992.

Williams JM. Cognitive behavior rating scales: manual research edition. Odessa, Florida: Psychological Assessment Resources, Inc., 1987.

Williams JM, Klein K, Little M, Haban G. Family observations of everyday cognitive impairment in dementia. Arch Clin Neuropsychology 1:103-109, 1986.

Williams M. The effect of past experience on mental test performance in the elderly. Br J Med Psychol 33:215-9, 1960.

Williams MA, Campbell EB, Raynor WJ Jr., et al. Predictors of acute confusional states in hospitalised elderly patients. Res Nurs Health 8:31-40, 1985.

Williams MA, Campbell EB, Raynor WJ, et al. Reducing acute confusional states in elderly patients with hip fractures. Res Nurs Health 8:329-37, 1985.

Wilson RS, Rosenbaum G, Brown G et al. An index of premorbid intelligence. J Consult and Clin Psych 46:6:1554-5, 1978.

Wilson G. Models of ageing and their relation to policy formation and service provision. Policy and Politics 19:37-47, 1991.

Wing JK, Cooper JE, Sartorius N. The measurement and classification of psychiatric symptoms: an instruction manual for the PSE and Catego Program. Cambridge University Press, 1974.

Wurtman R.J. Alzheimer's disease. Scientific American 251: 62-74, 1985.

Yerkes RM. The Binet versus the point-scale method of measuring intelligence. Journal of Applied Psychology 1:111-22, 1917.

Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiat Res 17:37-49, 1983.

Zarit SH, Miller NE, Kahn RL. Brain function, intellectual impairment and education in the aged. J Am Ger Soc 26:2:58-67, 1978.

Zhang M, Katzman R, Salmon D et al. The prevalence of dementia and Alzheimer's disease in Shanghai, China: impact of age, gender and education. Annals of Neurology 27: 428-437, 1990.

Zun L, Gold I. A survey of the form of the mental status examination administered by emergency physicians. Ann Emergency Med 15:916-22, 1986.

Appendix A

PRE-OPERATIVE QUESTIONNAIRE: INTERVIEW 1

1. ID#: ____________ 2. DATE OF INTERVIEW: ______________

3. INTERVIEWER: ____

4. PT'S NAME: ____

5. DATE OF BIRTH: ..= Unknown

6. AGE:

7. MARITAL STATUS: 0 = Married 2 = Divorced/Separated
1 = Widowed 3 = Never married

8. SEX: 0 = Male 1 = Female

9. RACE: 0 = White 1 = Black 2 = Other

10. LIVING LOCATION: 0= Pt's Home
1= Other's home (specify person)___________
2= Congregate housing
3= Nursing home

11. WHAT IS THE NAME OF YOUR NEXT OF KIN (or PRIMARY CAREGIVER):

____

12. WHAT IS HIS/HER AGE:

13. WHAT IS HIS/HER ADDRESS?

____
Number Street Apt#

____
City State Zip Code

14. TELEPHONE NUMBER OF ABOVE: (____)________________

15. RELATIONSHIP TO PATIENT:

0 = NO P-C Provider
1 = Spouse
2 = Daughter/In-Law
3 = Son/In-Law
4 = Brother/sister
5 = Other relative
6 = Friend/neighbor
7 = Other
9 = Unknown

16. MAY I ALSO ASK THIS PERSON TO COMPLETE A BRIEF QUESTIONNAIRE?

Yes No

Pre-operative Questionnaire

17\. WHAT OTHER CLOSE FAMILY DO YOU HAVE APART FROM (11._________)?
(For each person): HOW OLD IS HE/SHE AND WHERE DOES HE/SHE LIVE?

Relationship:

1=Spouse
2=Dghtr/In Law
3=Son/In Law
4=Brother/Sister
5=Other relative
6=Friend/Neighbour

Location

0=Lives with Pt
1=In Walking Distance
2=In Metro Area
3=Out of Metro Area

No.	Relationship	Age	Location	No.	Relationship	Age	Location
1.				6.			
2.				7.			
3.				8.			
4.				9.			
5.				10.			

18\. HOW MANY YEARS OF EDUCATION DID YOU COMPLETE?

0 = Through High School: 01 02 03 04 05 06 07 08 09 10 11 12
1 = College: 01 02 03 04
2 = Graduate: 01 02 03 04 05 06
3 = Other: 01 02 03 04 05 06

19\. WHAT HAS BEEN YOUR JOB OR PROFESSION DURING MOST OF YOUR ADULT LIFE?

Specify___________________________________

0 = Professional
1 = Technical/Secretarial
2 = Skilled Manual
3 = Unskilled Manual
4 = Non-labor force

20\. (If still in labor force) HAVE YOU RETIRED FROM THIS? 0=No Yes=1

(If yes), which year did you retire?_______________

Pre-operative Questionnaire

21. NOW I AM GOING TO ASK YOU IF YOU TAKE PART IN THE FOLLOWING ACTIVITIES AND FOR HOW MANY HOURS EACH WEEK?

Activity	Hours per Week
Paid employment (If yes), Occupation____________________	____________
Volunteer work	____________
Meetings of Organizations/Clubs/Church	____________
Hobbies (specify)________________________	____________
Social events/eating out etc.	____________
Sports (specify)_________________________	____________
Other (specify)__________________________	____________
Total hours:	____________
Total Activities:	____________

22. ON AVERAGE, HOW MANY WAKING HOURS DO YOU SPEND WITH OTHERS EACH WEEK?

0 = Less than 5 hours
1 = 5 to <15
2 = 15 to <30
3 = 30 to 50
4 = More than 50

23. HOW MANY PEOPLE HAVE YOU SPENT TIME WITH DURING THE LAST WEEK? (Not on the telephone)

0 = No one
1 = 1 to 2
2 = 3 to 6
3 = 7 to 10
4 = More than 10

24. ON AVERAGE, HOW MANY TELEPHONE CONVERSATIONS DO YOU HAVE WITH PEOPLE EACH DAY?

0 = None
1 = 1 to 2
2 = 3 to 5
3 = More than 5

25. IS THERE A PET IN THE HOUSE? 0 = No 1 = Yes

Pre-operative Questionnaire

NOW I AM GOING TO ASK YOU SOME QUESTIONS RELATING TO THE SURGERY AND YOUR MEDICAL AND PHYSICAL STATUS:

26. WHAT TYPE OF SURGERY ARE YOU GOING TO HAVE?

0 = Cataract
1 = Other Eye Surgery: (specify)________________________
2 = Prostate
3 = Other Urologic: (specify)___________________________
4 = Other Surgery: (specify) ___________________________

27. WHAT IS THE SCHEDULED DATE FOR THE SURGERY?_______________

28. DO YOU HAVE ANY OF THE FOLLOWING MEDICAL CONDITIONS?

Disease 0 = No 1 = Yes

Heart Disease ______________________________

Hypertension ______________________________

Diabetes ______________________________

Arthritis ______________________________

Chronic lung disease or Asthma ______________________________

Kidney problems ______________________________

Cancer or leukemia ______________________________

Effects of stroke ______________________________

Other neurological problems ______________________________

Gastric ulcers ______________________________

Other stomach or intestinal problems ______________________________

Problems with circulation (arms/legs) ______________________________

Thyroid or other glandular problems ______________________________

Glaucoma ______________________________

Cataracts ______________________________

Other (specify) ______________________________

Pre-operative questionnaire:

29. WHAT PRESCRIBED MEDICATIONS ARE YOU CURRENTLY TAKING DAILY?

Specify each medication mentioned:-

Cardiac: ______________________________
____
____

Anti-hypertensive: ______________________________
____
____

Diabetic: ______________________________

Anti-allergenic/ ______________________________
anti-histamine ______________________________

Tranquillizer/sleeping pills______________________________

Analgesics ______________________________

NSAIA's ______________________________

Anti-gout ______________________________

Antibiotics ______________________________
____

Hormones ______________________________

Cortico-steroids ______________________________

Gastro-intestinal ______________________________
____

Vitamin/minerals ______________________________

Eye drops ______________________________
____

Others & unknown: ______________________________
____
____

Pre-operative Questionnaire

30. HAVE YOU EVER HAD TREATMENT FOR A PSYCHOLOGICAL PROBLEM?

0 = No 1 = Yes 9 = Unknown

(If yes), Was a diagnosis made? 0=No 1=Yes 9=Unknown
(If yes), specify______________

Were you hospitalized? 0=No 1=Yes 9=Unknown

Did you receive medication? 0=No 1=Yes 9=Unknown
(If yes), specify______________

31. DO YOU NEED HELP WITH ANY OF THE FOLLOWING ACTIVITIES?

ADL 0= Can Do 1 = Needs Help 2= Cannot do

Bathing ______________

Dressing ______________

Toileting ______________

Transfer ______________

Continence ______________

Feeding ______________

Ambulation ______________

32. DO YOU NEED ANY ASSISTANCE WITH WALKING?

0 = No 1 = Yes

(If yes), what? No Assistance ______________

With Cane ______________

With Walker ______________

With Person ______________

Wheel Chair ______________

Bed-Ridden ______________

Pre-operative Questionnaire

33. DO YOU EVER DRINK ALCOHOL? 0 = No 1 = Yes 9 = Unknown

(If Yes), How many days per week do you drink ________
How many drinks do you have on those days? ________
Total each week ________

34. HAVE YOU EXPERIENCED ANY RECENT CHANGE IN SEXUAL FUNCTIONING?

In the last year? 0 = No 1 = Yes
In the last month? 0 = No 1 = Yes
(If yes), explain____
____

35. HAVE YOU EXPERIENCED ANY RECENT PROBLEMS WITH THE FOLLOWING?
0=No 1=Yes 9=Unknown

Concentration ________

Confusion ________

Depression ________

Memory ________

Other ________

NOW I AM GOING TO ASK YOU TO DO A FEW TESTS FOR ME AND RESPOND TO A QUESTIONNAIRE ON HOW YOU FEEL ABOUT THINGS:

36. MINIMENTAL STATE EXAM: Score: ________
Maximum Score: ________

37. VERBAL FLUENCY 1 ________
2 ________
3 ________
Total ________

38. SIMILARITIES ________

39. GERIATRIC DEPRESSION SCALE: Score: ________

Geriatric Depression Scale: Yesavage, 1983

Choose the best answer for how you felt over the past week

1. Are you basically satisfied with your life? yes / no
2. Have you dropped many of your activities and interests? yes / no
3. Do you feel that your life is empty? yes / no
4. Do you often get bored? yes / no
5. Are you hopeful about the future? yes / no
6. Are you bothered by thoughts you can't get out of your head? yes / no
7. Are you in good spirits most of the time? yes / no
8. Are you afraid that something bad is going to happen to you? yes / no
9. Do you feel happy most of the time? yes / no
10. Do you often feel helpless? yes / no
11. Do you often get restless and fidgety? yes / no
12. Do you prefer to stay at home, rather than going out and doing new things? yes / no
13. Do you frequently worry about the future? yes / no
14. Do you feel you have more problems with memory than most? yes / no
15. Do you think it is wonderful to be alive now? yes / no
16. Do you often feel downhearted and blue? yes / no
17. Do you feel pretty worthless the way you are now? yes / no
18. Do you worry a lot about the past? yes / no
19. Do you find life very exciting? yes / no
20. Is it hard for you to get started on new projects? yes / no
21. Do you feel full of energy? yes / no
22. Do you feel that your situation is hopeless? yes / no
23. Do you think that most people are better off than you are? yes / no
24. Do you frequently get upset over little things? yes / no
25. Do you frequently feel like crying? yes / no
26. Do you have trouble concentrating? yes / no
27. Do you enjoy getting up in the morning? yes / no
28. Do you prefer to avoid social gatherings? yes / no
29. Is it easy for you to make decisions? yes / no
30. Is your mind as clear as it used to be? yes / no