Epidemiology and Community Health, 1979, **33**, 32-47

Theories and measurement in disability

R. G. A. WILLIAMS

From the Medical Research Council Medical Sociology Unit, University of Aberdeen

In an earlier paper (Williams *et al.*, 1976) a model of disability was described which gave rise to an appropriate measurement technique. The model was presented as 'an intuitive construct or concept of disability', and attention was concentrated on validating the measurement technique in terms of it. The present paper, by contrast, attempts to draw out the theoretical assumptions of that intuitive construct, and to compare them with two other theoretical positions. The aim is, firstly, to develop and demonstrate the explanatory power of the model, and, secondly, to derive from its explanatory principles further implications for the measurement of disability. It will be evident that the original conception and the first results of the scaling technique already published have had a reciprocal influence on one another, and the theoretical development is not easily characterised as either deductive or inductive, but it is presented in a quasi-deductive form in order to test it against fresh data and analysis.

The scaling technique described in the earlier paper had both strengths and weaknesses. Its strengths lay in its objectivity, and in the very minimal assumptions it required. It concerned the way in which disabled people demonstrably behaved, yet in most cases their patterns of behaviour turned out to be such that uncontentious judgements of disadvantage could be made. Moreover, the scale had the unusual virtue of detecting those individuals to whom its assumptions did not apply.

The weaknesses were in part those of the modest level of rank-order measurement achieved. Certain questions still had to be left to arbitrary judgement, including the relative ranking of differences between ranks in the scale. An interval scale would probably be needed for this task. There remained also the question of how to assign cases to whom the scale assumption did not apply. Finally, it was not clear how the theoretical model would determine the relevant universe of disability items. This paper attempts to make progress in these areas.

It is usual to distinguish between impairment, disability, and handicap, but there is an intermediate area which requires the distinction of a fourth category. Briefly, a condition may be described as belonging to

1. A class of diagnoses ('impairment').
2. A class of restricted activities defined in relation to a limb or organ which performs them, such as the movement of joints, or the flow of air from the lungs (sometimes 'impairment' or 'impaired function' but also sometimes 'disability').
3. A class of restricted activities defined in relation to the purposes of ordinary daily life, such as dressing or cooking (sometimes 'disability' and sometimes 'handicap').
4. A rank of disadvantage assigned to a person on the basis of his performance on any of the first three descriptions, usually the third ('handicap').

This paper is concerned with disability in the sense of the third description, as a limitation on activities of ordinary daily life. It seeks to explain the patterns of disabilities usually experienced. It first considers them as products of impairment, in the sense of either of the first two descriptions above. Then, after an intervening section, it considers them as products of the way in which people rank disadvantages when confronted with impairment. Old terms are used in order to avoid coining neologisms, but the reader is asked throughout to bear in mind these special definitions.

The next section focuses on the apparently commonsense position that the pattern of disability is determined in part by the impairments suffered.

THE MECHANISTIC THEORY

There is a view which assumes that as disability is caused by impairment, so the pattern of disability is caused by the pattern of impairment. If this can be described as crude mechanism, the theory which I outline here is refined mechanism. Refined mechanism points out that the same impairment can cause disability in one case and not in another. A common example is the finger injury which ruins the career of a pianist or violinist but does not disable the ordinary person. To explain the pattern of disability suffered, one must look not only at the type of impairment but also at the customary

activities of the person impaired. But since the impairment remains basic, its presence in a limb or body system creates disability in the person who uses that limb or body system for an important activity. Consequently, if a disability appears at all, it is referable to the locus of impairment.

There is a range of examples which lend credibility to this theory. Characteristically, an amputation or an injury is cited, but several other conditions are so obviously local as to make good examples. With many impairments, however, it is more difficult to define a specific 'locus' in the same sense. What is the site of a stroke, or of a circulatory problem, in this context? There are at least two ways of answering such a question. One is to concentrate on the site of symptoms, the other is to pursue the origin of the impairment in diagnostic terms, and to concentrate on the site of the disease. Neither approach will, however, define uniquely the locus of many impairments. Symptoms are often widely distributed through the body, and the number of possible disease sites is so large that further grouping of them is necessary. A mechanistic theory of disability patterns would have to fix on crucial symptom sites: to rank symptom sites in order of importance for ordinary activities, so that patterns of symptoms which included several sites were classified under the most important. Alternatively, pursuing the diagnostic approach, the theory would have to fix on crucial disease sites: to specify the body systems on which different activities primarily depend, so that disease sites would be classified under the body system to which they mainly contributed.

Supposing that all this can be done with reasonable agreement, the mechanistic theory is nevertheless not nearly as obvious as it appears to be in its characteristic examples. Moreover, some impairments which have an unambiguous site in one sense, such as a damaged lung or heart, are not easily seen as impinging on a specific group of activities. These impairments obviously put an end to athletics, but for the most part they simply require that activities should be carried out at slower speeds according to the time and energy available.

Finally, the very obviousness of the link between impairment and disability may be suspicious, for the validity of a commonsense axiom depends on the purpose of its use. The usual context in which the link between impairment and disability is made is when someone is offering an 'account' of his failure to meet his normal duties (Scott and Lyman, 1968). He justifies a specific disability *on the grounds* of his impairment; and in some circumstances, particularly when he has a highly visible and locally specific impairment, his claim may be more easily adjudicated than when a general loss of energy is involved. Consequently, the locus of impairment becomes a well-known resource for adjudicating this kind of claim. However, when the purpose at hand is explanation, not justification, and when the whole pattern of disability is involved and not one salient activity which happens to be relevant in the context, the link between impairment and disability is axiomatic only in the very broad sense that impairment sets some form of limit to activity in general. It may not be surprising, then, that a recent survey of research on rehabilitation remarks that 'We are thus left with the uncomfortable conclusion that traditional clinical measures of outcome are unrelated to functional* capacity' (Nichols, 1975). The next section explores an alternative way of conceiving the relationship of impairment to disability which may account for this.

RATIONAL CHOICE THEORY

Increasing breathlessness, because of malfunction of the circulation or the lungs, is a good example of an impairment which acts mainly to slow down the speed at which the sufferer can do things, to force him to divide into several stages an activity which he would otherwise do in one, and hence to cause him to cut down the daily outgoings from his budget of time and energy. The analogy with a money budget is obvious. There is a fixed 'income' of time to be spent, perhaps also (although this needs more qualification) a fixed 'income' of energy. It is also plausible that there is a normal 'budget' of expenditure on different items of activity, which is reached by a complicated trade-off of the varying 'utilities' of these items of activity in relation to the time and energy each demands. In this rendering of reality, impairment is like an increase in income tax, to which the consumer must respond by cutting down on aspects of his activities that he can best afford to lose.

More precisely, a disabled person who tries to match the cost of his effort to the importance of his different activities is very like the manager of the simple economy with which most economics texts begin. First he produces his own activities in varying combinations. At any point he is able to say how many units of activity A he would add to his day by giving up a unit of activity B. He is able to state the ratio at which he would substitute A for B, whether (1) in terms of what they cost him in time and effort, or (2) in terms of what increment of A would just compensate him in

*Function here means behaviour.

utility for a loss of B, so as to leave him indifferent between the two combinations. He can state (1) a 'cost ratio' and (2) an 'indifference ratio'.

All possible cost ratios between two activities for a given supply of time and energy can be drawn on a graph as a continuous line YZ (Fig. 1). The

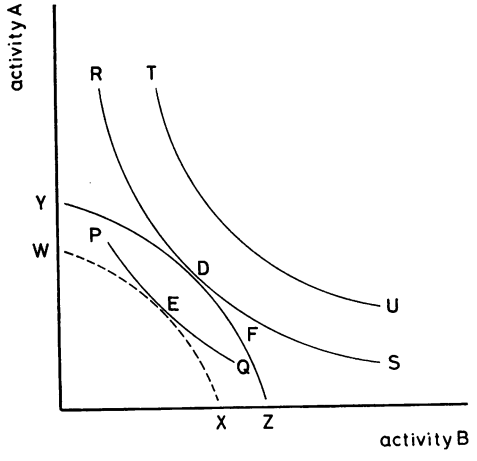

Fig. 1 *Rational choice model of disability.*

line could be straight, if both activities were performed equally efficiently by any combination of time and energy; or it could be curved, if for either activity a special combination of time and energy was more efficient than others. The curve chosen does not matter in the present context, and the line has been drawn with the convex curve usually found in economics.

A similar line PQ can be drawn of all possible ratios of the two activities which leave the disabled person indifferent compared with one given combination of the two activities. The line can be repeated for more and more plentiful combinations of the two activities, and the disabled person is assumed to prefer more plentiful combinations, that is, more of both activities, to less of both activities. The lines RS and TU represent increasingly plentiful combinations. Because a mixed pattern of activities is usually preferred to only one activity, this line is usually drawn with a concave curve, as in the lines PQ, RS, and TU in Fig. 1. Again, the shape of the curve does not matter here.

The point D, where the curve RS just touches YZ, is the point where the disabled person, using all the time and energy at his disposal, can just achieve one combination out of the most plentiful combinations which he would like equally well. This point is usually unique—the other plentiful combinations which he would like just as much are out of reach —but there could be several such points for all that it matters here. The important conclusion here is that at all other points on the cost curve the disabled person is still fully employed but to less advantage. Conveniently for what follows, the optimum point is characterised by a single ratio of substitution between the two activities which is the same for their relative costs in time and energy as for their relative value to the disabled person. In terms of the graph, the tangent of both curves at D is the same.

If impairment now reduces the total level of activity of the disabled person, what happens? In terms of Fig. 1 the cost curve shifts inwards, say to the dotted line WX. A new optimum point E is established. Both activities are lost in a certain ratio. The ratio in which activities are simultaneously reduced need not be the same as the ratio in which they are substituted for each other, but they will in fact be the same if the disabled person's indifference curves and cost curves remain parallel at the reduced level of activity. The practical meaning of this condition can be illustrated by an example. When healthy, a woman cooking eight hot meals a week and going out about six times a week may be able to trade one hot meal for one extra trip out of the house. If seven hot meals and seven trips appear just as good to her, the odds are that her optimum point is somewhere between these two combinations, and for the next hot meal, which she could give up for another trip out, she would want a little more than one extra trip to compensate her. Now impairment reduces her capacity both for cooking and for going out of the house. On the whole it would be expected that her best combination would still lie in a roughly equal division between cooking hot meals and going out, say five hot meals a week and four to five trips out of the house. This expectation is stated more precisely in the assumption that the cost curves and indifference curves of the two activities remain parallel at the reduced level of production. The woman might, of course, stick to cooking seven hot meals and reduce her trips out to two or three, but in that case something quite startling must have happened to her original willingness to trade one activity for another if it suited her. Perhaps she always cooked seven hot meals a week on principle, and therefore would not trade below that level in any circumstances. If it was not a matter of principle, but still of trading preferences, then either the relative costs of cooking and going out change quite sharply when only a little of them is done, or their relative value changes sharply.

A 'sharp' change means something precise. If a reduction of overall activity brings some reduction, however small, in every activity, the change in

relative cost or relative value is not sharp as the word is meant here. If any activity actually increases or remains unchanged, then the change is 'sharp'. An increase of an activity does not mean that it takes longer—that would be an increase in its cost. It means that more of it is performed—for example, eight hot meals are cooked instead of seven. Although this is an unlikely result, it is not impossible for relative costs or relative values to change sharply at low levels of subsistence. There is a standard economic example. As a man gets poorer, he may eat more potatoes or bread. Thus there may be activities which, like potatoes or bread, are an 'inferior good'. Watching television is probably one of these activities.

It will be assumed, for the moment, not that the cost curves and indifference curves remain parallel, which is too strict an assumption, but that they do not change 'sharply' in the sense of revealing any particular activity as an inferior good. Meanwhile there is one other situation in which a decrease in overall activity may lead to an increase in a particular activity. This situation is one in which a disabled person is employed in some forced combination of activities which is a considerable way from his optimum, and then, when his overall level of activity is reduced, is able to choose the combination which he himself judges best. This change is represented in Fig. 1 by, for example, a move from F to E. It will also be assumed for the moment that this situation does not hold.

The two assumptions thus made allow a simple calculus to be presented of the disability patterns which ought to be found in practice, if impairment does constitute a general increase in the costs in time and energy of all activities, and if people respond to this by budgeting losses in their activities according to their own preferences. The cost increase in this calculus is presented like a sequence of higher and higher 'tax' demands which must be paid with units of activity. The 'taxes' are in increments of one cost unit at a time, the unit being set equal to the cost of the cheapest activity unit being considered. The calculus has the following elements.

Firstly, for the sake of simplicity, each activity is reckoned to have a critical point above which a person is non-disabled (= 0) and below which he is disabled (= 1 or more). Thus the differentiation of disability patterns does not go beyond the case where all activities are disabled in this sense. If the critical point (as in many disability measurements) is the receipt of personal help, the present rational choice model would tend to predict that when all activities are disabled, the help received with all of them is greater the less the importance which the disabled person attaches to performing the activity himself. Broadly, the picture would be one of a little help received in everything, and nothing done entirely for the disabled person. The meaning thus given to disability in all activities needs to be borne in mind, because it is theoretical and may not correspond to the reader's experience. The focus here, though, will be on the transitional patterns of disability and whether these correspond to experience. If not, then this picture of overall disability will not do so either.

The transitional patterns might not be particularly interesting if they passed by rapidly. This would occur if activities were lost in very small units so that change was virtually continuous, as it is with most economic goods. But there is plenty of evidence that the transition is slow and that disability in all activities is extremely rare, and this could not be so if activities were traded in very small units. Moreover, it is well known that many activities are performed in 'packages' which are worthless if not completed. A trip to do the shopping is wasted if the shopping cannot be done. Hence it is necessary in studying a transition to focus more closely on the stepwise interchange of whole units of activity than is usual in economic applications of rational choice. Consequently the second element of the choice calculus presented requires losses of each activity to be paid in units, not fractions of a unit.

Thirdly, since a reduction in overall activity has been assumed to involve a reduction, however small, in every activity, a positive number can be assigned to each activity which represents the ratio at which the next unit is lost compared with the next unit of other activities. It is not necessary to know what the ratio is in fact; one need only represent the range of what it could be. It is easiest to present this figure as if it were a price: that is to say, if four units of A are lost for one unit of B, it is easiest to say that one unit of B is four times more valuable or costly than one unit of A. In the calculus B then appears with a weight of four and A of one. Taking as an example a world with only four activities, weights can be assigned to represent varying degrees of inequality between these activities. Two schemes are presented in Table 1, one with uniform weights, the other with unequal weights in arithmetic progression.

These two schemes of weighting naturally do not exhaust the variations in weights for disability which have so far been suggested in the literature. They do represent, however, the ordinary range produced by rules of thumb. One common rule of thumb is to assign each disability a score of one and treat the total as the disadvantage suffered. More sensitive schemes, on the other hand, partition degrees within one disability, assigning them successive scores of, for example, one, two, and

three; or distinguish major and minor items with scores perhaps of one for minor and two for major items. The general effect of such distinctions is to introduce unequal weighting for disabilities which will approximate at most to the ratios of an arithmetic progression. There are also theoretically sophisticated approaches which claim, though tentatively, that professional judges weight health status in similar ratios. Hence the two schemes shown here represent practice in assessing disability, and offer the justification which it requires in terms of a theory of consumer preferences.

Fourthly, the interest is in the *next* unit of each activity to be lost, starting at the point where all activities are just non-disabled. This means that, because some activities lose that unit before others, a snapshot is required of the process of transition up to the point when the next units have been lost from all activities. If the next unit of activity B is four times as valuable as that in activity A, then the calculus must normally allow payments of four units of A to be made but no more, at least until the next unit has been paid from activity B. At this point all activities are in any case disabled. Thus the calculus must normally show losses of activity in the stated ratio of preference. The word 'normally' is used to show the essential pattern, but reality is a little more confusing, because activities are lost in whole units. The point at which all activities are just non-disabled is a point in the middle of the process of loss, where it may be, for example, that four units of A have been lost since the last unit of B was lost. Even if B is four times more valuable than A, it is now B's turn to lose a unit. The disabled person will give up no more of A until he makes another sacrifice of B. The ratio in which he loses A and B remains the same, however, as that of another disabled person who has just sacrificed a unit of B, and who consequently will sacrifice no more until he has lost another four units of A. The variation in disability patterns resulting from this has limits which are described by a formula.*

It may seem arbitrary, at first sight, to say that the calculus starts at the point where all activities are just non-disabled. There are people who allow some activities to drop well below a 'normal' minimum while maintaining others well above it. The minimum 'non-disabled' level could of course be defined as that which each disabled person would trade all activities to maintain, so long as he also kept above his minimum level for any activity. Public definitions of disability could then be treated as mere approximations to, or 'averages' of, such perceptions. An alternative approach is to make use of the fact that personal help plays a large part in defining disability. The significance of such help lies in its being 'altruistic', that is, not fully rewarded at the expense of the recipient. It signifies that a social minimum has been reached by the disabled person which his relatives (or the taxpayer) do not want to leave unremedied even at their own cost. Because the disabled person can switch his resources of time and energy from one activity to another, his passing the social minimum on one activity will not stimulate help if he could switch resources to that activity without threatening to pass the minimum in others. But whatever the approach, the non-disabled level, even if it is defined separately for each activity, tends to be a minimum level of production beyond which any activity may receive help. This is where the calculus begins.

Table 1 represents the disability patterns which would rationally be chosen if impairment acts so as to increase the costs of all activities and the sufferer responds by budgeting losses of activities solely according to his preference. The weights represent the values or costs set on each activity, and the rows show the activity units lost so as to lose minimum value for each cost increase. The rules of the calculus which have been discussed are summarised as follows:

1. Cost increases must be paid in whole units of activity, not fractions.
2. Multiple units of one activity must be paid according to their preferred ratio of loss, allowing only for uncertainty about which activity has most recently lost a unit (see footnote below).

*The value of an activity unit P has been defined in relation to its rate of loss. If $Px_i + Px_j + \ldots Px_n = K$ where K is the total loss sustained and x, as multiplier of P, gives its rate of loss, then x_i is defined by the equation $x_i = \frac{P_j}{P_i} x_j$. However, because the x_i must be whole units, not fractions, uncertainty enters, and either of the following equations may be equally likely:

$$(1)\quad x_i = \left[\frac{P_j}{P_i} x_j\right]$$

$$(2)\quad x_i = \left[\frac{P_j}{P_i} (x_j - 1)\right]$$

where the square brackets signify the integral part of the value. The probability of all points in the interval between these two values of x_i appears, in the absence of any reason to deny randomness, uniform. Hence the limits of uncertainty can be stated as the following constraint:

$$x_i \geq \left[\frac{P_j}{P_i} (x_j - 1)\right]$$

and all solutions for K which conform to this requirement can be treated as equally probable.

Provisionally, it would appear from Table 1 that these rational choice schemes in which the only constraint on choice is a sort of budgetary limit set by impairment would generate surprisingly diverse patterns of disability. If the patterns are in fact more restricted, however, the task will be to identify the additional sources of constraint. The mechanistic theory described earlier may be able to show one source of constraint. The next section identifies another. It may be that a rational choice theory could apply only within the area of freedom left by these constraints.

Table 1 *Weighting disability. Two typical four-item schemes: (A) uniform and (B) arithmetic progression*

SCHEME A

Weights = 1 1 1 1

'Tax'	*Scaling pattern*					*'Tax'*	*Scaling pattern*				
1	0	0	0	1	*or*	3	0	1	1	1	*or*
	0	0	1	0	*or*		1	0	1	1	*or*
	0	1	0	0	*or*		1	1	0	1	*or*
	1	0	0	0			1	1	1	0	
2	0	0	1	1	*or*	4	From this point all items lose one or more units.				
	0	1	0	1	*or*						
	1	0	0	1	*or*						
	0	1	1	0	*or*						
	1	0	1	0	*or*						
	1	1	0	0							

SCHEME B

Weights = 4 3 2 1

'Tax'	*Scaling pattern*					*'Tax'*	*Scaling pattern*				
1	0	0	0	1		6	1	0	1	0	*or*
2	0	0	1	0	*or*		0	1	1	1	*or*
	0	0	0	2			1	0	0	2	*or*
3	0	1	0	0	*or*		0	0	2	2	
	0	0	1	1		7	1	1	0	0	*or*
4	1	0	0	0	*or*		1	0	1	1	*or*
	0	1	0	1	*or*		0	1	1	2	*or*
	0	0	1	2			0	0	2	3	
5	0	1	1	0	*or*	8	1	1	0	1	*or*
	1	0	0	1	*or*		1	0	1	2	*or*
	0	1	0	2	*or*		0	1	1	3	
	0	0	1	3		9	1	1	1	0	*or*
							1	1	0	2	*or*
							1	0	1	3	*or*
							0	1	2	2	*or*
							0	1	1	4	
						10	1	1	1	1	*or*
							1	0	2	2	*or*
							0	1	2	3	
						11	1	1	1	2	*or*
							3	2	0	1	*or*
							3	1	2	0	*or*
							4	2	1	0	
						12	1	1	1	3	*or*
							0	2	1	4	
						13	1	1	2	2	*or*
							1	1	1	4	*or*
							0	2	2	3	
						14	1	1	2	3	*or*
							0	2	2	4	
						15	From this point all items lose one or more units.				

As the amputated limb is the classic case for the mechanistic theory, and breathlessness for the rational choice theory, so there is another kind of classic case which suggests the third theory. I have referred to the role of impairment in giving an 'account', a justification, of disability. Local impairment offers a simple and visible resource for adjudicating such accounts. By contrast, the free play of rational choice may make such accounts very difficult. Rational choice would offer such a variety of disability patterns, based on such abstract principles, that adjudication by families, workmates, friends, nurses, and doctors would be made very difficult. These considerations suggest that an important constraint on the pattern of disability may be its conformity to convention, a convention which is necessary for communicating the legitimacy and ordinariness of the disabled person's failures to those around him. The development of this idea will require reference to the sociological theory of deviance.

THE DEVIANCE THEORY

Two valuable discussions of disability or handicap as forms of deviance (Freidson, 1972) or stigma (Goffman, 1963) emphasise the restrictions placed on a disabled person by the necessity of his being classified or 'placed' by others. In Goffman's analysis of stigma the 'others' are usually met in passing, in streets, at parties, at shops, in applying for a job; and the categories against which they try to match the disabled person are correspondingly crude. It is the failure of the disabled person to match one of these crude categories, and the attributes felt to be 'ordinary and natural' for members of it, which constitutes stigma. In Freidson's argument, by contrast, the 'others' who must classify the disabled person are primarily agencies of rehabilitation or care. Here it is not the failure of the disabled person to match a category but his success that is significant. Success means that he represents a certain sort of work for the agency; he receives a definite response from the professional, whereas he receives an embarrassed or equivocal one from the stranger. In addition to these 'others' with whom the disabled person must communicate, yet another group is involved in Freidson's argument, the 'public' or the 'community', who are tempted to label the clients of agencies in their own way, organising for them a segregated deviant role. Unlike Goffman's public, this public has a definite category for the disabled person, and definite notions of how to respond to him; he is to act the role of the blind man, the village idiot, or the cripple.

Goffman's public are strangers and Freidson's are neighbours. Between both authors, there emerges a nascent theory about the classification of disabled people by strangers, neighbours, and professionals. But among these social groups there is a conspicuous omission: the family. The

understanding of his position by members of his family is obviously critical for a disabled person; but in classifying, naming, and understanding his position they are, of all people, the least likely simply to follow the patterns so far set out. Of course, elements of those patterns may well be present; sexual stereotypes are so strong that husbands or wives may continue to be embarrassed by a sense of the spoiled identity of their partners; families are closely involved in defining their members as work for professional agencies; and families may also organise for themselves butts, scapegoats, invalids, and other deviant roles. But families in which a disabled person is seen simply as a spoiled man or woman, as a butt, or as somebody else's business, reveal a surprising crudity of classification and understanding. In what ways, then, should the more intimate knowledge which families have of their members be manifested?

The chief crudity of classification, which seems to belong more appropriately to encounters in the market place, the agency, and the street than to family relationships, is the tendency to dichotomise. It is only a highly specific relationship or a very simple social order which offers two alternative lines of response, acceptance or rejection, the offer of a fixed role or the avoidance of contact. More diffuse relationships, involving a complex of expectations, not all necessarily compatible with each other, allow the persons taking part to be matched against what is expected of them in degree rather than in kind. The argument is that a notion is held, primarily by family members, but secondarily also by professions, of *relative* deviance from usual expectations. This notion very clearly organises response to a disabled or sick person and, in doing so, directs the sequence of disabilities experienced in the course of illness and impairment.

For anyone recently recovered from an illness, it is instructive to review the steps that marked his recovery in the talk which passed between him and those looking after him. The steps are very frequently the product of a social activity which may be called 'constructing progress'. In this activity, a curriculum is set up which can bear, not only to the patient but more importantly to his nurses, the meaning of uniform advance. The logic by which this is achieved is that of cumulation on a single dimension. Just as children are socialised by a long curriculum in which the capacities gained earliest are expected to be retained while new capacities are added, so the sick or disabled deviant is socialised or rehabilitated by a curriculum in which activities are gained and retained in due order along a single dimension. Furthermore, in both cases, since the curriculum is a negotiated order, idiosyncratic pleasures of the unsocialised person will take second place to activities typical of his conventional position. The cumulative and conventional characteristics of this process naturally limit severely the variety of disability patterns available on a system of rational choice.

The benefits of the process lie in justifying and making intelligible the pattern of the disabled person's failures, and in restricting the immediate demands upon him to those he can properly manage. Just as a normal child is measured by his progress along the curriculum, so the deviance of illness is made normal with reference to a similar construction of progress. The adult patient and the child are both offered a succession of temporary identities with their own ordinary and natural attributes which are recognised by the members of their social group. At the same time, there is a further use of the curricular logic, and that is to create a measurement of distance between temporary and ultimate identity as a full adult member of society. (Of course childhood illnesses have their own curricula aimed at restoring the child to the activities appropriate to his age). The principal difference between the way distance is created for the child and for the adult patient lies in the direction of movement through the curriculum. The flow along the educational ladder is one-way, with a continuous succession of new cohorts at the bottom; the flow on the rehabilitative ladder forms a great variety of two-way movements, with succession of new cohorts from the top. Hence, distance on the educational ladder can be measured by average 'years' of attainment, and a child is normal for a six-year-old or is abnormal in having a mental age of four; but distance on the rehabilitative ladder is measured rather by dispersion below the norm represented by the healthy. Where ordinary experience is the norm, the distance of a deviation is the comparative rarity of its experience. This proposition does not make all rare experiences deviations; on the contrary, only those conditions of a person which are marked by others as undesirable, and which are ordered into a rehabilitative curriculum, can be said to be deviations (in their eyes) having a property of *distance* from the norm; and it is the distance of this kind of deviation only which it may be appropriate to construct on the basis of the relative rarity of its experience. That sick or disabled people make some computation of the frequency of their experience amongst others, and assess their good or ill fortune on that basis, could be illustrated in a number of examples. Age group enters into the computation, more specifically in later years, but

since older people tend to attribute a component of their distance from the norm to age, and a component to illness, the essence of the computation may be fairly uniform among adults, relating to what they know of the incidence of incapacity in all age groups.

The particular consequence for disability patterns of the present discussion is the hypothesis that failures in activity which are associated with one's identity as a normal person in the eyes of others will be found to cumulate along a single dimension which also corresponds to the frequency of occurrence of the component disabilities. Diagrammatically, a typical set of disability patterns and prevalences according to this theory is shown in Table 2. As before, an activity performed is O and an activity failed is 1.

The pattern in Table 2 is important because it corresponds to the shape taken by a Guttman scale. One of the ways of computing a Guttman

Table 2 *Deviance theory: typical disability patterns*

Disability pattern	*Activity items*			
	A	B	C	D
1	0	0	0	1
2	0	0	1	1
3	0	1	1	1
4	1	1	1	1
No. failing each item	20	40	60	80

scale is to order the items initially according to their prevalence, and this, too, corresponds to the requirements of the hypothesis. Similarly, for a Guttman scale to exist in a given body of data, items must be failed or passed in a cumulative order, and this order must be uni-dimensional for the whole group of people who are assessed. Consequently the first test of the deviance theory of disability is whether, when disability items are ordered by their prevalence, the patterns of individuals fall into a Guttman scale.

However it is not necessary for all individuals to follow the same scale, but for all who can be shown to share a common role within a common social group or culture to follow the same scale. No great precision can be shown about this. It has already been suggested that it is not merely families who are involved in constructing the position and progress of disabled individuals; families and professionals exchange cues on the subject, and friends and neighbours are additional witnesses, commentators, and agents of conformity. But if the relevant social group is much bigger than the individual family, it is by no means predictable how big it would be, nor where the typical construction of progress made by one group might differ from that made by another. All that can be said is that differences may be looked for, and if they are aligned with different roles or with different social or cultural groupings, that would be additional evidence for the deviance theory against its present rivals.

Again, it is not necessary that any items of activity which could count as disabled should form a cumulative scale. On the contrary, only those items should scale which threaten the conventional identity of an ordinary person. This definition, too, is a cultural one, although one would guess that activities associated with adult identity would vary less in their nature than in their relative priority.

Finally, the Guttman scale is by itself a static concept; but the argument has proceeded from an essentially dynamic notion of constructing progress. Change in disability patterns therefore offers a further test of the thesis. Should all change occur along the scale pattern? The argument would rather be that all *socially controlled* change should occur along the scale pattern. Rehabilitation is certainly controlled but deterioration may not be, although one would expect the organising forces of convention to bring a deteriorated pattern into order after a certain period of time. If this is so, cases which remain static after that period should also be organised into scale patterns.

Three theoretical approaches to explaining the pattern of disability have now been presented. The next step is to compare them against the results of old and new data.

DATA

Two sets of data have helped in evaluating the theoretical possibilities which have been outlined. The first set, from which some relevant results have already been published in the earlier paper, consists of home interviews with 157 women and 88 men aged 35 to 74 living in north Lambeth, London, in 1967. All were randomly sampled from a census of the area and identified as being disabled in one or more of four areas of activity: self-care, domestic duties, mobility, and occupation. General aspects of the survey have been reported elsewhere (Bennett *et al.*, 1970; Garrad and Bennett, 1971).

The second set of data consists of two extensive pilot studies carried out by the Institute of Social and Economic Research at York on disabled people selected for the purpose by a number of medical and social agencies in England and Scotland in 1973–74. Both studies include disabled people at home, in hospital, and in local authority accommodation. The first study involved two interviews separated by three months, thus introducing evidence on disability changes, and initial data was obtained from 154 people living at home and 111 in institutions. The second study, a single interview,

obtained data from 206 people living at home and 133 in institutions. Again, a report on the survey procedure has already been published (Wright, 1974).

DEVIANCE OR RATIONAL CHOICE

The data can first help to illuminate the varying inferences derived from the rational choice schemes and the deviance theory about the prevalence of varying patterns of disability. If rational choices about activities to be lost are made within a 'budget' on the two schemes so far constructed, the Guttman scale hypothesis should be rejected. If activities are lost in accordance with a symbolic construction of the impaired person's position and progress, the Guttman scale hypothesis should be confirmed.

The earlier report already referred to showed that Guttman scales obtained among men and women in Lambeth. However, it was noted that, although high levels of confirmation were achieved, the procedure could conceivably capitalise on chance alignments of the data and should ideally be replicated. Here, then, the replications performed with the two sets of York data are reported.

The choice and definition of items presented many difficulties. The Lambeth questionnaire had not been formulated with deviance theory in mind, and selection between the many coding distinctions it offered was made early in the development of the ideas presented here. But in retrospect the Lambeth items seem to have been selected consistently with the theory, although they were unlikely to have covered all the ground, and the mobility items in particular stood as proxies for ranges of activities which should ideally be specified in terms of ordinary purposes (Williams *et al.*, 1976). The coding distinctions available in the York data presented two new questions:

1. Which definitions fitted the Lambeth items most closely?
2. Which additional items were relevant?

The decisions on these two questions were not easy. On two items, it was necessary to take into account the distribution of answers in interpreting the fit of the first York study. There is a risk in such reasoning, because the deviance hypothesis partly concerns the relative prevalence of failures on each disability item. However, the analysis of the second York study could proceed automatically from decisions made on the first, so the risk was not serious.

The second problem presented by the York data was to decide which additional items should be included from those available. On the criterion sketched earlier, that items included should represent activities belonging to conventional adult identity, the following decisions were made:

Included—washing without help (question different from Lambeth).

Excluded—incontinence, cutting toe-nails without help, pursuit of personal interests.

The exclusion of incontinence, an item obviously bound up with identity, was made on the grounds of the partially autonomous nature of bowel and urine retention, disqualifying it as an activity in the sense of the other items.

When these decisions had been made, replication involved three things: firstly, the scaling properties of the York data with the same order imposed on the same items as in Lambeth; secondly, the scaling properties and comparative sequence of the same York items when ordered, as the Lambeth items originally were, according to the frequency of disabled responses; and, thirdly, the scaling properties of *all* the relevant York items, old and new, ordered according to the frequency of disabled responses.

The conventional criteria of a Guttman scale are the coefficients of reproducibility (values above 0·90) and of scalability (values above 0·60: the SPSS computation, used here, is conservative, underestimating the coefficient; see McConaghy (1975).

The coefficient of reproducibility averages item reproducibilities which are also quoted here, and which should not fall below 0·85. The relevance of many of these rules in a series of actual replications, rather than in estimating the replicability of a single instance, is dubious. In clinical and administrative dealings with disability it is more useful to study the case, the aggregate behaviour of groups of cases, and prediction about the future of the case, than the item response or the scale as a whole, and it is appropriate to extend the reporting of Guttman scale properties to meet these uses.

Initially, therefore, reliance is placed on the reproducibility coefficients and the coefficient of scalability as summary indicators of the extent of scaling; but after that it is much more useful to study the proportion of nonscale cases occurring, the proportion in which only one 'error' is observed, and whether there is any sign of repeated nonscale patterns.

The first test presented reckons the reproducibility and scalability of the Lambeth items when imposed in the Lambeth order on the York data (Tables 3 and 4). Seven to eight items were replicable for women, four to six for men, since not all the questions asked in the first York study were asked in the second. The three samples are referred to in the Tables as L (Lambeth), N1 (the first national study done from York), and N2 (the second York study).

Table 3 *Women: disability in Lambeth and national samples*

Order as in Lambeth	*Item reproducibilities* L	N1	N2
1. Eat without help	1·00	0·97	—
2. Get out of bed, with help	1·00	0·99	0·99
3. Sit and stand without help	1·00	0·92	0·97
4. Use WC (L: or commode) without help	1·00	0·97	0·97
5. Dress without help	0·94	0·91	0·96
6. Cook all meals (N1 and 2: usually)	0·90	0·92	0·93
7. Walk out of doors unaccompanied (N1 and 2: without help)	0·90	0·93	0·93
8. Do all own shopping (L: and cleaning, laundry)	0·96	0·98	0·99
Overall reproducibility	0·95	0·95	0·96
Coefficient of scalability	0·69	0·69	0·78
Nos. in each sample	157	119	161

Table 4 *Men: disability in Lambeth and national samples*

Order as in Lambeth	*Item reproducibilities* L	N1	N2
1. Eat without help	0·99	0·97	—
2. Get out of bed, with help	0·99	0·97	—
3. Use WC (L: or commode) without help	0·99	0·94	0·93
4. Sit and stand without help	0·98	0·91	0·98
5. Dress without help	0·88	0·97	0·98
6. Walk out of doors unaccompanied (N1 and 2: without help)	0·93	0·94	0·93
Overall reproducibility	0·95	0·95	0·94
Coefficient of scalability	0·71	0·63	0·45
Nos. in each sample	88	35	45

In terms of reproducibility and scalability, two replications well above the conventional levels for single instances were obtained for women with samples of very adequate size. For men sample size was problematic; also, a very high proportion in both N1 and N2 were not disabled on any item: in fact only 14 cases in N1 and the same number in N2 were disabled on any item. Results for men should therefore be regarded merely as straws in the wind, not as replications in themselves. N1 provides a fair omen; N2 has acceptable reproducibility but poor scalability. In the remaining tests scales for men will be ignored.

The second test presented orders the Lambeth items in each sample by the frequency of disabled responses which the item generates (Table 5). Reproducibility and

Table 5 *Women: disability in Lambeth and national samples*

Order by item frequencies L	N1	N2	*Frequencies of disabled responses* L	N1	N2
1	2	—	1%	3%	—
2	1	2	1%	2%	1%
3	4	3	7%	12%	4%
4	3	4	8%	6%	6%
5	5	5	17%	14%	12%
6	6	6	32%	28%	27%
7	7	7	33%	54%	45%
8	8	8	78%	82%	77%
Overall reproducibility			0·95	0·95	0·96
Coefficient of scalability			0·69	0·70	0·78

scalability would be expected to be high, since they were high for women on the first test. Here the interest lies in considering the stability of item order in replications. Item frequencies are also shown, and it will be noticed that in studies of disability several items are very rare. Because of this, the tail of the disability scale may be liable to instability until enough numbers and enough replications of fairly numerous samples can establish the pattern. In the present case N2 presents the same order as Lambeth, while N1 switches two pairs of adjacent items. Thus, although there is some instability in the tail of the scale, it is quite circumscribed, and suggests that larger samples might eliminate it.

The third test proposed adds to N1 and N2 the washing item which was considered different from the Lambeth item but relevant to the deviance theory. The York question ran 'Do you wash yourself without help from anyone'? The Lambeth question was very similar but it was followed up by questions on help with shaving, combing hair, etc., for which a separate code was supplied and included as a disabled response. When the York item was added to the scales for N1 and N2, ordered by item frequency, the coefficients obtained were as follows:

	N1	N2
Reproducibility	0·95	0·96
Scalability	0·62	0·74

In terms of the order shown in Table 5, however, the item was interpolated differently in N1 and N2, between items 4 and 5 in N2 and between items 4 and 3 in N1. Although in N1 the item remains between items 4 and 5 as in N2, the switch of items 4 and 3 which has already been described involves the present item in the anomaly.

The patterns of disability derived from deviance theory and from the two rational choice schemes can now be compared. The upshot is so far in favour of the deviance theory. In Table 1, the uniform weighting scheme would produce only 21% of its transitional patterns in scale form. The weighting scheme which follows an arithmetic progression would produce a greater proportion of scaling cases (57%), but that proportion is still a long way below the actual proportion observed if the first four Lambeth items are treated as a separate scale (75%). It is clear, therefore, that if disabled women budget losses of their activities according to their preferences, they do not behave as if the critical unit of the activities which they lose had a uniform value for them. Moreover, if these women, in deciding their preferences, give unequal weights to the critical unit of an activity the loss of which makes them disabled, the inequality is not as gentle as the ratios of an arithmetic progression. It looks at the moment as if a steeper progression may produce a result consistent with reality, and in this case it may be necessary to integrate (where that is possible), or else adjudicate, the claims of deviance theory and of rational choice. An alternative is that, at successive stages in recovery or deterioration, each of the activities considered in the data becomes an 'inferior good' in the sense defined earlier, that is, the activity is maintained or increased in substitution for other activities (whether mentioned in the data or not) which have been drastically curtailed. This is an important and interesting possibility in explaining many activities characteristic of the disabled, such as watching television or doing occupational therapy. But special evidence would be needed to allow its relevance for the activities in these data, which are usually presumed to be performed to a higher standard by healthy people. For the moment, then, deviance theory offers the most immediately employable explanation of disability patterns, although the rational choice theory has potential which has not been fully explored.

These findings mean that, if the mechanistic explanation of disability now proves to be false, and the behaviour of the disabled is explained by their values, then apparently most ordinary practice in assessing disability (which tends to imply something between uniform and arithmetically progressing numbers for the

items) does not have the justification presumed of it, of reflecting 'consumer' values. For if the consumer choice model of disabled behaviour is valid, it is valid only for some steeper progression yet to be tested.

DEVIANCE OR MECHANISTIC EXPLANATION?

The mechanistic theory did not entail a cumulative and unidimensional scale of disability, as it was propounded earlier. But it need not be inconsistent with such a scale. Indeed, a competent authority might be able to draw out a cumulative pattern in the impaired sites which are typical of major diagnostic categories such as arthritis or stroke, and it is not inconceivable that the patterns would have enough in common to allow a unidimensional and cumulative scale of impairment sites to be constructed. Attention is directed here, though, to testing the lesser but necessary claim that particular impairment sites coincide with particular levels of accumulated disability.

The Lambeth data included an extensive documentation of impairment which was not available in the York data. The organising principle was the diagnostic one referred to earlier, in which the disease site is named and related to a classification of body systems in the aspects which impinge on activity. Four groups of impairments were thus defined: locomotor, internal, sensory, and a residual category which was primarily mental. However, an examination of the data underlined the complexity of symptom patterns: 94% of Lambeth women had symptoms which could provisionally be assigned under two or more of these headings, and diagnostic skills were essential in arriving at the underlying classification of impairment site. If, however, it was argued that for the women their symptoms were more important than the seat of the disease in determining what they could not do, there would be a need for some different way of defining which symptoms had priority. This second classification has been attempted on the basis already discussed, that symptoms evidently assignable to specific limbs form the classic case in the mechanistic argument. Hence symptoms in upper or lower limbs, or in both, should coincide particularly clearly with specific levels of accumulated disability, and these symptoms are frequent; indeed they are found in the majority of the disabled women.

In Tables 6 and 7 both classifications of impairment site are related to cases which scale perfectly among the Lambeth women. These cases form the critical test for the ability of the mechanistic theory to explain the basic structure of disability patterns. The separate implications of nonscale cases are treated below. Symptom sites are unrelated to scale structure. The argument must therefore rest on the relation of impairment sites to accumulation of disability, and here the evidence of Table 6 is ambiguous.

A relationship between impairment site and scale structure certainly exists. It permits the prediction that locomotor impairments will more probably be associated with heavier disabilities, and that internal impairments will, on the whole, be associated only with the first level of disability. But the prediction is far more accurate in reverse—given the severest accumulations of disability, it is very probable that locomotor impairment is involved. It appears that impairment site as a causal agent, and therefore as a predictor, is somewhat unspecific. This finding gives very uncertain support to a mechanistic explanation.

Table 6 *Scale types by impairment site*

	Scale types: number of disabilities			
Impairment site	*1*	*2–4*	*5–11*	*N*
A. Locomotor	16	10	13	39
B. Sensory	4	3	0	7
C. Internal	20	9	2	31
D. Other, including mental	3	5	1	9
	43	27	16	86

χ^2 (A versus B + C + D) 10·223 (2 df).
$P < 0{\cdot}01$.

Table 7 *Scale types by symptom site*

	Scale types: number of disabilities			
Symptom site	*1*	*2–4*	*5–11*	*N*
A. Lower limbs	16	10	10	36
B. Mixed limbs	2	1	1	4
C. Upper limbs	7	5	1	13
D. Central	18	11	4	33
	43	27	16	86

χ^2 (A + B versus C + D) = 3·916 (2 df).
$P > 0{\cdot}10$.
χ^2 (A + B + C versus D) = 1·495 (2 df).
$P > 0{\cdot}50$.

The nature of the uncertainty may be clarified if the role of impairment in the three theories is recapitulated. From the rational choice and the deviance perspectives, impairment acts only so as to impose a general 'tax' on time and energy. From the mechanistic perspective, the site of the impairment is a specific agent creating specific patterns of disability. Table 6 seems to reflect a situation somewhere between these possibilities. The Table is consistent with the 'tax' explanation if it is held that the highest rates of 'tax' are nearly always imposed by locomotor impairment. This is the prediction from disability pattern to impairment which, as has been said, best suits the figures. The Table is, however, not easily made consistent with any adequately specific version of the mechanistic explanation. It predicts from impairment site to disability level with relatively poor probability, and it distinguishes only shallow levels of disability.

It may be that the fault lies in the classifications of impairment which are so far available. The best hope of improvement would seem to be some more subtle classification of locomotor impairment, but the obvious subdivision according to the particular limbs in which symptoms occur does not appear to be promising. Meanwhile, the deviance theory remains a fair provisional account of the structure of disability.

But can the deviance theory be confirmed by further evidence which shows not merely a structure of disability consistent with it, but also variations in the structure which are attributable to group definitions of progress in rehabilitation? Two additional tests were suggested earlier: the first, whether differences could be found in the disability patterns associated with different roles or cultural groups; and the second, whether rehabilitated or static cases show a more precise scale pattern than recently deteriorated cases.

That different scales should be needed for the different sexes is an argument for the social construction of disability patterns which it is easy to overlook. The difference might, however, be merely between the activities usually carried on; and this is no more than is already conceded by what I called 'refined mechanism'. A difference of order in similar items between the sexes would be more interesting. Such a difference was found in the earlier paper. In addition, differences were reported between special subgroups of women and the majority of women. One special subgroup was unrelated to site of impairment on either of the impairment classifications. It seemed rather to be related to the insecure economic position of women in the old working class culture of central London. These women, for whom the extended family is a major resource, were willing to accept help with their cooking rather than limit their mobility, by contrast with the priorities of most of the women. In this they accorded with the priorities of those Lambeth men who kept house for themselves.

A second subgroup of women preferred to accept help with cleaning rather than, as was more usual, with shopping. The interpretation here was more ambiguous, for the women were distinguishable both by symptom and by household situation, and symptom site and household situation were themselves interrelated. This example cannot count as evidence for either the deviance theory or the mechanistic theory, and the interpretation would have to be resolved on the merits of the two theories in the other contexts. For further details, see Williams *et al.*, (1976).

Finally, differences in the scale organisation of changing or long static cases became accessible in the York study N1. Changes after three months were recorded in women at home, and the results are presented in Table 8. Although numbers are small,

Table 8 *Disability changes after three months*

At three months	*Nonscale cases*	*%*	*All cases*
Static	1	2·5	40
Improved	1	4·2	24
Deteriorated	9	37·5	24

there is indeed a tendency for recently deteriorated cases to show more nonscale patterns, and the scaling of recently improved or static cases is remarkably precise.

The tenor of these additional tests, then, is in favour of the deviance theory.

AREAS UNEXPLAINED BY THE DEVIANCE THEORY

It remains possible that the deviance theory holds one key to the general structure of disability, while the mechanistic theory contributes towards explaining what the deviance theory cannot explain. The unexplained area is at its maximum defined by the nonscale cases. Can the evidence relate nonscale cases to the mechanistic perspectives?

The two aspects of impairment are tabulated against nonscale cases in Tables 9 and 10. Symptom sites are again unrelated to scalability. Impairment sites again show some relationship, and it is one which is intelligible

Table 9 *Nonscale cases by impairment site*

Impairment site	*Nonscale cases*	*Scale cases*	*All cases*
A. Locomotor	28	53	81
B. Sensory	2	8	10
C. Internal	6	38	44
D. Other, including mental	7	15	22
	43	114	157

χ^2 (A versus B + C versus D) = 6·609 (2 df).
$P < 0·05$.

Table 10 *Nonscale cases by symptom site*

Symptom site	*Nonscale cases*	*Scale cases*	*All cases*
A. Lower limbs	16	45	61
B. Mixed limbs	2	6	8
C. Upper limbs	4	17	21
D. Central	21	46	67
	43	114	157

χ^2 (A + B + C versus D) = 0·605 (1 df).
$P > 0·40$.
χ^2 (A + B versus C versus D) = 1·320 (2 df).
$P > 0·50$.

from the previous discussion. Locomotor problems are a little more resistant to scaling than internal problems. There are two possible explanations for this, which are not exclusive. Such problems may make conformity to a scale physically difficult, or they may constitute striking reasons for justifying non-conformity. The second possibility needs consideration as well as the first because of a further interesting and unexplained feature of the nonscale cases.

There is a far wider potential range of nonscale variation than in fact occurs. The majority of nonscale cases among women had a pattern where one unexpected ability was maintained amid the normal scale expected for someone with that number of items disabled. Represented graphically, an example on a five-item scale would be 01011, or 01101 or 01110 where 00111 was expected when three items were disabled. Nonscale cases of this kind will be referred to as 'one-error patterns'.

This restriction on the occurrence of nonscale patterns is not easily explicable as the result of the physical impossibility, through locomotor impairment, of performing specified activities. But if it is supposed that the restriction derives from a partial and as yet unsuccessful effort of those caring for the disabled person to bring his pattern of disabilities into conformity with his place on a socially constructed scale, then the remaining deviation he shows may be explicable by the acceptance of locomotor impairment as a licence for nonconformity.

At the same time, because the one-error pattern is likely to be seen as near to conformity, it would minimise misunderstanding and friction with those surrounding the disabled person. In this way the mechanistic theory may have a part to play in the context of the deviance theory.

IMPLICATIONS FOR THE ASSESSMENT OF HANDICAP

The earlier paper published on the Lambeth data showed that assessment of handicap was simplified by

the Guttman scale structure found in disability. The person with seven disabilities was always worse off than the person with six, because his seven disabilities included the other's six. Thus, arbitrary judgements were virtually eliminated from the main task of assessment, but they continued to be necessary in three areas: in assessing the representativeness of the items of disability chosen for study against the universe of relevant items; in assigning nonscaling cases; and (where this is necessary) in assessing the overall extent of change in an experimental and a control group where individuals shift from different initial levels of handicap to new levels. From the theoretical discussion and further evidence set out in this paper, it is now possible to suggest partial criteria for the relevance of disability items, for assigning nonscale cases, and for generating a ratio scale of disability, on the assumption, which remains provisional, that the deviance theory offers the best account of the structure of disability.

The universe of relevant disabilities has been defined as those which threaten the conventional identity of an adult member of society. This identity includes a range of activities which it is taken for granted that an adult voluntarily and consciously carries out, and the neglect of which creates embarrassment in others and a presumption that an account or justification is called for. Proper management of the body and proper participation in a private household have been cited as obvious examples which account for all the items used as data. It is easy, too, to exclude as irrelevant personal hobbies, or those services which are normally performed for each person professionally. But there is a wide margin of activities about which it would be possible to disagree, and considerable room for further clarification. The definition emerging from the present argument constitutes a start.

Some progress can also be made in assigning nonscale cases. Just as the deviance theory seems to account for the structure of scaling cases, so it seems to account in an important respect for the structure of nonscaling cases. One-error cases appear to represent attempted approximations to the conventional scale. On the face of it, this does not help much, for there are logically two points of the scale, one above and one below, to which a one-error case might be an approximation. A case taking the pattern 101, for example, might be an approximation to 111 or to 011. However, according to the argument previously set out, an error case approximates to a scale position only in the sense that the disabled person concerned is being encouraged by his friends and relatives to conform to a logic of recovery. It would contradict this logic to encourage him to get worse, that is, to move from 101 to 111. The conclusion must be that a person showing a disability pattern 101 was being encouraged to achieve the pattern 011, but for reasons inaccessible to us the pattern 101 was accepted as a substitute. Whatever the explanation of one-error variations from the scale, therefore, it seems reasonable to assess their intended value as being equal to the scale pattern which contains the same number of disabilities. Other types of error remain uninterpretable, and their assessment requires judgement. But these more variable error cases are a very small proportion of the disabled, usually amounting to less than 10% in all.

The criterion of a ratio scale is necessarily offered tentatively. Numerous approaches to deriving such a scale have been attempted, mostly by forms of rating and expressed preference. The utilitarian philosophy which underlies these attempts implies a choice of values, and consequently of representative disability items, which may not coincide with the perspective and the items relevant in an explanatory theory like the one sketched here. Besides, pain and the risk of death are considerations in a utilitarian classification of health states which are not taken into account in the present disability scale. However, if the premises advanced earlier are true, there are interesting implications for what a ratio scale of disability ought to look like. Although restricted to one perspective, these cannot be overlooked in more ambitious attempts to evaluate states of health.

The implication is that, in items which people order cumulatively so that the severest accumulations of disability are also the rarest, they are assessing acceptability of the disability as a function of its frequency. For the disabled person and his family, the frequency of his condition is experienced in terms of the fraction of their acquaintance who are similarly afflicted, including in the 'acquaintance' those of whom they are aware without necessarily having met them. There will be many local variations in this experience. But it can be approximated on average by the fraction of the general population with the condition. It is important to note that this is true only of scalable conditions.

Strictly speaking, on this argument, the frequency of a scalable condition represents not a value but a description: the degree of conformity of the disabled person to the norm. Conformity as a descriptive term consists in being like everyone else in some presupposed milieu, non-conformity in being the odd man out. If the general population is taken as an approximation to this milieu, then the degree of conformity of a condition is expressible as the proportion of the general population who experience that condition, and it varies from 0 (theoretical total non-conformity) to 1 (the possession of characteristics common to everyone). The deviance theory, however, supposes in addition that conformity is being treated as a value in the behaviour of the disabled. The priorities attaching to certain activities are negotiated by the disabled, their families, and neighbours and others, in a common subculture; and the value attaching to conformity ensures that this pattern of priorities is then adhered to. The next question is what relationship the descriptive scale of conformity bears to the evaluative scale. In the same way one might ask, for example, what relationship an evaluative scale of obesity bore to the scale of body weight in avoirdupois.

There is some plausibility in supposing flat-footedly that the relationship of acceptability to conformity is linear. A scale is produced by these means which gives weight to extreme disabilities in much the same way as judges have been shown to do in court decisions (Rosser and Watts, 1972). For example, a scalable condition experienced by 1% of the population is half

as acceptable as one experienced by 2%, and the ratio of acceptability between them is the same as that between scalable conditions experienced by half of the population on the one hand, and all the population on the other. But it is also probable that there is a limit on the diminution of acceptability as conformity drops towards zero, which is set by the size of the population about whom a disabled person normally gets personal information. Acceptability may reach zero, in fact, before conformity. Similarly, to be the same as most of the people one knows may be sufficiently 'normal', and acceptability may reach its maximum well before conformity approaches unity. These possibilities are represented by points A and B in Fig. 2. However, there are also reasons for being wary of supposing that the line AB is straight. To be unique among the people one knows may well be especially unacceptable compared with being one of two or three. As conformity diminishes, the diminution of acceptability may well be subjectively magnified (the dotted curve AB). On the other hand, perhaps it is mainly the first steps in deviance that count, in which case the continuous curve AB would be appropriate. Both these very simple curves could be defended.

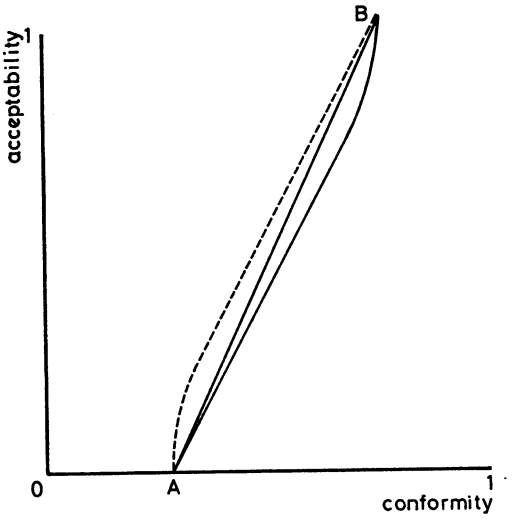

Fig. 2 *Relation of acceptability to conformity.*

In the light of these examples, it is safer to propose only that the evaluative scale of acceptability which apparently underlies the behaviour of the disabled is a function of the descriptive scale of conformity and has a true zero. Studies of how the descriptive scale is evaluated will be needed. But if the argument of this paper is valid, any proposed index of health states which includes valuations of disability would need to be considered in relation to the descriptive scale of conformity defined here, and the function of that scale which it represents would need, if it were other than the simple function described here, to be defended with additional reasons.

What emerges, therefore, is a possible criterion of validity for ratio scales of disability. It seems from the foregoing arguments that the true ratio scale of the acceptability of scalable disabilities must be a function of their perceived rarity. Comparison with other attempts to elicit a ratio scale may show that, with suitable precautions for the extreme values, acceptability may be treated as a linear function of conformity. But great caution is still necessary, as with all attempts to impose numbers on social life. Moreover, it is perhaps appropriate to stress here the other corollary of the arguments presented, that acceptability of disabilities also varies significantly with the perceived social network of the disabled person, with the role he fills, and with the negotiation of that role and its various priorities with the group to which he belongs. These considerations make it inappropriate, for many purposes, that a single scale should represent valuations of disability for a whole population. Rather, a family of scales is appropriate. Ultimately, therefore, it may be possible to generate a number of interrelated scales with varying degrees of refinement for different purposes. One scale could be produced which was valid within known limits for the general population. Those limits would be stated by the proportion of nonscale cases other than one-error cases, since the deviance hypothesis can account for one-error cases, and by the groups within the population which show distinct scaling patterns, such as the Lambeth women who are economically insecure. From this overall scale, it would be possible to branch out into a number of scales valid for these distinct groups, or for other groups where additional items were relevant such as occupation or schooling. With additional items, or with more precise knowledge of the size and morbidity of the reference group to which individuals conform, further distinct subgroups might emerge, and eventually a hierarchy of generalised valuations of disability could be established. At the most particular level, finally, would be the clinician attempting to predict from an individual's current disability pattern the next critical disability which that individual is likely to lose or regain. Depending on the generality of the statement required, then, different scales would be to the purpose, but they would have a known relationship to each other.

These possibilities are still some way from realisation. The purpose of this paper has been to give an airing to the theoretical ideas that indicate this line of development. The need is for better explanations of disabled behaviour before attempting more sophisticated evaluation of it.

I thank Professor A. E. Bennett and Mr. K. G. Wright for the use of the two sets of data referred to; Mr. J. Cairns, whose tabulations of the York data I have drawn on; and Mr. B. Sauvé for advice on computing.

Comment: A. J. CULYER
Reader in Economics
University of York

I believe Williams is the first to attempt a general theory of disability explaining, among other things, why it is that physical impairments may lead to the various changes in social behaviour that are here termed 'disability'. I propose to focus on Williams's third

sentence. I shall not refer to the details of measurement, Guttman scales, etc., but I shall attempt to show that, suitably extended, his second model (the so-called rational choice model) can be generalised to embrace the others as special cases and to provide a framework for discussing with greater clarity many of the famous questions concerning disability, handicap, and impairment, the relationships between them, and the implications, if any, for social policy and community medicine.

This more general model I term the 'choice' model in order both to emphasise the discretionary nature of the link between impairment and disability and, by dropping 'rational', to show that the basic ideas in no way depend upon special theories about what rational choice actually is.

We begin by assuming that the human activities of the sort that figure in Guttman scales are those of ultimate interest in disability, whether from a broad policy point of view or at the level of clinical or social work. With Williams, let us continue to assume that, in principle, these activities can (a) be measured as a daily, weekly, etc., rate on a ratio scale, and (b) be associated with particular intensities with which various limbs and other physical bodily functions are used. For example, running is, among other things, lung and leg intensive compared with reading, which is eye intensive. Given an individual's bodily functioning, the time he chooses for each activity (this may not be independent of the former), his talents, the value he places on various activities, the social pressures he faces from relatives, professionals, etc., in any period of time the individual will actually engage in a measurable amount of each activity.

In Fig. 3 I suppose that, in view of these various factors, the individual could either perform Oa of activity A or Ob of activity B. Assuming that we are dealing with two activities only, for convenience of exposition, then the feasible set of combinations of A and B which he might perform is determined by the area Oab. I have assumed ab to be a straight line although it might be convex from above—a relatively unimportant difference. It is unlikely to be concave, although if it were, some of the conclusions of 'indifference curve' theory may not follow. Let us suppose that individuals prefer activity to non-activity. In that case, an individual will be located at some point on the boundary ab of the feasible set: precisely where is of little interest at this level of generality.

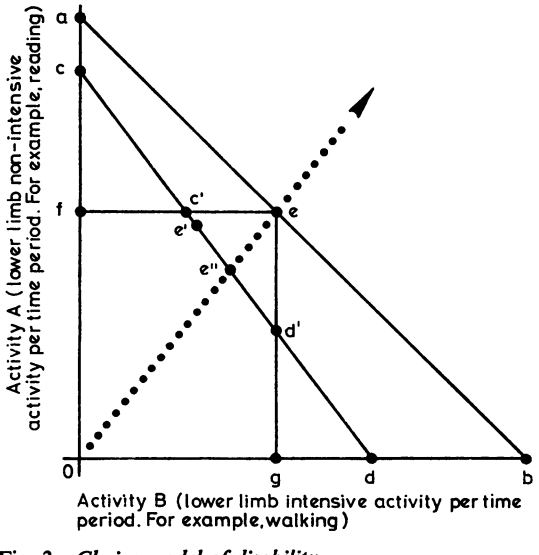

Fig. 3 *Choice model of disability.*

Let us now introduce an impairment of physical functioning, or an additional impairment to one that already existed when the individual was on ab. It may be the loss of a limb, or a lung. The maximum amount of B that can possibly be performed is now Od, and the maximum possible amount of A is Oc. The new feasible boundary is cd upon which the individual must now settle. Two features of cd are important in comparing it with ab: (1) cd lies entirely within ab, hence less of at least one activity must be chosen and (2) cd has a steeper slope than ab on any ray from the origin. Hence there will be a tendency to substitute activities in which the individual is less impaired for those where he is more impaired.

This prediction does not depend upon rationality. It is true that, if we assumed rationality of the sort postulated by Williams (essentially concave indifference curves), and added the assumption that each activity is 'normal' in the technical economic sense that neither activity is 'inferior', then less of activity B would be undertaken. But let us assume instead that individuals act irrationally—for example, that they act at random. Such irrational individuals would be distributed evenly along ab. On the average, they would be located at e, halfway along ab. Similarly, impaired persons confronted with cd would likewise be located on the average at e′, halfway along cd. Since e′ lies to the left of e, highly irrational people will behave (on the average) just the same as rational people.

Another form of irrationality might be that people act 'conservatively', attempting to maintain the same proportionate levels of different activities. This implies locating at the intersection of the ray Oe and cd, at e″. Which of these various models best describes actual events is a matter of empirical test. My own hunch would be that the rational choice hypothesis, as developed here, best applies to the young disabled, among whom one would expect to see a high variance of disability given basic impairments—depending on tastes, determination, parental wealth, social expectations, etc. The 'conservative irrational' model might be a useful starting point in modelling the disability of elderly persons.

Rational 'tastes' are, however, well worth bringing in because they greatly enrich the insights of this model of disability. Suppose, for example, that social pressures, or personal attitudes, shift 'tastes' systematically towards less, or more, of the activity that is intensive in the impaired function. The interesting possibility arises that the individual might move in the region cc′ on the new frontier, involving more reading but less walking than before; or, in the other case, in

the region d'd, involving more walking but less reading than before. The model includes the logical possibility that a person may engage in *more activities in which he is impaired than he formerly did.* This possibility is flatly denied by the mechanistic theory. So much the worse for that theory, since the world actually does contain at least one professional but one-legged footballer. The possibility clearly depends, however, on Od > Og or Oc > Of.

This more general theory includes the mechanistic and deviance theories in the following ways. The mechanistic theory forms the basis for describing the change in the feasible set, modified so that the possibilities are no longer 'all or none' but 'more or less', depending on severity. The deviance theory comes in as one, but only one, of the influences which determine where on the feasible frontier individuals will be located.

Another natural extension of this model is to use it as the basis for framing social policies designed to help the disabled. The following are among the questions that need more thought. Is policy more properly concerned with shifts in the feasible set or with points chosen within a feasible set? In layman's language, should self-help be penalised by less public help, or lack of self-help be rewarded by giving public help? Is it possible to devise relevant weights to combine levels of functioning and levels of ability to form overall indices of disability and impairment? The model makes it clear that these matters are not merely technical, or for professional ajudication only. Such weights are in a highly relevant sense 'political'.

Reprints from Dr. R. G. A. Williams, Institute of Medical Sociology, Westburn Road, Aberdeen AB9 2ZE.

References

Bennett, A. E., Garrad, J., and Halil, T. (1970). Chronic disease and disability in the community; a prevalence study. *British Medical Journal*, **3**, 762–764.

Freidson, E. (1972). Disability as social deviance. In *Medical Men and their Work*. Edited by E. Freidson and J. Lorber. Aldine: Chicago.

Garrad, J., and Bennett, A. E. (1971). A validated interview schedule for use in population surveys of chronic disease and disability. *British Journal of Preventive and Social Medicine*, **25**, 97–104.

Goffman, E. (1963). *Stigma: notes on the management of spoiled identity*. Prentice Hall: Englewood Cliffs.

McConaghy, M. J. (1975). Maximum possible error in Guttman scales. *Public Opinion Quarterly*, **39**, 343–357.

Nichols, P. J. R. (1975). Some psychological aspects of rehabilitation and their implications in research. *Proceedings of the Royal Society of Medicine*, **68**, 537.

Rosser, R. M., and Watts, V. C. (1972). The measurement of hospital output. *International Journal of Epidemiology* **1**, 361–368.

Scott, M. B., and Lyman, S. W. (1968). Accounts. *American Sociological Review*, **33**, 46–62.

Williams, R. G. A., Johnston, M., Willis, L. A., and Bennett, A. E. (1976). Disability: a model and measurement technique. *British Journal of Preventive and Social Medicine*, **30**, 71–78.

Wright, K. G. (1974). Alternative measures of the output of social programmes: the elderly. In *Economic Policies and Social Goals*. Edited by A. J. Culyer. Martin Robertson: London.