

WORLD HEALTH ORGANIZATION
REGIONAL OFFICE FOR EUROPE



ORGANISATION MONDIALE DE LA SANTÉ
BUREAU RÉGIONAL DE L'EUROPE

WELTGESUNDHEITSORGANISATION
REGIONALBÜRO FÜR EUROPA

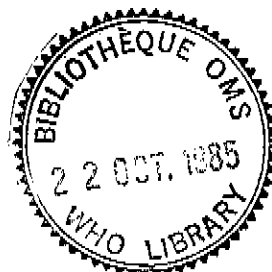
ВСЕМИРНАЯ ОРГАНИЗАЦИЯ ЗДРАВООХРАНЕНИЯ
ЕВРОПЕЙСКОЕ РЕГИОНАЛЬНОЕ БЮРО

Consultation on Field Studies for Risk Factor Assessment

INDEXED

ICP/HST 102 s02/9
8 October 1985
ENGLISH ONLY
(Unedited)

Bristol, United Kingdom, 29 October - 1 November 1985



EXPERIENCE GAINED FROM A LONG-TERM COHORT STUDY OF THE MEASUREMENT
OF MORBIDITY, IMPAIRMENT, DISABILITY AND HANDICAP

by

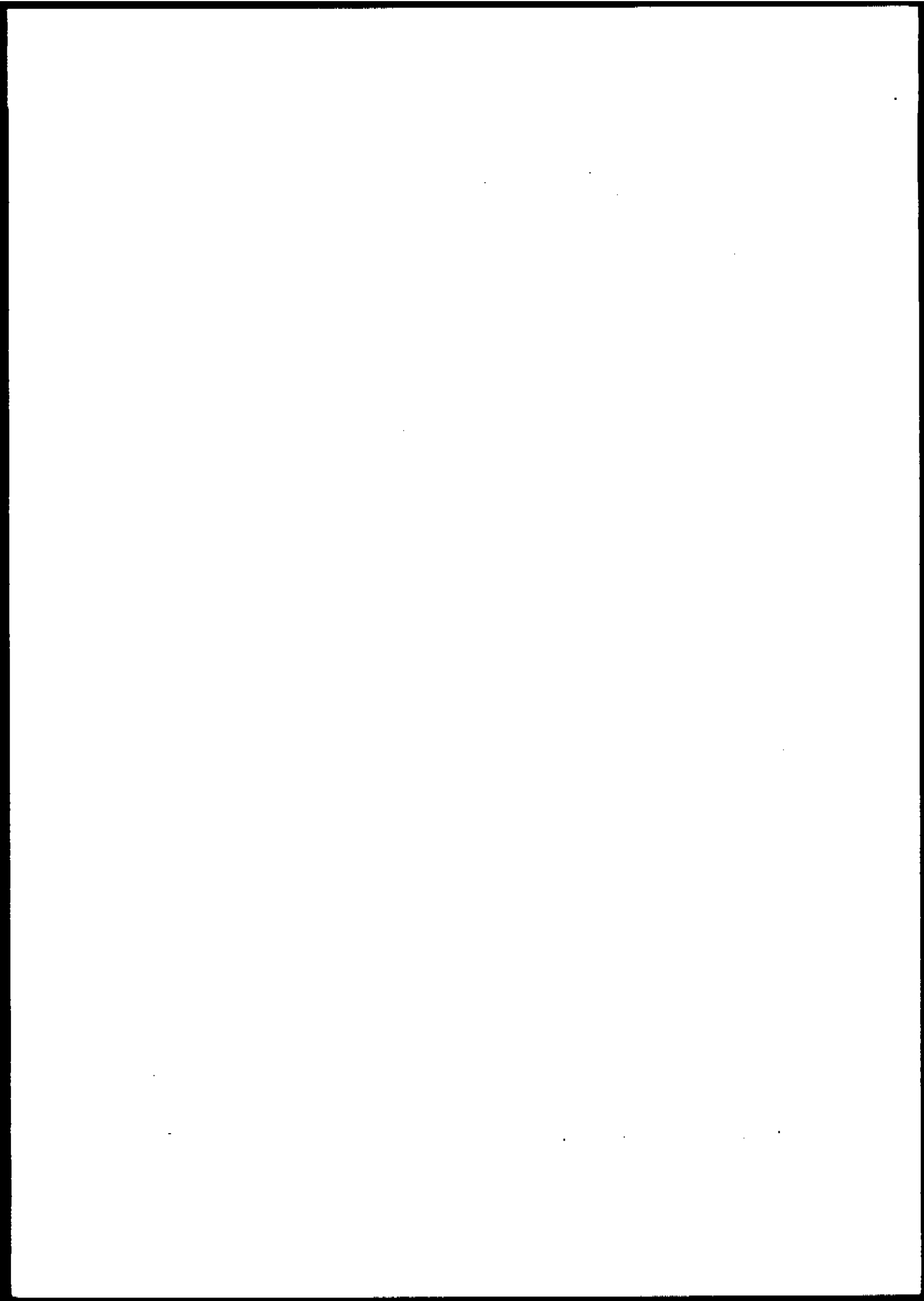
Mary N. Haslum
Department of Child Health Research Unit
Bristol

The issue of this document does not constitute formal publication. It should not be reviewed, abstracted or quoted without the agreement of the World Health Organization Regional Office for Europe. Authors alone are responsible for views expressed in signed articles.

Dieses Dokument erscheint nicht als formelle Veröffentlichung. Es darf nur mit Genehmigung des Regionalbüros für Europa der Weltgesundheitsorganisation besprochen, in Kurzfassung gebracht oder zitiert werden. Beiträge, die mit Namensunterschrift erscheinen, geben ausschliesslich die Meinung des Autors wieder.

Ce document ne constitue pas une publication. Il ne doit faire l'objet d'aucun compte rendu ou résumé ni d'aucune citation sans l'autorisation du Bureau régional de l'Europe de l'Organisation Mondiale de la Santé. Les opinions exprimées dans les articles signés n'engagent que leurs auteurs.

Настоящий документ не является официальной публикацией. Не разрешается рецензировать, аннотировать или цитировать этот документ без согласия Европейского регионального бюро Всемирной организации здравоохранения. Всю ответственность за взгляды, выраженные в подписанных авторами статьях, несут сами авторы.



WORKING PAPER FOR CONSULTATION ON FIELD STUDIES FOR RISK FACTOR ASSESSMENT

Experience gained from a long term cohort study of the measurement of morbidity, impairment, disability and handicap.

Dr. M N Haslum, Principal Research Officer 1970 Birth Cohort Study
Research Fellow, Department of Child Health, University of
Bristol.

Introduction

The 1970 Birth Cohort Study is a national and longitudinal study of the health and development of 15,000 children living in England, Scotland and Wales who were born between 5th and 11th April, 1970. Large cohort studies such as this offer considerable opportunities for both cross-sectional and longitudinal multivariate research into children's lives and their physical and cognitive development. Unfortunately, however, because such studies appear to be able to ask so many different questions their limitations can sometimes be forgotten. This paper, therefore, sets out to consider some of the difficulties of long term cohort studies in the measurement of morbidity, impairment, disability and handicap in children.

Penalties of asking a question more than once

Cohort studies often employ more than one questionnaire. A child's mother may be interviewed by a health service worker to obtain a social, family and medical history of both the mother and her child. On another occasion a clinical medical officer examines the mother and/or the child. It is unthinkable that a doctor would not take a medical history during a medical examination. Later, when the two medical histories are compared over the cohort for evidence of a particular morbidity there is never 100 percent agreement.

Three interesting examples of the consequences of asking the same question twice in a cohort survey occurred recently in the analysis of the 1970 Birth Cohort data from the 1980 follow-up study when the children were 10 years of age. In the first, the children's mothers and their teachers were asked if the child stammered or stuttered. 2.4 percent of the children were reported as having a stammer/stutter by their mothers but only 1.7 percent were reported as having a stammer/stutter by their teachers. When the children identified by their mothers as having a stammer/stutter, however, were compared with those identified by the teachers there was agreement in only 76 of 393 cases. What then is the

prevalence of stammer/stutter in 10 year olds? There are a number of answers: 0.7 percent (mothers and teachers agree); 1.7 percent (teachers' view); 2.4 percent (mothers' view), 3.5 percent (children identified by mothers or teachers). What then is the prevalence of speech problems in 10 year olds, 2.7 percent or 5.4 percent?

The second example involves the reported prescription of spectacles to the children:

- 980 (4.6%) took their spectacles to the survey medical examination;
- 388 (2.8%) did not take their spectacles to the examination but both parents and doctors agreed that the child had spectacles;
- 135 (1.0%) did not take their spectacles, the parents said the child had spectacles but the doctor said the child did not have spectacles;
- 107 (0.8%) did not take their spectacles, the parents said the child did not have spectacles but the doctor said the child did have spectacles;

Altogether 1610 (11.7%) children may have been prescribed spectacles.

This problem was dealt with at the analysis stage by taking the first two groups as definite spectacle prescriptions and the last two as possible prescriptions. Analyses were carried out on the first group and then repeated with the inclusion of the latter (Stewart-Brown, 1985).

In the third example, 12,908 children completed motor coordination tests administered in a standard manner with standard equipment by clinical medical officers who were then asked to rate the children's limb coordination as normal, questionably clumsy, mildly clumsy, severely clumsy. In other parts of the 10 year follow-up the children's mothers and their teachers separately rated a number of the children's motor skills. These ratings were subjected to principal components analyses to yield a clumsiness rating from the mothers and one from the teachers. The clinical medical officers, mothers' and teachers' ratings were compared with a count of the number of motor tests failed by the child. Failure was defined as a score in the bottom tenth centile following the practice of Gubbay (1975).

The clinical medical officers', mothers' and the teachers' assessments were significantly related to the number of tests failed but the level of the correlations was very low: Spearman's rho = .32, .06, .02, respectively (Haslum, 1985). A poor correlation between clinical medical officers' assessments of 'limb dyspraxia' and failure on survey tests was also reported in the Isle of White survey (Rutter et al, 1970) and from a study of 850 children in Cheshire (McInlay, 1985).

Whilst the correlation with failure on the survey tests implied that both mothers and teachers were identifying some children who failed the tests

as 'clumsy', the correlation between the mothers' and teachers' ratings ($\rho = 0.01$) was statistically non-significant. This indicates, therefore, that they were not identifying the same children as clumsy. This particular problem probably results as much from the fact that clumsiness, at present, is a rather vague concept in people's minds as the fact that the same question or enquiry was made of a number of different people.

Asking a health-related question of more than one person in a survey can produce highly discrepant answers. One way to avoid this is to ask the question only once deciding in advance what type of respondent is likely to produce the most reliable information.

The problem in the first example of stammer/stutter is that not only may mothers and teachers have different ideas of what constitutes stammering or stuttering but the children may actually behave differently at home than they do at school.

Discrepancies of reporting can be coped with at the analysis stage, as the second example shows, by running an analysis first on a group defined by agreement between reporters and then repeating the analysis with the inclusion of the discrepant reports.

In the third example the solution would seem to be to use failure of standardised motor tests as the definition of clumsiness. There are now at least twenty well known tests of motor proficiency in children and although some are far too long and require too much equipment for survey use, some are suitable and actually contain norms, notably, the Comanche Screening Test of Motor Organisation (Bradley 1982) which contains British norms for the Gubbay tests (1975, 1978).

Validation of medical diagnoses

In each of the national sweeps of the 1970 Birth Cohort in 1975 and 1980 information was collected on children's respiratory problems, convulsive disorders and hospital admissions and parents were asked for written permission for access to their children's medical records. Respiratory problems and convulsive disorders are notoriously difficult to classify and without recourse to the children's General Practitioners classification would be very difficult. This is usually done by requesting a child's General Practitioner to complete a very specific questionnaire about a particular morbidity such as respiratory problems or convulsive disorders from his medical records and his knowledge of his patient. It appears to work well.

Validation from hospital records and consultants' notes provides an interesting problem: "When is a piece of information validated?" If a mother says her son was admitted to hospital at age 6 for three nights

with a bad chest and the hospital notes say he was admitted at age 5 years 10 months for 4 nights with bronchopneumonia, is the mother's information validated? Should her information be changed to correspond with the hospital notes?

Whilst copies of hospital notes which cover admissions are relatively easy to obtain, information from Casualty and Outpatient Departments needed for the validation of children's accidental injuries have proved far more elusive and so far in the 1970 Birth cohort study validation of such injuries has been confined to accidents which resulted in hospital admissions.

Standardisation and Validity of Survey Instruments

One of the great advantages of cohort surveys is that they are large enough and sufficiently representative to provide their own standardisation sample for the instruments used. An example of this was the use of the Edinburgh Reading Test (Godfrey Thompson Unit, 1978) in the 1980 follow-up of the 1970 Birth Cohort. This test was originally standardised on a sample of 4,000 children with an age range from 6 to 17 years by the Godfrey Thompson Unit in Edinburgh (the test authors). Use of the test in 1980 follow-up provided a standardisation sample of 12,906 10 to 11 year olds. Analyses of reading attainment is carried out using the survey norms without recourse to original test norms. The validity of the test in relation to other tests of reading, however, is dependant on the intercorrelations established during the development of the original test.

The description of behaviour problems in a large survey poses particular problems of validity. Many surveys have used behaviour ratings scales. Such scales have usually been developed for a specific age range of children and validated against psychiatric opinion. All too frequently, however, they may be used outside the age range for which they were developed as nothing else seems suitable or available. Behaviour scales raise particular problems for cohort studies. If, for example, the Rutter Child Behaviour Disorder Scale (Rutter, et al 1970) is used in a survey of children aged 5 and, then 5 years later, when they are 10 years old, do the items mean the same at both ages? Does an item such as "Cries for little cause" imply the same behaviour difficulty at the two ages? What is the test-retest reliability over time?

One of the temptations which cohort studies put in the way of the investigator of behaviour problems in children is the idea that normative data may give a more useful picture and perspective of children's behaviour problems than clinical experience. Indeed Vikan (1985) has argued that behavior ratings are important in their own right since they signal a suffering in the child and/or his parents that will always be of some consequence to the child's development. Boyle and Jones (1985)

suggest that clinical judgement itself varies too much to provide a 'gold standard' against which to judge criterion validity of other behaviour ratings.

Operationalising definitions

"In the context of health experience, a disability is any restriction or lack (resulting from an impairment) of ability to perform an activity in the range considered normal for a human being".

International Classification of Impairments, Disabilities and Handicaps. WHO Geneva 1980 page 28.

Definitions such as that given above, the definition of disability from the WHO Experimental International Classification of Impairments, Disabilities and Handicaps (1980) are a heaven-sent gift to cohort studies. The definition of disability rests on the concept of normality of performance and the cohort study provides the opportunity to define such normality statistically.

In the 1980 follow-up of the 1970 Birth Cohort, when the study children were 10 years old, mothers rated the children's skills on fifty items which covered mobility, self-care, independence and communication. The ratings were converted to 'z scores' with means of 0 and standard deviation 1. Group mean scores were compared statistically for children with particular impairments on each item, for example, cerebral palsy or severe mental retardation, and standard errors were calculated to examine the variability of the means.

In addition, for any item, a score less than -3.1 standard deviations from the mean and therefore associated with a probability of occurrence of 1 in 1,000 was designated as indicative of severe disability. A score between -2.3 and -3.0 standard deviations, associated with a probability of less than 1 in 100 but greater than 1 in 1000, was designated as indicative of moderate disability, and a score associated with a probability of less than 1 in 10 but greater than 1 in 100 was designated as indicative of mild disability. Thus it was possible to attach labels to group mean scores for individual items and to suggest that the group of children with cerebral palsy in the cohort were moderately or severely disabled in certain skills and only mildly disabled in others. Children with diabetes mellitus were no different from children with no impairments on any of the skills items except one: "Using time to regulate their daily lives" and, in this, they excelled. In this way it was possible to examine the first part of the WHO model of disablement and to test the extent to which specific impairments were associated with particular disabilities. (Haslum et al 1983).

An attempt to examine the model further to include handicap is in progress

and the 1980 follow-up collected mothers', teachers' and examining medical officers' accounts of the effect of children's impairments and disabilities on their daily lives at home and at school. There are considerable difficulties, however, in trying to apply the WHO definition of Handicap to a cohort study where 'cultural norms' (ICIDH, 1980) have to be assumed and the 'value attached to his or her disability by a child or his peers' (ICIDH opcit) is not known.

ICIDH classification applied to surveys

The 1980 follow-up of the 1970 Birth Cohort was designed before the publication of the experimental ICIDH classification in 1980. It has proved too difficult to apply this classification in retrospect. In the present author's opinion, the classification codes of impairment, disability and handicap can only be designated by an examiner with the subject present and the ICIDH manual to hand. The complexity of the classification suggests that it may not be amenable to survey work. Its heavy dependence on observers' assessments of disabilities and handicaps suggests the need for careful training of the observers to be used in the survey and rigorous assessment of inter- and intra-observer reliability. When the survey is cohort-size and extends to thousands of children this becomes impracticable. In addition, whilst it may be possible to apply these classification codes to adult disablement their application to childhood disablement is fraught with difficulty.

Conclusion

It is obviously of paramount importance that the measurement of morbidity, impairment, disability and handicap should be comparable between the countries in the proposed long term assessment of risk factor studies. This will involve not only agreement on definitions of specific morbidities, such as asthma, febrile convulsions, and so on, but also the same extent of validation against medical records.

It should be remembered that asking the same question of a number of different people does not necessarily improve the answer.

The assessment of disability in the cohort studies should focus on specific skills which are required by children in their daily lives in all the countries. This could be enhanced by the standard presentation of tests of gross and fine motor skill.

The assessment of handicap should be kept as simple as possible: "Do you consider the child is at a disadvantage, compared with his peers, as a result of his impairment or disability?" and, better still, at an appropriate age, ask the child himself. It should be remembered that an individual's perception of his handicap, other people's perception of his handicap and the interaction between the two perceptions are likely to be

subject to many external and changing influences. It may be that in cross-cultural epidemiological cohort studies it is only sensible to attempt the identification and prediction of impairments and associated disabilities and to leave the problems of handicap for a smaller scale and different type of study.

References

Boyle M H, Jones S C (1985) point out clinical judgement itself varies too much to provide a 'gold standard' against which to judge criterion validity of other behaviour ratings.

Boyle M H, Jones S C Selecting Measures of Emotional and Behavioural Disorders of Childhood for use in General Populations. Journal of Child Psychology and Psychiatry 1985, 26, 1, 137-159.

Bradley G. The Comanche Screening Test of Motor Organisation with suggested activities for Clumsy Children (British Norms for the Gubbay Test) City of Manchester College of Higher Education 1982.

Godfrey Thompson Unit. Edinburgh Reading Test. Hodder and Stoughton, Sevenoaks, Kent, 1978.

Gubbay S S. The Clumsy Child. Sanders London 1975.

Gubbay S S. The Management of Developmental Aproxia, Developmental Medicine and Child Neurology 1978, 20, 643-646.

Haslum M N. Clumsiness in Children In: MacFarlane A. Progress in Child Health Volume 3 (In press).

Haslum M N, St.Claire L, Stewart-Brown S and Morris A E. A Profile Analysis of the Everyday Skills of Children with Impairments In: Child Health and Education Study. A Collection of Papers from the 10 year follow-up. Second Report to the Department of Health and Social Security 1983. Unpublished Report.

International Classification of Impairments, Disabilities and Handicaps. World Health Organisation, Geneva, 1980.

Mackinlay I, 1985 Personal communication to author.

Rutter M, Tizard J and Whitmore K. Education, Health and Behaviour, London, Longman.

Stewart-Brown S. Spectacle prescribing in 10 year old children. British Journal of Ophthalmology (In Press).

Vikan A 1985 has argued that behaviour ratings are important in their own right since they signal a suffering in the child and or his parents that will always be of some consequence to the child's development.

Vikan A. Psychiatric Epidemiology in a sample of 1510 Ten-Year-Old Children - 1. Prevalence Journal of Child Psychology and Psychiatry, 1985, 26, 55-7