Archives of Disease in Childhood, 1986, **61**, 620–624

Correspondence

Evaluation of a developmental screening system for use by child health nurses

Sir,

Dr Eu has carried out an interesting and useful evaluation of the Woodside developmental screening system,[1] but there seems to be an important error in her interpretation of the results.

The figure given in Table 3 for children passing the screening test is based not on the total number of children who were passed as normal in the study population (397) but on the sample of 77 (approximately 20%) who were tested with the Griffiths scales. If we assume that this sample was a fair reflection of the total population of 'normal' children, then of the 397 assessed as normal on the screening test, the Griffiths would be expected to identify 26 as abnormal and 371 as normal. When these figures are entered in Table 3 the sensitivity looks less impressive. In this study the screening system in fact identifies only half of the children who would be defined as having developmental problems by the Griffiths scales.

Table *Table 3 (Eu[1]) with modification ('doubtful' cases omitted as in original)*

Woodside	*Griffiths*		*Total*
	Abnormal	*Normal*	
Abnormal	22	5	27
Normal	26	371	397
Total	48	376	424

Sensitivity 22/48=0·46. Specificity 371/376=0·99.

The problem of low sensitivity might well be intrinsic to the whole concept of developmental screening 'and does not necessarily reflect any particular inadequacy of the Woodside system. It will always be difficult to achieve high sensitivity in screening for disorders that are by their nature difficult to define and have an unpredictable natural history.

DAVID HALL
St George's Hospital Medical School, London SW17 0RE

Dr Eu comments:

I am grateful for the opportunity to respond to Dr Hall's letter on my evaluation of the Woodside system as used by child health nurses.[1] His assumption that 'the Griffiths would be expected to identify 26 as abnormal and 371 as normal' is quite reasonable on the details given in the paper, but incorrect in point of fact. The 77 children who had been assessed as 'normals' on the screening assessment and were subsequently retested on the diagnostic Griffiths test represented a validation sample and not a norming or standardising sample. They were selected, in retrospect, either from child care centres for convenience or because parents were worried about their 'normal' child: a large proportion was from the lowest socioeconomic group where the prevalence of 'doubtfuls' and 'abnormals' was more than twice that for the other socioeconomic groupings (Table 4[1]). The value of 0·81 presented for sensitivity would thus have been appreciably higher if the validation sample of 77 had in fact been 'a fair reflection of the total population of 'normal' children, and probably only marginally lower if the figures had been adjusted proportionally as suggested by Dr Hall.

Interestingly enough, of the five children who were assessed as 'normal' on the Woodside and abnormal on the Griffiths, two were listed to be further followed up because they failed the hearing or vision assessment, a third child was recorded by the child health nurse as being 'uncooperative', and a fourth as being 'difficult to assess'. Such children should in fact have remained uncategorised—that is, neither as 'normal' or 'abnormal', nor even as 'doubtful'—if the nurse was unable to obtain a satisfactory developmental assessment.

Therefore, though I agree that it would be ideal for the *whole* population sample to undergo both the screening assessment and the diagnostic test, I would differ with Dr Hall's view that 'it will always be difficult to achieve high sensitivity' in developmental screening. Furthermore, I do not believe that developmental disorders are 'difficult to define' or that they have an 'unpredictable natural history'. In three useful longitudinal studies Werner[2] showed a good correlation between early assessment and reassessment after 20 years, Griffiths[3] in the validation of her psychological test found that testing 270 children over periods of between three and 62 months showed that the correlation coefficient for obtaining 'repeatable results' was r=0·77, and most recently Ross *et al*[4] have shown that in 94 children with a birth weight of <1501 g there was 'a significant correspondence between classification of children on the Bayley MD1 at 12 months and on IQ at 3 years (χ^2=40·9, p<0·001)'.

The two major forces in child development are biological and environmental. In my view the biological forces tend to be relatively stable and therefore predictable, and any difficulties that might arise in predicting final outcome tend to result from variations in the environmental forces.

References

[1] Eu BSL. Evaluation of a developmental screening system for use by child health nurses. *Arch Dis Child* 1986;**61**:34–41.
[2] Werner EE, Smith RS. *Kauai's children come of age*. Honolulu: University of Hawaii Press, 1977.