Short report

Interpretation of plantar reflexes: biasing effect of other signs and symptoms

J. VAN GIJN AND B. BONKE

From the Department of Neurology, Academic Hospital, 'Dijkzigt', and the Department of Medical Psychology and Psychotherapy, Erasmus University, Rotterdam, the Netherlands

SUMMARY Twenty neurologists were asked to judge a number of plantar responses on film. Each film was preceded by a slide with a fictitious abstract of history and examination (minus the plantar reflex). The main part of the presentation only served to disguise the fact that two films, both showing equivocal toe movements, were presented twice at the same sitting, but with opposing information as to the probability of a Babinski sign. Interpretation of these identical pictures differed significantly ($P < 0.01$), conforming to the information given. Thirty other neurologists who rated the films without previous data showed no such change of opinion.

'Most of us assume that, except in occasional borderline cases, the signs we observe are present and those we do not observe are absent', wrote Fletcher in 1952, and he went on to show a vast disagreement between different observers about the physical signs of emphysema. Observer variation has been studied especially in radiology (Garland, 1959), but was also found in the examination of plantar reflexes (McCance *et al.*, 1968).

Quite apart from these differences, a group of doctors may be influenced as a whole by the idea that a particular sign ought to be there or not, on the strength of previous information. The 'Emperor's clothes syndrome' is even said to be epidemic in neurology and cardiology departments (Gross, 1971). Suggestion was shown to have a powerful effect on the reporting of echoencephalograms (White *et al.*, 1969), but the infective agent has not been isolated from clinicians. In this study, we used filmed toe movements during plantar stimulation to investigate whether previous knowledge biased the answer 'up' or 'down'.

Methods

SUBJECTS

The experimental group consisted of 20 neurologists from two Dutch university hospitals. Most of them had a few years' experience in clinical neurology, some more, others less. For practical reasons they were examined in three groups, but

the instructions remained identical. It was explained that this was a study about the interpretation of plantar reflexes and that some of these reflexes were to be shown on film, accompanied by some other data to simulate the clinical situation as well as possible.

Thirty other neurologists from a third university hospital constituted the control group: they judged the films without additional data (no information: I_0).

FILMS

Nineteen short films, each showing a right foot from the medial side, formed the basis of the stimulus material. They were separated by title numbers. Plantar stimulation could be seen and was performed three times for every foot. Within the sequence of 19 films, two films were repeated: No. 4 was identical with No. 15 (film A), and No. 7 with No. 18 (film B). These two duplicated films were the actual object of study (test films). All three reflexes in either test film were equivocal: voluntary, slight up-and-down movements of the great toe. The other 15 films served to cover up the two repetitions (cover films), and consisted of downward, upward, equivocal, and absent hallux movements in about equal proportions (true or simulated reflexes).

ACCOMPANYING INFORMATION

A slide with a fictitious abstract of history and physical signs (minus the plantar reflex) introduced every film fragment for the experimental

group. The 19 abstracts were chosen from a stock of 36; three separate neurologists arranged these in nine groups of four, according to their 'Babinski values'. The two films that occurred twice were on one occasion preceded by 'high value information' (I_{up}), indicating, for instance, a right sided hemiplegia, and at the other presentation by 'low value information' (I_{down}) from the opposite end of the scale (for example, bizarre complaints and no signs). The order was reversed for the two test films, to control time effects: A(I_{up}), B(I_{down}), A(I_{down}), B(I_{up}). The remaining 15 abstracts were chosen from the whole range of 'values' and were assigned more or less at random to the cover films. We avoided mismatches between the aspect of the foot and the age or sex given on the slide.

RECORDING OF INTERPRETATIONS

All subjects were given a list of the patients' numbers, each number followed by five possible choices: unmistakably extensor, possibly extensor, neither flexor nor extensor, possibly flexor, and unmistakably flexor. They marked only one answer for each foot, according to their own opinion. At the end of every sitting the participants were asked to write down any comments.

Results

The subjects' marked choices for all films were converted into numerical scores. With regard to the cover films, the experimental group and the control subjects showed no significant difference in mean ratings (Mann-Whitney, $z=0.15$, not significant).

The scores for each test film were subjected to a two-factor analysis of variance. For both films there were significant interactions between the two subject groups and the two presentations ($F_A=5.36$, $F_B=7.23$, degrees of freedom (df)=1.48, both $p<0.05$). The significant interactions for both films were the reason for disregarding possible main effects for the two factors. The Figure illustrates the interactions and the various mean ratings.

All relevant differences between means were t -tested. For both films, I_{up} resulted in significantly more pathological markings than I_{down} ($t_A=3.10$, $df=19$, $p<0.01$ and $t_B=2.94$, $df=19$, $p<0.01$). Compared with I_0 (control group) at corresponding presentations, I_{up} also gave a significant deviation to the pathological side ($t_A=2.03$, $df=48$, $p<0.05$ and $t_B=2.59$, $df=48$, $p<0.01$). On the other hand, no significant differences were found between presentations with I_{down} and I_0 , although I_{down} tended to give more normal scores ($t_A=1.46$,

Figure Mean ratings for the two test films

———— = Film A; - - - - - = Film B; I_0 = no information; I_{up} = information suggesting extensor plantar response; I_{down} = information suggesting flexor plantar response

$t_B=0.74$, not significant). The differences between first and second presentation for I_0 did not reach significance ($t_A=0.28$, $t_B=0.95$).

Discussion

The results show that many neurologists have, to some extent, already made up their minds whether the great toe will go up or not, before they have even started to examine the plantar reflex. Comparison with unbiased subjects revealed that it is easier to convince oneself that an equivocal response represents the expected Babinski sign than to disregard out-of-place extensor movements of the hallux. This influence of foreknowledge may not surprise psychologists, but at the bedside it is a reason for considerable modesty in the interpretation of equivocal plantar reflexes. It also affects the conclusion of a recent report (Dohrmann and Nowack, 1973), that slow 'hockey-stick' stimulation is the best method to produce an extensor response, as the authors were their own measuring instruments and did not state what they expected to find.

How can one protect oneself against bias? Clinical routine appeared to be no safeguard in

this study, although the number of subjects was too small for conclusions about relative probabilities. First of all, the reasons whether or not a doubtful extensor movement 'counts' should be logical, not intuitive. It has been argued long ago but convincingly (Walshe, 1914) that reflex contraction of the extensor hallucis longus forms an integral part of the flexion reflex of the leg, and one of us showed recently that this can be turned to practical advantage (van Gijn, 1976). This synergy is largely unknown: only one of the 50 subjects in this study regretted that the films did not show the leg (a fair number complained they could not see the other toes, but the 'fan sign' (Babinski, 1903) has more traditional than practical value). In addition, one should be aware of the power of prejudice. We have no reason to assume that these effects are limited to the 20 neurologists in our experimental group.

We are grateful to all the doctors who willingly subjected themselves to this study. Professor F. Verhage, Dr H. J. Duivenvoorden and Dr H. van Crevel gave valuable advice. The latter also classified the abstracts, together with Dr A. R. Wintzen and Dr J. J. van der Sande (Leyden). Th. van der Biessen filmed all the toes.

References

- Babinski, J. (1903). De l'abduction des orteils. *Revue Neurologique*, **11**, 728–729.
- Dohrmann, G. J., and Nowack, W. J. (1973). The up-going great toe—optimal method of elicitation. *Lancet*, **1**, 339–341.
- Fletcher, C. M. (1952). The clinical diagnosis of pulmonary emphysema—an experimental study. *Proceedings of the Royal Society of Medicine*, **45**, 577–584.
- Garland, L. H. (1959). Studies on the accuracy of diagnostic procedures. *American Journal of Roentgenology*, **82**, 25–38.
- Gijn, J. van (1976). Equivocal plantar responses; a clinical and electromyographic study. *Journal of Neurology, Neurosurgery, and Psychiatry*, **39**, 275–282.
- Gross, F. (1971). The Emperor's clothes syndrome. *New England Journal of Medicine*, **285**, 863.
- McCance, C., Watt, J. A., and Hall, D. J. (1968). An evaluation of the reliability and validity of the plantar response in a psychogeriatric population. *Journal of Chronic Diseases*, **21**, 369–374.
- Walshe, F. M. R. (1914). The physiological significance of the reflex phenomena in spastic paralysis of the lower limbs. *Brain*, **37**, 269–336.
- White, D. N., Kraus, A. S., Clark, J. M., and Campbell, J. K. (1969). Interpreter error in echoencephalography. *Neurology (Minneapolis)*, **19**, 775–784.