THE DIOPTRON II-IN PRACTICE

Chris French and Ivan Wood[1]

IN OUR LAST article (THE OPTICIAN, August 14) we examined the design and operation of the Dioptron II. We remarked that the consumer response of our students appeared to be favourable, notwithstanding the observation that there remained a hard core minority who were unconvinced of its usefulness. We have no doubt that the students' judgements were in part based on claims they had seen in advertising literature and evaluations reported in journals. Nearly all these relate to studies carried out abroad, mainly in the United States. For some British optometrists the paucity of British studies may present a dilemma, and we hope that they will find the following results of

Accuracy

Whenever you are evaluating a test, procedure, method, technique or instrument you will want to know how accurate it is. The trouble is that when the average person talks about accuracy he tends to do so without precision. To different people it can mean different things. In part this is because there *are* indeed different ways of measuring accuracy, but also in part it is because they are confusing two different concepts. Accuracy can best be understood by separating it into two elements—reliability and validity. *Reliability* is the consistency with which measurements are made—their repeatability. *Validity* is the degree to which measurements measure what they say they measure.[2] Thus an imaginary instrument which gave exactly the same results every time would have 100 per cent reliability. This is obviously a good state of affairs, but it does not in itself imply that the instrument's measurements have high validity. To assess validity, one needs to make comparison with results which are 'correct' by definition. Our problem is that we do not know what *is* correct! There are, of course, various candidates in this area including a conventional retinoscopy and subjective refraction, but these are carried out by humans who are themselves prone to bias and error. Still, they are at present the norm and it is therefore only natural that they should be the main criteria against which the Dioptron should be validated. Of course if Dioptrons had been invented first then we might well be assessing the validity of retinoscopy and subjective refraction against the Dioptron rather than vice-versa, although in essence it would amount to much the same thing.

Human foibles

A difficult problem frequently arises when one tries to assess reliability and validity. Unfortunately, human beings have foibles and are very easily influenced, even when they believe they are being unbiased and objective. This is not just a failing of the layman but is also one of the scientist and clinician. 'Experimenter effects', as some manifestations of this phenomenon are labelled (see Rosenthal, 1966), are ubiquitous and very important.

For example, everyone has expectations as to what they are going to find out as a result of an investigation. Frequently we call these guesses 'hypotheses' because it sounds better. Well, if after doing the required calculations and computations you get the results expected there is the temptation to shout 'Hoorah' and call it a day. Consider, however, the situation when your results are not what you had expected. What do you do then? Of course you check and check again until you come across a mistake. If the correction of this mistake gives you what you had expected in the first place then of course you breathe a sigh of relief and label it *the* mistake and look no further. But of course all this is wrong! It leads to false reports and an accumulation of more 'positive' published findings than should be warranted. Good scientists — and good *optometrists* — should always double and treble check irrespective of their hopes and fears. It is for reasons like this that the ophthalmic optician looking for a new instrument should always take every study with a pinch of salt and not just depend on the findings of a single report—whether it be reporting good or bad results.

Even this report should not be above suspicion, although it is our belief that we have taken the utmost care. Initially we requested the students to transpose the Dioptron's findings on to computer coding forms, but we soon found that this operation had to be double-checked because the students were very careless in their transcriptions. Of course this brings out a particularly useful feature of the Dioptron—its print out. At least the computer does not forget what it finds before it writes it down, and when it does this it normally does it in a legible form.

Expectations and bias can cause problems in other ways. In some evaluation studies people have tried to validate Dioptron findings by comparing them, for example, with retinoscopies. Unfortunately, they have not carried out the latter 'blind' without knowledge of the machine's findings. They have had this knowledge and have either argued (i) that this has not influenced them or alternatively (ii) that this is the best comparison procedure as this is how one would use an automatic refractor in practice. Certainly one might well use a Dioptron this way much as one might make similar use of patients' previous records, spectacle prescriptions and retinoscopies and use them as a basis for one's subjectives. But the method, although perhaps providing interesting numbers, does not give us a proper test of an instrument's validity. All it tells you is how much a biased retinoscopy will differ from the Dioptron's print out. The closer they are to each other then the greater has been the degree of influence of the Dioptron on the optometrist (Reimers *et al*, 1973). So, in all our studies, refractions and retinoscopies were carried out blind and the participants were exhorted not to check on the Dioptron's findings until after they had recorded their own results.

Operator error

The Dioptron is fully automatic as we have described previously, but as with any instrument patient co-operation must be obtained and the operator must follow the standard procedure. The task is not difficult but any simple measurement procedure requires a certain amount of skill and savoire-faire. In the clinical situation one would expect the one or two operators involved to quickly acquire the necessary proficiency without any problems. In our studies we used operators in two different ways. On two occasions we used fully trained operators, but on the third we used undergraduates who had been instructed briefly by Coherent. Because of the short induction period and limited practice it might be felt that some of the measurements would not be up to scratch. Our personal opinion is that this would be an

continued on page 20

[1]*Ivan Wood and Chris French are lecturers at UMIST Ophthalmic Optics Department*

[2]*Sometimes people carefully refer to precision and accuracy in place of reliability and validity.*

The Dioptron II —in practice

continued from page 18

exaggeration. A good instrument should indicate when its findings are not dependable and the Dioptron does this in several ways as we have already described. In practice one would tend to repeat measurements where there were too many missed scans or the confidence factor was low, and normally this would result in an acceptable second time measurement. For technical reasons we did not ourselves follow this procedure. In this report we will only be presenting first time Dioptron findings but these will be ones which the instrument indicated it had relative confidence in.

Comparison procedures

Our results are based on our own experience with the Dioptron in 1979 and 1980. In the main, the data comes from the UMIST Open Clinic in the Ophthalmic Optics Department. The clinic is open two mornings each week for five months of the year, and final year students see patients (students and employees at the university) while being supervised by full-time members of staff or local ophthalmic opticians who undertake part-time teaching. Two sets of Dioptron readings (student operators) and a final subjective refraction were recorded for each patient. With the generous help of the staff from Manchester Royal Eye Hospital we also undertook two small supplementary studies involving different types of patients. In the first, which was carried out on members of the general public (visitors to a UMIST Open Day) we compared the Dioptron findings with those from the retinoscopies of an hospital ophthalmic optician, while in the second on outpatients at Manchester Royal Eye Hospital we compared the Dioptron findings with the final prescription.

Reliability

There appears to be a dearth of information on the Dioptron's reliability. In fact, we have asked Coherent USA to let us know what they think it should be on more than one occasion. Paradoxically, they have declined. One can only speculate on their reticence.

We only have reliability data for our Open Clinic patients where they were tested twice with an intervening conventional refraction. When you plot the sphere equivalents for the first Dioptron measurements against those for the second you obtain a narrow pencil of dots presenting a Pearson product-moment correlation coefficient of 0.995 for almost 250 eyes. The graph looks good and the coefficient sounds good but does either do the optometrist any good? If we square the correlation coefficient then we get a pretty good *psychological* idea of the degree of agreement between the two sets of measures—it is the number that you or we would attach to the pencil of dots. Thus 0.995 represents 99 per cent agreement. Despite this, for a clinician to really get to grips with these results and truly assimilate their meaning we require them in a different form. One way of doing this is to look at the *difference* between the pairs of Dioptron readings for each eye. When we do this we discover that for the sphere equivalent measures over three quarters of the differences are less than 0.25 dioptres, twenty-four out of twenty-five are less than 0.50 and one in a hundred are larger than 1.00. Another statistic which may be useful is the standard error of measurement (sem). This assumes that if measurements are repeated a number of times on the same subject the results will be normally distributed with the best estimate the mean of this distribution and the sem its standard deviation. Thus the smaller the sem the more reliable our data. In practice one does not normally measure the standard error of measurement directly. Instead one deduces

TABLE 1

RELIABILITY

Expressed in terms of the test-retest coefficient and the standard error of Pearson product-moment correlation measurement.

		Dioptron II		Dioptron I	Students
		test-retest r	sem	sem	sem
sphere equivalent	(n = 249)	0.995	0.17 DS	0.16 DS	0.26 DS
sphere component	(n - 249)	0.994	0.17 DS	0.18 DS	0.31 DS
cylinder component	(n = 249)	(0.912)	(0.15 DC)	0.11 DC	0.19 DC
axis	(n=100)	(0.988)	(6°)	6°	11°

Parentheses have been inserted around some of the above as these statistics assume normally distributed data and this assumption is particularly violated by the axis data.

Only those measurements where the Dioptron II itself indicated 'good' dependability were included. That is there had to be three or fewer missed scans and the 'confidence factor' had to lie between 0.01 and 1.00. For the axis calculations small cylinders of less than 0.50 DC were excluded.

The Dioptron II subjects were patients to the UMIST Open Clinic and the operators final year undergraduates.

The Dioptron I study was carried out by Sloan and Poise (1974) and relates to direct measurements on only four eyes. The student refractionist results also relate to direct measurements of the sem and are based on the performance of first year students on 410 eyes (circa 5000 refractions).

TABLE 2

DIOPTRON II RELIABILITY

Expressed in terms of the difference between two independent Dioptron measurements

		magnitude of differences		
		< ¼ D | < 5°	< ½ D | < 10°	< 1 D | < 20°
sphere equivalent	(n = 253)	76%	96%	99%
sphere component	(n = 253)	78%	95%	100%
cylinder component	(n = 253)	82%	97%	100%
axis	(n = 102)	59%	83%	99%

The percentage figures show the relative frequency with which differences of the magnitude indicated occurred between two independent Dioptron measurements carried out on the same eye.

Only those measurements where the Dioptron itself indicated 'good' dependability were included. That is there had to be three or fewer missed scans and the 'confidence factor' had to lie between 0.01 and 1.00. For the axis calculations small cylinders of less than 0.50 DC were excluded.

The subjects were patients to the UMIST Open Clinic and the operators final year undergraduates.

continued on page 23

continued from page 20

it from other statistics. The sem for the sphere equivalent data is 0.17 dioptres. This means that if our normality assumption is legitimate then two-thirds of our measurements will be within 0.17 DS of our best estimate.

Of course it is only natural for any clinician pondering these figures to still wonder what they mean. Optometrists just do not know how reliable their own retinoscopies or subject!ves are and do not have a personal criterion with which to compare. Frequently when they think they *do* know their consistency they are mistaken. Our reliability figures are clearly better than those reported by Hirsch (0.23 D to 0.25D) in 1956 and Safir *et al* (0.29D to 0.39D) in 1970 for retinoscopy, and Jennings and Charman for duochrome, simultan and laser refraction (0.32D) in 1973. Tables 1 and 2 give our corresponding figures for sphere and cylinder components and cylinder axis.

In 1974 Sloan and Poise measured the Dioptron I's sems directly but for only four eyes (a year later Poise added two eyes and took one away). If we average their figures we obtain the numbers also given in Table 1. It can be seen that their results for the first version of the Dioptron and a few eyes are quite compatible with ours for the Dioptron II and 249 eyes. In Table 1 we also present the results for novice student refractions with the sem measured directly.

As indicated before, we only used Dioptron findings which were dependable.[3] This may have resulted in our stacking the cards in the Dioptron's favour. Clearly some patients will have been excluded from our sample and these may have had difficult eyes. In this, the Dioptron's task was undoubtedly made easier but we feel the amount of bias involved was small. With student operators, around 70 per cent of the Dioptron measurements reached our criteria of dependability. If dependability was not related to patient difficulty then one would expect 91 per cent of paired Dioptron measurements to include at least one 'good' measure. We found that the proportion was close to this but a little lower at 89 per cent—confirming that patient difficulty was a small but significant factor as one would expect. Obviously if one wanted to obtain 'good' results for virtually all one's patients then one would need several Dioptron measurements on a few of them. In practice, of course, one would be unlikely to pursue this goal very far and instead one would be more likely to settle for 'useful' measurements and at least one of these would be expected by repeating fewer than 10 per cent of the Dioptron findings—See Tables 3 and 4 for details.

[3] *The criteria we use here are a compromise between stiffer ones which would give more accurate results than those we report but would result in the rejection of more Dioptron findings; and sloppier ones which while involving the need for fewer repetitions would give less accurate results. In particular it should be noted that in certain circumstances it might make sense to reject measurements involving as little as one missed scan and also reject those where the 'relation junction' (line 7) was high. Alternatively one might choose to relax the requirement that the confidence factor should not exceed 1.00.*

TABLE 3

DIOPTRON II 'CONFIDENCE FACTOR'

internal consistency or 'confidence factor'	verbal description of dependability	relative frequency
0.01 to 1.00	'good'	70%
1.01 to 10.00	'useful'	23%
10.01 to 20.00	'suspect'	2%
20.01 plus & & 0.0	'untrustworthy'	5%

The percentage figures show the relative frequency with which the Dioptron recorded measurements of the indicated 'confidence factor' which can range from 0 to 999. The figures are based on 968 eyes.

The subjects were patients to the UMIST Open Clinic and the operators final year undergraduates.

TABLE 4

PROBABILITY OF OBTAINING DEPENDABLE RESULTS

—when *two* Dioptron measurements are carried out

at least one 'good' or better measurement—89%
at least one 'useful' or better measurement—*99%*

The descriptions are those applied to the Dioptron's 'confidence factor'. The figures are based on 484 eyes measured twice.

The subjects were patients to the UMIST Open Clinic and the operators final year undergraduates.

Validity

When you plot the Dioptron sphere equivalents against,

i) the subjective findings of our students in the Open Clinic,
ii) the subjective findings on outpatients from a local eye hospital, and
iii) retinoscopy results on members of the general public,

one again obtains scatter diagrams with narrow ellipses of points which pass through the origin. The corresponding Pearson product-moment correlation coefficients are lower than those reported for reliability. These validity coefficients vary between 0.97 and 0.99—equivalent from 95 to 98 per cent agreement. Figure 2 shows the scatter diagram of Dioptron against subjective for the Open Clinic.

Once again, perhaps the best way of presenting this information is by considering the differences. For the non-hospital patients just under half the differences are less than 0.25 dioptres and just over three-quarters are less than 0.50 DS, while two or four per cent are larger than 1.00 DS. It is interesting to note that the retinoscopy-Dioptron and subjective-Dioptron differences are quite similar and that this similarity extends to the sphere and cylinder component comparisons and perhaps even that for the cylinder axis, despite the fact that we are dealing with different instrument operators, different validity criteria and different optometrists. The differences for hospital patients appear larger although one must be cautious here as we are only dealing with dependable measurements on fewer than 60 eyes and the criteria of validity for subjective refraction may itself be questionable in the presence of poor visual acuity.

Coherent USA's reticence in reporting reliability figures does not extend to validity where they do make claims without making it too clear as to the type of patient they have in mind. Our validity numbers are a little different from theirs but we do not feel that this discrepancy is important—particularly as we *must* continue to bear in mind the fallibility of the human validity criteria. Our validity figures for sphere and cylinder components, and cylinder axis are detailed in Tables 5 and 6.

Validity figures have frequently been reported, particularly for non-British groups. Our validity coefficients are similar to those published by Sloan and Poise (1974), and Poise (1975) for the Dioptron I. The former's report is particularly interesting as it includes figures comparing independent retinoscopy and subjective measures. These correlations are quite similar to ours for the Dioptron II—subjective comparison, but caution should be observed as we are dealing with correlation coefficients which are particularly sensitive to disruption by the odd 'rogue' observation, are dependent upon the range of refractions under scrutiny, and assume normal distributions. In passing it should be mentioned that the

continued on page 24

The Dioptron II —in practice

continued from page 23

data of our Open Clinic and Open Day groups of patients were negatively skewed tending towards myopia with an average prescription of -1 DS, and standard deviations of 3 DS and 2 DS respectively. The hospital group on the other hand tended to be hypermetropic with a mean of + 2 DS and a standard deviation of 5 DS.

TABLE 5
VALIDITY

Expressed in terms of the Pearson product-moment correlation coefficient between the Dioptron results, and subjective and retinoscopy findings.

DIOPTRON II vs SUBJECTIVE (i)			
sphere equivalent	(n = 825)	0.985	
sphere component	(n = 825)	0.985	
cylinder component	(n = 825)	(0.835)	
axis	(n = 372)	(0.947)	
DIOPTRON II vs SUBJECTIVE (ii)			
sphere equivalent	(n = 57)	0.99	
sphere component	(n = 58)	0.99	
cylinder component	(n = 57)	(0.86)	
axis	(n = 57)	(0.96)	
DIOPTRON II vs RETINOSCOPY (iii)			
sphere equivalent	(n = 163)	0.973	
sphere component	(n = 163)	0.968	
cylinder component	(n = 163)	(0.803)	
axis	(n = 68)	(0.96)	
DIOPTRON I vs SUBJECTIVE (iv)			
sphere equivalent	(n = 126)	0.982	[0.991]
sphere component	(n = 126)	0.978	[0.988]
cylinder component	(n = 128)	0.766	[0.764]
axis	(n = 66)	0.902	[0.968]
DIOPTRON I vs SUBJECTIVE (v)			
sphere equivalent	(n = 300)	0.980	
sphere component	(n = 300)	0.977	
cylinder component	(n = 300)	0.686	
axis	(n = 70)	0.935	

Only those measurements where the Dioptron II itself indicated 'good' dependability were included. That is, there had to be three or fewer missed scans and the 'confidence factor' had to lie between 0.01 and 1.00. For the axis calculations small cylinders of less than 0.50 DC were excluded with the exception of analysis (ii) owing to the rather small number of eyes involved.

Three groups of Manchester subjects were involved.:
(i) patients from the UMIST Open Clinic with operators and refractionists who were supervised final year undergraduates; (ii) patients from MREH outpatients with a trained operator and hospital ophthalmic optician as refractionist; (iii) patients were visitors to a UMIST Open Day with a trained operator and hospital retinoscopist.

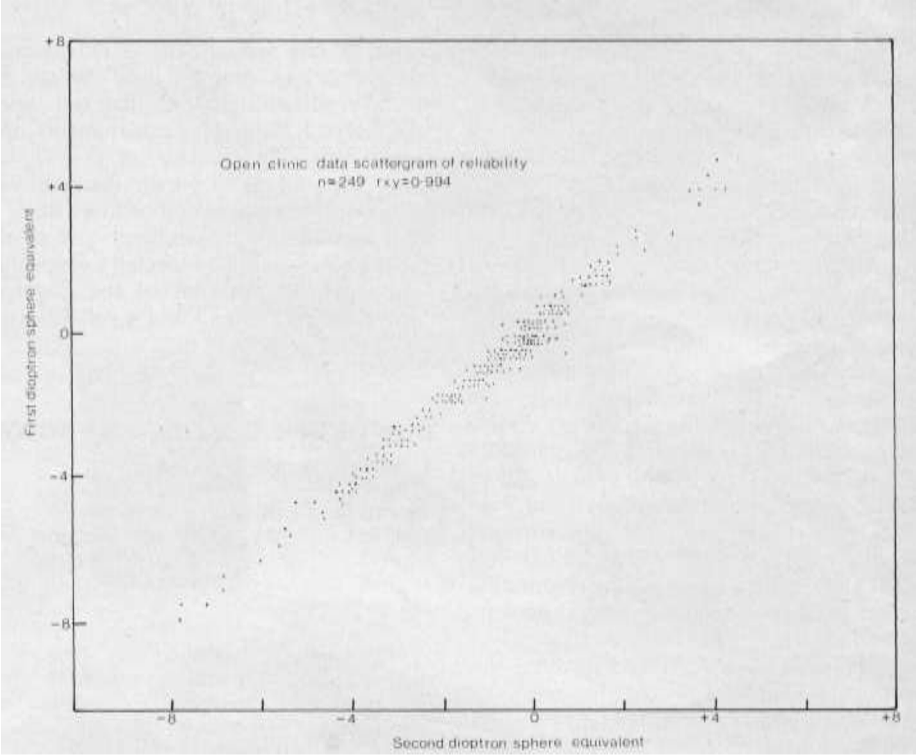

Figure 1: Reliability scatter diagram comparing independent Dioptron sphere equivalent readings for Open Clinic patients

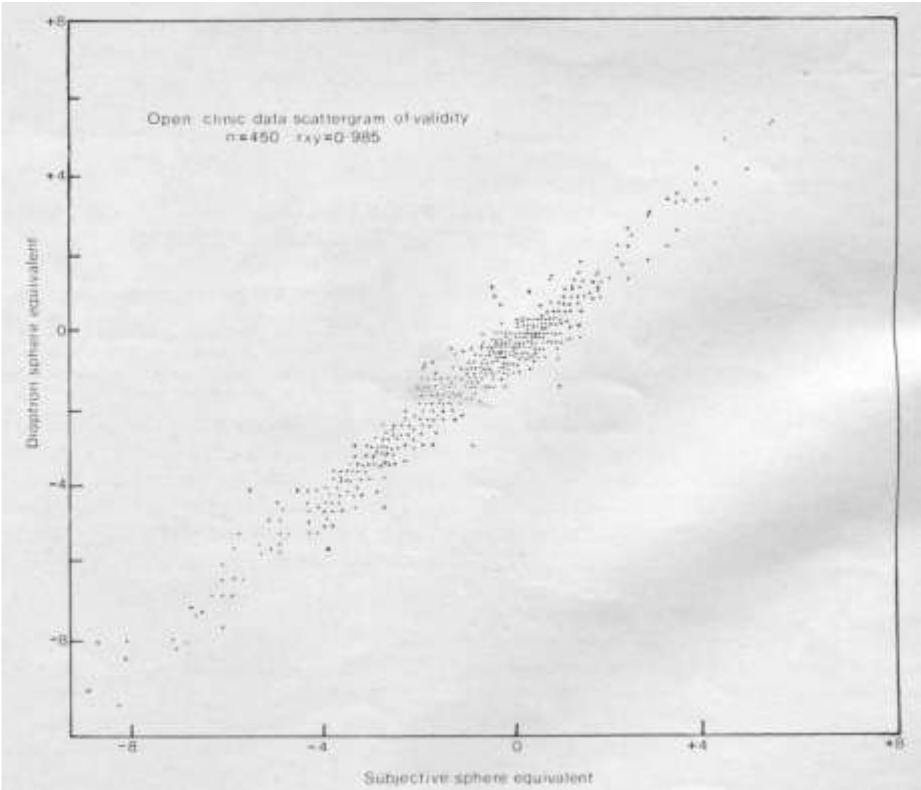

Figure 2: Validity scatter diagram comparing Dioptron and subjective sphere equivalent data from the Open Clinic

continued on page 30

The Dioptron II —in practice

continued from page 24

Parentheses indicate that the normality assumption of the Pearson product-moment correlation coefficient has been violated.

Study (iv) was reported by Sloan and Poise in 1974, and (v) by Poise in 1975. Both used a Dioptron I. The figures in square brackets are the correlations for a blind comparison between retinoscopy and subjective.

TABLE 6

DIOPTRON II VALIDITY

Expressed in terms of the difference between the Dioptron's and conventional subjective and retinoscopy findings.

	magnitude of differences		
	¼D	½D	ID
	5°	10°	20°
DIOPTRON vs SUBJECTIVE (i)			
sphere equivalent (n = 825)	48%	77%	96%
sphere component (n = 825)	51 %	79%	97%
cylinder component (n = 833)	61%	85%	97%
axis (n = 372)	40%	61%	87%
DIOPTRON vs SUBJECTIVE (ii)			
sphere equivalent (n = 57)	39%	72%	96%
sphere component (n = 58)	38%	66%	93%
cylinder component (n-57)	49%	81%	95%
axis (n = 57)	39%	61%	89%
DIOPTRQN vs RETINOSCOPY (iii)			
sphere equivalent (n = 163)	47%	76%	98%
sphere component (n= 163)	44%	79%	96%
cylinder component (n= 163)	61%	87%	99%
axis (n = 68)	34%	64%	89%

The percentage figures show the relative frequency with which differences of the magnitude indicated occurred between the Dioptron and given findings.

Only those measurements where the Dioptron itself indicated 'good' dependability were included. That is, there had to be three or fewer missed scans and the 'confidence factor' had to lie between 0.01 and 1.00. For the axis calculations small cylinders of less than 0.50 DC were excluded with the exception of analysis (ii) owing to the rather small number of eyes involved.

Three groups of subjects were used:

(i) patients from the UMIST Open Clinic with operators and refractionists who were supervised final year undergraduates;

(ii) patients from MREH outpatients with a trainer operator and hospital ophthalmic optician as refractionist;

(iii) patients were visitors to a UMIST Open Day with a trained operator and hospital retinoscopist.

Conclusion

In conclusion we can say that our own experience with the Dioptron II has been broadly favourable. Our results are not too different from the North American reports concerning the validity of the Dioptron measurements. Previous studies have tended to neglect the question of reliability. We can say that the Dioptron compares most favourably in terms of reliability when compared with other known refraction methods.

References

HIRSCH, M j The variability of retinoscopic measurements when applied to large groups of children under visual screening conditions. *American Journal of Optometry and Archives of American Academy of Optometry*. 1956. 33. 410-416.

JENNINGS, J A M and CHARMAN, W N A comparison of errors in some methods of subjective refraction. *The Ophthalmic Optician*, January 6, 1973.

POLSE, K A An automatic objective optometer. *Archives of Ophthalmology*. 1975. 93. 225-231.

REIMERS, P L, COHN, T E and FREEMAN, K D The influence of bias upon retinoscopy. *American Journal of Optometry and Archives of American Academy of Optometry*. 1973. **50.** 647-651.

ROSENTHAL, R *Experimeter Effects in Behavioural Research.* New York: Appleton-Century-Crofts. 1966.

SAFIR, A, HYAMS, L: PHILPOT, J and JAGERMAN, L S Studies in Refraction 1.The precision of retinoscopy. *Archives of Ophthalmology*. 1970. 84. 49-61.

SLOAN, p G and POLSE, K A Preliminary clinical evaluation of the Dioptron. *American Journal of Optometry and Physiological Optics*. 1974. **51.** 189-197.