MIST (Melbourne Initial Screening Test) Evaluation Studies

February 1998

Part 1: Reliability and validity testing of the MIST
Part 2: A negative predictive study of the MIST
Part 3: A positive predictive study of the MIST

Authors

Linda Santamaria, MAppSc, DipAppSc, DOBA
Shayne Brown, MAppSc, DipAppSc, DOBA
Ian Story, PhD, BBSc

Institution

Department of Clinical Vision Sciences
La Trobe University
VIC   3086
Australia

Address for Correspondence

Linda Santamaria
Department of Surgery, Monash University
Monash Medical Centre
Level 5, Block E
246 Clayton Rd
Clayton, VIC, 3168

MIST (Melbourne Initial Screening Test) evaluation studies

ABSTRACT

The aim of these three studies was to assess the test-retest reliability of the MIST (Melbourne Initial Screening Test), a new vision screening tool designed for use by vision screeners of pre-school children and to determine the positive and negative predictive value.

In the first study, 471 children participated in two sessions, either with the Sheridan Gardiner Singles (SGS) or the MIST. The children were allocated to four groups - Group A, test-retest with SGS, one orthoptist; Group B, test-retest with MIST, one orthoptist; Group C, test-retest with MIST, two different orthoptists; or Group D, test-retest with MIST, an orthoptist and a nurse. The test-retest reliability of the MIST demonstrated correlations of 0.79 and 0.78 for the right and left eyes respectively, in comparison to the SGS score of 0.71 and 0.79. The inter-rater reliability of the MIST demonstrated correlations of 0.60 and 0.74 for orthoptists, and 0.63 and 0.60 with orthoptists in comparison to nurses. These results confirm that the MIST is a reliable screening tool for pre-school vision screeners.

In the second study, 201 children participated in a gold standard orthoptic and ophthalmic examination after passing their MIST screening. The negative predictive value of the MIST vision screening program was found to be 97.5%. Five children were found to have failed the gold standard examination. Four of the children had an astigmatic error of 1.50 dioptres in one eye, one of whom also failed the visual acuity criterion. One other child had an intermittent esotropia. This result means that 2.5% of the children who pass the MIST vision screening test may actually have strabismus, amblyopia or a refractive error, something to be noted in any screening program. Both nurses and parents must be aware of this information when vision screening is performed.

In the third study, forms were returned for 3,854 children referred from the MIST vision screening program over a three-year period. Visual assessment results of 2,623 of these children were obtained from examining clinicians. The positive predictive value of the MIST was found to be 44.4%. Given the less than 5% prevalence of amblyopia in children of this age group, this over-referral rate would be considered acceptable. A balance must be reached between the positive and the negative predictive values. The follow-up examination of each child is a cost to the health care system to be considered in balance with the early detection of a visual problem and any long-term costs associated with visual disorders.

The MIST (Melbourne Initial Screening Test) has been designed as a simplified vision screening test for 3.5 to 4.5-year old children, to be performed by maternal and child health nurses in the state of Victoria.[1] It was introduced in 1998 as part of the visual surveillance program, combined with observation of ocular position and corneal reflections and questions regarding family history and ocular problems. As previously described by Brown and Story[1] the MIST, a letter matching test, has a pass/fail method of assessment rather than a threshold test of visual acuity, with five test letters of 3/5 size. Initial results showed that it was easy to administer by the nurses and had a high compliance rate.

The maternal and child health nurses perform vision screening as part of a general developmental assessment, allowing earlier screening at minimal cost to the health system. As part of a wider assessment this allows vision screening in the context of the child's whole development. This surveillance program provides not only assessment, but also gives the chance to raise parental awareness and provide incidental education on many health and development issues.

Since the implementation of the MIST in 1998 there has been much public debate concerning the value of preschool vision screening in the absence of randomised control studies investigating the age effect of amblyopia treatment and outcomes, with the recommendation that programs should be discontinued.[2] Various authors then emphasised the need for further research and it may be suggested that in the absence of appropriate evidence on such aspects as the natural history of amblyopia, the age effect of amblyopia treatment outcomes, or whether amblyopia is disabling, that preschool vision screening should continue in the most cost-effective way.[3-5]

The MIST forms part of the routine general developmental assessment in a similar manner to that reported by Thorburn and Roland[6]. Previous reports have suggested that orthoptists are more effective at vision screening than other health professionals,[7, 8] however studies have shown that screening can be performed effectively by others providing that training is given to ensure high quality assessments.[6]

The accepted criteria for a screening test are that it must be simple, reliable and valid.[9] The purpose of the evaluation project, funded by The Department of Human Services, was to evaluate the effectiveness of the MIST as a valid vision screening tool, with three separate studies, the reliability and validity testing of the MIST, a negative predictive study and a positive predictive study.

3

Part 1: Reliability and validity testing of the MIST

INTRODUCTION

The reliability of any screening test must be evaluated not in the clinic, but in the hands and venue of the screening personnel. In this context, test-retest results could be considered in two ways, either as the repeatability of an individual test score or of a pass/fail classification only.[10, 11] The former may provide more detail of individual test results, but in the context of vision screening, pass or fail is the information required to make a referral decision.

There have been few studies on the test-retest reliability of available visual acuity tests. One issue to be considered is the number of children who are able to complete a test at the first presentation. Hered and colleagues[12] found with the HOTV chart at 3 metres, that 93% of 3 to 5-year olds were testable at the first session compared to 97% at a second session, with 85% of 3-year olds and 98% of 4-year olds testable at the first session. Similar results were reported by other studies.[11, 13] A study, with single optotypes at 6 metres reported a lower testability rate of 64% in children less than 3 years old, but 93% in 4-year olds.[14] Another, using isolated surrounded optotypes at 4 metres, found 67% of 3-year olds and 87% of 4-year olds testable.[15] These and other studies show an increase in testability with increased age, decreased testing distance and repeat testing.[16]

Test-retest reliability studies have reported widely varying results for correlational values. Hered and colleagues,[12] grading each child as pass, fail or untestable, reported an F value of 0.54 for the HOTV chart with 3 to 5-year old children, or a value of 0.71 after excluding all children untestable at the first session. In comparison, Friendly[11] found a rho value of only 0.16 for a slightly wider age group, and Sprague et al[13] a correlation coefficient of 0.398, using continuous variable measurements. A more recent study, using isolated surrounded HOTV letters, reported a correlation of 0.82 in a group of 2 to 7-year old children.[15] McGraw et al[17] studied the test-retest reliability of a single letter test at 3 metres and found a coefficient of repeatability of two lines difference in a group of children with mean age 5.3 years. In adult populations, the correlation coefficients have been reported to be much higher at 0.98[18] or 0.84[19] for the Bailey-Lovie LogMAR chart. Comparisons between studies are difficult as methodology and statistical analyses vary widely.

For a screening test to be efficient it must have a high compliance rate, so that a result can be obtained from the majority of participants. As a screening program is designed only to identify those at risk, an actual measurement of visual acuity is not required. This may in effect mean a balance between ease of administration and accuracy of the result. The referral practice would tend to be towards referring any child who shows any risk of failing the test. The aim of this study was to assess the test-retest reliability of the MIST. The test-retest reliability of the Sheridan Gardner Singles (SGS) was assessed as it is a standard clinical test, commonly used in clinics and vision screening programs. The MIST was then evaluated for its reliability in comparison to this clinical test, both with experienced eye health personnel and within its screening context with maternal and child health nurses.

METHOD

Participants
Children were recruited from kindergartens and child care centres in the Melbourne metropolitan region. After an initial approach to centres, written consent was obtained from those wishing to participate. Parental Information and Consent forms were then distributed, to be signed and returned prior to testing. The study was approved by the Human Ethics Committee, La Trobe University, and the Department of Human Services Ethics Committee.

4

Procedures

Vision was tested with either the SGS at 6 metres or with the MIST at 3 metres, both letter matching tasks. The right eye was tested first as a routine procedure. The time taken for the vision tests was recorded.

The aim was to assess each child twice, by either the same or a different tester. Each centre was allocated to one of four groups. The SGS test-retest (Group A) and the MIST test-retest (Group B) were retested by the same orthoptist. Two other orthoptists were involved in the MIST inter-rater test by different orthoptists (Group C). Seven maternal and child health nurses were involved in the MIST inter-rater test of orthoptists in comparison to nurses (Group D).

The results of each test were graded as pass, fail or untestable. As the data was ordinal measurement, correlational analysis was performed with Spearman's rho (ρ). Due to the extreme skew in the distribution of the responses, where the agreement is predominantly in one cell, measures of agreement such as Cohen's kappa will tend to underestimate the true agreement. In this study, where the children were from a normal population, the vast majority of results would be in the pass/pass cell, with only a few in the fail/fail cell, so it is not possible to statistically test the amount of disagreement.

RESULTS

An orthoptist or a nurse, with either the SGS or the MIST, tested a total of 583 children. As some were absent at the time of one of the testing sessions, 471 children participated in the full test-retest study. The children ranged in age from 35 to 67 months. The number of children, their mean age and the period of time between the two testing sessions are presented in Table 1.

Table 1   Number of children in each test-retest group

	Test N = 583	Retest N = 471	Mean age (SD) (months)	Mean retest time (SD) (days)
Group A, SGS, one orthoptist	137	111	48.2 (6.15)	23.3 (9.19)
Group B, MIST, one orthoptist	123	107	47.2 (5.87)	23.6 (10.93)
Group C, MIST, two orthoptists	148	118	46.5 (5.37)	23.8 (11.74)
Group D, MIST, orthoptist/nurse	175	135	47.3 (5.72)	23.0 (8.76)

Group A: Test-retest reliability of the SGS, conducted by one orthoptist

In this group of children, Orthoptist 1 tested the children using the SGS at both sessions. A referral is recommended in cases where the vision for either one or both eyes is classed as a fail, therefore the results are classed as a pass if the child passes with each eye or a fail if the child fails in either eye. The contingency table presents the result for each child, for the test and retest (Table 2).

5

Table 2   Contingency table of the pass/fail results for SGS retest, conducted by one orthoptist

	Test 1: Pass	Fail	Untestable	Totals
Test 2: Pass	92 (83%)	4 (3.6%)	6 (5.5%)	102 (91.9%)
Fail	3 (2.7%)	1 (0.9%)	2 (1.8%)	6 (5.5%)
Untestable	0 (0%)	0 (0%)	3 (2.7%)	3 (2.7%)
Totals	95 (85.6%)	5 (4.5%)	11 (10%)	111

The SGS vision score was categorised as the line of visual acuity achieved, from 6/60 to 6/6. Using Spearman's correlation coefficient of this score, a significant relationship was found between the SGS test and retest results for both the right and left eyes ($\rho = 0.71$, $\rho = 0.79$, respectively, $p = 0.0001$).

Group B: Test-retest reliability of the MIST, conducted by one orthoptist
In this group of children, both testing sessions were performed using the MIST by the same orthoptist (Orthoptist 1). The contingency table presents the pass/fail/untestable results for the MIST test-retest (Table 3).

Table 3   Contingency table of the pass/fail results for MIST retest, conducted by one orthoptist

	Test 1: Pass	Fail	Untestable	Totals
Test 2: Pass	88 (82%)	8 (7.5%)	1 (0.9%)	97 (90.7%)
Fail	2 (1.9%)	2 (1.9%)	0 (0%)	4 (3.7%)
Untestable	2 (1.9%)	0 (0%)	4 (3.7%)	6 (5.6%)
Totals	92 (86%)	10 (9.4%)	5 (4.6%)	107

The MIST is scored as the number of correct letters. Using Spearman's correlation coefficient of the MIST score, a significant relationship was found between the MIST test and retest results for both the right and left eyes ($\rho = 0.79$, $\rho = 0.78$, respectively, $p = 0.0001$).

Group C: Inter-rater reliability of the MIST, conducted by two different orthoptists
In this group of children, both testing sessions were performed using the MIST, but with a different orthoptist at each session. Three orthoptists were involved with the MIST testing procedure in different combinations (Orthoptists 1, 2 & 3).

The contingency table shows the pass/fail/untestable results for the MIST for each child (Table 4).

Table 4   Contingency table of the pass/fail results for MIST retest, conducted by two different orthoptists

	Test 1: Pass	Fail	Untestable	Totals
Test 2: Pass	100 (85%)	6 (5.1%)	4 (3.4%)	110 (93.2%)
Fail	0 (0%)	1 (0.8%)	1 (0.8%)	2 (1.7%)
Untestable	2 (1.7%)	0 (0%)	4 (3.4%)	6 (5.1%)
Totals	102 (86.5%)	7 (5.9%)	9 (7.6%)	118

6

Using Spearman's correlation coefficient of the MIST score, a significant relationship was found between the MIST test and retest results for both the right and left eyes (ρ = 0.60, ρ = 0.74, respectively, p = 0.0001).

Group D: Inter-rater reliability of the MIST, orthoptist and nurse
In this group of children, both testing sessions were performed using the MIST, an orthoptist at one session and a nurse at the other. One orthoptist (Orthoptist 1) and seven nurses were involved with the MIST testing procedure. Seventy-nine of the children (59%) were seen by the orthoptist initially, followed by a nurse. Regardless of tester, 86% of children passed this vision test at the first session, increasing to 93% by the second test. The number of children who could not be tested decreased from 6.8% to 1.5% on retest.

As the aim in this group of children was a comparison of the orthoptist's and the nurses' results with the MIST, analysis was done comparing these two results rather than those of Test 1 against Test 2. The contingency table shows the pass/fail/untestable results for the MIST test-retest sessions, orthoptist compared to nurse (Table 5).

Table 5   Contingency table of the pass/fail results for MIST retest, orthoptist versus nurse

Nurse:	**Orthoptist:** Pass	Fail	Untestable	Totals
Pass	113 (84%)	3 (2.2%)	1 (0.7%)	117 (86.7%)
Fail	6 (4.4%)	2 (1.5%)	0 (0%)	8 (5.9%)
Untestable	6 (4.4%)	3 (2.2%)	1 (0.7%)	10 (7.4%)
Totals	125 (92.6%)	8 (5.9%)	2 (1.5%)	135

Using Spearman's correlation coefficient of the MIST score, a significant relationship was found between the orthoptist and nurse MIST results for both the right and left eyes (ρ = 0.63, ρ = 0.60, respectively, p = 0.0001).

Testing time
The time taken to perform the SGS and the MIST is presented in Table 6.

Table 6   Time taken to perform each test

	Test 1 Mean time (SD) (seconds)	Test 2 Mean time (SD) (seconds)
SGS Orthoptist 1	147 (42.2) N = 129	123 (41.6) N = 109
MIST Orthoptist 1	116 (40.6) N = 120	79 (22.9) N = 102
MIST Nurses	259 (122) N = 58	130 (32.5) N = 69

7

DISCUSSION

Test-retest and inter-rater reliability of SGS and MIST
The first stage of this study was to determine the test-retest reliability of the SGS. If visual acuity was recorded as the line achieved, then it could be seen that SGS test-retest reliability shows a correlation of 0.71 in the right eye and 0.79 in the left eye. This high correlation showed good but not perfect reliability of this test, with some variation in the test-retest results expected for children in this age group. One reason for the test-retest reliability being a little higher in the left than the right eye may be due to a learning effect, as the right eye was always tested first, therefore learning and confidence may increase within the testing session, tending towards a higher correlation for the left eye. This study has shown that the SGS test is a reliable clinical test for the test-retest of visual acuity with children in the 3.5 to 4.5-year age group. These results are difficult to compare with any previous studies as the methods of analysis differ, with the previously reported correlation results varying from 0.16 to 0.54 for vision charts,[11-13] versus 0.82 for a surrounded single test.[15]

The MIST test-retest correlation, when tested by the same orthoptist, was 0.79 for the right eye and 0.78 for the left eye. This result demonstrates a level of reliability similar to that of the SGS, with a slightly higher result in the right eye for the MIST than the SGS. It may be considered that the MIST, being a simpler test than the SGS is less learning dependent within the session than the SGS, and so shows a more equal correlation with each eye. However, this is not particularly shown in the other two MIST test-retest groups. In the context of repeat testing by one eye health care professional, the MIST produces as good a test-retest correlation as the SGS.

With two different orthoptists performing the MIST, the inter-rater reliability values were 0.60 with the right eye and 0.74 with the left eye. This moderate to good correlation, indicates a substantial relationship, though a little less than with one tester. This was achieved with a combination of three orthoptists, in different combinations of first and second tester. The use of three rather than two orthoptists for this assessment may have decreased the inter-rater reliability, but allowed an evaluation within the context of different clinical testers.

The inter-rater relationship between the MIST results gained by an orthoptist and a nurse also produced a moderate to good correlation, with 0.63 in the right eye and 0.60 in the left eye. This was achieved with the use of seven nurses as testers, which may reduce the reliability result, but would provide a more realistic evaluation of the test in the screening context of many different nurses.

Pass/fail results for test-retest of SGS and MIST
The aim of a screening test is to determine those children who may have a vision problem and require further assessment. One factor required for the success of a screening program is that the test has a high compliance rate so that the majority of the children can be successfully graded as pass or fail. The number of children achieving a pass level at the first test was identical for both the MIST and the SGS, at a level of 86% with an orthoptist, increasing to around 92% on retest. This increase in testability of a clinical test from the first to a second session confirms previous reports.[12, 13, 16] When comparing the test completion rates of the nurses to that of the orthoptists, it was found that more children were successfully tested by an orthoptist (93%) than by a nurse (87%).

The number of children who were untestable on the first test with the SGS was 10%, decreasing to 2% on the retest. For the two MIST retest groups, the number of children untestable was around 5% to 7%, with no real change from the first to the second test. This higher rate of compliance at the first test might be explained again by the simpler design of the MIST, with learning having a lesser effect from the first to the second test. However, this is comprised of a higher fail rate at the first session with the MIST than with the SGS. One hypothesis to explain this may be that the simpler design of the MIST means that the test is more in the control of the

8

child being tested than in the skill and experience of the tester. In this case the less cooperative children are more likely to fail in the first session. A similar number of children in both the SGS and the MIST retest groups, 9% and 8.4% respectively, failed or were untestable at the first test and then passed at the second test, showing a learning effect. In a screening context these children would represent false positive results.

One group of interest in a screening program are the children who were untestable at the first test and then determined as a fail at the retest session. As the numbers of these children in each of the groups was small, it was not possible to determine any increased risk factor, but the question would remain after a single testing, whether these children are unable to complete the task due to their cognitive ability, requiring sustained concentration, or to their visual ability. It could also be seen that a small number of children actually passed the first test and then failed or were untestable at the retest. This occurred in both the SGS and MIST retest groups. All of the five children who changed from a pass to a fail on retest were aged 45 months or younger. It must be noted, that in the normal population at this age a small number of children may be expected to be unable to complete the task and some variation in cooperation from one session to another may be expected.

In the orthoptist/nurse MIST retest group there was a small group of children who were passed by one clinician, but were failed or untestable by the other. Three children were passed by a nurse but failed by an orthoptist. Two of these children were passed on their second test, one passed their first test. Six children were passed by the orthoptist, but failed by the nurse, three were failed on their first test and three on the retest. As some were failed at each session, the learning effect cannot be considered the major reason for these discrepant results. These children with conflicting results from an orthoptist and a nurse pose several questions. There were a larger number of children failed by a nurse but passed by an orthoptist, which may be due to the relative lack of experience of the nurses in vision testing of small children.

It can be seen that for testing by an orthoptist, the mean time required to perform the MIST is less than that required to complete the SGS. The nurses required more time to complete the MIST than did the orthoptist, but again less time for the repeat test. This is understandable given the difference between an experienced eye health care professional and the nurses for whom the test had been relatively recently introduced. It can be inferred that the nurses would take much longer to perform the SGS than the MIST, so the MIST is a quicker screening test to perform, an advantage given that the vision screening is performed within the context of a total developmental screening session. This compares favourably with Thorburn and Roland's report[6] that vision testing with single letters by health visitors took 8.1 minutes within the routine developmental check.

The results from this study would confirm those of other studies that both the testability rate and number of children passing a second test is increased.[12, 13, 16] The MIST may appear to be biased towards a higher fail rate, particularly in the first session, but this cost of the false positive results must be balanced by the risk associated with false negative results from a vision screening program. In a screening program a decision must be made whether to refer after one screening test or to retest each child. This study would confirm the recommendation to retest each child who fails or is untestable at the initial vision screening,[12, 13, 16] however, this must be considered in the light of the nurses' workplace, and the fact that the 3.5-year assessment is the last for children within this system. If there is a high level of false referrals, the issue of costs to the public health system, and the financial and emotional costs to parents whose child is referred on for further assessment must be considered. However, this must be balanced by the consideration of the nurses' time and organisation required for a second screening session and the number of children who may not return for a retest, with the risk that they actually have a fail result and so do not gain further follow-up.

CONCLUSION

In conclusion this study has demonstrated that the MIST has a pass rate as high as that of the SGS, with 86% of children passing each at the first test, with the nurses showing a similar rate to the orthoptists. The number of children who were untestable with the MIST was in the order of 5% to 7%, lower than that of the SGS, but this was complemented by an increased fail rate at the first test. The MIST has been demonstrated to be quicker to perform than the standard clinical test of SGS, taking approximately 2.25 to 5 minutes for the nurses to perform on children for the first time, with a mean time of 4.3 minutes.

The test-retest reliability of the MIST demonstrated as strong a relationship as that of the SGS, confirming the reliability of this new vision screening test in comparison to one of the standard clinical tests. In the context of a vision screening program, in the hands of maternal and child health nurses, rather than the clinical context of eye health professionals, the inter-tester correlation is not quite as strong, but a substantial relationship was still found. It may be expected that the orthoptist/nurse inter-rater reliability would improve as the nurses become more experienced with the test.

Part 2: A negative predictive study of the MIST

INTRODUCTION

The incidence of amblyopia in the population is cited as between 1.2% and 5.6% by different studies, with strabismus ranging from 2.7% to 6%. The combined prevalence of amblyopia and strabismus is accepted to be about 5%, high enough for a screening program to be justified.[2, 20, 21] With any screening programs, but particularly for a condition of such low incidence, it is important to determine the negative predictive value (NPV) of the screening test to determine the likelihood of the condition truly being absent.[21] A few studies have reported the NPV of vision screening programs using tests of visual acuity.[21-25] This information is important as the implication parents receive from a child passing a vision screening program is that their child has no vision problems, and so may not receive any further visual surveillance. Visual acuity (VA) is considered to be the most effective test for amblyopia screening, and 3.5 years is considered to be the minimal age suitable to gain good compliance levels for this subjective test.[26]

In order to assess the NPV a gold standard examination with pass/fail criteria must be predetermined in order to consistently assess the visual status of those children who pass the screening test. De Becker et al[21] commented that there was no clear consensus on the definition of normal visual function and that their standard was defined such that children who passed it were not considered to require any further follow-up or treatment. They defined the criteria for failing a gold standard examination for 4.5 to 5.5-year old children. As age has been shown to be a significant factor in VA assessment, it is apparent that some of the criteria needed to be modified for the younger children in the present study.

The proportion of children able to be tested with a letter matching task varies with age.[12, 13, 27] This would support the removal of the 'inability to complete the gold standard examination' as a failure criteria for children of this younger age group. A justification for passing the gold standard could be made on a judgment of those tests completed. The complexity of the task is the next consideration. It is suggested that cognition plays a greater part in crowded acuity than in single optotypes[28] and other studies have reported fewer 3-year old than 4-year old children successfully completing a linear test.[12, 13, 29] The acceptance of occlusion may be another indicator of the compliance level of children for VA testing, with studies showing a greater compliance with patching by 4-year old than 3-year old children.[27, 28, 30]

The level of VA accepted as normal in this age group must be decided. For vision screening, the accepted level of VA in Australia is 6/9.[1] In relating the referral criteria to a gold standard for 3.5 to 4.5-year old children in particular, and accounting for the different testing distances, various levels have been cited in the literature. Some accept 3/6 as the acceptable level for either linear or single optotypes[10, 23, 31, 32] others 3/4.5[11, 20, 30] and even 3/3,[26, 33] referring any child with VA less than these levels. These studies would support the decision to set the gold standard for VA for 3.5 to 4.5-year old children at 3/4.8 using the LogMAR letter matching chart. As a proportion of children were expected not to be able to complete testing with the chart, the acceptable level for single optotypes was set at the stricter level of 3/4. The criteria of acceptable difference in acuity between the eyes of less than two lines[21] would be supported by previous studies, with 96 to 98% of children having no more than one line difference.[28, 30]

A further standard that required consideration for this age group was the level of stereoacuity. Two studies set the strictest screening standards of less than 80 seconds using the Titmus stereo test,[23, 34] one with older children. Earlier studies with Titmus tests reported that only 50% of children younger than 4 years gained stereoacuity better than 200 seconds of arc,[35] and 83% of 4-year olds gained 100 seconds or better.[36] With the Randot E, Simons[37] recommended a screening criteria of 250 seconds, though 168 seconds has been said to be the most sensitive threshold.[10] The referral criteria for two screening programs using the Randot test have the acceptable level of stereoacuity at 200 seconds of arc, referring any children worse than this.[21, 34] There is a reported

11

proportion of children who are unable to be tested with stereopsis tests, even in the presence of binocular vision, varying from 2.1%[38] to 4%.[10] Given these different findings, the gold standard criteria for this study was set at a pass level of better than 200 seconds of arc.

The aim of this study was to examine 200 children who had passed the MIST vision screening by the maternal and child health nurse, giving them a full orthoptic and ophthalmic examination in order to detect any ocular or visual problem that may have been missed by the screening program. Ethics approval was granted by the Human Ethics Committee, La Trobe University and the Department of Human Services Ethics Committee.

METHOD

Participants
Four testing centres from different regions of metropolitan Melbourne were selected – outer eastern, outer southern, inner Melbourne and inner western regions. This selection allowed a representative sample of both children and maternal and child health centres. All children who passed the screening test in these municipalities during the recruitment time were asked whether they were interested in participating in the project, a systematic sampling method. However, as interpreter facilities were not available, there was a necessary restriction to parents who had an appropriate level of english language. If parents expressed an interest, they were given an Information Form, then contacted by the researchers and an appointment was made for the assessment. All testing was performed within a few months of screening.

Procedures
All clinical assessments were performed according to standard procedures and the data sheet protocol. In total, there were four ophthalmologists and five orthoptists who performed the assessments. At all sessions, each was 'blind' to the results of the other.

Prior to the assessment, each parent/guardian read and signed an Informed Consent Form. The first stage of the assessment consisted of the collection of general information. The orthoptic assessment consisted of
- cover test near (1/3 metre) and distance (6 metres)
- four dioptre test near and distance
- prism cover test near and distance
- ocular movement assessment
- convergence near point
- VA, right then left eye, at 3 metres. Acuity was performed initially with LogMAR reversible letter chart and matching card, followed by Sheridan Gardiner Single letters (SGS) if chart acuity was not possible. If LogMAR vision of 3/3.8 was not achieved with either eye, further testing was done with single letters
- fusional vergences, near and distance
- stereoacuity, Randot Stereotest

After the orthoptic examination was complete, Cyclopentolate 1% was instilled in each eye. Each child was seen by the ophthalmologist 30 minutes later. The ophthalmic examination included macroscopic inspection, slit lamp examination, retinoscopy performed at the examiner's standard working distance and indirect ophthalmoscopy.

Gold Standard Criteria
The criteria for failing a gold standard examination in this study of younger children were defined after adaptation of those by De Becker et al.[21]
- VA of 3/4.8 (-3) or less in one or both eyes, (LogMAR <0.24) <u>or</u>
- VA of 3/4 (-2) or less in one or both eyes, SGS
- a difference in VA of two lines or more between eyes (LogMAR difference >0.18)
- stereoacuity of 200 seconds of arc or less

12

- any constant or intermittent heterotropia
- refractive error of
    - 0.75 dioptres or more of myopia
    - 3.50 dioptres or more of hypermetropia
    - 1.50 dioptres or more of astigmatism
    - 1.00 dioptres or more of anisometropia
- any anomaly judged significant enough to require follow-up.

All data from the examinations was coded and entered into a computer statistical program (Statview) for analysis. VA was recorded on the data sheets as the number of letters correctly matched on each line, an interpolated LogMAR score. Single optotype acuity was coded according to the number of letters correctly matched.

RESULTS

Over the recruitment period, Parent Contact Forms were received from the nurses for 296 children, of whom 201 (100 males and 101 females) participated in the examination. The mean age of the children at testing was 45.7 months (3 years, 9.7 months) (SD 2.93 months) with a range from 40 to 55 months. The 201 children were each seen at one of four testing centres, by one of five orthoptists and one of four ophthalmologists. The numbers assessed by each can be seen in Table 7.

Table 7   Numbers of children examined by each tester (N = 201)

Ophthalmologist	Number of children	Orthoptist	Number of children
1	72	1	54
2	45	2	48
3	44	3	43
4	40	4	41
		5	15

Family History
From the family history, 61 of the children (30%) had one or both parents who wore glasses. Seventeen (8.4%) of the children had a relative with a known strabismus and another 14 (6.9%) reported a family member with a 'lazy eye'.

Ocular deviation
Only five of the children were noted to have an exophoria for distance fixation (mean deviation of 5.2 prism dioptres). Thirty-eight children were shown to have some latent deviation for near fixation, an exophoria noted in 35 (mean 6.1 prism dioptres) and an esophoria noted in three (mean 5.3 prism dioptres). One child (#162) was found to have an intermittent small esotropia.

Convergence near point (CNP) was normal in all children, with a mean CNP of 3.9 cm (SD 0.96), ranging from 1 to 7 cm.

Binocular function
All 201 children demonstrated a binocular response with the 4-dioptre prism test and so were considered to show no central scotoma.

An acceptable prism fusion vergence range for near fixation was obtained from all children, with the mean convergence break point being 38.7 prism dioptres (SD 6.46) and the mean divergence 16.1 (SD 2.78), similar to that previously reported.[35] Prism fusion vergence range for distance fixation was obtained on all children except one, mean convergence 16.6 (SD 4.34) and mean divergence 11.8 (SD 2.32).

13

One-hundred-and-sixty-six children (82.5%) obtained a stereoacuity result of at least 70 seconds of arc. The minimum standard defined for this current study was stereoacuity of better than 200 seconds of arc, 95% achieved this. Ten children (5%) did not obtain this level, four obtained 200 seconds of arc and two obtained 400 seconds of arc. This left four children from whom a result was not obtained, due to a lack of comprehension of the test. All but one of these 10 children were demonstrably binocular with good prism fusion vergences. The child previously described (#162) with an intermittent esotropia gained stereoacuity of 400 seconds.

Visual acuity
The preferred test of VA was the 3 metre LogMAR chart with matching letters. If the children were unable to cope with this complex task, the simpler SGS was used. Using the chart, VA of at least 3/3.8 (≥3 letters) (LogMAR ≤ 0.14) was achieved in the right eye by 121 (61%) and in the left eye by 119 (59%) of the children. The gold standard level of VA to be achieved in this study was 3/4.8 (≥3 letters) (LogMAR ≤0.24), and the number of children achieving this level increased significantly to 162 (81%) in the right eye and 160 (80%) in the left eye.

In the right eye, of the 39 children who did not achieve chart acuity of 3/4.8, 33 achieved a single acuity level of 3/3 and six achieved a level of 3/4. This meant that all of the 201 children gained either 3/4.8 linear acuity or 3/4 single optotype acuity in the right eye. In the left eye, of the 41 children who did not achieve 3/4.8, 36 achieved a single optotype acuity level of 3/3 and four achieved a level of 3/4. This meant that 200 of the 201 children gained either 3/4.8 linear acuity or 3/4 single acuity in the left eye.

One child did not achieve a satisfactory level of VA. This child (#115) achieved an unequal VA result of 3/4.8 (2) (LogMAR 0.26) in the right eye and 3/9.5 (3) (LogMAR 0.54) in the left eye with the linear test, and 3/3 (1) and 3/6 being achieved in the right and left eyes respectively with the SGS.

Visual acuity difference
Of the children achieving a result with the chart, only two children had an acuity difference of two lines or greater. One child (#115) described previously had a difference of two lines and one letter on the chart, and seven letters on single acuity. Another child had a difference of two lines on the chart, but achieved equal VA of 3/3 with the single optotypes, so was considered to have passed the gold standard using the single letters.

Refractive error
The refractive error was recorded as the retinoscopy result calculated for the working distance, using minus cylinders, with no calculation being made for the effect of the cyclopegia. For the 201 children the mean spherical error was +1.18 dioptres (SD 0.703) in the right eye, ranging from 0.00 to +3.25 and +1.17 dioptres (SD 0.685), ranging from 0.00 to +3.00 in the left. The mean astigmatic error was 0.17 dioptres (SD 0.286) in the right eye and 0.16 dioptres (SD 0.263) in the left eye, both ranging from 0.00 to 1.5 dioptres. All children passed the gold standard for spherical equivalent refractive error. Four children failed the gold standard criteria, with an astigmatic error of 1.50 dioptres in one eye (#6, #19, #115, #158).

The mean anisometropic difference between the two eyes was 0.07 dioptres (SD 0.119), ranging from 0.00 to 0.75 dioptres. All children passed the gold standard for anisometropia.

Ophthalmic assessment
On macroscopic inspection, only two children were reported to have an ocular anomaly, one with a small right lid chalazion, the other with a slight right ptosis. On slit lamp examination three children were noted to have some observable differences. One child had a small area of focal endothelial changes in the peripheral cornea, two other children were noted to have epiblepharon, one mild and one minimal. On dilated ophthalmoscopy, two children were noted to have physiological variations, one with physiological cupping of the discs and another with buried

optic nerve head drusen. None of these conditions were considered clinically significant, nor to require any further follow-up.

Failure of the gold standard criteria
In this age group it is expected that a small number of children will be unable to complete the full gold standard assessment due to factors such as the level of cognitive development and ability to maintain concentration on the tasks for the length of time required.[12, 13, 27] It is therefore necessary to interpret the results to the gold standard assessment in this light, with the incompletion of any test viewed in association with the other results obtained.

Only five children (2.5%) were unable to complete the examination, each missing one test. For each of these children confirmation was available from all other tests, both objective and subjective, that ocular status and visual function were normal. Measurement of distance prism fusion vergence was not obtained on one child, but the child was orthophoric for both near and distance fixation, with near prism vergence, stereopsis and VA all normal. Measurement of stereoacuity was not obtained on four children, all of whom were orthophoric with normal prism vergences, VA and refraction, but were noted to have either a poor understanding of english or no comprehension of the test. All of these children were considered to have passed the gold standard examination.

With stereoacuity, six children did not achieve the pass level of 140 seconds of arc. If these children are included with those unable to do the test, then this meant that 5% of the children were not able to complete testing of stereoacuity, similar to that of previous studies.[10, 38] Of these six children, five were orthophoric, with normal VA, prism vergences and refraction.

On reviewing the results to each test, five children were considered to have failed the gold standard assessment. The details of these children are shown in Table 8.

15

Table 8   Children who failed the gold standard assessment

Particpant	Visual acuity	Binocular status	Refraction
#6	*linear* R 3/4.8 L 3/3.8	orthophoric good prism vergences stereoacuity 50″	*R +1.00/-1.50 x 180° L +1.00/-0.75 x 180°
#19	*linear* R 3/6 L no cooperation *single* R 3/3 L 3/3 (-1)	orthophoric good prism vergences stereoacuity 40″	*R +2.50/-1.50 x 170° L +2.25/-1.25 x 20°
#115	*linear* R 3/4.8 (3) *L 3/9.5 (2) *single* R 3/3 (1) *L 3/6 (1)	orthophoric good prism vergences stereoacuity 70″	R +2.25/-1.25 x 90° *L +2.25/-1.50 x 90°
#158	*linear* R 3/6 L 3/6 (-1) *single* R 3/3 L 3/3 (-1)	orthophoric good prism vergences stereoacuity 70″	R +1.75/-1.25 x 180° *L +1.75/-1.50 x 180°
#162	*linear* R 3/7.5 L 3/6 *single* R 3/3 (2) L 3/3 (2)	*intermittent RET good prism vergences *stereoacuity 400″	R +2.75 L +2.75

* failed criterion

One child (#162) was found to have a small intermittent esotropia, evident only after continued dissociation and a stereoacuity of 400 seconds of arc, consequently failing the gold standard on two criteria. Results to other tests were normal, with good prism vergences, VA and only a small hypermetropic spherical refractive error, all indicating good control of the deviation, so this would not be considered a significant concern. On follow-up assessment at a later date that there was no sign of the esotropia.

With VA testing only one child failed the gold standard assessment. This child (#115) failed the VA level in the left eye with both the chart and single optotypes, had a VA difference of greater than two lines and an astigmatic refractive error of 1.50 dioptres in the left eye, so failed the gold standard on three criteria.

Three other children failed the gold standard for the criterion of astigmatic refractive error (#6, #19, #158). Each of these passed the VA standard, with only one line or less difference.

Negative predictive value
Of the 201 children tested who had all passed the vision screening, five failed the gold standard assessment. This resulted in a negative predictive value of 97.5%, with 95% confidence limits of 95.4% to 99.7%, for the MIST vision screening program.

DISCUSSION

The NPV of the MIST vision screening program of 97.5%, with 95% confidence limits of 95.4% to 99.7%, is similar to previous values reported. De Becker at al[21] with a gold standard assessment of 157 children, reported a value of 93% for the detection of any ocular problem, but a value of 98.7% (95% confidence limits of 95.4% to 99.85%) for amblyopia, strabismus and high refractive errors, which are the usual target conditions of a vision screening program. The Enhanced Vision Screening Program[21] has referral criteria of VA less than 6/9 or stereoacuity less than 200 seconds of arc. Kennedy et al,[23] with a gold standard assessment of 241 children, reported a value of 94% for their screening program, though it must be noted that the VA referral criteria was less than 6/12, which may be considered rather low for kindergarten children and so may assist in explaining the lower value. Enzenauer[22] reported a NPV of 97% for public health screeners over a wider age group. Marsh-Tootle,[24] using single optotype E-test and De Becker et al's[21] criteria, reported a NPV of 77.8%. The children in this study were younger than De Becker's, which may explain the lower NPV. None of the false negatives were considered to require treatment, merely increased surveillance. Newman,[25] in a retrospective study testing linear acuity two years after single optotype orthoptic preschool screening reported a NPV of 99.6% (95% CI, 98.7% to 99.9%) for the acuity test alone, but increased to 100% when other tests of ocular alignment were included.

De Becker et al[21] commented upon the NPV for vision threatening conditions, calculating this at 97.6%, with four of their children classed as having potentially vision threatening disorders and the other seven with minor ocular problems. In the present study, one of the five children would be classed as having amblyopia on the first assessment, with a VA difference of more than two lines. Another had an intermittent strabismus, with normal visual function. The remaining three children with an astigmatic refractive error all had VA results within the normal range for either the chart or single letter test, with a VA difference of no more than one line. The astigmatic error was only 1.50 dioptres in each case, so failing the gold standard according to De Becker et al's criteria.[21] However one other study set this level at greater than 1.50 dioptres, in which case these three children would have passed, giving a NPV of 99%.

The level of VA achieved by the children in this study was comparable to that reported in the literature. All of the 201 children were able to perform the letter matching task, a result similar to other studies.[13, 27] If the complexity of the task was considered, it was found that 80% of the children in this study could satisfactorily complete the linear test, similar to other reports.[12, 13] Considering the level of VA, 3/3.8 was achieved by 60% of the children on the LogMAR letter matching chart, with 81% achieving 3/4.8, results similar to those reported using the SSAS linear test.[28]

Methodologically, one bias introduced by the use of volunteer participants is that many parents participated as there was a known family history of ocular problems and so wanted a vision assessment. This may have meant that the sample contained a higher incidence of children with a family history.

The examiners were aware that each of these children had passed the screening test, so some examiner bias towards expecting normal results may have been possible, as discussed by De Becker et al.[21] As the aim, as emphasised in the assessment protocol, was to determine whether any visual problems had been undetected by the screening, and so perform a gold standard examination to determine this, it was felt that this was not a concern. The gold standard assessment, which involved a combination of objective and subjective tests and two different examiners with measurements of each aspect of ocular and visual function, would be expected to detect any problem, regardless of bias. The issue of the use of different testers for the assessment of each of these children must be considered. The importance of the performance of all tests according to standard procedures and the data sheet protocol was discussed prior to the commencement of testing in order to minimise any effects. The comparison of each child's test

results against the gold standard criteria was done at a later stage by the principal researcher, not by the clinicians.

CONCLUSION

This study has found that the negative predictive value of the MIST vision screening program was 97.5%, a similar value to those previously reported.[21-23] This means that there is a risk that 2.5% of children who pass the MIST 3.5-year old vision screening program may actually have either amblyopia, strabismus or refractive error. Any screening program must maintain a balance between the positive and negative predictive values. It must be noted that any attempt to increase the negative predictive value of the MIST would mean a higher incidence of false referrals.

The importance of this finding must be part of any vision screening protocol and the information provided to parents. It must be emphasised that even though a child has passed the screening test, there is a still a very low risk that a potentially vision threatening problem may have been missed. Both screeners and parents must be aware of this when vision screening is performed.

Part 3: A positive predictive study of the MIST

INTRODUCTION

The positive predictive value (PPV) is a measure of the proportion of the population which fails the screening test and is actually found to have the disorder confirmed on diagnostic testing,[34] the true referrals. It must be noted that the prevalence of a condition has an effect on the false positive rate. A condition with low prevalence, even with high sensitivity and specificity, may still have a relatively large number of false positives.[39]

It has been suggested that screening programs conducted by orthoptists, or experienced eye health care professionals, have higher sensitivity and specificity than those by other personnel.[40, 41] Orthoptist-based screening programs have reported a PPV from 74% to 94%,[40, 42, 43] with nurses or other health personnel from 25% to 69%, but generally around 40% to 50%.[21, 22, 24, 34, 44-49] Some studies compare referral rates, reporting orthoptists as referring less children, 4.2% to 5.1%,[40, 49] compared to 6.5% to 8.6% by nurses.[34, 49] The total efficiency of a screening program needs to be considered. Though orthoptists may provide more accurate referrals from vision screening, this needs to be offset by the advantage of the more general developmental assessment that can be provided by a community based nurse or health worker.

The efficacy of a screening program is dependent upon the attendance rate.[50] The uptake for preschool vision screening programs has been reported, varying from 20% to 99.5%.[25, 40, 44, 46, 51, 52] The MIST program generally attracts around 50% of the children in this age group.[53] The attendance rate at school entry screening would be a much higher level than this. However, the major concern with the detection of amblyopia is that vision screening at school entry may be too late for effective treatment to be implemented, particularly now that children may be turning six or seven years of age at the time of school screening.[51] A further issue to consider with any vision screening program is the follow-up by parents when a child fails the screening test, the need to obtain a specialist ocular assessment. Follow-up rates have been reported from 61% to 65%.[45, 54, 55]

Over-referral is a problem for the acceptance of any screening program. If it is noted that there are too many children referred from a vision screening who do not have a problem, then this may be perceived by both the parents and the eye health professionals that the test is worthless.[56] Over-referral results in a cost to either the individual or the community, though this must be balanced against the cost of undetected visual problems if there was no screening program. It must also be noted that parents are good observers of their children and 'are rarely wrong when they think that there is a problem with their child's eyes or vision'.[56] In this case any child who passes a vision screening test, but their parents are concerned, is likely to be referred on for full assessment. This is compounded by the tendency to refer if there is any doubt about the child's visual function, as a screening test is not able to exclude all visual problems.

The problem with determining the PPV of any program is that the diagnosis of true or false referral is generally not set against a gold standard but is reliant upon clinical reports. These may be received from any number of eye health professionals whom the children have attended after referral from the vision screening program. This is particularly noted by De Becker et al[21] in their negative predictive study of a vision screening program. When a sample of children who had failed the screening were given a gold standard ocular examination, the positive predictive value was found to be much less than what was defined by clinical reports, reducing to 50% from an originally cited 72%.[34] The criteria for classifying results into true or false positive outcomes needs to be set independently of the clinicians, as clinical bias will result in different interpretations of the screening outcome.

19

METHOD

This stage of the evaluation project aimed to analyse the results of the ocular examination of all children who were referred to an eye health professional following testing with the MIST, to assess the level of true or false positive results. Ethics approval was granted by the Human Ethics Committee, La Trobe University and the Department of Human Services Ethics Committee.

Procedures

The MIST kits were supplied to all Maternal and Child Health Centres in Victoria. During the training sessions the evaluation projects were explained to the nurses and their assistance requested. At the time of implementation of the program, information was sent to ophthalmologists, optometrists and orthoptists advising them of the program and requesting assistance in the return of the forms. A triplicate copy pad of referral forms was provided with the MIST kits.

For any child who failed the screening test the nurse completed the referral form, including the child's identifying details, the nurse's details and the MIST score. A section was available for the addition of any further comments or observations of appearance or visual behaviour. One copy was returned by the nurse to the Project Officer. Two copies were then given to the parent. The referral process included the explanation to parents of the MIST results, and the need to obtain a full ocular examination to determine whether this was a true result and if so, the reasons for the reduced vision.

The next stage of the process for the parent and child was to obtain an appointment for an eye examination. Upon completion, the clinician was requested to complete Section 2 of the form and return one copy to the Project Officer. This included the clinician's details, the assessment results, diagnosis and any action taken. Parents gave signed consent for this information to be passed on.

The researchers then matched the nurses' referrals with the returned specialist forms. As copies of the referral form were expected to be received from nurses for all children who had failed the MIST during the 3-year time period, this was the basis of the sample of children. If no specialist form had been received, those parents who had given consent were contacted. For those parents who could be contacted, the project was explained, the parents were reminded about the screening test, and were asked whether they had yet attended a vision assessment. If the clinician's details were known, then a form was sent requesting the results. If the clinician's details were not recalled, then an attempt was made to assess the outcome of the examination by asking the parents about the result and this information was added to the form, annotated as parent information.

All data from each child's MIST and visual assessment was coded and entered into a computer statistical program (Statview) for analysis. The return referral form requested measurements of VA; a yes/no response for the diagnoses of amblyopia, refractive error, strabismus and other; a yes/no response for treatment, no treatment and review. The classification of a true or false referral was then determined by the Project Officer.

As VA of 6/9 is considered a pass for children of this age group,[1] the criterion of 6/12 or worse was set as a true failure of the screening test. Any child diagnosed with a strabismus, or a VA difference of two lines or more[21, 28, 30] was also considered to be a true failure. VA of 6/9 or better was classified as a false referral. Some children had VA recorded at a level between 6/9 and 6/12, usually due to such factors as the test optotypes available and the testing distance. These children were recorded as a pass result, as their acuity was better than 6/12.

There were a number of children for whom VA was not recorded, mostly because a reliable result was not achieved. The other information on the form was then used to classify the result for these

children, including the diagnosis according to the clinician and whether any treatment or review was recommended.

RESULTS

Over a 3-year period, referral forms were returned for 3,854 children (2,015 males and 1,839 females). The mean age of the children was 44.3 months (SD 3.17), range 36 to 60 months. Of these children, 98% of parents or guardians gave signed consent for the clinical information to be returned. Eye health professionals returned 2,264 (59%) of these forms without any requests or reminders required.

Attempts were made to contact the remaining 1,590 children. Contact was made with 793 of these parents. Of the parents who were contacted, 535 were able to provide the specialist's name. Requests were then sent to the clinicians for information, of which 367 (69%) were returned. Fifty-nine parents reported that their child had been seen, but were unable to recall the details of the clinician. It was found that 199 parents had not followed up the failed vision screening. This means that of all the children whose visual assessment results were known and reported by either the parent or the clinician, a total of 3,057 children, 199 (6.5%) were known not to have continued on for an eye examination. This is a measure of the known response of the number of children who were not taken for a full ocular assessment. The outcome is unknown for 797 children who were unable to be contacted, or who had not given their consent for further contact. If the unlikely presumption was made that every one of these children did not follow up with an eye examination, then the maximum rate of non-attendance following vision screening failure would be 996 of the 3,854 children (26%).

The proportion of the population who attend the 3.5-year assessment varies around 49% to 52% each year. Over the 3-year period, the referral rate of children who failed the vision screening by the nurses was 6.5%.[53]

Reasons for referral after vision screening

The vision screening protocol includes vision testing with the MIST and an observation of the ocular structures, including corneal reflections. Most of the children were referred because they had failed the MIST (3,395 children, 88%) but a number were referred for other reasons. A number of children actually passed the MIST, but were still referred (235 children, 6.1%). Reasons were provided for some of these; behaviour whilst performing the test (45%), strabismus (11%), family history (10%), parent request (3.5%), pathology/observation (1.5%). The remaining 224 children (5.8%) were those not able to complete the MIST, with similar reasons given for referral.

The MIST referral procedure recommends an eye examination by an optometrist, ophthalmologist or orthoptist, either privately or through the public health system. The majority of children attended an optometrist (78%). Others attended an ophthalmologist (14%), an orthoptist (4%) or both (4%). The mean time lapse between the vision screening test and the eye examination was 6.3 weeks (SD 9.97), ranging from one to 135 weeks. Most children (88%) were assessed within three months.

Analysis of referral outcomes

Collation of the returned referral forms allowed analysis of the level of true or false referrals arising from the MIST visual surveillance program. There were 2,623 forms returned from the specialists, providing the clinical information from the examination. Each result was graded into one of seven categories for classification as true or false referrals (see Table 9).

21

Table 9   Classification of outcome of ocular examination

Category	Ocular outcome	True or false referral
Reduced VA or strabismus	VA of 6/12 or worse, two lines or more difference, or a strabismus	A **true** referral
VA 6/9, given treatment	VA of 6/9 or better (pass result), but given glasses or some form of treatment	A borderline **true** referral, actually passed the VA criterion, but given treatment
No VA, given treatment	No VA established, but given glasses or some form of treatment	A borderline **true** referral, no VA established, usually due to poor cooperation, but given treatment
Reduced VA, no problem	VA of 6/12 or worse, but with no refractive error or other reason detected, probably developmental	A borderline referral, likely to be **false**, but unconfirmed
No VA, no problem	No VA established, but no refractive error or other problems detected	A borderline referral, likely to be **false** but unconfirmed
Uncooperative, unconfirmed	Uncooperative for ocular assessment, no confirmation possible	A borderline referral, likely to be **false** but unconfirmed
Passed all tests	Passed all tests VA of 6/9 or better than 6/12, no strabismus	A **false** referral

The results have been analysed in total, but are also considered in the light of whether the child actually failed the MIST or was referred for some other reason. It can be seen that the confirmed true positive rate was 42.3% of all the children referred (see Table 10). There was a further 2.8% who may be true referrals, those who received some form of treatment, but with a VA of 6/9 or with no VA obtained. Another finding of interest was the 3.2% of children for whom testing was a difficulty even with the clinician, those for whom a VA was not established. The incidence of children diagnosed with strabismus was 4.2%.

Table 10   Referral outcomes for all children with an ocular report (N = 2,623)

Category	N	%	Referral outcome	N	%
Reduced VA or strabismus	1,110	42.3%	True referral	1,110	42.3%
VA 6/9, given treatment	59	2.3%	Borderline true		
No VA, given treatment	12	0.5%		71	2.8%
Reduced VA, no problem	194	7.4%	Borderline false		
No VA, no problem	55	2.1%			
Uncoop, unconfirmed	17	0.6%		266	10.1%
Passed all tests	1176	44.8%	False referral	1176	44.8%

For the 2,623 children with an ocular report, the positive predictive value (PPV) calculated for those children who had a confirmed reduced VA or strabismus was 42.3% (see Table 11). If the PPV was calculated only for those children who failed the MIST, a slightly higher value of 44.4% was found. If these children were combined with those who did not complete the test, which would mean the inclusion of all children who did not successfully complete the MIST, then the PPV was also a slightly higher value. It can be seen that the PPV calculated for those children who were referred even though passing the MIST was significantly lower than for any other children.

Table 11   Positive predictive value (PPV) (reduced VA or strabismus)

	Confirmed true positive	False positive or unconfirmed result	Total number of children	Positive predictive value	95% Confidence limits
All children	1,110	1,513	2,623	42.3%	40.4% – 44.2%
Failed MIST	1,032	1,294	2,326	44.4%	42.4% – 46.4%
Did not complete MIST	50	92	142	35.2%	27.4% – 43.0%
Failed or did not complete MIST	1,082	1,386	2,468	43.8%	41.9% – 45.8%
Passed MIST	28	127	155	18.1%	12.0% – 24.1%

Visual acuity

The MIST is designed as a pass-fail test only, not as a measure of threshold VA. In order to calculate the level of VA defects that have been detected by the screening test, the level of the worst eye for each child was noted. These results are shown in Table 12.

Table 12	Level of visual acuity recorded by clinicians (N = 2,537)	
Visual acuity	Number of children	Percentage (%)
Better than 6/12	1,267	49.9%
6/12 to < 6/18	682	26.9%
6/18 to < 6/36	432	17.1%
6/36 or less	156	6.1%

Treatment outcomes

According to the returned clinicians' forms, 30.8% of the children were prescribed some form of treatment, with another 47.9% receiving no treatment, but requested to return for review at a later stage.

DISCUSSION

This study of the PPV of the MIST vision screening program is based on a 3-year period from April 1998 to March 2001, with 3,854 referral forms returned. This included approximately 65% of the children referred by the maternal and child health nurses according to Department of Human Services, figures.[53] As this was an extra task requested it might be expected that there would not be a complete return rate, complicated by the large number of nurses, more than 550, who are distributed over more than 800 centres throughout 78 municipalities. The 6.5% referral rate by the nurses is within the expected range for the detection of amblyopia, which has an accepted incidence of less than 5%. Other nurse based screening programs have reported similar referral levels.[34, 49]

A number of children (6%) were referred even though they had passed the MIST, another 6% because they had not been able to complete the test. The major reason cited for referral in both these groups was the child's behaviour while doing the test, with strabismus being suspected in 11% of the children who passed. The proportion of children who were untestable was consistent with other studies.[13] It has been recommended that repeat testing reduces the rate of false referrals, and that all children should be retested prior to referral.[11-13, 40, 54] This recommendation may have been confirmed by the reduced true positive rate in this group of children, however this must be balanced against the management and time constraints of the nurses in the screening situation. This group of children also had a higher rate of non-completion of the testing procedures reported by the eye health professionals.

Of interest were those parents who did not follow up the vision screening failure with a full ocular assessment for their child. The design of this study did not allow this figure to be truly determined, but could only be calculated for those parents who were contactable. The rate of known non-follow-up was 6.5%, but it may have been greater if all children could have been traced. This figure is lower than that reported by other studies, which varied from 35% to 9%.[54, 55, 57] The follow-up time for the children in this study was also good, with more than half the children being seen within one month and the majority (88%) within three months of the vision screening.

Other authors have outlined the major reasons for not following up on the screening results.[57] The screening process issues have been stated as lack of confidence in the process, lack of information about the screening, and lack of communication. Other factors include confusion regarding the choice of eye professional, difficulty in scheduling appointments, and the cost of

assessment and glasses, if these are required. One criteria for a successful screening program includes effective, available treatment.[9, 34, 58] It appears from this current study that most of these factors must be adequate in the maternal and child health program and in the availability of eye health professionals for assessment. The fact that the majority of children attended an optometrist is understandable given the health care system, where optometric assessments are bulk-billed and require no medical referral. Some parents who had not followed up the screening test said they would do so when the child was ready for school, or when the parent next had their appointment. There may be some need to further emphasise the importance of the timing of assessment and treatment for amblyopia, though in general this appeared to have been done.

The Positive Predictive Value
The PPV of the MIST as a vision screening tool for those children who failed the test was 44.4%, with 95% confidence limits of 42.4% to 46.4%. This value is similar to that of other studies which have nurses or other health professionals as the primary screening personnel, ranging from 25% to 69%,[21, 24, 34, 44-49] with several being in the range of 40% to 50%.[21, 44, 45, 47, 49] A comparison with other studies is difficult as not all were based on a defined standard of VA, acuity difference or strabismus, but were based on the clinicians' reports of whether a problem was diagnosed. This allows clinical bias to alter the criteria of definition of a 'problem' and so may vary the rate of true or false outcomes. However, it can be seen that even with the stricter criteria, the PPV of the MIST is within the range of that cited by other screening programs.

Previous studies have shown that experienced eye health professionals may provide more accurate referrals from vision screening,[7, 8] but this must be balanced against what may be seen as an ineffective screening program that relies on specialists providing screening services. This would mean that screening personnel would cost more and would involve only one aspect of the child's general development, so would require multiple screening programs. The implementation of the MIST as the vision screening tool by maternal and child health nurses means that the vision test is part of a general assessment and so the results can be viewed in the context of the child's whole development. As part of the 3.5-year assessment the MIST does not add to the cost of the process, except for a slight increase in the time required.

A PPV of 44.4% means an over-referral rate of approximately one in two children being referred unnecessarily. De Becker et al[21] stated that this may be considered acceptable with a condition of such low prevalence as amblyopia. This study has shown the MIST to have an acceptable PPV, with general acceptance by the nurses and a good follow-up rate by parents, all contributing to the community awareness. In terms of the economic cost to the community, there is minimal cost involved for the nurses' assessment, but the costs of over-referral will be distributed to the general health care cost. The emotional costs to parents brought about by a failure on the vision screening test and the necessity to attend for an ocular examination also needs to be considered. Parents need to understand the importance of following through with the process, without raising unnecessary anxiety.

The level of VA documented by the clinical assessment was of interest. Of the 2,537 children with a recorded visual acuity, 6% had 6/36 vision or less in at least one eye, another 17% had less than 6/18, significantly reduced levels of VA, which were detected by the vision screening program.

CONCLUSION

This study has found that the PPV of the MIST vision screening test was 44.4%, with 95% confidence limits of 42.4% to 46.4%, comparable to other vision screening programs which have other non-eye care professionals as the primary screening personnel.[21, 24, 34, 44-49] This value means that for each child referred correctly, there is another child who may have been referred unnecessarily. However, given the less than 5% prevalence of amblyopia in children of this age group, this over-referral rate would be considered acceptable.[21]

The negative predictive value was found to be 97.5%. In any vision screening program, a balance must be achieved between the positive and negative predictive values, between the likelihood of a false referral and the importance of obtaining a full examination to determine whether there is a problem.

In the MIST vision screening program it could be recommended that all children who fail the test should be returned for a repeat test before being referred. This may result in an increased positive predictive value, but must be balanced against the constraints of the maternal and child health nurses' working environment and the chance that some children may be lost from the system if they do not return for the repeat test.

Currently there appears to be a very good follow-up rate after failure of the vision screening test, demonstrating that there appears to be good parental acceptance of the worth of the program, with parents realising the importance of obtaining a full eye examination.

ACKNOWLEDGEMENTS

REFERENCES

1. Brown S, Story I. A new approach to visual acuity screening for pre-school children. Ophthalmic Epidemiology 1998;5(1):21-27.
2. Snowdon S, Stewart-Brown S. Preschool vision screening. Health Technology Assessment 1997;1(8).
3. Williams C, Harrad R, Sparrow J, Harvey I, Golding J. Conclusions for or against services are invalid without appropriate research evidence (letter). British Medical Journal 1998;316(21):937.
4. Rahi J, Dezateux C. The future of preschool vision screening services in Britain (editorial). British Medical Journal 1997;315(15):1247-1248.
5. Hall D, Stewart-Brown S. Screening in child health. British Medical Bulletin 1998;54(4):929-943.
6. Thorburn R, Roland M. The effectiveness of preschool vision screening. British Journal of Community Nursing 2000;5(1):41-44.
7. Bolger P, Stewart-Brown S, Newcombe E, Starbuck A. Vision screening in pre-school children: comparison of orthoptists and clinical medical officers as primary screeners. British Medical Journal 1991;303(23):1291-1294.
8. Bray L, Clarke M, Jarvis S, Francis P, Colver A. Preschool vision screening: a prospective comparative evaluation. Eye 1996;10:714-718.
9. Simons K. Preschool vision screening: rationale, methodology and outcome. Survey of Ophthalmology 1996;41(1):3-30.
10. Fern K, Manny R, Garza R. Screening for anisometropia in preschool children. Optometry and Vision Science 1998;75(6):407-423.
11. Friendly D. Preschool visual acuity screening tests. Transactions of American Ophthalmological Society 1978;76:383-480.
12. Hered R, Murphy S, Clancy M. Comparison of the HOTV and Lea Symbols charts for preschool vision screening. Journal of Pediatric Ophthalmology & Strabismus 1997;34(1):24-28.
13. Sprague J, Stock L, Connett J, Bromberg J. Study of chart designs and optotypes for preschool vision screening. American Orthoptic Journal 1988;38:18-23.
14. Merritt J, Game S, Williams O, Blake D. Visual acuity in preschool children: the Chapel Hill-Durham Day-care vision study. Journal of the National Medical Association 1996;88(11):709-712.
15. Holmes J, Beck R, Repka M, Leske D, Kraker R, Blair R, et al. The amblyopia treatment study visual acuity testing protocol. Archives of Ophthalmology 2001;119(9):1345-1353.
16. Kohler L, Stigmar G. Vision screening of four-year-old children. Acta Paediatrica Scandinavica 1973;62:17-27.
17. McGraw P, Winn B, Gray L, Elliott D. Improving the reliability of visual acuity measures in young children. Ophthalmic and Physiological Optics 2000;20(3):173-184.
18. Lovie-Kitchen J. Validity and reliability of visual acuity measurements. Ophthalmic and Physiological Optics 1988;8(Oct):363-370.
19. Siderov J, Tiu A. Variability of measurements of visual acuity in a large eye clinic. Acta Ophthalmologica Scandinavica 1999;77:673-676.
20. Ehrlich M, Reinecke R, Simons K. Preschool vision screening for amblyopia and strabismus. Programs, methods, guidelines, 1983. Survey of Ophthalmology 1983;28(3):145-163.
21. De Becker I, MacPherson H, LaRoche G, Braunstein J, Cottle R, McIntyre L, et al. Negative predictive value of a population-based preschool vision screening program. Ophthalmology 1992;99(6):998-1003.
22. Enzenauer R, Freeman H, Larson M, Williams T. Photoscreening for amblyogenic factors by public health personnel: the Eyecor Camera System. Ophthalmic Epidemiology 2000;7(1):1-12.
23. Kennedy R, Sheps S, Bagaric D. Field trial of the Otago photoscreener. Canadian Journal of Ophthalmology 1995;30(4):193-197.

24. Marsh-Tootle W, Corliss D, Alvarez S, Clore K, Daum K, Gordon A, et al. A statistical analysis of modifed clinical technique vision screening of preschoolers by optometry students. Optometry and Vision Sciences 1994;71(10):593-603.

25. Newman D, East M. Preschool vision screening: negative predictive value for amblyopia. British Journal of Ophthalmology 1999;83(6):676-679.

26. Beardsell R. Orthoptic screening at 31/2 years by Huntington Health Authority. British Orthoptic Journal 1989;46:7-13.

27. Egan D, Brown R. Vision testing of young children in the age range 18 months to 4 1/2 years. Child: care, health and development 1984;10:381-390.

28. Salt A, Sonksen P, Wade A, Jayatunga R. The maturation of linear acuity and compliance with the Sonksen-Silver acuity system in young children. Developmental Medicine and Child Neurology 1995;37:505-514.

29. Whitton M. Visual acuity testing in pre-school aged children - what can be expected? Australian Orthoptic Journal 1998;33:55-59.

30. Sonksen P. The assessment of vision in the preschool child. Archives of Disease in Childhood 1993;68(4):513-516.

31. Ingram R, Holland W, Walker C, Wilson J, Arnold P, Dally S. Screening for visual defects in preschool children. British Journal of Ophthalmology 1986;70:16-21.

32. American Academy of Pediatrics. AAP issues guidelines for vision screening in infants, children and young adults. American Family Physician 1997;55(1):346-349.

33. Dholakia S. The application of a comprehensive visual screening programme to children aged 3-5 years. Can a modified procedure be devised for visual screening by ancillary staff? Ophthalmic and Physiological Optics 1987;7(4):469-476.

34. Macpherson H, Braunstein J, La Roche G. Utilizing basic screening principles in the design and evaluation of vision screening programs. American Orthoptic Journal 1991;41:110-121.

35. McKenzie L. Lincoln Institute kindergarten screening. Australian Orthoptic Journal 1977;15:29-31.

36. Brown S, Jones D. A survey of the incidence of defective vision and strabismus in kindergarten children - Sydney 1976. Australian Orthoptic Journal 1977;15:24-28.

37. Simons K. A comparison of the Frisby, Random-Dot E, TNO, and Randot Circles stereotests in screening and office use. Archives of Ophthalmology 1981;99(March):446-452.

38. Heron G, Dholakia S, Collins D, McLaughlan H. Stereoscopic threshold in children and adults. American Journal of Optometry and Physiological Optics 1985;62(8):505-515.

39. Bailey R. Assessing the predictive ability of the test-positive findings of an elementary school vision screening. Optometry and Vision Science 1998;75(9):682-691.

40. Newman D, Hitchcock A, McCarthy H, Keast-Butler J, Moore A. Preschool vision screening: outcome of children referred to the hospital eye service. British Journal of Ophthalmology 1996;80:1077-1082.

41. Newman D, East M. Prevalence of amblyopia among defaulters of vision screening. Ophthalmic Epidemiology 2000;7(1):67-71.

42. Jarvis S, Tamhne R, Thompson L, Francis P, Colver A. Preschool vision screening. Archives of Diseases in Childhood 1990;65:288-294.

43. Wormald R. Preschool vision screening in Cornwall: performance indicators of community orthoptists. Archives of Disease in Childhood 1991;66:917-920.

44. Allen J, Bose B. An audit of preschool vision screening. Archives of Disease in Childhood 1992;67:1292-1293.

45. Juttman R. The Rotterdam Amblyopia Screening Effectiveness Study (RAMSES): compliance and predictive values in the first year. British Journal of Ophthalmology 2001;85:1332-1335.

46. Kvarnstrom G, Jakobsson P, Lennerstrand G. Screening for visual and ocular disorders in children, evaluation of the system in Sweden. Acta Paediatrica 1998;87:1173-1179.

47. Poterio M, Cardillo J, De Senne F, Pelegrino R, Jose N, Norato D, et al. The feasibility of introducing a visual screening test for children during vaccination campaigns. Journal of Pediatric Ophthalmology & Strabismus 2000;37(2):68-72.

48. Robinson B, Bobier W, Martin E, Bryant L. Measurement of the validity of a preschool vision screening program. American Journal of Public Health 1999;89(2):193-198.

49. Spowart K, Simmers A, Tappin D. Vision testing in schools: an evaluation of personnel, tests and premises. Journal of Medical Screening 1998;5:131-132.

50. Moseley M. Preschool vision screening. British Journal of Ophthalmology 2003;87(8):931.

51. Donahue S, Johnson T, Leonard-Martin T. Screening for amblyogenic factors using a volunteer lay network and the MTI photoscreener. Ophthalmology 2000;107(9):1637-1656.

52. Williamson T, Andrews R, Dutton G, Murray G, Graham N. Assessment of an inner city visual screening programme for preschool children. British Journal of Ophthalmology 1995;79:1068-1073.

53. Department of Human Services. Children attending the 3.5 year general developmental assessment. Personal communication, 2001.

54. Lim H, Quah B, Balakrishnan V, Lim H, Tay V, Emmanuel S. Vision screening of 4-year-old children. Singapore Medical Journal 2000;41(6):271-278.

55. Mark H, Mark T. Parental reasons for non-response following a referral in school vision screening. Journal of School Health 1999;69(1):35-38.

56. Romano P. Vision/eye screening: test twice and refer once. Pediatric Annals 1990;19(6):359-367.

57. Yawn B, Lydick E, Epstein R, Jacobsen S. Is school vision screening effective? Journal of School Health 1996;66(5):171-175.

58. Hall D. Health for all children. Third ed. Oxford: Oxford University Press; 1996.