Purpose. To determine population age norms in the first three years of life for binocular and monocular grating visual acuity (VA) obtained with Vistech–Teller Acuity Cards (TAC).

Methods. TAC was used to estimate grating acuity in 646 healthy infants and children born at due date ± 2 weeks, all of whom underwent ophthalmologic and orthoptic evaluation. The sample consisted of 20 age groups from 0 to 36 months. Sixty-nine percent of the children attended day care centers in the city of Sao Paulo. The sample was composed of white (63.0%), mulatto (25.2%), African-Brazilian (11.0%), and Asian (0.8%) infants and children, most of whom (97%) were from low-income families. Tests were conducted by eight highly trained testers, six of whom were orthoptists.

Results. Binocular and monocular norms for grating VA are presented in terms of tolerance limits for 90% of the population with 95% probability. The range of tolerance limits is approximately 2.5 octaves at most ages. There were no statistical differences among scores obtained by the different testers. There were no differences in VA due to race, sex, and first or second eye tested. The results on binocular (99.3%) and monocular (96.2%) testability and on mean test duration (13 minutes for one binocular and two monocular measurements) confirm the clinical applicability of TAC.

Conclusions. The binocular and monocular grating VA norms obtained in this large-sample study are different from the preliminary norms published with the TAC. Results from this and other studies (see Mayer et al, page 671, this issue) strongly point to a need for redefinition of the preliminary VA norms. Invest Ophthalmol Vis Sci. 1995;36:657–670.

Behaviorally measured grating visual acuity (VA), which reflects the ability to resolve the elements of a pattern, develops rapidly from birth to 6 months and then at a much slower rate up to adult acuity at 3 to 5 years of age.1–4 There is no question about the clinical importance of assessing visual acuity at an early age. Early detection and proper correction of visual abnormalities, such as congenital cataract and strabismus, can prevent permanent visual impairment.5–9 Reversibility of impairment has been shown to occur in animal models of early abnormal visual experience,10 and clinical application of these findings has been confirmed.11

Behavioral grating visual acuity can be measured successfully in infants in the laboratory using a variety of psychophysical procedures12–17 based on pattern preferences first described by Fantz18 in 1958. Measurement of grating acuity in the clinical setting required the development of a rapid measuring technique. The acuity card procedure (ACP) developed in 1985 by McDonald and colleagues19–21 provided a potential answer to this problem. It involved the construction of commercially available equipment (Vistech–Teller Acuity Cards [TAC]) and used an informal but rapid testing procedure. Preliminary binocular norms presented with the test were compiled from data from six different studies19–24 with a total sample of 249 children. Preliminary monocular norms were based on data from two studies from the same laboratory,20–21 with a sample of 72 children. Acuity norms were presented at 1- to 6-month intervals.25
Since its release, ACP has been used extensively both in research and in clinical practice. Data based on ACP are already available on various methodologic aspects of its application (use in private practice, test–retest reliability, intertester agreement, special procedures for use with visually impaired children), on normal visual acuity development, on comparison with refractive errors, on application to different ocular pathologies, and on assessment of visual acuity in children with neurologic impairment. This large applicability generated the need for a review of the original preliminary norms widely used in clinical practice and clinical research at present and the need for large scale studies specifically designed for the purpose of establishing more statistically valid age norms.

To establish normative data, two conditions are basic. The first consists of ruling out the possibility of including pathologic cases in the sample of normal children. This can be accomplished by submitting every tested child to ophthalmologic and orthoptic evaluation, a requirement not fulfilled by several studies that aimed at defining norms. The second is to have a sufficiently large sample that can be representative of the population. The present study met these conditions, and its purpose was twofold: to establish precise tolerance limits for grating acuity measured with TAC in healthy children from 0 to 36 months of age and sample size per age group, as in Tables 1 and 2. Age range was ±1 month for all age groups above 6 months of age. The children were tested in three different places: the Department of Ophthalmology at the Paulista School of Medicine–Hospital São Paulo (N = 123; 19%), the Department of Experimental Psychology at University of São Paulo (N = 75; 12%), and nine public day care centers in the city of São Paulo (N = 448; 69%).

Racial categories were: white (N = 407, 63.0%), mulatto (N = 163, 25.2%), African–Brazilian (N = 71, 11.0%), and Asian (N = 5, 0.8%). The present sample was classified according to and follows the distribution of the racial groups that exist in the Brazilian population according to the last census published in 1980. Most (N = 626, 97%) were from low-income and the remaining (N = 30, 3%) were middle-to-high income families.

Subjects had no known neurologic or medical abnormalities and had no previous ophthalmologic assessment. Complete ophthalmologic examination, including external eye inspection, cycloplegic retinoscopy, indirect ophthalmoscopy, and corneal reflexes, was conducted. All children were also submitted to orthoptic evaluation, which consisted of assessing ocular alignment and motility.

The final sample was composed of infants and children that met the following criteria:

1. Birth date coincided with due date or was ±2 weeks from it. In the latter case, the age was corrected.
2. Emmetropia or refractive error was within ±5 spherical diopters of hyperopia, ≤3 spherical diopters of myopia, ≤3 cylindrical diopters of astigmatism, and ≤2 diopters of anisometropia (spherical equivalents).
3. There was an absence of fundus alterations, apparent ocular deviation, and ocular diseases.

An additional 80 children were excluded from the final sample either because they did not undergo the eye examination (N = 47), they did not meet our criteria on the eye examination (N = 22), they were not born within 2 weeks of due date (N = 9), or because they did not complete the binocular test of acuity (N = 2).

The tenets of the Declaration of Helsinki were followed in the present study. For children who were brought individually to the testing place, the procedures were explained and informed consent was obtained from each subject’s parent(s) or guardian(s) before testing. In public day care centers, a meeting with directors, nursemaids, and parents or guardians was arranged to explain the reasons for early visual assessment and the type of procedures to which the child would be submitted. Only children whose parents or guardians attended the meeting were tested. Three parents or guardians did not agree with the procedures, and they were asked to leave a written statement.

Apparatus and Procedure

Grating acuity was assessed with a variant of the acuity card procedure binocularly and monocularly. The first measurement was always binocular. The first eye tested monocularly was randomly selected (OD was the first eye tested in 338 cases –54.2%). Monocular measurements were made with an adhesive eye patch (OFTAM, AMP, São Paulo, Brazil). Infants up to 20 days of age were tested only binocularly. Details of the
Age Norms for Vistech–Teller Acuity Cards

### TABLE 1. Binocular Norms for Vistech/Teller Acuity Cards

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>N</th>
<th>Test Distance</th>
<th>Start Card (cyc/deg)</th>
<th>Mean Acuity (cyc/deg)</th>
<th>Lower Limit* (cyc/deg)</th>
<th>Closest Card (cyc/deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>17</td>
<td>38</td>
<td>0.50</td>
<td>0.66</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>38</td>
<td>0.30</td>
<td>2.02</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>38</td>
<td>0.44</td>
<td>3.89</td>
<td>1.59</td>
<td>1.1</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>38</td>
<td>0.44</td>
<td>5.48</td>
<td>2.14</td>
<td>1.7</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>38</td>
<td>0.44</td>
<td>7.44</td>
<td>3.38</td>
<td>3.1</td>
</tr>
<tr>
<td>8</td>
<td>34</td>
<td>55</td>
<td>1.3</td>
<td>9.81</td>
<td>5.10</td>
<td>3.1</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>55</td>
<td>1.3</td>
<td>11.59</td>
<td>6.05</td>
<td>4.7</td>
</tr>
<tr>
<td>12</td>
<td>40</td>
<td>55</td>
<td>1.3</td>
<td>11.08</td>
<td>4.13</td>
<td>3.1</td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>55</td>
<td>1.3</td>
<td>13.04</td>
<td>5.63</td>
<td>4.7</td>
</tr>
<tr>
<td>16</td>
<td>41</td>
<td>55</td>
<td>1.3</td>
<td>13.08</td>
<td>5.91</td>
<td>4.7</td>
</tr>
<tr>
<td>18</td>
<td>35</td>
<td>55</td>
<td>1.3</td>
<td>12.39</td>
<td>5.10</td>
<td>4.7</td>
</tr>
<tr>
<td>20</td>
<td>37</td>
<td>55</td>
<td>2.4</td>
<td>13.81</td>
<td>6.60</td>
<td>4.7</td>
</tr>
<tr>
<td>22</td>
<td>36</td>
<td>55</td>
<td>2.4</td>
<td>14.76</td>
<td>7.68</td>
<td>6.4</td>
</tr>
<tr>
<td>24</td>
<td>44</td>
<td>55</td>
<td>2.4</td>
<td>14.64</td>
<td>6.20</td>
<td>4.7</td>
</tr>
<tr>
<td>26</td>
<td>37</td>
<td>55</td>
<td>2.4</td>
<td>16.66</td>
<td>8.92</td>
<td>6.4</td>
</tr>
<tr>
<td>28</td>
<td>42</td>
<td>55</td>
<td>2.4</td>
<td>15.29</td>
<td>7.25</td>
<td>6.4</td>
</tr>
<tr>
<td>30</td>
<td>35</td>
<td>55</td>
<td>2.4</td>
<td>17.24</td>
<td>8.61</td>
<td>6.4</td>
</tr>
<tr>
<td>32</td>
<td>25</td>
<td>55</td>
<td>2.4</td>
<td>17.36</td>
<td>9.96</td>
<td>6.4</td>
</tr>
<tr>
<td>34</td>
<td>37</td>
<td>55</td>
<td>2.4</td>
<td>19.19</td>
<td>9.50</td>
<td>6.4</td>
</tr>
<tr>
<td>36</td>
<td>32</td>
<td>55</td>
<td>2.4</td>
<td>17.82</td>
<td>7.12</td>
<td>6.4</td>
</tr>
<tr>
<td>Total</td>
<td>646</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* In the first three age groups, tolerance limits are calculated for 90% of the population with 90% probability.

Acuity card apparatus have been described elsewhere. The subject was seated in front of a gray screen (Vistech, Dayton, OH). A rectangular opening in the middle of the screen was used to present the acuity cards. Children 0 to 6 months of age were positioned 38 cm away from the screen, and children from 7 to 36 months of age were positioned 55 cm from the screen. The luminance of the cards was measured with a photometer (Sekonic Auto Lumi L-158) and was greater than 10 cd/m² (grating acuity is not affected by changes in luminance above this value).

A standard set of 16 acuity cards (Vistech, Dayton, OH) with half-octave steps (0.23 to 38 cyc/cm) was used. The cards were arranged face down on a table behind the screen, with low spatial frequencies (wide stripes) at the top of the stack and progressively higher spatial frequencies below. The tester received each card from an assistant who selected the start card based on the age of the child according to the specification from the TAC Handbook (see Tables 1 and 2). The tester was not informed about the start card value and was blind to the left-right position of the grating on it.

The task of the tester was to judge from the subject’s behavioral reactions—such as eye movements, head movements, and pointing—which side of the card displayed the grating. Three responses were available: left, right, and no response (child cannot resolve grating). Cards were presented by the tester with several reversals of left-right position as many times as the child found necessary for a judgment. A correct judgment was considered a positive response; an incorrect judgment or the inability to resolve the grating was considered a negative one. The assistant provided the tester with feedback as to the correctness of the judgment and moved on to the next card.

A modified staircase procedure was used to determine the grating visual acuity threshold. Cards were presented from lower to higher spatial frequencies in 1-octave steps up to the threshold region, and then at 0.5-octave steps around the threshold. After a negative result (tested reported “not seen” or made a wrong judgment), step size was reduced from 1-octave to 0.5-octave, and a lower spatial frequency was presented until a positive response was obtained. Testing was continued until two consecutive staircase reversals occurred. VA threshold was defined as the spatial frequency of the last card that received two positive responses. If this criterion was not met in six trials (counting from the first negative response), threshold was defined as the highest spatial frequency card that received the greatest number of positive responses. The tester classified confidence of each grating acuity assessment according to a 5-point subjective scale, ranging from 1 (very low) to 5 (very high).

A test consisted of one binocular followed by two
TABLE 2. Monocular Norms for Vistech/Teller Acuity Cards

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>N</th>
<th>Test Distance</th>
<th>Start Card (cyc/deg)</th>
<th>Mean Acuity (cyc/deg)</th>
<th>Lower Limit* (cyc/deg)</th>
<th>Closest Card (cyc/deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8</td>
<td>38</td>
<td>0.44</td>
<td>2.31</td>
<td>0.42</td>
<td>0.6</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>38</td>
<td>0.44</td>
<td>3.09</td>
<td>1.18</td>
<td>1.1</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>38</td>
<td>0.44</td>
<td>4.15</td>
<td>1.57</td>
<td>2.2</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>38</td>
<td>0.44</td>
<td>7.18</td>
<td>2.78</td>
<td>2.2</td>
</tr>
<tr>
<td>8</td>
<td>34</td>
<td>55</td>
<td>1.3</td>
<td>8.31</td>
<td>4.31</td>
<td>3.2</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>55</td>
<td>1.3</td>
<td>10.88</td>
<td>4.26</td>
<td>3.2</td>
</tr>
<tr>
<td>12</td>
<td>40</td>
<td>55</td>
<td>1.3</td>
<td>9.82</td>
<td>4.00</td>
<td>3.2</td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>55</td>
<td>1.3</td>
<td>10.35</td>
<td>4.82</td>
<td>3.2</td>
</tr>
<tr>
<td>16</td>
<td>41</td>
<td>55</td>
<td>1.3</td>
<td>10.07</td>
<td>4.70</td>
<td>3.2</td>
</tr>
<tr>
<td>18</td>
<td>35</td>
<td>55</td>
<td>1.3</td>
<td>9.95</td>
<td>5.21</td>
<td>4.7</td>
</tr>
<tr>
<td>20</td>
<td>37</td>
<td>55</td>
<td>2.4</td>
<td>11.12</td>
<td>3.91</td>
<td>3.2</td>
</tr>
<tr>
<td>22</td>
<td>35</td>
<td>55</td>
<td>2.4</td>
<td>12.09</td>
<td>4.30</td>
<td>4.7</td>
</tr>
<tr>
<td>24</td>
<td>43</td>
<td>55</td>
<td>2.4</td>
<td>12.31</td>
<td>4.03</td>
<td>3.2</td>
</tr>
<tr>
<td>26</td>
<td>37</td>
<td>55</td>
<td>2.4</td>
<td>13.00</td>
<td>6.62</td>
<td>4.7</td>
</tr>
<tr>
<td>28</td>
<td>42</td>
<td>55</td>
<td>2.4</td>
<td>12.70</td>
<td>5.61</td>
<td>4.7</td>
</tr>
<tr>
<td>30</td>
<td>34</td>
<td>55</td>
<td>2.4</td>
<td>14.12</td>
<td>7.09</td>
<td>4.7</td>
</tr>
<tr>
<td>32</td>
<td>24</td>
<td>55</td>
<td>2.4</td>
<td>15.22</td>
<td>7.99</td>
<td>6.4</td>
</tr>
<tr>
<td>34</td>
<td>37</td>
<td>55</td>
<td>2.4</td>
<td>14.97</td>
<td>7.46</td>
<td>6.4</td>
</tr>
<tr>
<td>36</td>
<td>32</td>
<td>55</td>
<td>2.4</td>
<td>14.98</td>
<td>5.86</td>
<td>6.4</td>
</tr>
<tr>
<td>Total</td>
<td>624</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* In the first two age groups, tolerance limits are calculated for 90% of the population with 90% probability.

monocular grating acuity measurements. Total test time was measured from the beginning of binocular measurement to the removal of the occluder in the end of the second monocular test. It should be noted that the 2-week-old group would be expected to have the lowest total test times because they were tested only binocularly.

After grating acuity assessment of the subjects, the orthoptic and ophthalmologic examinations were performed, with both professionals unaware of each other's results. The ophthalmologist was unaware of the grating acuity results. The cycloplegic agents used were cyclopentolate (1 drop/eye) plus tropicamide 0.5% (1 drop/eye).

Training of the Testers
A team of eight acuity card testers was formed, two of whom had extensive experience with TAC. One of them and an author of this article (DFV) gave theoretical and practical training to the other six testers—four orthoptists and two psychology undergraduate students. During training, each tester had to measure binocular and monocular grating acuity in five normal children and five children with different ocular disorders, ranging in age from 0 to 36 months. No age group was privileged in this phase. To be included as a tester, at the end of the training period each participant had to display errorless performance in the “scrambled cards” procedure.44 The number of subjects/tester were: AB - N = 236 (36.5%); EG - N = 178 (27.5%); MM - N = 108 (16.7%); VF - N = 54 (8.3%); SS - N = 25 (3.9%); RE - N = 21 (3.3%); FC - N = 21 (3.3%), and RO - N = 3 (0.5%).

Retest
A retest was performed when one or more of the following conditions occurred: the subject did not cooperate during the test; acuity scores were under the age norm by more than 1 octave; there was an interocular acuity difference (IAD) equal to or greater than 1 octave or a confidence of grating acuity estimate lower than 3 octaves (fair). The retest was always performed by a different tester who, though aware that it was a retest, received no information about the previous test score or the reason for the retest. It should be noted that the norms were based on the test results of all subjects from the final sample, regardless of the fact that part of the subjects were retested.

Statistical Analysis
All data analysis was conducted using log values of the VA scores. Norms were calculated based only on test results. Retest data were not included.

Norms were described using tolerance limits,46 which assume normal distribution of the function in the population. These are intervals that cover a previously fixed percentage of the population for the function under study, with an attached probability. For the present study, tolerance intervals for 90% of the population with a probability of 95% were used. The exception was for the 2-week, 2-month, and 3-month-old groups, for whom a probability of 90% was
adopted because there is a steep rise in acuity in this phase and the N was small in these groups. After adherence to the normal curve was tested and confirmed for the group with the largest N, the determination of tolerance limits was made using an N close to 38 subjects in each age group, as required for tolerance limits with the specified probabilities (N for each group is presented in Tables 1 and 2). Tolerance limits were calculated with the equation: $TL_k = \text{mean} - k \times SD$ and $TL_u = \text{mean} + k \times SD$, where $TL_k$ and $TL_u$ are, respectively, the lower and upper tolerance limits; $k$ is a constant based on the value of N (number of subjects), the desired probability, and the percentage of the population covered; and SD is the standard deviation of the mean.46

Analysis of variance (ANOVA) was performed using VA scores to evaluate agreement among testers, including age as a factor, among VA scores from both eyes obtained by the eight TAC testers separately for monocular and binocular VAs; test IADs across the 19 age groups; evaluate possible differences among racial categories, sex, and first and second eye tested in monocular VA; and compare differences between binocular and monocular VA in the 20 age groups. The Pearson coefficient was calculated to verify correlation between refractive errors of OD versus OS and refractive errors versus VA in the entire sample. All analysis performed used test results. Retest data were used only to compare with test data on the same children.

RESULTS

Population Norms

Tolerance Limits. Data from the final sample of 646 children were used to calculate binocular tolerance limits. The monocular tolerance limits were based on data from 624 of these children. Both limits were determined from VA scores taken from only one eye per subject, randomly selected between first and second eyes tested in each child. The rationale for randomly selecting VA between eyes is that both eyes should not be used because they are not independent,47 and using only the first eye tested is also inappropriate because practice and fatigue might affect variability.48 The first eye tested was randomly selected so that left and right eyes would be equally represented in the sample. Binocular and monocular VA means with respective SEs and tolerance limits are shown for each age group in Tables 1 and 2 in cyc/deg and in Figure 1.

A one-way ANOVA for binocular VA ($F_{18,623} = 89.72, P = 0.01$) and monocular VA ($F_{18,623} = 31.54; P = 0.01$), followed by the Scheffé procedure, identified pairs of age groups that were significantly different at the 0.05 level. For monocular and binocular VA, 2-month-olds had significantly lower acuity than 4-month-olds, who, in turn, had significantly lower acuity than 6-month-olds ($P < .05$ for all). After 6 months, there was a gradual increase in acuity with no significant differences between adjacent age groups ($P > .05$ for all). Tables 1 and 2 present, respectively, binocular and monocular norms and the closest card to the lower limit in cyc/deg for clinical use. Tolerance limits are also presented, along with the preliminary norms44 in Figure 2 for binocular and monocular VA scores. A steep increase in binocular and monocular VA is observed from birth to approximately 6 months, followed by shallow growth thereafter. The present results exhibit much higher monocular VAs than the preliminary norms in the first 18 months of life. The lower limit was higher by about 1 octave between 8 and 14 months and by less than 1 octave from 4 to 6 months. As to the upper limit, the difference was close to 2 octaves from 2 to 24 months. From 28 to 34 months, the lower limit was lower than the preliminary norms by less than 1 octave, and for the last age group the difference was slightly higher than 1 octave.

Interocular Acuity Differences. Mean IADs (OD minus OS) for the 19 age groups are shown in Table 3, along with SDs and tolerance limits for IADs. Absolute IAD values are also shown. Differences ranged from −0.15 octaves (OS better) to 0.16 octaves (OD better). In absolute IAD values, the range was from 0 to 3 octaves. These differences were not restricted to specific age group(s). A one-way ANOVA based on the raw IADs showed no significant differences among the 19 age groups ($F_{18,623} = 0.81, NS$).

Multifactor Comparisons. VA scores from the first and second eyes tested were compared on the basis of sex and the four racial categories in two separate one-way ANOVAs to verify the possibility of differences from any of the groups. Fourteen of the 16 possible groups were tested; because there was only one child in the Asian female group, this category could not be included in the analysis. The ANOVA result was significant for first eye tested ($F_{6,623} = 2.65; P = 0.02$), second eye tested ($F_{6,623} = 2.47; P = 0.02$), and binocular data ($F_{6,623} = 2.49; P = 0.02$). However, the post hoc Scheffé test did not reveal the source of significance in any of these cases; no two groups were significantly different at the .05 level.

Binocular Versus Monocular Results. Mean difference between binocular and monocular VAs were compared across the 19 age groups. Differences were not significant for all groups ($F_{18,623} = 1.04, NS$).

Refractive Errors

The incidence and type of refractive errors found in the 646 subjects was almost always identical for the two eyes, as shown by the calculated correlations for the whole sample, which were based on spherical

corrected age (months)

FIGURE 1. (A) Mean binocular acuity (±1 SEM) and tolerance limits for 90% of the population with 95% probability based on data from 646 healthy infants and children as a function of age. (B) Mean monocular acuity (±1 SEM) and tolerance limits for 90% of the population with 95% of probability based on data from 624 healthy infants and children as a function of age.

 equivalents ($r = 0.97; \ P < 0.05$). Refractive errors ranged from $-0.50$ to $+4.75$ D in spherical equivalent (mean $= +1.25; \ SD = 1.18$ D). Most eyes were hyperopic (49.5%) and had compound hyperopic astigmatism (23.8%); 8% had emmetropia. Other types of refractive errors including myopia, hyperopic astigmatism, myopic astigmatism, compound myopic astigmatism, and mixed astigmatism. Each occurred in fewer than 5% of the subjects.

No correlation was found between refractive error and visual acuity. The Pearson coefficients for this comparison in OD and OS were, respectively, $r = -0.18$ for OD and $r = -0.15$ for OS (NS).

Testability and Test Time

Testability and test time were based on the 726 children from the initial sample. Binocular testability was 99.3%, and monocular testability was 96.1%.

Average test time of the entire sample, with the exception of the youngest age group, was 12.61 minutes (SD = 5.28 minutes), with a median of 11 minutes and a mode of 10 minutes. This test duration included one binocular and two monocular measurements, in addition to eye patching time. Minimum mean test time (minutes) was $11.42 \pm 3.55$ for the 26-month-old group, and maximum was $15.88 \pm 7.83$ for the 2-month-old group.

The 2-week-old group was tested only binocularly, and its average test time was $6.91 \pm 3.63$ minutes.

Agreement Among Testers

Monocular VA results in the TAC obtained for each of the eight individual testers in the 646 children of the final sample are shown in Figure 3 along with the preliminary norms. To evaluate agreement among observers, a two-way ANOVA including age as a factor was performed combining data across the 20 age groups, pooled into 7 age groups, according to similar-
FIGURE 2. (A) Binocular tolerance limits of grating VA for 90% of the population with 95% probability, measured by Vistech–Teller Acuity Cards, as a function of age. The preliminary norms for the same instrument are presented for comparison (dashed lines). (B) Monocular tolerance limits of grating VA for 90% of the population with 95% probability, measured by Vistech–Teller Acuity Cards, as a function of age. The preliminary norms for the same instrument are presented for comparison (dashed lines). VA = visual acuity.

ity in mean acuity. For binocular data, age groups were: 0.5, 2, 3, 4, 6 to 18, 20 to 24, and 26 to 36 months old. The monocular age groups were: 2, 3, 4, 6, 8, 10 to 24, and 26 to 36 months old. The results for both binocular (F_{28.645} = 1.17 NS) and monocular (F_{29.623} = 1.37, NS) VAs showed no significant differences among testers for any of the age groups.

Test–Retest

Retests were performed only in special cases (see Methods). Twenty-six boys and 23 girls, with corrected age from 1 to 36 months, were uncooperative and had low scores or high IADs. They represent 7.6% of the total sample (49/646).

Monocular VA pairs of test and retest scores were obtained in 49 children (98 eyes). Visual acuity in the retest was higher in 38 pairs (39%), lower in 36 (37%), and identical to that of the test in the remaining 24 (24%). Differences in VA between test–retest ranged from 0 (N = 24) to 3 octaves (N = 1).

The percentage of acuity scores differing by a given number of octaves in the 98 monocular test–retest pairs is presented in Figure 4. Differences >1 octave were more frequently observed in children from 0 to 6 and 19 to 24 months of age. In these, the differences were approximately 35% of the retested children.

Interocular acuity differences (IAD), or OD minus OS, in the test were compared with those in the retest. The IAD in the test ranged from 0 (N = 2) to 3 octaves.
The subjective degree of confidence in each grating Vistech-Teller Acuity Cards in a large sample of scale in 95% of the tests. Excellent (86%) and good acuity judgment was scored by each tester in a 5-point

The present study provides binocular and monocular

DISCUSSION

Age

TABLE 3. Inteocular Acuity Differences and Respective Tolerance Limits for the Vistech Teller Acuity Cards

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>N</th>
<th>OD--OS Difference (oct)</th>
<th>Absolute SD (oct)</th>
<th>OD--OS Difference (oct)</th>
<th>Absolute SD (oct)</th>
<th>OD Better</th>
<th>OS Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8</td>
<td>0.54</td>
<td>0.78</td>
<td>0.06</td>
<td>0.97</td>
<td>2.72</td>
<td>2.6</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>0.32</td>
<td>0.46</td>
<td>0.27</td>
<td>0.49</td>
<td>1.32</td>
<td>0.78</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>0.21</td>
<td>0.31</td>
<td>-0.01</td>
<td>0.38</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>0.31</td>
<td>0.29</td>
<td>0.04</td>
<td>0.43</td>
<td>0.96</td>
<td>0.88</td>
</tr>
<tr>
<td>8</td>
<td>34</td>
<td>0.22</td>
<td>0.31</td>
<td>0.01</td>
<td>0.39</td>
<td>0.83</td>
<td>0.81</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>0.49</td>
<td>0.48</td>
<td>0.16</td>
<td>0.67</td>
<td>1.56</td>
<td>1.24</td>
</tr>
<tr>
<td>12</td>
<td>40</td>
<td>0.58</td>
<td>0.44</td>
<td>-0.05</td>
<td>0.58</td>
<td>1.14</td>
<td>1.24</td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>0.43</td>
<td>0.55</td>
<td>0.12</td>
<td>0.69</td>
<td>1.57</td>
<td>1.33</td>
</tr>
<tr>
<td>16</td>
<td>41</td>
<td>0.27</td>
<td>0.30</td>
<td>-0.01</td>
<td>0.41</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>18</td>
<td>35</td>
<td>0.33</td>
<td>0.38</td>
<td>-0.04</td>
<td>0.45</td>
<td>1.01</td>
<td>1.09</td>
</tr>
<tr>
<td>20</td>
<td>37</td>
<td>0.45</td>
<td>0.50</td>
<td>0.02</td>
<td>0.68</td>
<td>1.43</td>
<td>1.39</td>
</tr>
<tr>
<td>22</td>
<td>35</td>
<td>0.46</td>
<td>0.45</td>
<td>0.01</td>
<td>0.65</td>
<td>1.36</td>
<td>1.34</td>
</tr>
<tr>
<td>24</td>
<td>43</td>
<td>0.34</td>
<td>0.47</td>
<td>-0.15</td>
<td>0.56</td>
<td>0.99</td>
<td>1.29</td>
</tr>
<tr>
<td>26</td>
<td>37</td>
<td>0.35</td>
<td>0.46</td>
<td>0.01</td>
<td>0.59</td>
<td>1.23</td>
<td>1.21</td>
</tr>
<tr>
<td>28</td>
<td>42</td>
<td>0.26</td>
<td>0.29</td>
<td>0.03</td>
<td>0.36</td>
<td>0.85</td>
<td>0.72</td>
</tr>
<tr>
<td>30</td>
<td>34</td>
<td>0.32</td>
<td>0.37</td>
<td>-0.05</td>
<td>0.49</td>
<td>0.98</td>
<td>1.08</td>
</tr>
<tr>
<td>32</td>
<td>24</td>
<td>0.43</td>
<td>0.39</td>
<td>0.08</td>
<td>0.58</td>
<td>1.37</td>
<td>1.21</td>
</tr>
<tr>
<td>34</td>
<td>37</td>
<td>0.26</td>
<td>0.3</td>
<td>-0.01</td>
<td>0.38</td>
<td>0.69</td>
<td>0.89</td>
</tr>
<tr>
<td>36</td>
<td>32</td>
<td>0.20</td>
<td>0.32</td>
<td>-0.03</td>
<td>0.38</td>
<td>0.78</td>
<td>0.84</td>
</tr>
</tbody>
</table>

* In the first two age groups, tolerance limits are calculated for 90% of the population with 90% probability.

(N = 1), and in the retest they ranged from 0 (N = 21) to 1.5 octaves (N = 5). In 34/49 (70%) children, IADs were lower in the retest.

Confidence of Grating Acuity Estimates

The subjective degree of confidence in each grating acuity judgment was scored by each tester in a 5-point scale in 95% of the tests. Excellent (86%) and good (12%) confidence was reached in the majority of binocular and monocular measurements. The remaining 2% were classified as fair.

DISCUSSION

The present study provides binocular and monocular tolerance limits for grating acuity measured with the Vistech-Teller Acuity Cards in a large sample of healthy children from 0 to 36 months of age. Norms for interocular acuity differences are also presented. The study showed that testers can be trained to measure VA within a reasonably close agreement and that a more stringent procedure than that typically used with the TAC is can be used with short testing times compatible with clinical or screening purposes. This is the first study to provide both binocular and monocular norms in a large sample of infants and children primarily from low-income families of different racial origins.

Below we discuss the use of tolerance limits as norm statistics and its adequacy for clinical interpretation of VA and IAD findings. We compare our findings with those of previous normative studies, mostly with preliminary norms. The norms derived by Mayer et al (this issue, page 671) from a sample with a very different socioeconomic composition are compared to the present norms. Complementary aspects concerning data (binocular versus monocular VA, refractive errors, and VA) or procedure variables (agreement among testers, test-retest in special cases, binocular versus monocular VA, testability, and test duration) are discussed.

Tolerance Limits as Norm Statistics

Visual acuity estimates can vary for the same age in different children and can also vary in the same child in different measurements. The first variability is inherent in the function being measured, and the second is inherent in the precision with which it can be measured. Assuming both are normally distributed, it can be said that a given result is normal if it falls within the interval delimited by standard deviations above and below the mean. This interval represents the confidence limits with which the mean obtained can be described. Confidence limits cover a population parameter (the mean) with a certain confidence. This can be translated to the present situation in this manner: If the VA value is within the confidence interval, there is a defined probability (90%, 95%, 98%, or
Age Norms for Vistech-Teller Acuity Cards

FIGURE 3. Individual monocular grating VA scores obtained with the Vistech–Teller Acuity Cards by the eight testers, identified by the initials in the upper right corner of each graph. Dashed lines are preliminary VA monocular norms. $\text{•} = \text{OD}; \bigcirc = \text{OS}. \text{VA} = \text{visual acuity}.$

whatever one chooses to define) that it is accepted as being equal to the mean.

One can, on the other hand, choose to obtain a range of VA values that encompasses a fixed portion (e.g., 70%, 90%) of the population with a specific probability. These intervals are called tolerance intervals, and their extreme points are called tolerance limits.46 In this case, the interval represents the normal VA range for a defined percentage of the population (any percentage that one defines). Translating the concept to the present study, this means that if a given VA value is within the tolerance limits fixed for 90% of the population, it can be said, with a certain defined probability, that the statement "VA is within normal range" is true. The greater the portion of the population covered, the greater will be the number of "ab-normal" cases in the interval. In the limit, if a 99.9% tolerance interval was chosen, almost any VA would have been classified as "normal." The objective of defining norms is to draw a line that will exclude the pathologies that exist in the population.

In the present study, the purpose was to establish, with a certain probability, intervals of normal values of VA for a specified percentage of the population. The interval chosen here, which covers 90% of the population, excludes 5% above the upper limit and 5% below the lower limit. In practice, for clinical purposes, it is the 5% below the lower limit that are of concern to the clinician. It is evident that the choice of a tolerance interval will depend on the number of referrals one can afford or find profitable. It affects the balance between false-positive and false-negative results. Tolerance limits were also used by Heersema and van Hof-van Duin29 in their normative study of VA in toddlers. In that study, they showed that VA is an approximately linear function of the logarithm of age. However, there might be clinical situations in which this condition is not met by the data. Because tolerance limits do not depend on an assumption of linearity with relation to age, they would be applicable in these cases.

Present Norms and Other Normative Studies

The first norms available for TAC presented by Teller et al25 are based on McDonald et al studies19–21 in the case of monocular norms and are distributed with the TAC.44

The finding of higher VAs than those presented in the norms above has been the trend in recent studies,3,29,48,49 and Mayer et al, this issue, which show consistent results despite the use of different procedures. It has been suggested that McDonald et al19–21 could have underestimated the mean acuity because even though photorefractive screening was performed, a full ophthalmologic assessment was not.

A discussion of the causes for the higher VAs found in the first year of life should take sample size and composition into account. Even though sample size was large, thus reducing the variance, and the sample was likely to be homogeneous, as shown by the absence of difference between any two groups when tested with the Scheffé procedure, these factors could not have elevated the averages. A small sample would result in an irregular function, but not in a uniformly lower (or higher) one. A larger sample simply contributes to a smoother function, with lower variance around each point.

Possible explanations for the higher VA scores found in the 4- to 12-month group must be related to procedural factors. The majority of the data here (69%) were collected in an environment familiar to the child—the day care unit. This might have contribu-
Norms for Interocular Acuity Differences

Knowledge of IAD is critical for the clinical diagnosis of amblyopia and of monocular disorders. Birch and Hale find that interocular difference offers a better criterion for monocular acuity deficit, with higher sensitivity than the mere application of the norms. Differences found in the sample used for the present norms are within the range of those found by others (Mayer et al, this issue, page 671).

Comparison of Studies: Children From Middle- to High- Versus Low-Income Families

As reported, this study was conducted in parallel with a similar study conducted by Mayer et al (this issue, page 000). Each had as its objective the establishment of norms for the TAC, and each used a large number of subjects. Some differences between the two studies are number of testers, racial and socioeconomic composition of the sample, testing procedures, number of age groups, and age intervals. The present study used eight testers in a low-income sample, whereas Mayer et al used two testers in a middle- to high-income sample. Racial differences between the two samples were also pronounced. Threshold determination differed in that the Mayer et al study used the classical procedure, whereas in this a modified staircase procedure was adopted (see Methods).

Upper and lower limits of both sets of results are compared in Figure 5. It is important to point out how similar the lower limits are, despite the very different composition of the samples. It should be stressed that although children of low-income families composed most of the sample, they attended public day care centers in São Paulo and were not undernourished. Low socioeconomic level might be responsible for lower stimulation, greater difficulty in obtaining medical care when needed, and diminished primary medical care. These factors, if present, did not seem to have any effect on the visual acuity of these children.

Complementary Aspects

Concerning Data. Binocular versus monocular VA. Birch and Hale’s report on forced-choice preferential looking shows slightly higher binocular VA scores after 6 months, but the corresponding statistical analysis shows that the differences are not significant for children from 15 to 60 months of age. The present study finds no statistical difference between binocular and monocular VA, but it also shows slightly higher binocular VA after 5 months.
Refractive errors and VA. Not many studies of VA determine refractive error. Ophthalmologic and orthoptic examination are an essential prerequisite of normative studies because it is absolutely necessary that any known pathology be excluded from the normal sample. In the present study, these factors were responsible for the exclusion from the original sample of 14 children with refractive error below the cutoff criteria (see Methods). In the sample used for the norms \( n = 646 \), refractive error, within the narrow range accepted for the sample, was not correlated with VA. This confirms other reports showing that refractive error is not causally related to the development of VA. Nevertheless a relationship might exist if extreme cases had been included.

Concerning Procedural Variables. Testability and test duration. Testability and test durations found in the present study showed that VA assessment of most children could be completed in a short time. Binocular and monocular testabilities reported in the literature ranges from 89% to 100%.\(^{19-21}\) The testabilities found here (99.3% for binocular and 96.2% for monocular measurements) are at the high end of this range.

In the present study, test durations for three tests, two monocular and one binocular, were 15 minutes on average \( (SD = 5.66 \text{ minutes}) \), and it could be estimated that each test took about 4.5 minutes. This average refers to the whole sample, but it must be kept in mind that younger subjects took a longer time to be tested \((\sim 7 \text{ minutes for infants younger than 1 month and 5 minutes for 2- and 3-month-old infants})\).

When comparing test durations reported from different studies, the procedure used should be taken into account. Test time is dependent on the psychophysical method. The modified staircase procedure adopted here might introduce a more precise evaluation of the grating acuity threshold at the cost of lengthened test time. In fact, test times found by Vital-Durand and Hullo\(^{22}\) using TAC were about half as long as ours. Times found by other authors using TAC\(^{21,27-30}\) were slightly shorter but more comparable to ours, despite the fact that they did not use a staircase procedure. The high success rate also contributed to the lengthening of the average test duration because every child was tested regardless of how long it took.

Agreement among testers. Screening programs require large groups of testers to be effective within budget and time constraints. One of the concerns of this work was to find out whether we could train testers to produce data with good intertester agreement. The largest numbers of testers reported in single studies of acuity card measurements were, respectively, five\(^{27}\) and four,\(^{28}\) in two studies using infants with perinatal complications; agreement was similar to that reported previously but was not as good as found for normals. Quinn et al,\(^{28}\) who used three testers in a clinical setting, report significant intertester differences and recommend, in line with the two former studies, that interobserver reliability be tested when using TAC.

The eight testers used in the present study were carefully trained before the beginning of data collection. This, and the continuous monitoring of the tests by one of us (SS), was responsible for the fact that no significant differences in average acuity values were found among testers. It must be pointed out that the comparison performed here is not very sensitive and
that a better approach would have been to collect data from the same eyes. However, it would have been difficult to submit the same child to eight different testers.

*Test–retest in special cases.* Retests, performed in special cases, showed higher VAs about as often as there were of lower VAs (39% versus 37%) and equal scores in the remaining cases. The reduction of interocular differences found in 70% of the cases indicates that the retest scores might be more reliable. However, because retests were only performed in special cases, the tester might have taken extra care knowing that something went wrong on the first test. In addition, reduction of interocular differences might be due to the phenomenon of regression of the mean—the tendency for normal subjects with extreme values to have less extreme values on retest. The implication is that VAs in the test might have been underestimated as frequently as overestimated, which is to be expected because these children were uncooperative, and testers reported low confidence in their estimates.

Because test–retests were performed only in especially difficult cases, it is not appropriate to analyze the results as representing agreement among testers. A measure of agreement depends on the assumption that the child does not change from test to retest. The expectation of the testers was the opposite—to achieve better cooperation or confidence on retest.

**CONCLUSIONS**

The present norms provide binocular and monocular data mostly from infants and children of low-income families, whereas Mayer et al’s study (this issue, page 671) provides monocular norms for children primarily from middle- to high-income families. Regardless of these differences and of the fact that in the present study a more stringent procedure for VA estimates was used, the norms obtained in both studies are remarkably similar. The two studies, therefore, complement each other and strongly point to the need for adoption of new norms for the Vistech–Teller Acuity Cards. Norms for interocular acuity differences, which are of important clinical application, are also reported.

In addition, we suggest that the earliest phase of development be focused in future work. In this phase, the growth in VA is rapid, and it would be convenient to have binocular and monocular norms for Vistech–Teller Acuity Cards based on shorter intervals. Testing of a larger number of children during the first 3 months of life would be required for the completion of an adequate set of norms. It should also be noted that binocular data may be clinically valuable, mainly in the first month, during which it is difficult to obtain monocular measurements.

**Key Words**

infant vision, grating visual acuity, visual acuity development, Vistech–Teller Acuity Cards, visual acuity age norms

**Acknowledgments**

The authors thank Velma Dobson, PhD, for her constant support throughout the study and for her visit to Brazil at the beginning of this study. The authors also thank Davida Teller, PhD, for her critical review of and suggestions for this manuscript; Jair Licio Ferreira Santos, PhD, from the Public Health Faculty of the University of São Paulo for his competent advice on all aspects of the statistical treatment of the data; Adriana Berezovsky, Eny Gitelman, Marcia Maia, M. Valéria Ferrari, Raquel Eliezer for their collaboration in performing orthoptic evaluations and testing the children; Fernando Cunha and Regis Oliveira for testing the children; and of Emilio de Haro-Munoz, MD, and César Lipener, MD, for performing ophthalmologic examinations.

**References**

Age Norms for Vistech-Teller Acuity Cards


