Objective assessment of Visual acuity in infants

by

Amithavikram Rugvedi Hathibelagal

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Vision Science

Waterloo, Ontario, Canada, 2013

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Purpose

Early detection of abnormal visual acuity (VA) is crucial in the identification and management of ocular and visual abnormalities in infants. Currently, the Teller Acuity Cards (TACs) are considered the gold standard for clinical testing and are effective in obtaining a quick estimate of an infant's VA, but they have certain drawbacks. They rely on a subjective assessment of the baby's looking behavior. Despite this, TACs have been found to have good validity and repeatability.

The current study investigates a new method to objectively assess visual acuity in infants, which is uses a video gaze tracker (GT) and computer-generated stimuli, developed in the lab of M. Eizenman at the University of Toronto. The purpose was to validate this method in adults and infants against current clinical VA tests. Visual scanning patterns were measured by the GT system that requires minimal subject cooperation in adult and infant populations. The targets were judged as seen when the relative fixation time on the grating exceeded a pre-determined threshold, as compared to the fixation time on the luminance-matched background.

Methods

Experiment 1: In 15 uncorrected myopic adults, binocular grating VA was measured. The targets were square-wave gratings of spatial frequency ranging from 2.3 to 37 cpd presented randomly in one of four positions on the screen. There were 6 objective protocols (in which

VA was judged by fixations). The subjects were naïve, as the only instruction given to the participants was to look towards the screen. The experimenter, who presented the gratings also acted as an observer by making judgments of seen/not seen responses using the objective information provided by the software. Objective GT VA was compared with VA measured with subjective responses using the same stimuli and with Teller Acuity Cards (TACs).

Experiment 2: Binocular grating VA for horizontal gratings was measured in 20 typically-developing infants aged 3 to 12 months. Spatial frequency ranged from 0.32 to 42 cpd and VA was measured on two visits with both the GT and TACs. A staircase protocol was used to obtain the VA threshold in the GT. The experimenter controlled the staircase method and an observer used the objective information of visual fixations using the software to judge if the grating was seen or not. Video cartoons were shown between stimulus presentations to keep the infant's attention towards the screen.

VA was also measured with the TACs held in the vertical orientation, so that the gratings were horizontal, similar to the GT method. A TAC stage was specially designed with a vertical slot in which the cards could be presented. The observer was masked regarding the participant's age and the starting spatial frequency. The study co-ordinator determined the choice of the start card which was randomized between participants so as to give an equal number of participants with each start card. The same start card was used for the second session of each infant. The threshold was defined as the highest spatial frequency for which the infant gave a clear, correct look and an unclear/inconsistent look for the next higher level.

The observer, who was masked regarding the absolute spatial frequency, increased or decreased the spatial frequency until this threshold was determined.

Results

Experiment 1: The mean age of participants in the adult study was 28.47±7.93 yrs and their mean uncorrected logMAR acuity was 0.9±0.2. There was no obvious difference among the mean acuities obtained by 6 objective GT protocols, the subjective GT protocol and the TACs. The GT showed agreement of 93% and 100 % within half an octave compared with the subjective protocol and TACs (horizontal gratings) respectively. There was 100% agreement within 1 octave of the objective GT with both the subjective protocol and the TACs (horizontal gratings). The objective gaze tracker VA showed significant correlation with uncorrected refractive error ($r = 0.87$, $p < 0.001$).

Experiment 2: The mean age of participants was 7.9±2.5 months. In both visits, the testability of the TACs was 100% across all infants. GT had 100% testability on the first visit and 95% testability on the second visit. The mean TAC acuity over two visits for all the infants was 0.7±0.23 log cycles per degree, while the mean log GT acuity over two visits was 0.86±0.30. Infant GT VA acuity estimates were within 1 octave of the TACs 90% and 79% of the time for the first and second visit respectively, while GT VA estimates were within half octave of the TACs 63.2% and 47% of the time for the first and second visit respectively. Eighty-seven percent of the GT VAs and 72.5% of TACs were within one octave of the mean age norms, although on average the GT gave better acuities than the TACs. There was an increase in GT

VA with increasing age ($r=0.80$, $p<0.005$ for the first visit and $r=0.77$, $p<0.005$ for the second visit). Both the TACs and the GT had repeatability of 89.5% within 1 octave between visits and 84.2% and 79% within half octave between visits respectively.

Conclusions

In adults, the gaze tracker gave VA thresholds which were equivalent to the TACs and were not significantly different from subjectively determined grating VA. The agreement of the GT with TACs in infants and with norms in the infant literature established good validity for the GT. Finally, the significant correlation with age confirmed the validity of the measurements of the gaze tracker. The repeatability of the gaze tracker was similar to that of the TACs, demonstrating the quality usefulness of the test.

These results demonstrate the potential for an automated test of infant visual acuity, which could be a powerful and useful tool for visual acuity assessment in infants and other population groups who cannot respond verbally. The staircase protocol established in the study could be fully automated in an objective version of the test. The raw data of eye movements obtained in this study such as the pattern of first fixations, time taken for first fixations, time spent fixing the stimulus, typical stimulus duration and time between presentations could be used to develop algorithms for fully automated testing of VA in infants.

Acknowledgements

I would like to extend my sincere thanks to my Supervisor, Dr. Susan Leat for her continued guidance and meticulous support, which played a significant role in finishing my thesis successfully. I am deeply indebted to her for the help rendered throughout my course.

I would like to thank my committee members Dr. Elizabeth Irving for her valuable suggestions and Dr. Moshe Eizenman for giving me the opportunity to work with the equipment he designed and for his valuable inputs to this thesis.

I would like to thank Dr. Krithika Nandakumar for her help during the project.

I would like to thank my lab members Balsam Alabdulkader and Mohammed Althomali for their words of encouragement and support.

I would like to thank Rose Morrison, who helped in calling the participants' parents and thus helping in recruitment.

I would like to thank all the past and current graduate students in Vision Science for their great support and for a wonderful time here throughout. Special thanks to my roommates Raj and Varadhu for a great time here in Waterloo.

I would also like thank Kim Tremblay and Krista Parsons for their help in the administrative work.

I would like to thank all my friends back home in India for their great support. My dear friend, Lijo deserves a special appreciation for the incredible support and for all the good wishes that kept me going.

I would like to thank all my family members back home for their great love and support. Without them, this would not have been possible.

Dedication

This Thesis is dedicated
to the Almighty GOD for his blessings
and
to my dear parents and my dear brother (Achutha)
for their great love and unflinching support

Table of Contents

List of Figures

List of Tables

Chapter 1- Introduction and Literature Review

1.1 Visual acuity

1.1.1 Definition

Visual acuity (VA) is the measurement of the ability to resolve detail. In the general sense, VA is the smallest object (detection) or aspect of an object (resolution, recognition) that can be resolved. The threshold of resolution refers to the minimum angle of resolution that allows the human eye to identify two points or two lines as two distinct stimuli. Visual acuity can be specified by the reciprocal of the threshold of resolution in minutes of arc. It is measured in adults clinically with a letter chart, wherein the patient's task is to recognize the smallest letters that they can on the chart[1]. The most common letter chart used is the Snellen chart. The size of the targets is based on the letter "E", in which each bar subtends one min of arc. The measure of acuity is expressed in Snellen notation, which is basically a fraction. The numerator is the testing distance. The denominator is the distance at which the detail of the target (bars of the letter E) subtends one min of arc. In a "normally sighted eye" these values are the same, as normal resolution is expected to be one minute of arc. i.e., VA is 6/6 (meters) or 20/20 (feet). Decimal notation is obtained by expressing the Snellen fraction as a decimal. The decimal notation for 6/60 would be 0.1. VA is also measured with the log MAR chart, which is considered the gold standard for clinical testing in adults and is denoted in log MAR notation. MAR refers to the minimum angle of resolution subtended at the eye. MAR is basically the reciprocal of the Snellen fraction and log MAR is obtained by taking the logarithm of the MAR. For example, for a Snellen acuity of 6/60 acuity, the MAR is 10

minutes of arc and log of MAR gives a logMAR value of 1. A logMAR chart has an equal number of optotypes in each line. The optotypes used have equal legibility and there is proportional spacing between letters and between lines. Moreover, the progression is logarithmic. The difference between each line in the chart is 0.1 logMAR.

1.1.2 Types of visual acuity

Minimum visible acuity refers to the smallest stimulus that can be detected i.e. whether it is present or not. It is also called detection acuity. The measurement of this acuity involves the task of detecting stimuli (as shown in Fig 1a) of decreasing size. The best threshold that can be obtained by the adult human eye is approximately 1 sec of arc[1].

Minimum resolvable acuity refers to the smallest angular separation between two close targets that can be resolved[2]. The best resolution threshold that can be obtained by the human eye is approximately 30 sec. of arc. This acuity is also known as the minimum separable acuity (as shown in Fig b) and most commonly clinically measured as grating acuity. This type of acuity is measured in many preferential looking tests (described below).

Minimum recognizable acuity refers to the ability to identify form and orientation. An example of this acuity is Snellen acuity (shown in Fig 1c) or any letter/picture matching test. The main difference between recognition acuity and resolution acuity is that the latter involves interpretation and is more susceptible to contour interaction. Contour interaction is

one cause of the crowding phenomenon which reduces recognition of a single letter in the presence of nearby letters. This is more accentuated in amblyopic eyes[3]. Thus recognition acuity is more sensitive for detecting amblyopia than resolution acuity. The typical normal values for recognition acuity are 1 min of arc or better.

Minimum discriminable acuity refers to the ability to determine the relative location of two or more stimuli with respect to each other. The measurement of this acuity involves the task of identifying the offset between stimuli as shown in the Fig 1-1[2]. The magnitude of the best threshold that can be obtained is 1-3 sec of arc. This is also known as hyperacuity[1].

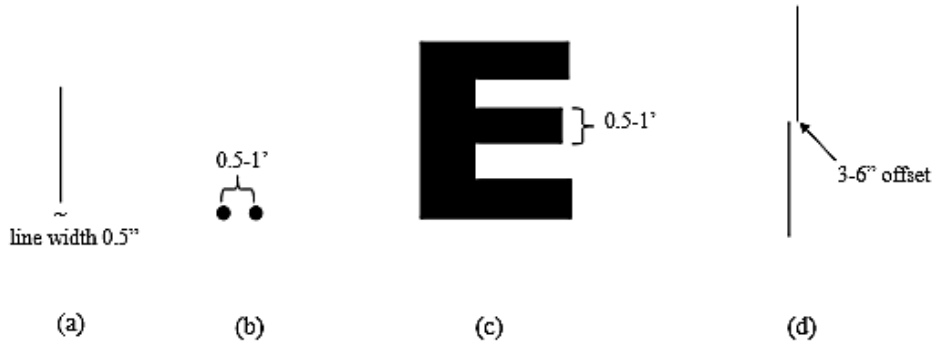


Figure 1-1. Types of Visual acuity

Figure 1.1 (a) Detection task (b) resolution task (c) recognition task (d) hyperacuity task. Based on the Fig 7.1 in Chapter "Visual Acuity" by Kathryn Saunders in "Assessing Child's Vision" by Leat et al.[2]

1.2 Methods of VA measurement in infants

Visual acuity measurement in infants is important in order to detect abnormalities. Infants cannot cooperate in subjective testing for VA as can people of other ages. A number of different ways have been developed to measure infant visual acuity. The most commonly used method is preferential looking. The other ways of measuring VA in infants are with visual evoked potentials (VEP) and optokinetic nystagmus (OKN). The latter two will be described first and then preferential looking and its development will be described in detail.

1.3 Visually evoked potential

The VEP technique estimates the electrical activity directly from the scalp above the visual cortex using surface electrodes[4]. The recording active electrode is usually placed 3 cm above the inion. The inion is the most prominent projection of the occipital bone at the posteroinferior part of the skull and is above the primary visual cortex. A reference electrode is placed in the midline, 30% of the distance between the inion and the nasion (the dip at the top of the nose, below the bony protuberance of the brow). An electrode clipped to the ear or placed on the forehead serves as a ground electrode. Visual stimulation causes electrical activity, whose amplitude and duration can be measured. The stimulus is repeated and the response is averaged in synchrony with the visual stimulus to separate the VEP from the general EEG activity and to improve the signal to noise ratio.

1.3.1 Types of VEP responses

1.3.1.1 Transient VEP

The transient VEP response refers to the electrical activity generated by a stimulus whose frequency of presentation allows brain responses to return to the pre-stimulus state before the next presentation. The rate of stimulus presentations is generally 2 per second. A transient VEP response can be generated in response to three types of stimuli, namely a flash stimulus, an alternating checkerboard or grating and an onset-offset checker board or grating.

1.3.1.2 Steady state VEP

Steady state VEP stimuli are unlike the transient VEP stimuli as they are presented faster than 8 alternations per second. The time between presentations is not enough for a complete VEP waveform. The brain responses reach a steady state and the responses begin to appear as series of sine waves. VEP amplitude and time characteristics are defined by power of the response and phase respectively.

Steady state VEP is used to measure VA. In this method, the participant is presented with gratings or checkerboards that range from low to high spatial frequency or check size which decreases from large to small size checks. Once 50-100 steady state VEP responses are recorded for each pattern size, the average VEP amplitude is plotted against check size or spatial frequency and the straight line fit is extrapolated to find the check or grating size that

matches to zero amplitude or baseline noise amplitude[4]. This is defined as the visual acuity limit.

1.3.1.3 Sweep VEPs Technique

Sweep VEPs (sVEP) are based on VEP extrapolation techniques. It essentially utilizes a steady state VEP in response to a rapidly changing or swept stimulus parameter, i.e., the pattern of elements changes rapidly over a short time period. For acuity assessment, a grating target is swept through a series of increasing spatial frequencies to beyond the acuity limit. The main advantage of sweep VEP is that it requires shorter recording times. Measures such as grating acuity and vernier acuity can be obtained by this technique[5].

Overall, there is good agreement in acuities among the different studies which have used VEP in infants, as shown in Table 1.1 and also these studies show how acuity develops with age.

Table 1-1. Acuities obtained by VEP according to the age

Study	Age range	VEP acuity (in cpd)
Norcia et al.[6]	1 month	4.5
Norcia et al.[7]	2 months	2.5-9
Hamer et al.[8]	0.5 -2.5 months	6
Hamer et al.[8]	5-7.5 months	14
Norcia et al.[7]	> 7.5 months	10-20
Norcia et al.[6]	8-12 months	20

1.3.2 Limitations of VEP

Any measure of VEP used to determine visual acuity is affected by varying behavioral states, although this is also true of any method used in infants. VEPs need to use a stimulus that is alternating. The alternating stimuli may be different temporal frequencies for the different VEP measuring techniques. Therefore there is a limitation in that visual acuities may vary for different temporal frequencies of the stimulus. Horizontal nystagmus can affect pattern reversal VEP recordings[9] which may be due to interaction of the retinal image motion with the stimulus[10]. Horizontal stripes[9] may be more effective for VEP recording in the case of horizontal nystagmus. Poor accommodation can result in abnormal VEPs[11]. The other general limitation of this measurement is that it measures resolution acuity and not recognition acuity. Although these are listed as limitations of VEP, it must be noted that the

same drawbacks exist for any other measures of infant visual acuity that use gratings as stimuli.

1.4 Optokinetic Nystagmus (OKN)

OKN is an involuntary, rhythmical, conjugate eye movement that occurs in response to a visual environment that is moving within the field of view. OKN consists of a slow following movement in the direction of the moving field alternating with a quick return saccadic eye movement. The purpose of OKN is to reduce the effect of retinal motion of the retinal image of moving fields.

OKN is elicited by continual movement of a large area of the retinal image. The signals from the retina reach the visual cortex through the lateral geniculate nucleus. The region of the visual cortex corresponding to detection of motion responds and sends a signal to the oculo-motor nuclei, which in turn act on the extra-ocular muscles to produce the eye movements. The eye movement velocity is less than the retinal image. This difference is called retinal slip velocity and serves as a continuous stimulus for further OKN. A mature OKN reflex has a latency of 1/8 of a second and it will be in response to any stimulus in any orientation/direction.

1.4.1 Development of OKN

Monocular OKN in infants less than 3 months old is asymmetric with respect to the direction of movement. The velocity and frequency of OKN is more prominent in the temporal to nasal direction compared to the nasal to temporal direction. Naegele et al.[12] found that by 5 months of age, OKN becomes fairly symmetric, if vision develops normally.

1.4.2 OKN Visual acuity procedure

This test works on the principle that an OKN response will be elicited by a moving field, if the detail in that field can be resolved. The presence or absence of OKN to different grating spatial frequencies has been used as an indicator of VA. Clinical OKN is performed using a drum with a fixed spatial frequency. To obtain different spatial frequencies, the drum would have to be moved to different distances. Typically, however, it is only used at one distance to demonstrate the presence of form vision. In research studies, OKN is performed differently compared to clinical methods and there are a variety of ways that is has been utilised. In the study done by Naegele et al.[12] eye movements elicited to horizontal moving vertical gratings were characterized as OKN. The eye movements were recorded by EOG by placing electrodes bitemporally. EOG (Electro-Oculogram) is a technique in which the electrical potential between the cornea and the retina is recorded. The potential difference between the cornea and retina sets up an electrical field between them. Eye movements are recorded by the change in the field vector as the eye moves. In the Naegele et al.[12] study, the infants were facing towards a large semicircular screen on which the vertical strips of gratings were

projected using a film projector. They found asymmetry of OKN (as described above) in infants' up to 5 months old[12]. The Catford drum is basically a drum that is motor driven and it contains circular targets of different size displayed in an aperture[13]. The smallest target size for which nystagmus cannot be evoked is considered as the threshold[13]. In a study done by Hopkisson et al[14]. using the Catford drum, vernier targets were used as stimuli instead of the usual circular targets. The acuity obtained was compared with Snellen acuity in children between 3 and 12 years of age[14].

1.4.3 Limitations

a. The measure of VA with a moving field may be different from visual acuity as typically obtained under static conditions.
b. If the Catford drum is used, it measures detection acuity; if gratings are used, as is usual, it measures resolution acuity.
c. Binocular OKN can be generated in children with cortical blindness because there is a sub-cortical pathway that is responsible for generation of nasal wards OKN, which dominates the cortically-driven temporal wards OKN in binocular viewing. Van Hof-Van Duin et al. found that monocular OKN was nearly always asymmetrical, showing prominence in the temporal to nasal direction in children with neurological disorders[15]. Of the 21 patients with neurological disorders having a positive visual function, 24% showed symmetry in binocular OKN while 57% showed asymmetry and in 19%, OKN could not be elicited[15].

1.5 Preferential looking

1.5.1 Development of preferential looking tests

The basis of preferential looking (PL) is the infant's visual preference for a patterned stimulus compared to a plain background[16]. Fantz[17] introduced the concept of preferential gaze for testing visual acuity in infants. He showed that infants under 5 days of age consistently looked more at black and white patterns than at plain surfaces. Different pairs of stimuli were presented to the infant. The infant's eye movements and fixation times were observed through a peep hole between the two stimuli. The infant's first fixation was noted and also the number of fixations and the amount of time spent fixating each grating was also noted[16,18]. Fantz et al.[19] showed that pattern vision improved in the first 6 months of life. They showed that at less than 1 month infants had a mean VA of 0.75 cpd and it improved to about 3 cpd by 5 months of age[19]. Fantz et al.[20] overcame infants' periods of inattention by developing a technique that allowed each infant to see 10 grating versus homogenous field pairings. He then combined the data from a group of infants of the same age to establish a mean acuity for that age group[20].

Teller et al.[21] later modified the PL concept into a 2-alternate forced choice (2AFC) test. It was called a forced choice preferential looking (FPL) test. They advocated that the observer should make a decision of whether the grating pattern was on the right or left depending on various cues given by the infant, such as the first fixation, duration of fixation or facial expression etc. The finest pattern at which the observer was able to correctly judge the

position of the grating stimulus 75% of the time was taken as the infant's visual acuity[21]. The FPL technique was also called a double psychophysical method, because it involved both the observer and the infant as subject[21]. The infant was the actual subject who was responding to the stimuli and the observer was also a subject as he/she had to make a forced choice decision for every stimulus. This was a breakthrough in obtaining the acuity in individual infants in a laboratory setting and laid the foundation for infant visual psychophysics. However, it was not useful in a clinical testing, as it was too time consuming to be included into a clinical eye examination. The FPL method is described in more detail below.

1.5.2 Forced choice Preferential looking – FPL

The infant was seated on the lap of an adult facing towards the stimuli. Black and white gratings were presented on a grey screen for the study of visual acuity. The person holding the child was masked to the location of the grating. The infant was seated in such a way that the baby's back rested against the holder. The observer had to make a forced choice decision regarding the position of the stripes. The observer was given feedback from the experimenter (who controlled the stimuli) regarding the correctness of his/her judgment for each trial. The feedback helped the observer to understand the infant's cues better. The role of the experimenter was to organize the stimulus spatial frequency and position for each presentation. The experimenter randomized the different grating spatial frequency levels and documented the responses. Randomization/counter balancing of stimulus conditions across

trials was important for testing the infants. The reason was that infants lost attention and performance decreased if they had a series of difficult stimuli i.e. stimuli close to threshold. The observer had a probability of getting 100% correct for the widest stripe width (lowest spatial frequency) and this reduced to chance level for the narrowest stripe width (highest spatial frequency). In between those two levels, there were decreasing levels of correct judgment towards high spatial frequencies. The results were plotted in the form of a psychometric function. The stimulus value for which there was 75% percent correct was calculated and taken as threshold.

1.6 Evolution of acuity card procedure from FPL

Teller et al.[22], using the FPL technique, showed that visual acuity for vertical and horizontal gratings in human infants was similar, despite the fact that many infants have significant astigmatism[23–27]. The orientation of required looking response was horizontal, requiring horizontal eye movements. In this study, eight individual infants in the age range 1 to 6 months of age were tested. Twenty psychometric functions were obtained between all 8 subjects[22]. Each session lasted up to 1 hour and the infants were seen over a 2 week period. The number of trials obtained per session was 200, although during the earlier sessions they obtained 80 or 100 trials[22]. Thus FPL provided valuable information for research purposes, but was too time consuming to be incorporated in a clinical setting[22]. Banks & Salapatek[28] (1976) and Atkinson et al.[29] extended the PL procedure to study contrast sensitivity in infants 1 month to 3 months old.

1.6.1 Acuity card procedure

The Acuity card procedure (ACP) was basically a variant of FPL technique[30]. It was different from FPL in terms of the observer's judgment. In the acuity card procedure, a subjective judgment of the baby's overall looking behavior was used to make the decision of left or right and the observer decided on the acuity level which gave a clear, correct looking response. In FPL only eye movements were used to make that decision[31], it was truly a forced choice decision and a psychometric function was used to obtain the threshold[32]. McDonald et al.[32] developed the acuity card procedure (ACP) which was a more rapid test for infant acuity testing. ACP visual estimates were obtained within 3 to 5 minutes, as compared to FPL acuities and OPL, which typically took 15-45 minutes. The acuity card procedure was later developed into the commercially-available Teller acuity cards (TACs), described below.

1.7 Teller acuity cards

1.7.1 TAC II (TAC Handbook, Stereo Optical Inc.2005)

The Teller Cards are seventeen in number and each of the cards measures 25.5 x 55.5 cm. They have peep hole at the center of diameter 0.4 cm. The cards have a grey background of approximately 35% reflectance. The gratings are 12 x 12 cm in size, centered in one half of the card and with a contrast of 60-70%. The background luminance is matched with the average luminance of the grating. They were manufactured to avoid edge artifacts by using a half width of a black or white stripe at the edge. The spatial frequencies available in the

Teller cards are 0.32, 0.43, 0.64, 0.86, 1.3, 1.6, 2.4, 3.2, 4.8, 6.5, 9.8, 13.0, 19.0, 26.0 and 38.0 cpd. The step size of the cards is 0.5 octave steps. The cards should be illuminated so that they have a minimum luminance of 10 cd/m^2.

The Teller Acuity Card stage was designed to minimize distractions for the child during acuity testing. It consisted of a large grey screen 79.5 x 74 cm in size with two 79.5 x 36.5 cm grey side panels. At the centre of the stage was a 20 x 47 cm aperture behind which the acuity cards are held during testing. A 35 x 55 cm shield was sometimes suspended in front of the screen, above the child's head, to block the parent's view of the cards. This prevented any bias from the parent, such as unintentionally moving the child towards the patterned stimulus[33].

A second version of the TACs, the TAC II, differed from the original TACs in two important ways. First, the face of the card was laminated, which presumably produces lower contrast. Second, better production technology eliminated the "edge artifact" that appeared as a visible line around the patch of grating on the card in the original TACs, which was a problem at the higher spatial frequencies. This "edge artifact" permitted detection of the location of the grating when the grating itself could no longer be resolved[33]. Clifford et al.[33] presented the results of a study comparing acuity obtained with the Original Teller Acuity cards and the Teller Acuity cards II in 60 children, with 20 children at each of 3 ages: 3.5 months, 11 months, and 30 months. The results of this study suggested that normative grating acuity values obtained by Original Teller Acuity cards™ need to be adjusted toward lower acuity

values by approximately 0.5 octaves to be appropriate for use with the new Teller Acuity cards[33].

1.8 Testing procedures

1.8.1 Operant PL

Visual acuity testing in infants can be challenging, especially in children aged about 1.5 years, who are easily distracted and have short attention span[34]. Therefore, to keep the attention of young children during acuity testing, Mayer et al.[35] introduced the operant PL method. Operant PL combined an audiometry technique that was used to test auditory sensitivity with FPL procedure[35]. The audiometry technique used visual reinforcement for auditory and speech testing in infants 5 to 18 months old[36]. Later, Mayer et al.[37] in another study adapted the same technique to obtain visual acuity for squarewave gratings in 50 children between 5 months and 5 years of age. The methods was as follows: Two identical animated toy animals, one to the left and one to the right of the screen, served as reinforcers[36] (Wilson et al. 1976). They were shown to the child only when the observer responded correctly, indicating that the child gave a correct looking response. Mayer et al.[37] also suggested that OPL can be used as a single procedure to follow acuity development across the age range 5 months to 5 years that they tested. The study showed that visual acuity improved with increasing age over that age span[37].

1.8.2 Procedural variations

There are several variations of the acuity procedure depending on the knowledge given to the observer of the details in the procedure. It can potentially cause bias if the observer knows all the details of the procedure, such as location of the grating, the order and the absolute values of the grating spatial frequencies. Different approaches have been used to try to control bias to a greater or lesser extent, but studies that control for bias the best are more complex to administer. Although there are different methods, the results agreed moderately with PL methods. Some of these procedures are described below.

1.8.2.1 Informed procedure

In this procedure, the observer knew the actual size and position of the grating on each card. A subset of nine cards was kept in the order of relative spatial frequency. The observer noted the position of the grating after presenting it to the child. Visual acuity was the highest spatial frequency for which the observer made a correct judgment of the grating position[38].

1.8.2.2 Dobson procedure

This was similar to the informed procedure, except that the location and size of the grating was not available to the observer. The observer knew only that the gratings increased in 0.5 octaves steps[38]. This makes the procedure less biased for testing compared to the informed procedure.

1.8.2.3 Random procedure

This procedure contained fewest sources of bias. In this procedure, the cards were arranged in a random order, including the blank card. The location of the gratings was marked on the back side of the card. After showing the cards to the infant, the observer arranged the cards such that the spatial frequency went from "seen" response to "non-seen" response as decided from the cues of the child[38].

1.9 Testability

Testability refers to whether a test can be administered sucessfully. A test is considered good for clinical use, only if it has good testability rates. The high rates of successful testing (as shown in the Table 1-2 below) showed that ACP is a potentially useful test to be used in clinical populations.

Table 1-2. Testability of various acuity card studies

	Study	Population	N	Age	Testability
1	McDonald et al.[32]	Typically developing children	64	1-6 months	100
2	Preston et al.[31]	Typically developing children (monocular)	9	24 months	75
3	Hertz[39]	Children with severe cognitive disability	19	22-86 months	95
4	Preston et al.[31]	Infants with ocular disorders (monocular and binocular)	20	2-8 months	100
5	Hertz[39]	Children with Down syndrome (binocular)	33	22-86 months	97
6	Hertz[39]	Children with Down syndrome (monocular)	33	22-86 months	85
7	Mohn et al.[40]	Children with severe cognitive disability and neurological at risk and normal infants (monocular and binocular tests)	510	14 months to 24 yrs	93
8	Hertz et al.[41]	Children with cerebral visual impairment	11	2-12 yrs	82
9	Hertz et al.[42]	Children with cerebral palsy	33	8 to 17 yrs	88
10	Marx et al.[43]	Non-communicative elderly	15	74-96 yrs	87
11	Dobson et al.[44]	Pre term infants (Monocular)	814	12 months	95
12	Hertz et al.[45]	Children with cerebral palsy	77	1-8 years	99

N = number of participants in the study

1.10 Validity

The validity of any given test can be assessed by comparing the results with a standard procedure, whose previous results are already established and considered accurate. Validity of a test can also be assessed by testing with known facts or expectations. For example, it is well known visual acuity improves with increase in age. Similarly it is know that VA is decreased in infants/children with ocular disorders. Tests that show agreement with these expectations can be considered valid. The quality and usefulness of a test can also be measured in terms of repeatability i.e. does the test give the same result when repeated on different occasions or by different testers. This includes intraobserver agreement, interobserver agreement and agreement between similar methods[46].

1.11 Validation studies; Acuity with age and in ocular disorders

1.11.1 FPL procedure

Early FPL studies[37,47] showed that VA improved with age. That finding was one way of validating the FPL procedure. Validation of PL was also undertaken by investigating whether acuity was poorer in infants with known eye disease. Van Hof van Duin et al.[47] studied development of visual acuity using the PL technique in 91 full-term and 36 pre-term infants with minimum perinatal complications. The pre-term infants, when uncorrected for gestational age lagged behind in VA, compared to the full term infants up to 6 to 8 months of

age. When they were compared according to gestational age, the acuities were similar in both groups[47]. They concluded that pre-term infants might have a slight acceleration of development of behavioral acuity compared to normal infants due to early visual experience. Mayer et al.[48] conducted a PL study in 343 pediatric patients ranging from 11 weeks to 5 years old who had a variety of ocular disorders. In general, test results from pediatric patients with structural ocular abnormalities were consistent with the severity of the disorder[48]. Thus, both these studies help to evaluate the validity of FPL technique.

1.11.2 Acuity card procedure

Validation of ACP was done by comparing the results against those obtained with FPL testing as shown in the Table 1-3. ACP involved fewer presentations than FPL but even so, the overall results were found to be comparable to the FPL procedure.

Table 1-3. Agreement between ACP and PL studies

S No.	Method	Age	% within 0.5 octave	% within 1 octave
1	McDonald et al.[32,49,50]	1-12 months and 18-36 months		95-100
2	Preston et al.[31]	2-8 months		75
3	Mohn et al.[51]	3 weeks to 22 years		100
4	Lewis et al.[38]	15-30 months	44-67	66-75

Previous FPL studies[37,47] had shown that VA improved with age and similar results were found with ACP[32,49,52,53] confirming its validity. Jacobson et al.[54] showed poor VA in infants with ocular disorders measured with ACP. Similar results were shown later by Mohindra et

al.[55]. They showed that there was considerable reduction in acuity in children with ocular disease and in contrast, good VA was obtained in unaffected eyes[55]. Preston et al.[31] also used ACP and obtained poor VA in infants with ocular disease. Thus all these studies validate the ACP by showing an agreement with FPL and a relationship between VA and the presence of ocular disorders. The validation of TACs was done by comparing the results obtained with the general acuity card procedure. Many TAC studies[56–58] have established that VA improved with age, which was similar to what was found in ACP studies[32,49,52,53].

1.11.3 Validity of TACs

Similarly to the general ACP, the validity of the TACs was established by comparing the acuities obtained in children with ocular disorders with controls. Luna et al.[59] studied the development of grating acuity in infants who had regressed stage 3 ROP and compared it with healthy pre-term infants. The results showed that the acuities between the two groups were significantly different when children with cerebral visual impairment were included with the group of infants who had ROP[59]. Birch et al.[60] studied the visual outcome in 135 infants (n=185 eyes) with cicatrical ROP. Eyes assigned to the normal/regressed and peripheral retinal changes categories (n = 120) had normal posterior poles. For analysis, infants with amblyogenic or neurologic conditions were removed from the group with normal posterior poles[60]. They found that grating acuities in this group with normal posterior poles were slightly lower than those of age-matched healthy full-term infants[60]. Katsumi et al.[61] studied visual acuity by preferential looking in three different groups of infants with ROP. Normal acuities were obtained in infants with normal maculae or minimal macular displacement while infants with dragged maculae on average showed acuity of 1 to 6 octaves

poorer than their normal counterparts[61]. Progressively worse acuities were obtained in children who had received vitrectomy post-retinal detachment[61]. Thus TACs show good validity in infants with ROP and the visual acuity estimates obtained with the Teller Acuity Cards correlate well with the functional status of the eye.

Good validity can also be established if VA is found to be poorer in children with amblyopia. Drover et al.[62] studied the effectiveness of TACs in detecting amblyopia in 126 children aged 7.8±3.6 years. They compared grating visual acuity measured with TACs and optotype acuity among 3 three different groups namely: a) children with amblyopia (n=45), b) children at risk of amblyopia (n= 44) and c) children with normal vision (n=37). Although grating acuity was better than optotype acuity in amblyopic eyes, TACs yielded a sensitivity of 80% for detecting amblyopia[62]. Similar results were obtained in study by Paik et al.[63] in a group of non-amblyopic and amblyopic children. Thus slightly poorer acuity obtained in children with amblyopia validated TACs.

TACs have been shown to be able to measure acuity in premature infants and populations with atypical development. Lanzi et al.[64] conducted a study involving 35 infants with periventricular leukomalacia. They found that twenty-three infants (66%) presented with visual impairment. Of these, 9 (26%) were totally or nearly totally blind and 14 (40%) were children with low vision. The other 12 (34%) had normal (2) or near normal (9) vision[64]. They concluded that TACs were useful for detecting potential visual impairment and for

improving both the clinical diagnosis of these disorders and the therapeutic approach to these patients[64].

TACs have also been validated by use in populations with visual impairment. Kushner et al.[65] conducted a study of 69 literate patients (age range: 15.5±6.8 yrs.) who had various levels of visual impairment. They concluded that TACs had good specificity but poor sensitivity for detecting visual impairment. TACs might overestimate VA in the presence of visual impairment of 6/21 or poorer and TACs overestimated VA in presence of amblyopia[65].

1.12 Interobserver agreement in ACP

Interobserver agreement refers to the agreement between acuities obtained by two or more different observers in the same population, using the same testing equipment and procedures. High interobserver agreement in ACP showed that the test has good reliability. Since the ACP testing relied on subjective assessment by the tester, it was important to study the interobserver and intraobserver agreement in acuity card studies. These measurements were made for both binocular (shown in Table1-4) and monocular acuities (shown in Table 1-5) in infants and children.

Table 1-4. Interobserver agreement in acuity card studies – binocular acuity

Study	Population	N	Age	% = 0.0 octave	% ≤ 0.5 octave	% ≤ 1 octave
McDonald et al.[32]	Normally developing children	64	1-6 months			92
Preston et al.[31]	Children with visual abnormalities	20	2-28 months	45	95	100
Hertz et al.[41]	Children with cerebral visual impairment	14	2-12 years			77.77
Hertz et al.[42]	Children with cerebral palsy	59	8 to 17 yrs			83
Hertz et al.[66]	Children with cognitive disabilities	44	2 to 7 years			88.6
Heersema et al.[53]	Normally developing children	50	1-4 years		82	92
Dobson et al.[44]	Pre term infants /perinatal complications	52	-7-31days	29	69	85
Marx et al.[43]	Non-communicative elderly	9	74-96 yrs		100	100
Dobson et al.[67]	Pre term infants /perinatal complications	59	-7-31days		70	90
Hertz et al.[45]	Children with cerebral palsy	77	1-8 years			79
Mash et al.[68]	Healthy pre term infants	76	1 month		66	87
Getz et al.[69]	Children with cognitive disabilities	25	3-38 months	36	76	96
	Healthy pre term infants	25		40	72	96

N= number of test pairs

Table 1-5. Interobserver agreement for monocular acuity in acuity card studies and FPL

Study	Population	N	Age	% = 0.0 octave	% ≤ 0.5 octave	% ≤ 1 octave
McDonald et al.[49]	Normal developing children	72	18-36 months			86
McDonald et al.[50]	Normal developing children	66	1-12 months		88	
Preston et al.[31]	Children with visual abnormalities	40	2-28 months		88	95
Dobson et al.[70]	Pre term infants /perinatal complications	382	4-12 months		58	80
Dobson et al.[70]	Pre term infants /perinatal complications	1015	-7-31days		63	85
Mash et al.[68]	Pre term infants /perinatal complications	1918	4-48 months		67	87
Mayer et al.[57]	Normal developing children	460	1-48 months	38	88	99
Getz et al.[69]	Healthy preterm infants (controls)	58	3-38 months	34	79	95
	Children with cognitive disabilities			26	78	91
Atkinson et al.[71]	Normal developing children	12	4 months			100
Maurer et al.[72] #	Normal developing children	57	6-12 months		86	93
Maurer et al.[72] #	Normal developing children	135	18-36 months		86	96
Maurer et al.[72] #	Aphakic children	101	6-36 months		57	79

N- number of test pairs and # refers to FPL studies

1.13 Intraobserver Agreement in ACP

Intraobserver agreement refers to agreement between the acuities obtained by the same tester taken at different times of the day or on different days. High intraobserver agreement also shows the reliability of the test (as shown in the Table 1-6).

1.14 Predictive Validity of TAC

Predictive validity is measured when a test measure is compared with another obtained at a specific time in the future. It is helpful in determining whether the child might go on to develop normal or abnormal acuity depending on the earlier estimate. Courage et al.[56] performed a VA assessment study in 140 healthy infants and children aged 1 week to 36 months using TACs. This was a cross sectional (all participants), as well as a longitudinal study (27 participants). The earlier acuity measurements made at one, three, six months were not predictive of the later estimate at 1 year of age[56]. In contrast, a study done by Mash et al.[73] showed that TACs were able to predict normal versus abnormal acuities. Normal TACs scores at earlier ages were predictive of normal TAC and letter acuity measured with the HOTV test at 48 months in 73 - 95% of eyes. Abnormal TACs scores predicted below normal acuity poor acuity in 39-80% of eyes[73]. Similar results were found by Hall et al.[74] who assessed grating acuity with TAC in 346 infants and children in the age range 3-36 months at risk for visual disorders and a sub-group of 76 children who were reassessed with

TAC after 3-8 years. They found that the initial TAC score had low sensitivity for identifying children with visual disorders but had high specificity in identifying children without visual disorders[74]. To conclude, the predictive validity of TAC is still poorer than the ideal, though it seems that TACs predict normal acuity better than abnormal acuity.

Table 1-6. Intraobserver agreement in acuity card studies

Study	Population	N	Age	% ≤ 1 octave
McDonald et al.[32]	Normally developing children	64	1-6 months	88
Hertz[39]	Children with severe cognitive disability	17	22-86 months	47
Hertz[39]	Children with Down's syndrome (binocular)	25	22-86 months	80
Hertz[39]	Children with Down's syndrome (monocular)	22	22-86 months	64
Hertz et al.[42]	Children with mild motor handicap	39	2-7 yrs	83
Hertz et al.[42]	Children with severe motor handicap	20	2-7 yrs	65
Hertz et al.[66]	Children with cognitive disability	42	8 to 17 yrs	83
Hertz et al.[66]	Children with cerebral palsy	120	1-8 yrs	66,72
Mash et al.[75]	Pre term infants/perinatal complications	233	2.5-18.5 months	91

N refers to number of test pairs

1.15 Limitations of the preferential looking procedure

FPL technique is susceptible to a major limitation. An infant's fixation on a stimulus can be taken as the child being able to resolve the grating. On the other hand, an infant's non-fixation on a target cannot always be taken to mean that the target is unresolvable, as it might be a combination of the infant's attention, cognitive or motor aspects that makes them uninterested in the grating target. Thus the lower limits of VA obtained must be interpreted with caution.

There is a potential for bias in judgment of seen and not seen responses if the tester knows some information regarding the patient e.g. the refractive status or a previous acuity card test result. Because the acuity card procedure is conducted at relatively short distances, it is less sensitive to detecting reduced VA due to myopic refractive errors than distance letter acuity for children. The orientation of the card affects the visual acuity measurement in astigmatic infants. Thus, TACs alone are not a very effective screener for all refractive errors.

The principle of the acuity cards is that the front surface of the card should appear uniformly gray when the grating is not resolved. Stains/imperfections on a card can result in an unwanted eye movement towards it, which could be interpreted incorrectly by the observer to indicate that the child resolved the grating or did not detect it, depending on the position of the blemish. Unfortunately, the surface of the original TACs was such that cards easily

became tarnished by dirt, finger prints or scratches, which meant that each acuity card had to be examined frequently and flawed cards needed to be replaced. However, TAC II cards were laminated on the front surface and thus were less vulnerable to marks [46].

As with VEP and OKN, FPL also has a disadvantage that it measures only resolution acuity and not recognition acuity. Resolution acuity is less sensitive in detecting amblyopia. The advantage FPL has over VEP and OKN is the stimulus/target, which is a static grating, unlike the other two methods.

Clifford et al.[34] studied the effect of using the testing stage on visual acuities norms obtained by TACs. Eighty children were tested, 20 each at age 3.5 months (94 to 117 days from due date), 11 months (320 to 363 days from due date), 17 months (502 to 534 days from due date), and 30 months (882 to 944 days from due date). They reported that differences between acuity obtained with and without the testing stage were significant in the 17 month age group. Infants around 1- 2 years tend to be easily distracted and the better acuity levels with the testing stage support that view. They concluded, however, that in all the other age groups there was no difference with or without the stage[34].

1.16 Development of visual acuity

There are many studies which have investigated the development of VA in the first year or first few years of life. Most of the studies showed that VA develops rapidly in the first six

months and then more slowly. Results using the acuity card technique are shown in Table 1-7. The different study populations, sample size, exact age range, method used, whether the acuity was monocular or binocular may have contributed to the differences in the acuities obtained in different studies. But the overall trend of improving acuity with increase in age is evident from all these studies.

1.17 Large sample acuity norm studies

Clinically useful norms of the TACs can only established if the studies include a large population with a wide range of age groups. Courage et al.[56] measured binocular TAC acuities in children between 1 week and 36 months of age. There were 7 groups; each consisted of 20 healthy infants. They found that newborn infants had poor vision (mean of 0.9 cpd) and VA rapidly increased up to 6 months. There was an improvement of acuity of 1 octave every 3 months during that period. VA improved further by 0.5 octaves at yearly intervals after that, at age 12, 24 and 36 months[56].

Mayer et al.[57] measured monocular acuities in 460 healthy children aged between 1 and 48 months of age. Mean acuity improved rapidly from 1 month to six months. Thereafter acuity developed slowly between 6 and 12 months of age. The mean acuity at 1 month was 1 cpd and at six months it was about 6 cpd, which was a 2.6 octave increase.

Table 1-7. Mean acuities (in cpd) and SD (in octaves) obtained according to study and age group

	New borns	Age (in months)										
		1	3	6	8	9	10	12	18	24	36	48
Courage et al.[56]	0.9±0.5	1.1±0.6	2.6±0.6	5.9±0.6				9.6±0.3		13.2±0.5	18.6±0.5	
Dobson et al.[30]	0.7±0.5											
Brown et al.[52]	1.0±0.5											
McDonald et al. a[32]		1.1±1.1		4.7±0.8								
McDonald et al. b[50] #		0.8±0.7		5.3±0.5				6.3±0.7				
Van Hof van Duin et al.[47]		1.3±1	4.1±0.6	7.8±0.5				10.2±0.5				
Mayer et al.[57]#		0.9±0.4	2.2±0.4	5.7±0.47		6.8±0.4		6.4±0.3	8.6±0.39	9.6±0.3	21.8±0.4	24.8±0.3
Salomao et al.[58]			3.9	7.4	9.8		11.6	11.1	12.39	14.64		
Kohl et al.[76]#								6.4±0.3		20.9±0.4		
Heersema et al.[53]								17.3±0.3		18.8±0.4	28.3±0.4	
McDonald et al.[49]										14.9±0.6	27.7±0.5	

- Monocular acuities

The rate of increase per month during the first five months was about 0.5 octaves per month, whereas over the age 6 months to 48 months it was 0.05 octave improvement per month. The mean acuity at 48 months of age was about 25 cpd[57].

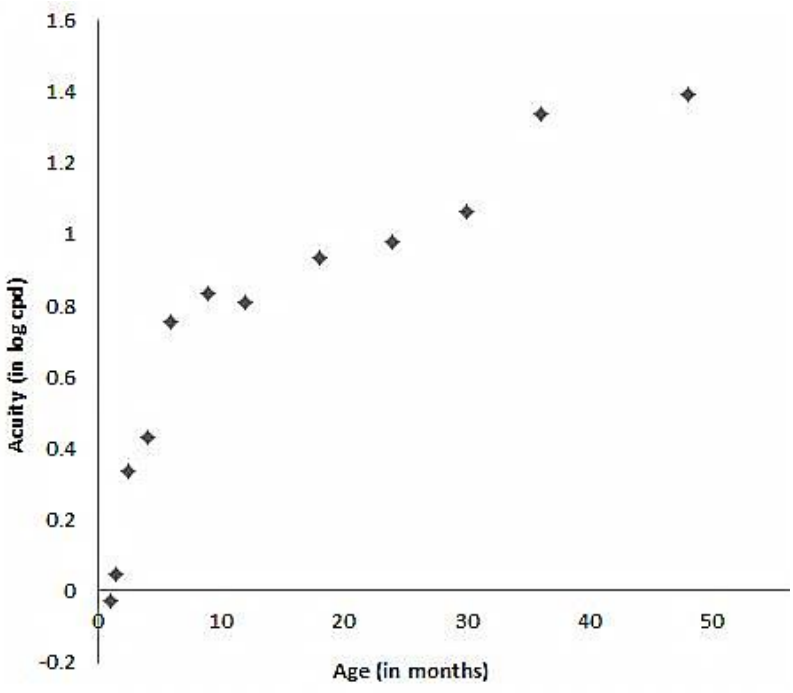


Figure 1-2. Visual acuity development with age

(derived from Mayer et al.[112])

Salamao et al.[58] measured binocular and monocular acuities in 646 healthy children aged between 0 and 36 months of age. Both monocular and binocular VA showed a sharp improvement in acuity from birth to 6 months of age and slower growth thereafter. There were significant differences in acuity between the 2 and 4 month age groups and as well between the 4 and 6 month groups. There were no significant improvements after 6 months

and among the higher age groups[58]. In a review of the literature, Leat et al. [77] concluded that VA became adult-like in children between age 5 years and mid teenage years.

1.18 Comparison of VA development between acuity methods

All the techniques, including VEP, OKN and PL, show that VA increases from birth to 6 months of age. The amount of improvement shown was different between the behavioral methods and the VEP method. VEP shows an improvement of 5 octaves over this age range, whereas PL and OKN show 3 and 2 octaves improvement respectively[78]. Most newborns tested with OKN demonstrated at least 1.5 cpd and by 2 months of age all infants demonstrated acuity of 2 cpd[79]. There was a rapid period of improvement between 2 months and 6 months, when it reached 6 cpd[19]. The PL procedure demonstrated similar levels of VA and improvement to OKN acuity within the first six months[78]. Marg et al.[80] showed that infants responded to VEP stimuli and demonstrated similar acuity of about 1.5 cpd at one month of age but this developed more rapidly than OKN or PL acuity to reach 30 cpd at 5 and 6 months of age. Sokol et al.[81] conducted a study in infants to compare the VEP and FPL data in 26 infants aged 3 months. VEP showed 1 to 2 octave higher acuity compared to FPL estimates.

Riddell et al.[82] also found that sVEP acuity developed at a higher rate than TAC acuity. Additionally, they found that TAC acuity reached sVEP acuity by 14 months. TACs and sVEP results were in agreement in showing no difference in acuity between pre-term and full-term infants. However, the agreement was poor between VEP and PL methods in

children who had neurological disorders[83]. Children with periventricular leukomalacia, the most common type of brain damage in preterm infants, had slightly poorer VEP acuities compared to TAC acuities[83].

Differences between these techniques may be because the information available for the brain to process at the various levels of the visual pathway is different for each of the methods[82]. The VEP response is obtained from the visual cortex, whereas the PL response would also include non-visual areas of the brain[82]. The maturation of the different parts or pathways of the visual system at different times might also explain why each of these methods gives rise to different acuities and different rates of development[78]. The discrepancies between FPL acuity and VEP acuity can be explained as follows:

1.18.1.1 Stimulus Differences

The important difference between behavioral measures and VEP is the stimuli presentation. Behavioral acuity with PL is usually assessed with stationary gratings, while the VEP, by necessity, uses temporally modulated stimuli. Sokol et al.[84] showed that PL acuity obtained with stationary gratings in infants was similar to VEP acuity with gratings of temporal frequency 3.5 or 7 Hz. Sokol et al.[85] in another study, found that acuity at either of these frequencies was 0.5 to 1 octave higher than the stationary gratings or VEP acuity with higher temporal frequencies[85]. Sokol et al.[86] also found that VEP and PL developed at a different pace, but reached similar levels of acuity at 12 months of age. They reported that the PL acuity of infants older than 5 months was at least one octave higher for stimuli reversing at 7

Hz than for stationary stimuli, although VEP acuity was even higher in older infants[86]. They also found that at 2 years, PL grating acuity was no longer tuned temporally[86].

1.18.1.2 Threshold

Behavioral measures of acuity generally apply a rather strict threshold criterion of 70 or 75% correct responses (of the observer), while the zero μV amplitude criterion used for VEPs is clearly more generous[87]. Behavioral measures improved reasonably by adopting a 50% criterion[87], although this is a questionable criterion, being at chance level. Mayer and Dobson[37] found slopes of the psychometric functions were steeper until three years of age and inspection of their data suggested acuity estimates would have changed much less with age for a 55% criterion than the adopted 70% criterion.

1.18.1.3 Retinal Area

There is evidence that VEP acuity is determined by a very small retinal area of no more than 2^0 even in two or three month old infants[88]. On the other hand, behavioral acuity of infants with foveal abnormalities had VA only slightly lower than that of normal infants, and often within the normal range[89]. This suggested that normal PL acuity may reflect parafoveal function[89].

1.18.1.4 Motivation

VEP provides an objective indicator of cortical functioning, whereas PL acuity reflects what the infant choses to look at. There is direct evidence that infants stopped responding behaviorally to stimuli which still elicited a clear VEP in the same child[90]. There was a difficulty of eliciting a PL response in older infants which led to the development of operant method of preferential looking[87].

1.18.1.5 Neural Basis

It seems reasonable to assume that the basic sensory neural mechanism underlying VEP and behavioral acuity are the same[87]. However the behavioral responses clearly involve additional areas of central nervous system beyond the primary visual cortex[87]. This may also explain some of the differences in acuity obtained by the acuity card procedure and VEP.

1.19 Development of other aspects of visual function

1.19.1 Contrast sensitivity

There were many studies which have investigated contrast sensitivity development in infants. Banks and Salapatek[28] used the PL technique to assess the contrast sensitivity function (CSF) in infants. The final threshold was obtained by the lowest contrast stimulus which the infant fixated reliably for a range of spatial frequencies[28] They found that at 2 months of age infants have poor contrast sensitivity. The area under CSF was small compared to adults[28] and the

peak of the response shifted towards the left, i.e. towards lower spatial frequencies[28]. Atkinson et al.[29] also used PL to measure CS and showed that the shift of the peak towards higher spatial frequencies occurred during the period when there was rapid development of VA, i.e., between 5 weeks and 8-12 weeks of age. Although there was a rapid improvement in acuity and contrast sensitivity, infants at age 8-12 weeks have not yet reached adult level[29]. At high spatial frequencies, infants who had high contrast sensitivity were able to respond to a contrast level of 7%, while adults were able to respond to a contrast level of 0.5%[29].

VEP techniques showed similar development in CSF as observed with the preferential looking techniques. Norcia et al.[7] showed that contrast sensitivity at all spatial frequencies improved by a factor 4 or 5 between age 4 and 9 weeks of age. They demonstrated that CS remained constant for low spatial frequencies and continued to increase for high spatial frequencies over the first 9 weeks of age[7]. CSF continued to develop at least until about 7 months of age[7]. In summary, CSF develops rapidly in the first six months and continues to develop gradually in the next few months as the visual system slowly increases in sensitivity towards the higher spatial frequencies[2]. The contrast sensitivity in children matures fully to adult like between 8-19 years of age[77].

1.19.2 Saccades

Saccadic eye movements are those eye movements which allow quick fixation on objects of interest in the visual field, so as to place the image on the fovea. They can be voluntary as

well as involuntary. Saccadic eye movements are not fully mature until 12 months of age[91], although Irving et al.[92] have shown that saccadic velocity and latency continue to develop through childhood and early adolescence (up to age 14 years). Infants make small amplitude fast eye movements in the right direction, but not of sufficient amplitude to fixate the new target. These movements are called hypometric saccades[2]. Multiple hypometric saccades are made by infants <3 months of age for large object jumps. Hypometric saccades are commonly also seen in healthy adults and in infants greater than 1 year, but they cover at least 90% of the distance from the object[91] although some saccades cover 100%. This development of the saccadic system to enable a change of fixation to an eccentric target by a single saccade is referred to as normometria[93]. During the first 7 months of life, infants tend to progress towards normometria. Saccadic latency is up to 1 second in healthy infants, whereas in adults it is up to 200 ms[91].

1.20 Conclusion

In conclusion, clinical visual acuity assessment in infants is critical in identifying any ocular disorders. VA is significantly affected in children with ocular disorders. Infants do not respond subjectively to VA, so therefore it is important to have a test that accurately measures VA objectively. TACs can be a versatile tool to assess visual acuity in various population groups in typically developing infants and children and also in infants and children with atypical development such as preterm infants or infants with retinopathy of prematurity, amblyopia and other ocular disorders. Nevertheless there are some

disadvantages with TACs. It still involves subjective components and there is a possible bias involved in the procedure. Therefore, there is a need for a less subjective/more objective method for assessing VA in infants.

Chapter 2 - Aims and Hypotheses

2.1 Purpose of the study

Teller acuity cards (TACs) were developed on the principle of preferential looking [17]. Although, TACs are widely used for measuring VA in infants, there is a disadvantage inherent with the use of preferential looking techniques. The interobserver agreement for measuring VA can be variable due to different clinical skills and experience of the examiner. A novice examiner may overestimate VA compared to an experienced tester. The use of the stage (which reduces distractions) is not common in the clinical testing with acuity cards. Therefore, there is another potential source of bias - if the examiner is able to see parent's eye movements and this might bias the examiner regarding the position of the grating. Quinn et al. (1993) showed that there was a slight disagreement in acuities between three testers in a clinical setting. So there is a potential need for a more objective method for measuring VA in infants.

The Gaze tracker (GT) system, developed by M. Eizenman (University of Toronto), tracks an infant's fixations during the presentation of a grating. The fixation information provided by the software gives objective information of the infant's gaze in space, which assists the observer in deciding whether the infant is fixating the grating target or not. The system as currently used is not completely objective, both in terms of judgment of fixations and in presentation of the stimuli. Potentially, however, both of these aspects could be fully

automated with development of specific algorithms, so that fully objective VA testing in infants is feasible in the future.

The current study was designed with the following objectives

1. To validate the gaze tracker measurements of VA by comparing them with the TACs in adults and to identify the best and most efficient protocol for obtaining VA in adults. The study was initially done with adults as they are more cooperative and have longer attention spans than infants, thus allowing the collection of more data. If unreliable results are obtained in adults, it would be unlikely that the system would work in infants.
2. To optimize the system for infants and demonstrate that the system can obtain a VA measurement in infants.
3. To validate the GT instrument in infants by comparing the GT measurements with TACs.

2.2 Hypothesis of the study

Adult study

1. There will be no difference in measured visual acuity obtained by the two methods namely, Gaze Tracker and TACs.
2. There will be a significant correlation between uncorrected refractive error and visual acuity in the adult population.

Infant study

1. There will be no difference in measured visual acuity obtained by the two methods namely, Gaze Tracker and TACs.
2. GT measurement of VA will have no difference in percentage of repeatable measurements as compared to TACs.
3. Visual acuity, as measured by the GT, will improve with age in the infant group.

2.3 General approach

This was an experimental study. Participants were healthy infants and young adults. VA was measured with GT as well as with TACs. VA was measured on 2 occasions in infants, so as to determine test-retest repeatability.

Adult study

Adults were kept naïve regarding the purpose of the study and given no instructions except to look at the screen, in order to mimic, as far as possible the behavior of an infant. However, as they are capable of subjective responses, it was possible to validate the GT against their subjective responses. Similarly both objective (VA based on naïve looking responses) and subjective acuity was measured with the TACs. The participants were all uncorrected myopes, so as to obtain different acuities which can be correlated with the refractive error.

Infant study

An analysis was done to determine agreement and correlation between GT and TACs. An analysis was done to examine the repeatability of GT measurements in comparison with the repeatability of TACs. The analysis also included correlation plots between GT and age.

The long term goal is to automate the VA testing in infants, for example using software which uses predefined algorithms to detect the eye fixation information and be able to obtain VA thresholds that are unbiased by subjective judgments. The data collected in this study is expected to be useful in developing these algorithms.

Chapter 3 – General methods

3.1 General principle

The instrument that we used for obtaining an objective measure of VA was a remote non-contact gaze tracker, developed by Guestrin and Eizenman[94]. It estimates point of gaze (POG). POG is the point in space imaged on the fovea of each eye. It works on the principle of tracking the eye using video cameras and then the eye tracking software captures the eye movement information. Light sources that illuminate the eye produce corneal reflections. The front surface of the cornea is illuminated by 3 infra-red lights and the virtual images of light sources appear as corneal reflections. There are two important reasons for using IR light. Firstly, the participant is unaware of the light sources and they do not cause any discomfort and secondly, tracking is not affected by the lighting condition[95]. One or more video cameras image the center of the pupil and the corneal reflections[96]. These are the landmarks used for eye tracking plus the iris-pupil border (Fig. 3-1.)

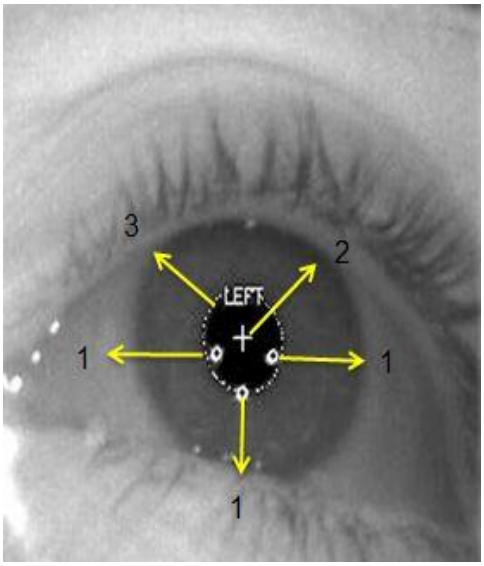


Figure 3-1. Tracking landmarks

1. Corneal reflections, 2. Centre of the pupil and 3. Iris-pupil border

The image co-ordinates of corneal reflections and the pupil center are captured by the video cameras[97]. These image co-ordinates are then converted into real-world co-ordinates to determine the position of the visual axis[97]. The average co-ordinates of the parameters, such as the corneal reflections and pupil center of each participant, are used to measure the eye parameters such as the corneal curvature, distance between the pupil centers and the corneal curvature, and the horizontal and vertical angle between the visual and optic axes[97]. These parameters are eventually used to determine the POG[97].

3.2 Previous literature - Eye tracking used to measure VA in adults

Using a similar gaze-tracker, Sturm et al.[98] objectively measured the visual acuity of 9 naïve adult participants using an objective assessment of visual scanning parameters as the participant was looking towards a uniform grey field with black and white square-wave gratings in one quadrant. A four alternate forced choice (4AFC) subjective test was performed later. Seventy two percent of the relative fixation time on the gratings was found compared to other regions of the screen for the gratings that were 100% reliably discriminated subjectively[98]. The probability density functions of the relative fixation times were computed to calculate the average VA[98]. There was 100% testability in the trials done in adults. The authors suggested that quantitative analysis of visual scanning parameters is more objective and may be more accurate than the traditional clinical PL procedure for measuring VA in pre-verbal patients[98].

3.3 Technical information of the remote non-contact gaze tracker

The gaze tracker, developed by Guestrin and Eizenman[94], was shown to have accuracy of ± 0.5 degrees for a participant sitting 70 cm in front of a 21" display. The spatial resolution of the gaze tracker was better than 0.1 degrees. The sampling frequency of the gaze tracker is 30 Hz. It allows head movements in the range as shown in the Table 3-1.

Table 3-1 Range for head movement which allows tracking in the eye tracker

Axis	Minimum value (in mm)	Maximum value (in
X (laterally)	-300	300
Y(vertically)	-250	250
Z (forwards/backwards)	-150	150

It works on the principle of tracking the eye using video cameras and then the eye tracking software captures the eye movement information as described above. The eye movement data are used to make judgments for obtaining a visual acuity threshold.

The instrument consists of two 21 inch LCD monitors, 7 infra-red light sources and 2 video cameras. Monitor 1 (marked as 1 in Figure 3-2) displayed the eye tracking software information which gave information about the stimulus, eye tracking and fixations. Monitor 2 (marked as 2 in Figure 3-2) displayed the stimulus (horizontal gratings) for the experiment, which could be alternated with videos in between stimuli. Monitor 2 was clamped to a movable table (marked as 6 in Figure 3-2) so that the distance between the cameras and the monitor could be adjusted. There were 7 potential light sources (marked as 3 in the Figure 3-2), although only three were used in the current study. There were two video cameras which were located 70cm (marked as 4 in the Figure 3-2) from the participant and below their line of sight. This distance was fixed. The same instrument was used for both adult and infant studies. Initially, the validity of the instrument in adults was determined and the results of

this study are presented in Chapter 4. Later, the validity in infants was tested and results are discussed in Chapter 5 and 6.

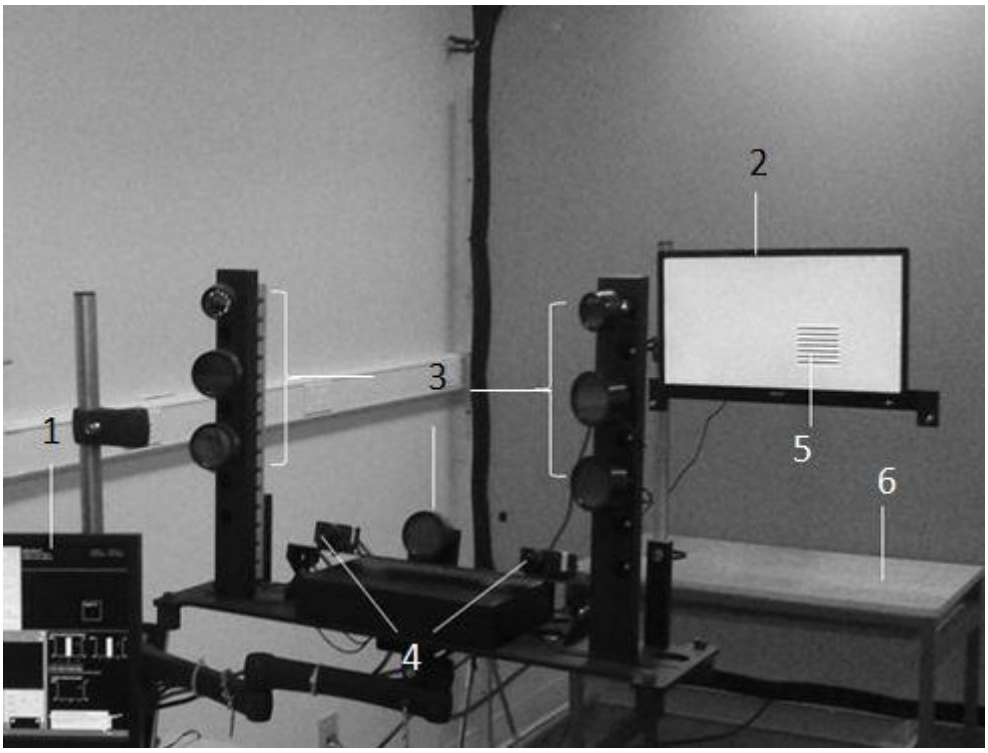


Figure 3-2. Gaze tracker instrument

1. Observer screen 2. Stimulus presentation screen 3. Infra-red light sources 4. Two video cameras 5. Grating stimulus 6. Movable table

There is a program called VAST that actually launches the eye tracker software. Initial data such as name, age and sex of the participant were entered. Then the test distance and size of the targets had to be selected. There were two different grating areas for the target. Small targets at 210 cms were chosen for the adult study while large targets at 70 and 120 cm were chosen for the infant study. In the infant study, the target area in the GT as a proportion of

the screen was 5 to 9% larger than the target area of the TACs as a proportion of the card. The larger targets (compared to those used for adults) were used for the infants in the GT, in order to ensure that they made a looking response.

The observer screen gave information about the location of the eyes with respect to the view of the eye tracker. This information was used to determine if the participant was at the right height and distance from the eye tracker (as shown 3, 4 and 5 in Fig.3-3). If the subject was at the right distance and height from the eye tracker, then the tracking began. This was seen in upper part of the monitor, which displayed instant video captures of both eyes by the two cameras namely camera 0 and camera 1 (as shown in the 1 and 2 of Fig.3-3). This gave an indication to the observer of whether the instrument was detecting the eyes, i.e., landmarks for eye tracking were determined. Once the eyes are detected by the tracker, calibration is done. The status of the calibration was displayed on the right of the screen (as shown 6 in the Fig.3.3).

The observer monitor also contained two important pieces of information used to judge if fixation was made on the grating as shown in Fig.3-4. This was the information of real time tracking of the eyes (red and blue dots) and the green-white histogram information, which gave information about the percentage of time the participant spent fixating the grating compared to the time spent on the blank region of the screen.

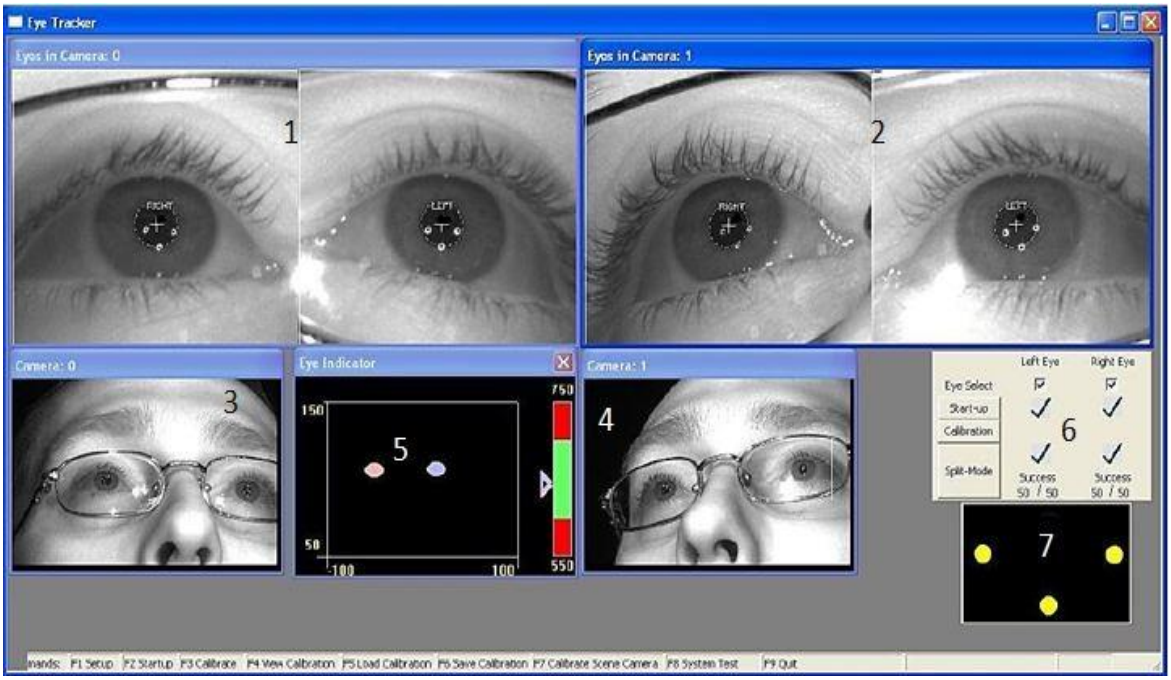


Figure 3-3. Eye tracker experimenter screen with an adult participant

1 and 2. Two eyes tracked by two cameras 0 and 1 respectively. 3 and 4. Position of the eye with respect to the view of eye tracker. 5. Markers for right and left eye. The arrow mark shows the distance of the participant from the cameras. The green zone indicates the range of distance within which the eye tracker can track successfully. 6. This window indicates if the eye tracker is able to detect and track the eyes and it also indicates if the calibration was done. 7. The last window represents the number of light sources that are used. Image courtesy: Eye tracker manual

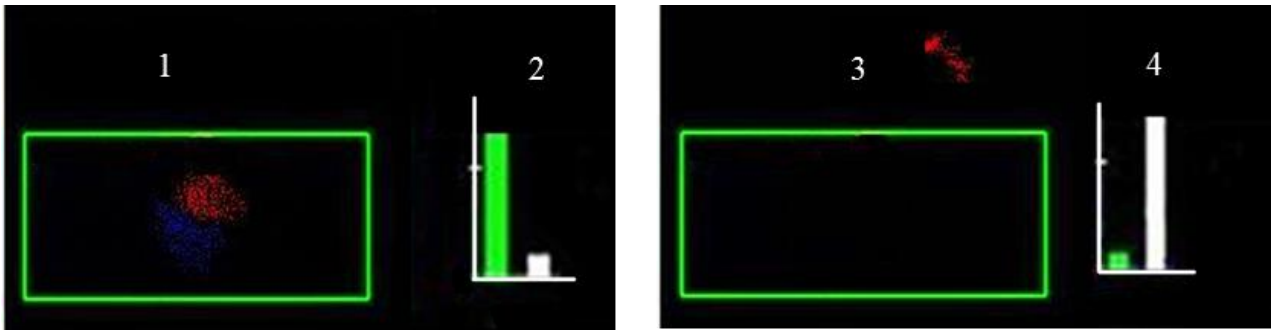


Figure 3-4. Gaze tracker information display

1 and 3 gives real time gaze tracking information of both eyes while 2 and 4 are histograms which show the percentage of time spent on the grating compared to the other regions of the screen. The blue and red clusters represent fixations of the right and left eye respectively. For the histogram, green bars represent the relative time fixating on the grating compared to the other regions of the screen. The white bar represents the time fixating outside the grating compared to the grating. The actual criteria for judgement using this information will be discussed in Chapter 4, 5 and 6.

3.4 Calibration of eye tracking in infants

Calibration procedures require participants to fixate on definite point(s) at definite periods in time. The calibration used in the gaze tracker is unique compared to other eye trackers, as it estimates the angle kappa between the visual and optic axes to determine POG[99]. Previously in the literature, studies have not estimated angle kappa, but either assumed that the visual and optic axes coincide[100] or the midpoint of intersection of the two optic axes[101] were considered to estimate POG. However, it has been shown that either of those methods leads to error in estimation of POG[102].

For the infants, the calibration uses 5 different small cartoons which moved to 5 different positions sequentially. The calibration procedure stops as soon as a good calibration is

achieved and repeats if the criterion for calibration is not achieved. So for some infants, only 3 or 4 of these calibration targets may displayed and for others all the five repeated sequentially until POG was determined. The calibration basically works on the probabilistic approach that an infant will be more likely to fixate on the target rather than the uniform background which makes it uniquely suitable for infants[99]. Infants do not need to fixate continuously and accurately throughout the time of calibration presentations[99]. In this method, POG can be calculated accurately even if the infant fixates 50% of the time on the calibration targets[99] Researchers in Eizenman's team[99] have developed an algorithm to determine the point of gaze by identifying clusters of fixations. The algorithm identifies a cluster by observing consecutive eye movements towards the same point or/and fixation on the same point for several consecutive time samples when a calibration target is presented[99]. A response which is analyzed as a fixation assumes that the fixation made by the infant was towards the calibration stimulus[99]. The probability of algorithms identifying a false cluster distribution correlating with a seen response is 0.01 %[99].

3.5 Stimuli and Spatial frequency

The stimuli used were horizontal square-wave gratings. The resolution of the computer monitor was a limitation for creating vertical gratings. In the adult study, the grating area was either approximately square or a vertical strip grating. The grating area was either rectangular or a vertical strip in the infant study. The target areas used were slightly larger in the infant study compared to the adult study. The details of range of spatial frequency used and size of

the stimuli used in the adult and infant study is tabulated in the Table 3-2 and Table 3-3 respectively.

Table 3-2. Set of spatial frequencies used for adult study

Sp.f. (in cpd)	Size (in cms)
0.58	25.4 x 3.2
0.82	8.9 x 8.9
1.16	9.6 x 9.6
1.64	9.0 x 9.0
2.31	25.2 x 3.7
3.22	8.0 x 8.0
4.62	8.2 x 8.7
6.73	8.3 x 8.3
9.25	8.4 x 8.2
12.33	26.8 x 3.6
18.50	8.6 x 8.6
24.67	8.5 x 8.5
37	26.8 x 3.7
74	26.8 x 3.8

Table 3-3. Set of spatial frequencies for infant study

Testing distance = 70 cms		Testing distance = 120 cms	
Spf (in cpd)	Size (in cms)	Spf (in cpd)	Size (in cms)
0.2	25.1 x 8.1	0.18	24.1 x 8.2
0.28	21.3 x 8.8	0.25	25.6 x 8.1
0.4	21.1 x 9.4	0.34	25.2 x 8.1
0.55	26.8 x 8.1	0.48	19.6 x 8.9
0.78	21.3 x 11.1	0.67	25.0 x 8.7
1.13	21.3 x 10.8	0.94	19.5 x 9.0
1.56	26.2 x 8.1	1.37	19.6 x 9.3
2.27	21.2 x 11.0	1.84	19.5 x 8.7
3.12	21.2 x 11.0	2.65	26.1 x 8.1
4.15	21.3 x 11.0	3.85	19.5 x 9.2
6.23	21.3 x 11.0	5.3	19.5 x 9.2
8.31	21.1 x 11	7.07	19.6 x 9.2
12.46	26.8 x 8.4	10.6	19.6 x 9.0
24.93	26.8 x 9.2	14.13	19.5 x 9.6
		21.2	26.8 x 9.1
		42.4	26.8 x 9.2

3.6 Summary

The fact that the gaze tracker allows for some head movements and has quick calibration[97] makes it distinctively suitable for use in infants. The GT appears to have the right capabilities in terms of hardware and software programming to track an infant's eye and thus be capable of objective measurement of VA in infants. These capabilities are; a) ability to calibrate by the use of probability functions, which means that it is not necessary for the infant to fixate the calibration stimulus all the time, but still allowing an accurate estimation of the POG. b) ability to track eye movements within a tolerance of head movement in range as mentioned in the Table 3.1. We are not aware of any other systems which have this combination of attributes and are so suited to the purpose. It has been shown that the gaze tracker can be used to measure VA in adults and also shown that it is possible to be determine the point of gaze in infants[94]. Thus the general method of using the gaze tracker as an objective measure for measuring VA in infants is feasible. The actual protocols to obtain VA in adults will be described in detail in the Chapter 4 and infant study protocols will be described in Chapter 5 and 6.

Chapter 4 – Adult study

In the current study, the aim was to validate the gaze tracker against the standard grating acuity test used in infants (TACs) but with adult participants, and to compare VA measured objectively with the gaze tracker against a subjective method. The secondary aim was to compare grating visual acuity obtained by different objective protocols and their relative efficiency. Adult participants were used this study, to demonstrate the potential of gaze tracking; if VA measurements do not agree with TACs in adults, they are unlikely to do so in infants.

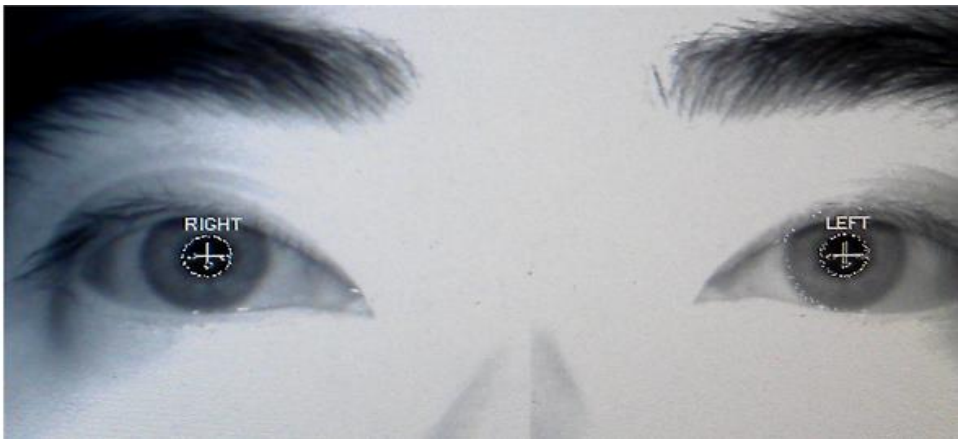


Figure 4-1. Landmarks for eye tracking in an adult participant

4.1 Methods

4.1.1 Subjects

Fifteen naïve adults (six women and nine men) were recruited from among the students and staff at the University of Waterloo, Canada. Subjects' ages ranged from 22 to 47 years (mean 28.47 years). This study was reviewed and obtained clearance through the Office of Research

Ethics at the University of Waterloo before commencement, and all patients gave informed consent before participating in the study. Subjects with a history of ocular disorders other than refractive error were excluded. Inclusion criteria were corrected visual acuity of at least 20/20, no strabismus, no history of ocular disease or surgery, and myopia with a spherical equivalent ≥ 1.50D to 10 D. The test was done binocularly in the uncorrected refractive state to obtain a range of acuity values. It was performed binocularly, as it was anticipated that the testability and validity of the GT in measuring infant VA would be assessed binocularly. Therefore, adults were also tested in binocularly. Myopes were chosen, as they will not accommodate in the uncorrected state and therefore the error of focus is known.

4.1.2 Experimental Set up and Visual Stimuli

Horizontal square-wave grating stimuli were presented on the 21 inch LCD computer monitor using the Gaze tracker as described in Chapter 3. The spatial extent of each grating stimulus consisted of either a square or vertical strip of grating. The sides of the square targets ranged from 8 to 9.5cms while the size of the vertical strip targets ranged from 3.7 x 25.6 cms to 3.7 x 26.6 cms. The vertical strip target areas were used for higher spatial frequencies in order to avoid edge artifacts (the strip extends to the top and bottom of the screen). The vertical strips were also occasionally used for lower frequencies to accustom the participants to seeing it. The square gratings were presented in one of the four quadrants (see Fig. 4-2.) and the vertical strip gratings were presented at one of four regions namely far left, far right, just right and just left of the center of the screen (see Fig. 4-3.). The gratings randomly appeared at one of the 4 positions. Luminance was measured with a luminance

meter (Minolta Chroma Meter CS 100). The Michelson contrast of the gratings was 98.5% and the mean luminance matched the background, which had a luminance of 67 cd/m^2. At this distance, the range of spatial frequencies used in the gaze tracker was similar to those in the TACs; 0.32 to 38.0 cpd in half octave steps (see Chapter 3 for the full range of spatial frequencies).

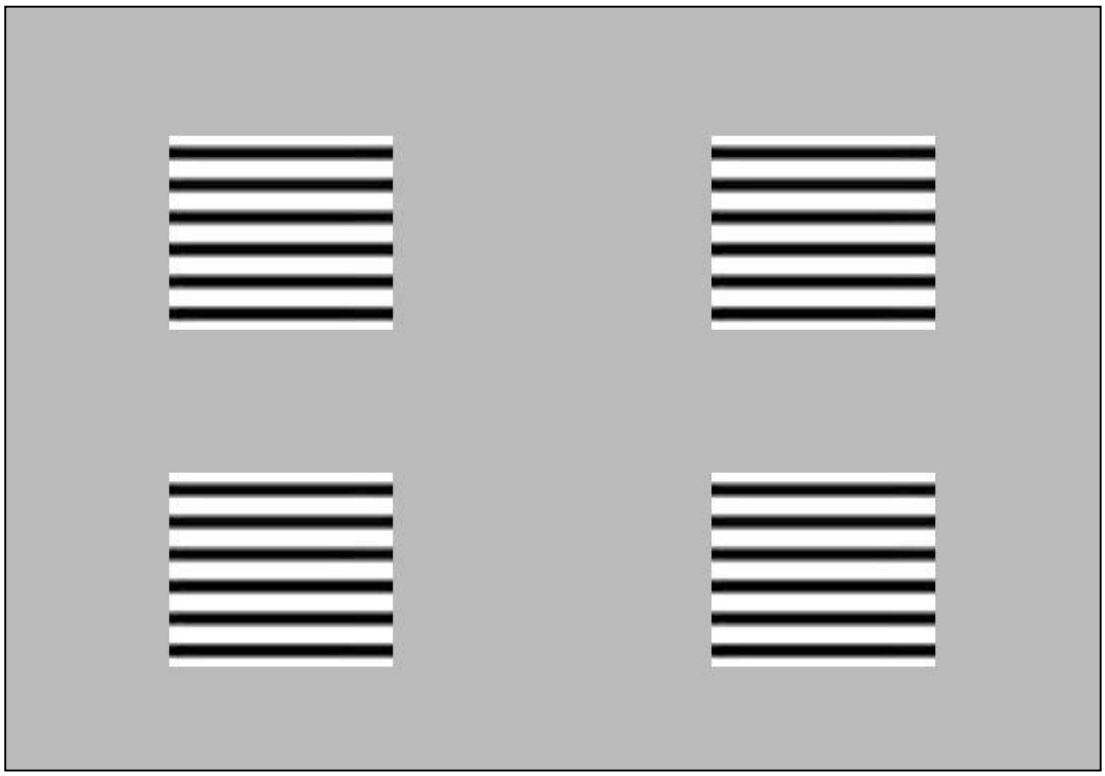


Figure 4-2. Presentation screen with square gratings.

The stimulus would appear randomly in one of these 4 positions.

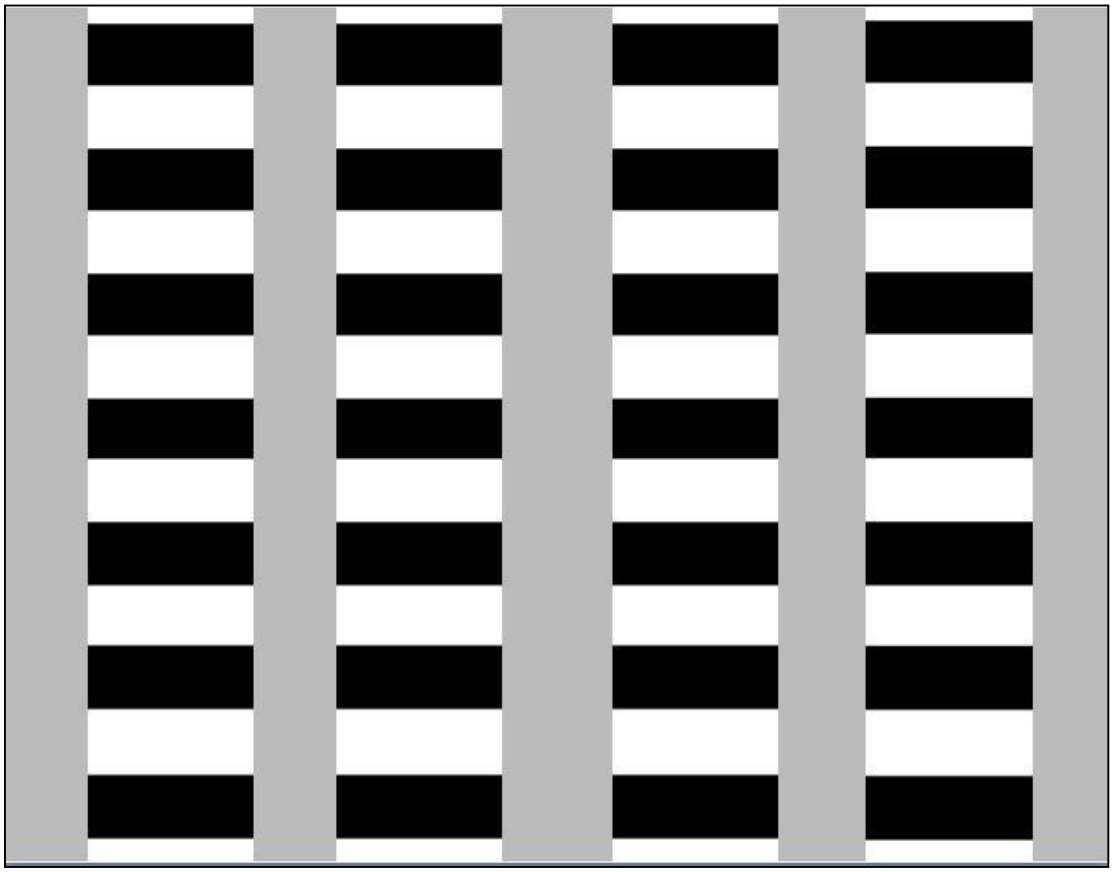


Figure 4-3. Presentation screen with vertical strip gratings.

The stimulus would appear randomly in one of these 4 positions.

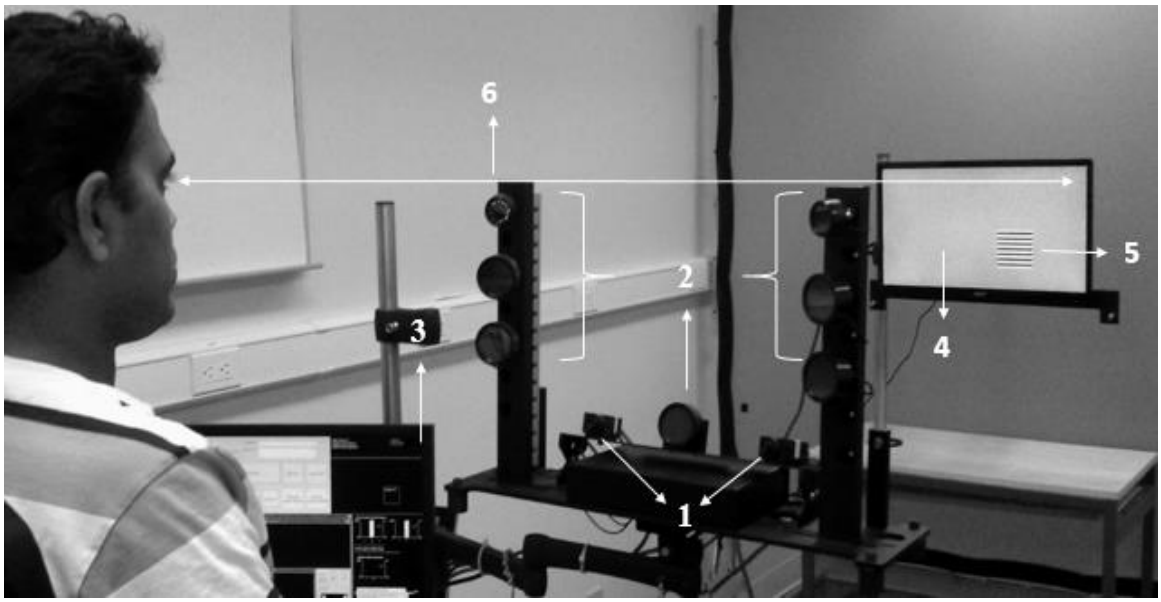


Figure 4-4. Setup of an adult participant with the Gaze tracker.

The numbers indicate the following: 1. Two video cameras 2. Infra-red light sources 3. Observer screen 4. Participant viewing screen 5. Grating stimuli 6. Viewing distance - 210 cms

4.1.3 Experimental Procedure

The testing distance between the observer and the monitor was 210 cm (see Fig. 4-4.), although the eye-tracker cameras remained at 70cms. The room lights were turned off during the time of the experiment. The experimenter presented the grating targets and also acted as an observer, using the eye movement fixation information to judge if the grating was seen or not. Prior to actual testing for VA, calibration of the eye tracker was done for each participant. During the calibration, participants looked towards different cartoon characters

moving across the screen. The description of how calibration is achieved is discussed in Chapter 3.

The adult participants were naïve for the objective measures i.e. they were asked to look at the screen, but given no further instructions. The objective measures were all completed before the subjective psychophysical measures, and visual acuity in adults was measured using the protocols in the order described below. For the subjective measures, the participants were asked verbally to indicate the position of the gratings.

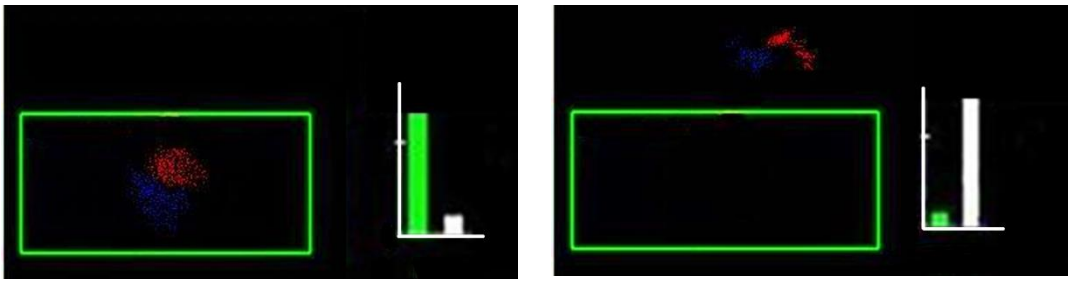
4.1.4 Criterion for threshold for first six protocols

For each presentation, the experimenter made a judgment of whether the participant was able to resolve the grating based on the fixation information provided by the software.

4.1.4.1 Seen/non-seen judgments in adults

The gaze tracking feature of the software gave information about the eye movements. The left eye and right eye's fixation was represented by the movement of red and blue colored dots respectively. Thus the movement of the colored dots gave information about the fixation patterns. On the screen viewed by the experimenter, the grating area was represented by a green outline. Therefore, clustering of the colored dots within the green target outlines was considered as a seen response (as shown in the Fig. 4A). Occasionally a few of the dots might wander just outside the border and these were included in the "seen" fixations. Fixation was also judged by the histogram which gave information of the ratio of the time

spent fixating the grating compared to the time spent outside the grating. When this rose to 75% or more, the target was judged as seen (as shown in the Fig 4B). If the clustering of dots was not within the green outline (as shown in the Fig. 4C) and the histogram showing the time spent on the grating was less than 75% (as shown in the Fig.4D) it was considered a not-seen response. There were several potential scenarios when a target was presented.



4A 4B 4C 4D

Figure 4-5. Eye tracker information depicting "seen" and "unseen" response.

Figure 4-5 shows a snapshot of the software, which provides information about the eye movements. In Fig. 4A and 4C the red and blue cluster of dots represents the left eye and right eye fixations respectively. The green outline represents the grating target area. In Fig 4A the red and blue cluster of dots is within the green box indicating this was a seen response as judged by the experimenter. In Fig 4C, the dots are not within the grating area, so this was a not-seen response. In Fig 4B and D the green histogram bar represents the percent time spent fixating within the grating and the white bar represents the percent time spent on the rest of the screen.

A. If the first fixation was on the target

1. If the fixation remained on the target for at least 2 seconds and if the histogram rose to 75%, it was judged as seen and the stimulus was terminated.
2. If the fixation did not remain on the target for 2 seconds, the stimulus remained on for up to 10 seconds. If there was a fixation of at least 2 seconds on the target within this period

it was judged as seen. There were a few occasions when the stimulus remained on for slightly longer than 10 seconds (approximately 10% of presentations).

B. If the first fixation was on an adjacent area (non–target area)

1. If fixation did not enter the target area within the first 4 seconds, it was judged at that point as not-seen.
2. If the fixation did move into the target within the first 4 seconds but did not stay for as long as 2 seconds, then the target presentation continued up to 10 seconds. If there was at least a 2 second fixation on the target, it was judged as seen. If not, it was judged as not-seen. The histogram was not used for this judgment, as it would not rise fast enough when the participant did not start off looking at the target.

The grating of a particular spatial frequency level was considered to be resolved if it were judged as seen according to these criteria. The VA threshold criterion was considered as 3 correct out of 4 presentations at one level with less than 3 out of 4 seen at the next higher spatial frequency. The grating occupied a maximum of 8% of the screen's surface area. This would mean there is an 8% probability of judging by chance that the grating was seen, which was less than the criterion level (at least 20%) set for judgment of seen/not seen responses – a 2 second fixation within ten second period would mean there is a 20% probability that it is judged as seen by chance.

4.1.5 Objective Gaze tracker protocols

4.1.5.1 Original TAC method

This method was performed first as this method closely follows the procedure used for clinical TACs, so this was the primary comparison that we wished to make. The starting

spatial frequency was mid-range i.e. 6.71 cpd which was expected to be above the threshold for VA for the participant and then increased in octave steps initially, but after the first reversal 0.5 octave steps were used. Each spatial frequency was presented as many times as required, up to 4 presentations, and the experimenter judged whether the participant saw the grating or not. If the experimenter was not sure if there was a fixation (a clear look), or the target was fixated after apparent exploratory eye movements, then the experimenter repeated the presentation. The spatial frequency was decreased or increased as necessary, to obtain the threshold.

4.1.5.2 Ideal Eye Movement (IEM) method

This method involved four presentations at each level and was expected to be the most accurate objective determination of threshold for adults among these protocols. Note that this protocol was not expected to be used with infants, but was included as a more accurate determination of threshold that could be used with adults, in order to compare the results of a more accurate method with the other protocols. This protocol was undertaken second so that it would be less influenced by fatigue or inattention than if it was performed later. The starting spatial frequency was two octaves lower in spatial frequency than the threshold determined from the Original TAC. Four presentations were shown at every spatial frequency increasing in 0.5 octave steps from the low starting spatial frequency to higher spatial frequencies until the threshold was determined.

4.1.5.3 Random IEM method

This was the same as the IEM, except that the spatial frequencies were presented in random order. It was thought that this might help maintain the participants' attention, as the stimulus would be less predictable.

4.1.5.4 Staircase TAC procedure

This was essentially like the Original Teller card procedure, except that the procedure started with a lower spatial frequency i.e. 3.21 cpd as typically a staircase method would start at a level well above the threshold[48].

4.1.5.5 Staircase TAC procedure plus video

This method was similar to the staircase TAC method, but a video was presented between each presentation in an attempt to improve attention.

4.1.5.6 Halving of Range (HOR)

This protocol was included to determine if the TAC method or staircase method could be improved upon for efficiency. The experimenter started at a low spatial frequency, expected to be above threshold. If a correct response was obtained, then the experimenter presented the highest spatial frequency available, which was expected to be below threshold i.e. not seen. Assuming that the participant was not able to resolve the highest grating, the experimenter presented a spatial frequency half way between these two frequencies (during the whole experiment, none of the participants was able to resolve the highest grating). Depending on the result obtained for that middle spatial frequency, the direction of the halving would

proceed. If a grating was resolved, then we tested half way between the grating that was resolved and the grating that was not resolved and so on, until threshold was reached.

4.1.6 Subjective gaze tracker protocol (subjective GT)

In this protocol, the participants were asked to give verbal responses to the presentations regarding the positions of the grating. It was therefore a four alternative forced choice procedure. Ten presentations were shown at each spatial frequency. The order of spatial frequencies was the same as the IEM method.

After these GT protocols were completed, the participant's visual acuity threshold was determined using the original TACs at the same distance (210cms). The acuity was first measured objectively (no instructions given and relying on fixations, as with an infant) and then based on the verbal response of the participants. It was undertaken with the TACs held in the usual orientation (card held horizontally, vertical gratings) in order to compare the gaze tracker VA with that obtained clinically with TACs. It was also undertaken with the TACs held in a vertical orientation, so that the orientation of the gratings was horizontal, matching that of the eye tracker. It would be expected that there would be more concordance of acuity when the gratings have the same orientation, especially in cases of high astigmatism in which the blur would be different in the two meridians. Finally at the same viewing distance, letter acuity was measured using the ETDRS log MAR chart, using a by-letter scoring[103]. The participant was asked to continue to attempt to name the letters until fewer than 2 per line were correct.

4.1.7 Analysis

The data were analyzed using correlation coefficients between TAC and GT measures of VA for each protocol. Gaze tracker VA was compared in a similar way with letter acuity and with uncorrected refractive error in the meridian perpendicular to the orientation of the gratings. We also looked for the percentage of agreement of VA within half or one octave, as is common in the infant literature[38].

4.2 Results

The thresholds obtained by the six objective gaze tracker methods were identical and therefore all these methods will be referred to as Objective GT hereafter. The results obtained by the objective GT were in good agreement with the other methods as shown in Table 4-1. Eighty-seven percent or higher were within half an octave and hundred percent were within one octave.

There was no obvious difference in the mean thresholds from the Objective GT protocols, subjective GT protocol, horizontal and vertical orientations of the TACs as shown in the Fig. 4-6. Objective GT thresholds were significantly correlated with the subjective GT protocol and TACs with horizontal gratings ($r= 0.91$ and 0.92, respectively; $p<0.001$) as shown in Fig. 4-7 and 4-8. This shows that all the OGT data points except one were within half an octave of the subjective GT results and all the OGT data points were within half an octave of the

TACs (horizontal gratings). This indicates very good agreement and correlation between the OGT and the TACs. Similarly, there were good correlations between the other measures of grating acuity, as can be seen in Table 4-1. And these were all significant at the $p<0.001$ level.

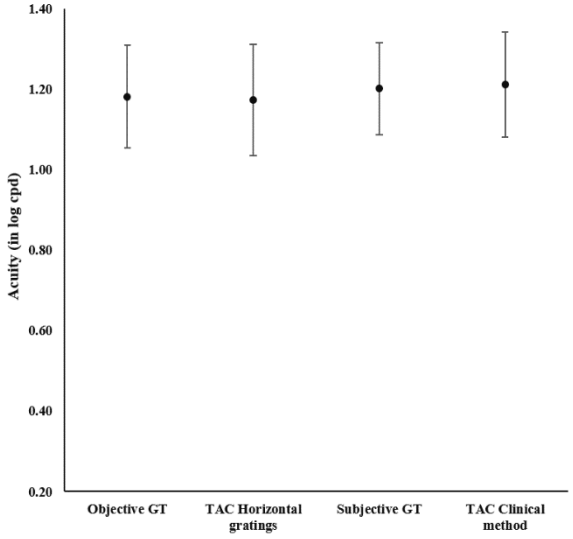


Figure 4-6. Error bars showing the mean and ± 1 SD for the different methods

The black solid circles indicate the average and the top and bottom solid lines on either side of the average indicate ± 1 standard deviation

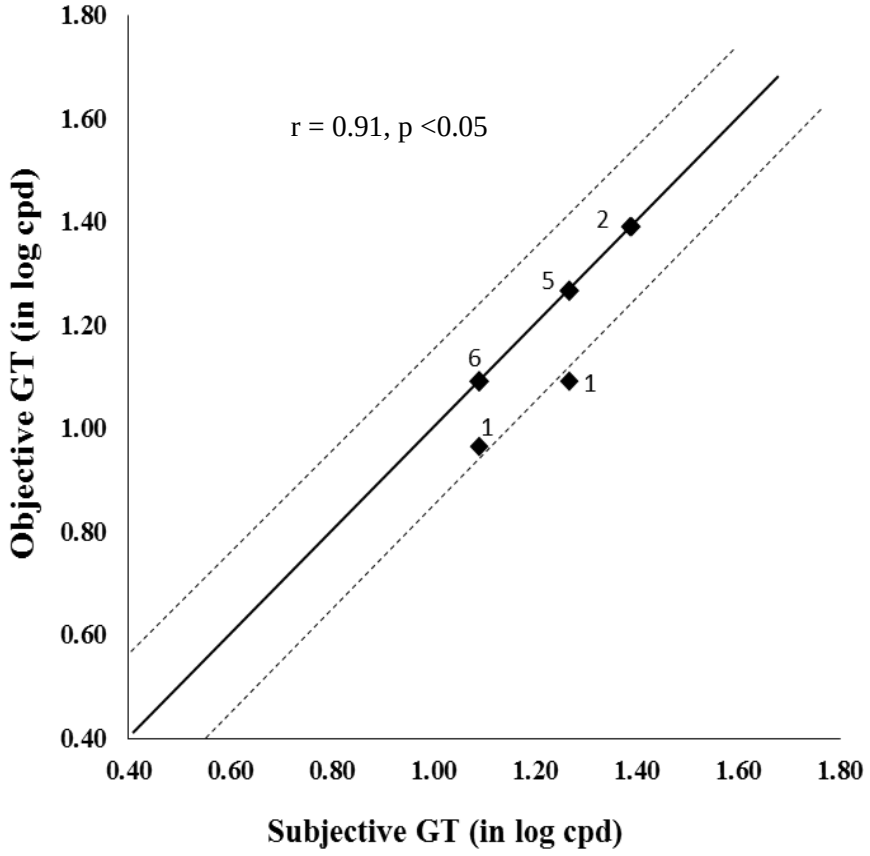


Figure 4-7. Scattergrams OGT against thresholds obtained with the subjective GT protocol.

The solid line is the line of equality and the dotted lines show plus and minus half octave from equality. The numbers in the charts represent the # of data points when there is more than one data point in the same location.

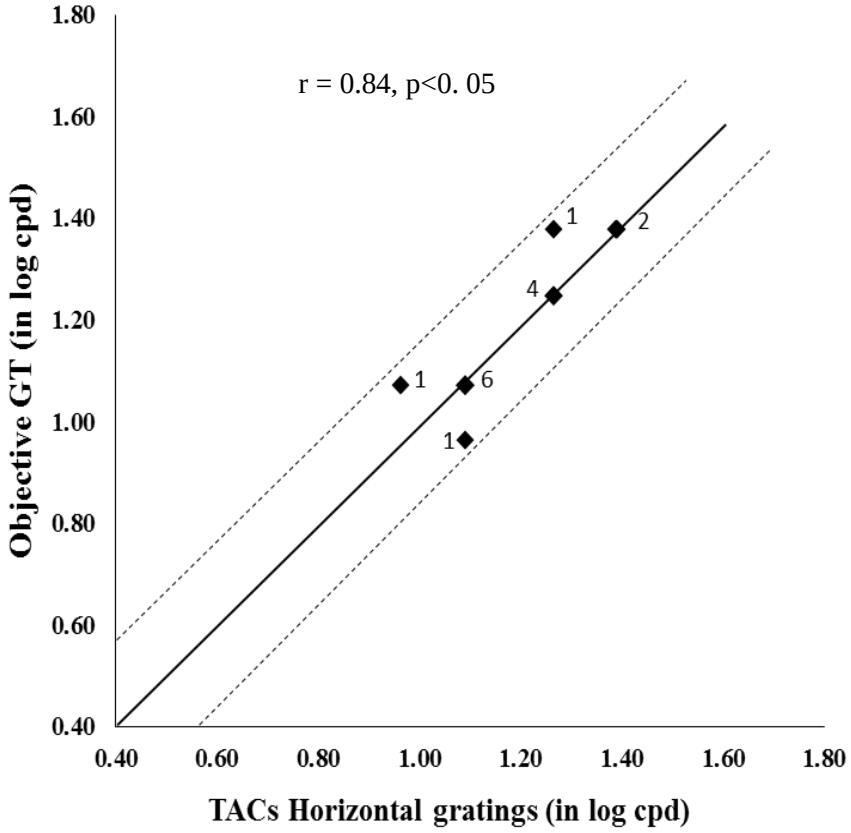


Figure 4-8. Scattergrams OGT against thresholds obtained with the TACs –horizontal gratings.

The solid line is the line of equality and the dotted lines show plus and minus half octave from equality. The numbers in the charts represent the # of data points when there is more than one data point in the same location.

There was a moderate correlation between letter acuity and grating acuity as measured either by the TACs or Objective GT (r=0.75, p=0.001 and 0.73, p=0.002 respectively, see Fig. 4-9. and Fig. 4-10). However, it can be seen that the grating acuity results are higher than the letter acuity, i.e. above the line for equality, indicating that grating acuity overestimated letter acuity. There is, therefore, a moderate correlation, but not good agreement.

Table 4-1.Agreement and correlation coefficients for different methods with the Objective GT

	Objective GT	Subjective GT	TACs		
			Clinical (vertical gratings)	Horizontal gratings	verbal responses
Objective GT	N/A	93%, r =0.91*	87%, r=0.84*	100%, r =0.92*	87%, r =0.84*
Subjective GT		N/A	93%, r =0.88*	93%, r =0.88*	93%, r =0.88*
TACs - clinical			N/A	87%, r =0.86*	93%, r =0 .99*
Horizontal Gratings -TACs				N/A	87%, r =0.86*

% within 0.5 octave/correlation coefficient, all the correlation values are significant at the p<0.05 level

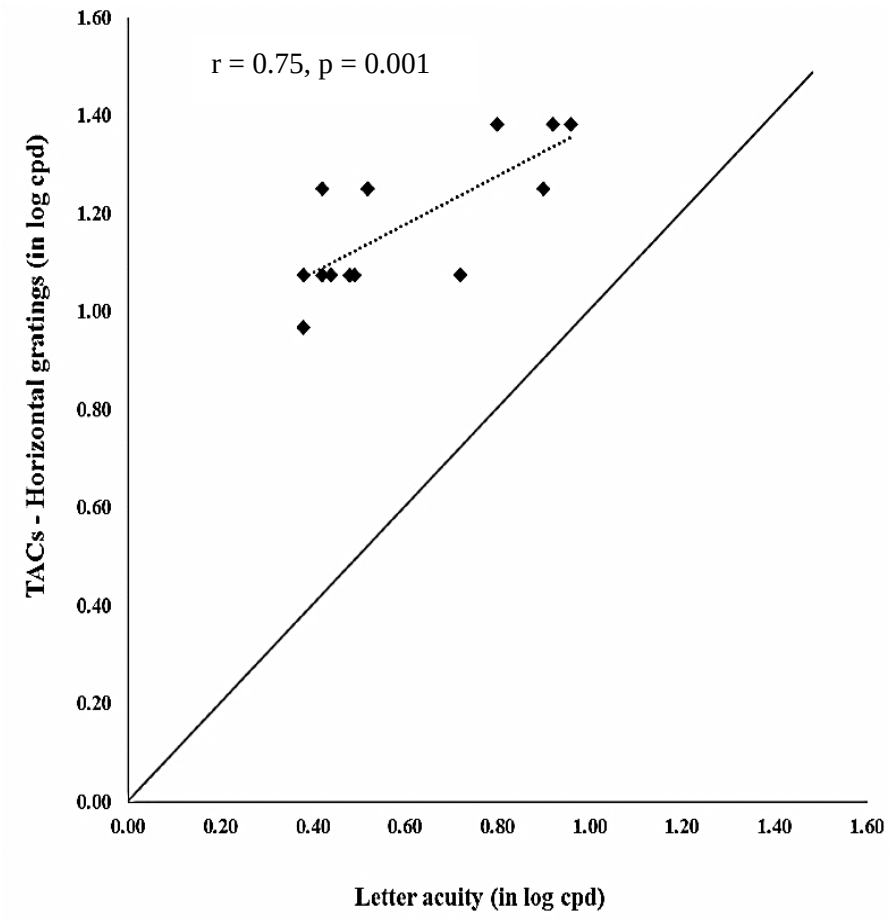


Figure 4-9. Scattergrams of correlation of TACs - horizontal gratings acuity against letter acuity

The solid line is the line of equality

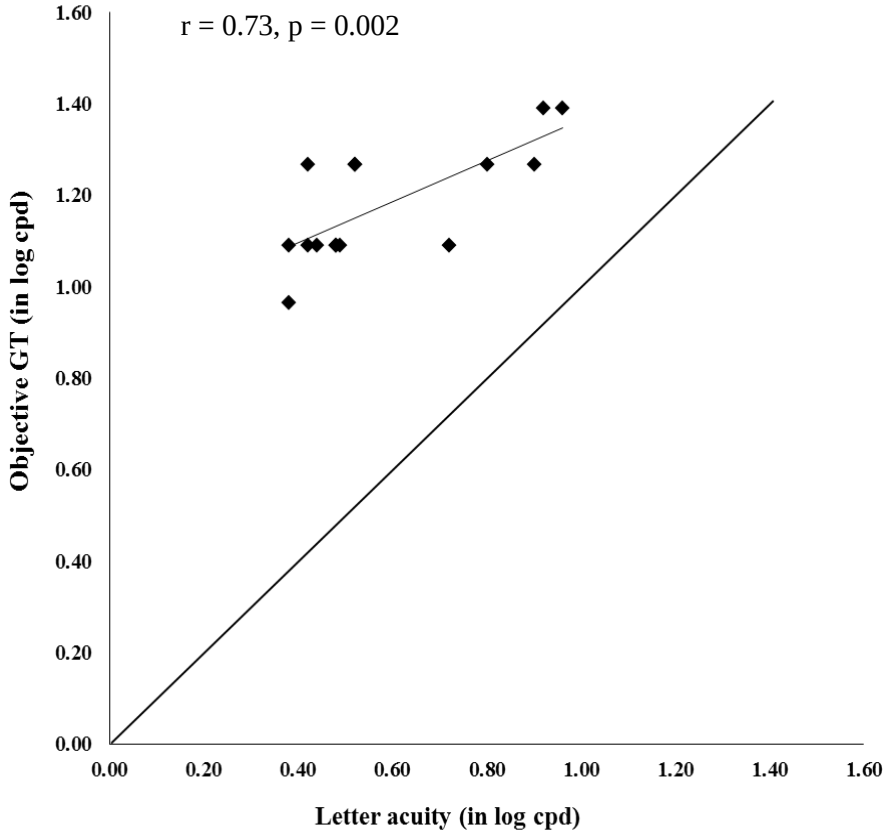


Figure 4-10. Scattergrams of correlation of Objective GT grating acuity against letter

The solid line is the line of equality.

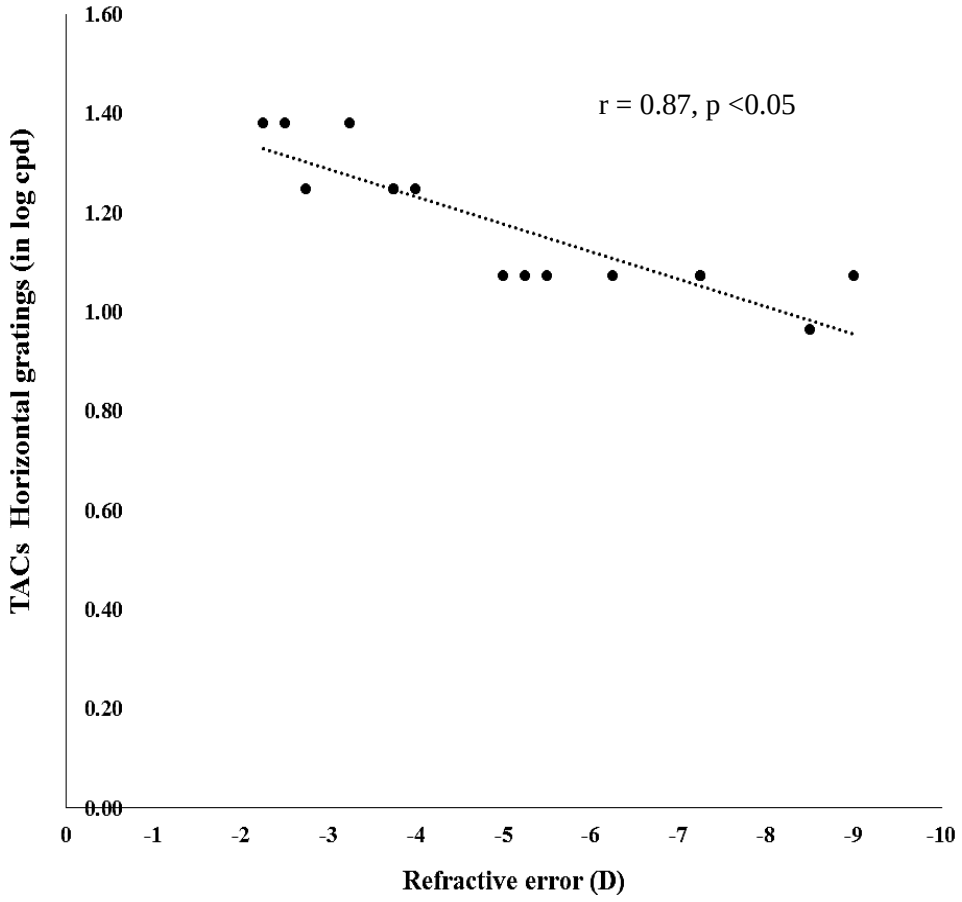


Figure 4-11. Scattergram of TACs (horizontal gratings) acuity against refractive error

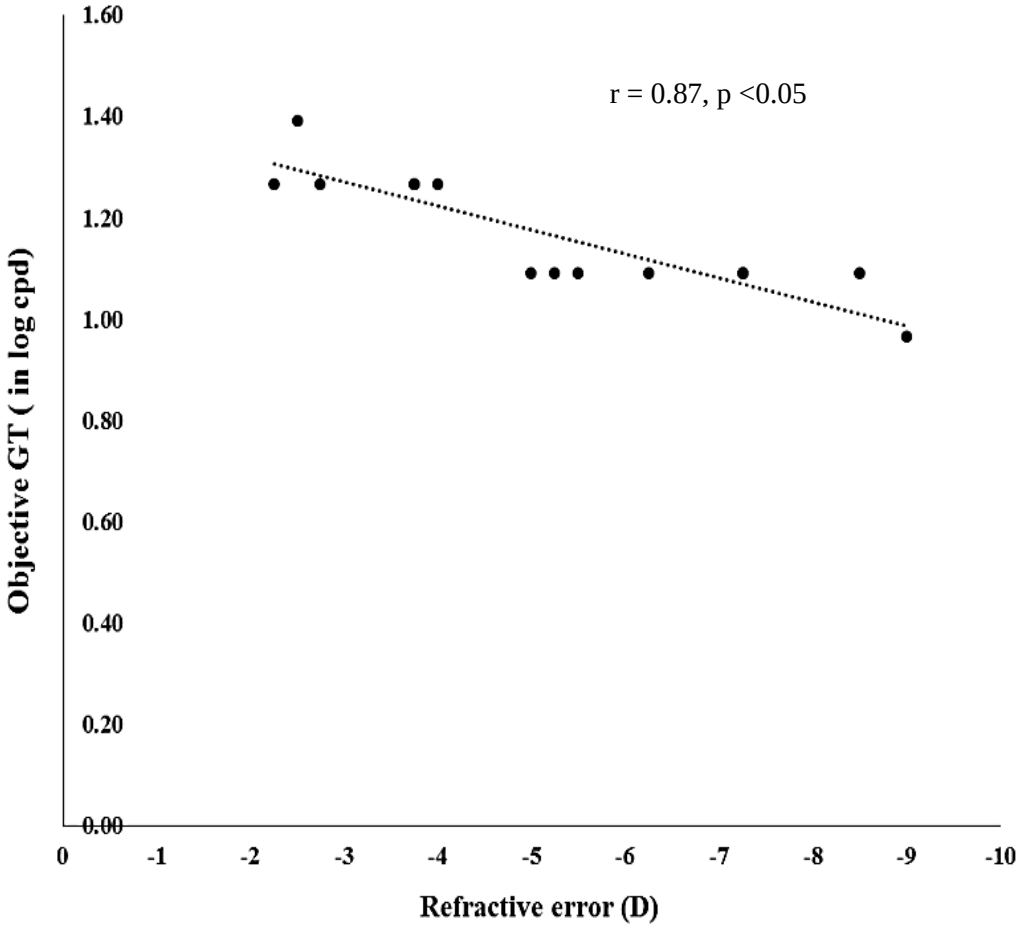


Figure 4-12. Scattergram of Objective GT acuity against refractive error

The correlation of grating acuity obtained by both TACs (horizontal gratings) and Objective GT with refractive error in the vertical meridian was r=0.87 (as shown in Fig. 4-11. and 4-12.). The correlation between letter acuity and spherical equivalent refractive error was similar and also significant (r=0.86, $p<0.001$).

The time taken for each of the protocols is shown in Figure 4-13. Repeated measures ANOVA showed that there was a significant difference between the durations ($p<0.0005$). Post hoc testing with Bonferroni correction showed that the Staircase TAC with video was significantly faster than the first 4 methods (Original TACs, IEM, Random IEM and the Staircase TAC) but was not significantly different from the HOR. All the OGT protocols yielded identical threshold results, but the Staircase TAC procedure with video and the HOR were more efficient. The mean time for the Staircase TAC with video was 86.76±16.49 seconds.

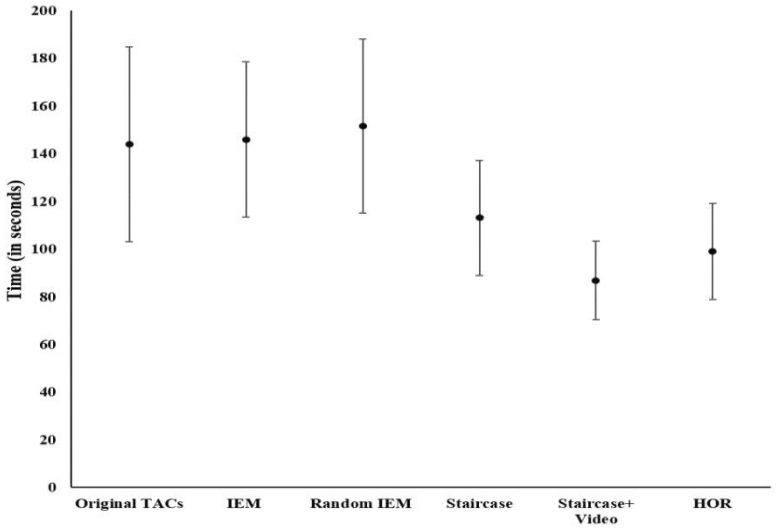


Figure 4-13. Error bars showing the mean time and ± 1 SD for different methods

The black solid circles indicate the average and the top and bottom solid lines on either side of the average indicate ± 1 standard deviation

4.3 Discussion

The main purpose of this study was to explore the validity of the gaze tracker in adults, with the intent that it may be used in future for infants. For this purpose, acuity measured with the gaze tracker was compared with the TACs. There was a good correlation and agreement for grating acuity between these methods. One hundred percent of GT results were within one octave of TACs and there was the same agreement with subjective GT values. These results were compared with the infant literature comparing different measures of VA in infants, as the next step using the same instrument would be measuring infant VA. The results agreed well with the literature. McDonald et al.[32,49,50] showed 95-100% agreement within one octave between the Acuity card procedure (ACP) and the Preferential looking (PL) procedure in infants aged 1 to 12 months and between 18-36 months. Similarly Mohn et al.[51] showed

100% agreement within one octave for binocular acuity between the ACP and PL procedure in normal and neurologically abnormal infants and children in the age group 3 months - 22 yrs. Therefore, GT showed similar agreement with TACs as the repeatability of TACs shown in infant literature.

The uncorrected refractive error correlated very well with grating acuity as measured by both GT and TACs and also with letter acuity. The grating acuity was affected less by uncorrected refractive error than the letter acuity. This was as expected from the literature and a detailed discussion will be found in the Chapter 7.

4.4 Conclusion

The hypotheses in the adult study were that there would be no difference in measured visual acuity obtained by the two methods namely, Gaze Tracker and TACs and there would be a significant correlation between uncorrected refractive error and visual acuity in the adult population. The above hypotheses were confirmed in this study. The validation of the gaze tracker visual acuity with adults is the first step before validating the measurements in infants. The two methods, HOR and Staircase with video, were the most efficient protocols and these were therefore chosen to be used in infant study. To conclude, we have shown that the gaze tracker using naïve eye movement responses gives a valid measure of grating acuity in adults, and therefore may have potential for measuring VA in infants. Future studies are needed to demonstrate its validity with infants or young children.

Chapter 5 – Preliminary Infant trials

5.1 Introduction

As described in the previous chapter, the GT showed good agreement with TACs in adults. Therefore, the next step in the process of validating the GT was to test infants. Ultimately, validity and repeatability should be measured in infants, but first it was important to check the feasibility and optimize the protocols in infants. The purpose of these preliminary trials was to optimize the use of the GT in infants and compare two different protocols of the GT with the Teller acuity cards (TACs). Towards this end we tested infants to obtain results for testability and to compare two different GT protocols, which were based on the most efficient protocols in adults.

5.2 Methods – study design

Visual acuity was measured with the GT and TACs on two occasions to determine the validity (comparison of GT with TACs) and test-retest repeatability of both tests. The two sessions were scheduled within a specified maximum time interval, which was 7 days for 3 to 6 month olds and 10 days for infants of 6 months up to 12 months[57]. This was because infant acuity is rapidly developing in the first year of life, which would become a confounder in the repeatability determination. This pilot study was designed such that the future infant study would use a similar experimental design.

5.2.1 Participants

Participants were recruited through flyers posted in the Grand River Hospital Birth Clinic, local bulletin boards and doctors' offices. There were also some referrals from word of mouth spread by the previous participants' parents. Participants were also recruited through patient records available at the Pediatric and Special Needs Clinic at the University of Waterloo School of Optometry and Vision Science. A staff member phoned the parent and briefly explained the study and then asked if they were interested to be contacted by the study investigator with further details. If so, a letter of information about the study was mailed or emailed to the parents. If they were willing for their child to participate after reading the letter of information, a few questions were asked by phone or email to determine initial eligibility. These were the child's birth date, estimated gestational age at birth, general health, history of ocular problems and whether the child had attained the normal milestones according to their age.

The tenets of the Declaration of Helsinki were followed and the study received clearance through The Office of Research Ethics at the University of Waterloo. Written informed consent was obtained from the participant's parent or guardian. Remuneration for participation was $20 per session.

All the acuity observations (TACs and GT) were made by a single, experienced observer, so as not to confound the results with inter-observer differences. She was a pediatric optometrist, who has experience in acuity testing in infants. The order of testing (GT or

TACs first) was randomized between participants, with the constraint that an equal number of participants started with the GT as with the TACs. Having been randomized, the same order of testing was followed in both the visits for each participant. At the end of the first session, the observer performed the eye screening test (described below) to determine that the complete inclusion criteria were met. This was undertaken at the end of the session to ensure that the VA testing was performed when the child was maximally and equally attentive in both sessions. The observer was masked to the absolute spatial frequencies of the gratings, the final TAC acuity and GT acuity until all visits were complete.

5.2.2 Initial assessment and eligibility for the study

5.2.2.1 Eligibility criteria

To be eligible for inclusion in the study, participants met the following criteria: (1) gestational age between 37 and 42 weeks at birth; (2) no major known medical problems; (3) normal development by parental report and observation; (4) completion of basic eye screening examination which included, Hirschberg and unilateral cover test to check for the presence of strabismus, broad H to test for incomittancy and refractive error by Mohindra retinoscopy.

Spherical refractive error exclusion criteria were based on non-cycloplegic retinoscopy results by Gwaizda et al (1993)[104] and are shown in Table 5-1. Infants with astigmatism >2.50D in any meridian were excluded[105].

Table 5-1. Refractive error inclusion criteria

Age group	Lower Limit	Upper Limit
≤ 6 months	-2.00 DS	+4.00 DS
>6 – 12 months	-1.50 DS*	+3.00 DS

* The lower limits were adjusted based on personal clinical experience (S. Leat)

5.3 Testing with TACs

Infants were tested with both Teller Acuity Cards (TACs) and as well as the gaze tracker. The TACs were used in both horizontal and vertical orientations. The TACs are usually used in the horizontal orientation, i.e., showing vertical gratings. During this testing, TAC stage was used. The stage basically consisted of a slot/opening in which cards were presented. Infant's attention was kept towards the center of the opening by use of audio visual stimuli such as toys and puppets. However, the gratings in the GT were horizontal, therefore to make TACs stimuli comparable to GT, we also measured acuity with the TACs held vertically (horizontal gratings). The order of testing between the vertical and horizontal orientation of the cards was also randomized with the constraint that both the methods were performed first an equal number of times.

The setup of TACs was that the infant was seated on the parent's lap who was seated on an adjustable chair. The distance between the infant and the TACs was 55cms. The TACs were performed as follows. First, the observer showed a blank card to see the infant's response

when the stimulus could not be resolved. Next the observer presented the lowest spatial frequency, to determine how a “clear” look appeared. Then she tested in one octave steps, showing one card at each level as long as a clear, correct look was observed. Once a less clear look or an incorrect judgment was made, testing was done in half octave steps. Previously presented cards could be tested again at the discretion of the observer. Testing was considered complete when the observer was satisfied that a threshold of the highest spatial frequency to give a clear, correct look had been obtained. The observer made sure to present a card at least twice at a spatial frequency level above threshold, at which the grating was not resolved. This was in accordance to the testing recommended in the TAC manual. Thus the threshold was obtained. The observer was unaware of the threshold obtained for the infant in the first session until both sessions had been completed. However, she was aware of the spatial frequencies that she was showing during the testing.

5.4 Testing using the Gaze tracker

The setup of gaze tracker is shown in Figure 5.1 below. The parent/attendant sat on an adjustable chair, whose height can be modified. Then, the infant was seated on the parent’s lap. The experimenter started the eye tracking program. The height and distance of the participant was adjusted until the eyes were detected by the program. The distance between the cameras and participant’s eyes was kept constant at 70 cm. The distance between the infant and the presentation screen was increased if the infant was able to see the penultimate spatial frequency. The distance between the cameras and the monitor was increased by moving the monitor back.

Figure 5-1. Setup of the gaze tracker with the infant

5.4.1 Judgment of seen/not seen

The infant's fixation response to each grating stimulus was judged as "seen", "not seen" or "unknown". As described in Chapter 4, the grating area was represented by a green rectangle, either in one of the 4 quadrants or as linear vertical strip in different lateral positions on the screen. Clustering of the colored fixation points within the green target outline was considered as a seen response. Occasionally the fixations clustered just outside the border or within that quadrant of the screen and these were also judged as "seen" responses. Seen and unseen judgments were made similar to the judgments made in adult study as described in chapter 4.

The observer was allowed one more option that was a "repeat/unsure response". The "unknown response" was used when the infant was fussy or not looking consistently at any position on the screen throughout the presentation time or if the observer was unsure i.e. if it

was not a clear seen or unseen response. The presentation would be considered valid as seen or not seen only if the infant attended towards the screen and there were fixation traces throughout the presentation. The observer verbally indicated the responses to the experimenter. The protocols used in visual acuity testing in infants with the gaze tracker were the staircase and halving of range (HOR) methods. These two methods chosen were similar on the earlier adult study results described in Chapter 4. These two methods were the most time efficient methods which also results that were equally accurate in adults. The order of protocols was randomized between participants but was kept constant for both visits for each participant with the constraint that both the protocols were performed first an equal number of times.

5.5 Gaze tracker methods

5.5.1 Staircase method

This method involved starting grating presentations in octave steps from a very low spatial frequency, which was above the threshold. A reversal refers to change in response from seen to not seen or vice versa. After the first reversal, the experimenter presented stimuli in 0.5 octave steps, but still presented one presentation at each spatial frequency, until there were three reversals in the staircase. In the staircase method for the first three infants, three presentations at each level were presented after the third reversal. Initially, two correct out of 3 presentations was taken as threshold. This criterion was used because it resulted in the same threshold as 3 out of 4 correct in adults and would also reduce the testing time compared to 3 out of 4. However, after experience was gained performing the GT on three

babies, it was decided to make the threshold criterion 3 correct out of 4 presentations so as to give more stable and reliable responses in the infants. In all infants, 2 incorrect responses were obtained from a level 0.5 octave higher, before deciding on the threshold. An example of the staircase method is shown in Table 5-2 below. The threshold obtained in this example was 4.1 cpd.

Table 5-2. Staircase method used with GT

Spatial frequency (cpd)	**Result**	**Description**
0.391	Seen	One presentation at each level initially
0.769	Seen	
1.54	Seen	
3.08	Seen	
6.15	Unknown	
6.15	***Not Seen***	***1st reversal***
4.1	***Seen***	***2nd reversal***
6.15	***Not Seen***	***3rd reversal***
4.1	Seen	Start testing with 4 presentations at each level until threshold is obtained threshold criterion met
4.1	Seen	Final threshold

5.5.2 Halving of Range (HOR)

This method consisted of starting with grating presentations at a very low spatial frequency (which is above the threshold i.e. easily seen) and then jumping to a very high spatial frequency (which is expected to be below the threshold and not seen). After the initial very high and very low spatial frequency, gratings were presented of spatial frequency half way between seen and not seen levels, until two reversals in response were obtained. Until the two reversals were obtained, only one presentation was presented at each level, including initial stimulus. During the testing for the first 3 infants, four presentations were presented after 2 reversals, moving up or down in spatial frequency until the criterion of the highest frequency to give 3 correct out of 4 presentations was met. However, as experience was gained in testing infants, the number of reversals was changed from 2 to 3 to be consistent with the staircase protocol. In both the cases, 2 incorrect responses from a 0.5 octave level higher were obtained, before deciding on the threshold. In the example in Table 5-3, the threshold was 7.03 cpd which was the highest spatial frequency giving 3 correct responses out of 4 presentations.

Table 5-3. Halving of Range method for GT

Spatial frequency (cpd)	**Result**	**Description**
0.474	Seen	One presentation
21.1	Unknown	
21.1	***Not Seen***	*1st reversal*
10.6	***Seen***	*2nd reversal*
14.1	***Not Seen***	*3rd reversal*
10.6	Not Seen	Start testing with 4 presentations at each level until threshold criterion met
10.6	Not Seen	
7.03	Unknown	
7.03	Unknown	
7.03	Seen	Final threshold
7.03	Seen	
7.03	Seen	

5.6 Results

There were 7 infants in this preliminary study, whose mean age was 6 ± 1.8 months. The Teller cards used in the usual horizontal orientation (gratings vertical) seemed to give rise to better acuities than when held in the vertical orientation. The mean acuity for vertical gratings averaged over the two visits was 0.72 ± 0.19 log cpd which was significantly higher (t-test, p =0.05) than the mean horizontal grating acuity (0.57± 0.29 log cpd). It was thought

that the use of the TAC stage for the vertical grating acuities may be the reason for this difference. It reduces distractions for the babies and also provides a method for attracting their attention towards the center of the card.

These were the protocols we aimed to use in this study. Since this was a preliminary infant trial study, uniformity in protocol could not be maintained in some cases. This was the first time infants were formally tested using the GT in the Waterloo laboratory. As experience was gained, some adjustments in methods were made. These are explained below. During the gaze tracker experiment with the infants, attention was maintained by playing videos in between the grating presentations.

5.7 Modifications to the GT and TAC protocols as a result of the preliminary infant study.

The following adjustments were made;

5.7.1 TACs

- A TAC stage was prepared with a vertical aperture, so that the TACs could be tested in the vertical orientation. This was similar to the stage used for the cards held in the usual horizontal position.
- It was decided that the observer should be as naïve as possible regarding the spatial frequency of the gratings. So the start cards for TACs would be randomised for all the babies, ensuring that the start card was above the threshold i.e. expected to be seen. The

label which states the actual spatial frequencies was covered, so that the examiner did not know the absolute spatial frequency and would only know the relative spatial frequencies, as is typical in other studies using acuity card procedures[32,56,58].

5.7.2 Gaze tracking

- As with the TACs, the observer would be unaware of the starting spatial frequency, so as to reduce the bias in judgement.
- There were some infants in the preliminary testing, who were >6 months and were able to see the penultimate spatial frequency at 70 cms. The next level is one octave higher than that, which means the acuity between those levels was not measurable. If the testing was stopped and recommenced after moving the screen, the child's attention was lost. Therefore, the starting test distance of the gaze tracker would be chosen according to age: 70 cms for infants less than 6 months and 120 cms for infants greater than 6 months.
- With some infants the calibration procedure either took a long time or was not completed. In these cases the number of image co-ordinate estimates for each target (as described earlier in chapter 3) was reduced to 75 samples instead of 100. This saves time as the instrument would take less time to reach 75 estimates, helping to ensure that infant did not become habituated to the calibration stimuli.
- The judgement of a seen response should be made if the fixation is stable and within the stimulus area, close to the edge of it, or in that quadrant of the screen, for at least 2 or 3 seconds within the first 6 seconds or the histogram rises up to the 75% level. The

observer would make the decision of seen or not seen i.e., between 6-10 seconds after onset of each stimulus. This change was made as it was noted, that if the infant is going to fixate the stimulus, they generally do so in the first few seconds and are unlikely to do so late in the trial.

- A different video would be played during infant and gaze tracker set up as that used during the trial.
- Initially, the observer verbally told the responses to the experimenter in the preliminary study. Sometimes, it was noticed that this distracted the infants. Therefore, a tactile mode of communication was set up between the observer and experimenter.
- A barrier screen was set up between stimulus display screen and program control screen. This was also done to avoid infant's attention moving towards the program control screen.
- Infants in this study performed 4 protocols with GT and TACs and invariably, they lost attention on the last protocol or became fussy towards the end. Therefore tests done at the end did not have good reliability in this preliminary study. Infants habituated very quickly to the GT stimuli and they lost attention and interest after 2 or 3 protocols (GT and TACs). Therefore, only 2 protocols would be tested in future trials (one GT and one TAC), to maximise the infants' attention with the expectation that this would make the test results more reliable.
- An additional stimulus was used to regain infant's attention. A battery operated bright flashing stimulus with on-off switch was attached to the second monitor which displayed the presentations. It was attached such that flashing stimulus was just above the monitor.

The observer had control over the switch and could use it when deemed necessary to regain infant's attention towards the screen. It was not used during stimulus presentation, but between.

This preliminary infant study gave us insight and experience on working with infants using the GT. Based on this experience, the protocol for the main infant study was designed and was made more optimal.

Chapter 6 - Validation of GT VA measurement in infants

6.1 Introduction

The validity of visual acuity measurements with the gaze tracker (GT) in adults has been established as described in Chapter 4. The GT measures of VA in adults were validated in several ways. Firstly, the GT measurements were in good agreement with the Teller Acuity cards (TACs) and secondly, there was a decrease in grating acuity with increasing refractive error, as would be expected.

The next step in development of the instrument was to validate it in infants, which was the aim of the experiment described in this chapter. The validity of the GT was measured by comparing with the Teller Acuity cards. The validity of the GT was also measured by comparing the VA with age, as it is expected that VA should improve with age for infants between 3 and 12 months. The reliability of the GT was measured by determining the repeatability of the test and this was compared with the repeatability of the TACs, both as reported in the literature and as measured in the current sample of infants. Infant testability for the GT was also determined.

Hypotheses

1 There will be no difference in measured visual acuity obtained by the two methods namely, Gaze Tracker and TACs.

2 GT measurement of VA will have no difference in percentage of repeatable measurements as compared to TACs.

3 Visual acuity, as measured by the GT, will improve with age in the infant group.

6.2 Methods

The study design, eligibility for the study and participant recruitment was similar to the descriptions in Chapter 5.

6.2.1 Sample size calculation

There are few previous data on which to base a sample size calculation. McDonald et al.[32] reported a correlation coefficient for intraobserver reliability of acuity card procedure in infants of 0.66. Based on this, and in order to obtain a significant correlation for repeatability, the sample size was found to be a minimum of 16 (power of 0.80 and alpha level set as 0.05). The correlation coefficient for intraobserver reliability of TACs (horizontal position) during the preliminary infant trial for 7 infants was 0.79, which would give a slightly lower sample size estimate of 10. We therefore aimed for a sample of at least 16.

6.2.2 Procedures

6.2.2.1 Test materials and protocols

TACs

The physical characteristics of TACs were the same as described in Chapter 1. A new TAC testing stage was constructed from three panels of hardboard painted grey to exactly match the shade of the background of the TACs and the commercially-available stage. The central panel contained a vertical opening of the same size as that for the commercially-available stage and behind which the acuity cards were presented. To ease the testing process the hardboard was mounted to a movable table. Side panels blocked the view for the infants to

avoid distractions (see Figure 6.1 and 6.2). This allowed the TAC procedure to be undertaken with the cards held in the vertical orientation, thus testing acuity for horizontal gratings which was different from the normal clinical procedure but which corresponded to the orientation of GT gratings. The average luminance of the acuity cards when placed in the apparatus was 1.65 log cd/m^2, measured with the Minolta luminance meter (Chroma Meter CS 100).

The observer had knowledge of each participant's name but was not informed of the starting spatial frequency or the participant's age. The start card was one of 4 spatial frequencies; 0.31, 0.43, 0.63 or 0.85 cpd. The choice of the start card was randomized by the study co-ordinator before the participant arrived and the same start card was used for the second session of each infant. The starting spatial frequency was randomized between participants to give an equal number with each start card. The observer was also masked to the absolute spatial frequencies of the gratings, the final TAC acuity and the acuity obtained by the GT when that was undertaken first.

.

Figure 6-1. showing the specially designed TAC stage (from outside – observer's view)

Figure 6-2. showing TAC stage (from inside i.e. infants view)

The grey color background matching the grey color used in clinic TAC stage)

All testing was done with the acuity cards placed behind the opening in the central panel of the TAC apparatus. Infants were seated or were allowed to stand (supported) on the parent's lap for testing. Before the test, the acuity cards were arranged in order of spatial frequency with right-left location of the grating randomized. The test distance was 55 cms for all the infants

When presenting each card, the observer did not know the position of the grating and this was verified as correct or incorrect by the study co-ordinator. Thus the observer did not see the grating herself and remained masked regarding the exact spatial frequency but only knew of relative spatial frequencies, which were indicated on the back of the card with a number which covered the true spatial frequency. The start card used for each infant was numbered 1 and the next half octave step card was numbered 2 and this continued with sequential numbering to the highest spatial frequency. Otherwise, the protocol used was the same as for the preliminary study described in the Chapter 5.

Gaze tracker

The GT equipment was the same as described in Chapter 4. For the gaze tracker, the infant was seated on the mother/father's lap on an adjustable chair see Fig 5-1. The infant was then adjusted to be at the right distance and height so that the GT tracked his/her eyes. An infant's video played during this setup. The infant's eye movements were calibrated using various small targets (cartoon characters with sound) which moved to different positions on the computer screen (see description in Chapter 3).

The experimenter controlled the order of spatial frequencies in the protocol and also decided the starting spatial frequency for each participant. The observer and the experimenter were different people, in order to avoid bias in the judgments of thresholds. The starting spatial frequency for the 120 cm distance was one of 0.25, 0.34, 0.48, or 0.67 cpd and was randomized between participants. For the GT at 70 cm the starting spatial frequencies was 0.28, 0.40, 0.55 or 0.78 cpd. The observer informed the experimenter whether the target was seen, unseen or not known/needed repeating, based on the infant's response as seen by the eye movement information provided by the GT software. The judgment for the infant's fixations (seen, unseen, unknown) were the same as described in Chapter 5. The observer was masked regarding the actual spatial frequencies by a black card that obscured the left hand side of the screen, in order to reduce bias. The grating presentations were square wave horizontal gratings as described in Chapter 4.

The protocol chosen for the infant study was based on the adult staircase method, which was one of the most efficient methods in reaching threshold in the adult study, as previously discussed in Chapter 4. The test distance was 70 or 120 cm for infants <6 months and >6 months respectively. These distances were chosen based on the preliminary infant trials as described in Chapter 5. At this distance, the range of spatial frequencies used in the gaze tracker was similar to those in the TACs as shown in Table 6-1.

Table 6-1. Upper and lower limits for GT and TAC for various test distances

	Testing distance	Lower limit/starting spatial frequency	Upper limit
TACs	55 cm	0.32 cpd	38 cpd
GT	70 cm	0.28 cpd	24.92 cpd
	120 cm	0.25 cpd	42.40 cpd

The protocol involved starting with a low spatial frequency grating which was expected to be above the threshold. Initially one stimulus was presented at each spatial frequency level, increasing in one octave steps, until there was a reversal. A reversal refers to change in response from seen to not seen or vice versa. After the first reversal, the experimenter presented stimuli in 0.5 octave steps, but still presented one presentation at each spatial frequency, until there were three reversals in the staircase. Then, a total of four presentations were presented at the previous level, moving up or down in spatial frequency until the criterion of the highest frequency to give 3 out of 4 presentations seen was achieved. Thus the end point was when there were three correct responses at the threshold level and at least 2 incorrect responses at the next higher spatial frequency. The GT test was paused if the infant was distracted or fussy. An additional stimulus was used to bring the infant's attention back towards the screen when s/he looked away. This was a small colorful, battery-operated flickering ball attached to the computer monitor that was controlled by the observer.

The experimenter timed both the TACs ad GT tests using a stopwatch. Small breaks during the testing and rechecking the test distance were included in the overall time. The interval

between the TACs and GT tests was not included. Similarly, long breaks taken while feeding or changing a diaper were not included in the overall time.

6.3 Analysis

The data were analyzed for the percentage of agreement of VA within half or one octave for repeatability of TACs and GT, as is common in the infant literature. Similar analysis was undertaken for the measurement of agreement between the TACs and GT for each visit. The association between GT VA and age was assessed with scatterplots and correlation coefficients. A repeated measures 2x2 ANOVA was used to assess effects of order or method on time taken. All the analyses were done using Excel and Statistica (StatsSoft Corp., USA).

6.4 Results

6.4.1 Demographics of study sample

The initial sample was made up of 20 infants, 55% (11) of whom were females. The age range of infants tested was between 3.2 and 11 months with a mean of 7.9 and standard deviation of 2.5 months.

6.4.2 Testability

The success rate for completion was 100% for TACs on both of the two sessions, while the GT had 100% completion rates for first session and 95% for second session. This lower testability in the second session was because one 5 month old baby boy could not finish the

second GT protocol. The rest of the analysis was undertaken with 19 participants (58% female) who completed both tests. They had a mean age of 8.1 months and SD of 2.43 months.

6.4.3 Validity

The validity of the GT was tested by comparing the acuities with the TACs.. Scatterplots of GT acuity against TAC are shown in Fig.6-3.and Fig.6-4. The correlation between the GT and TACs in the first and second visit was 0.69 and 0.62 respectively ($p<0.05$ in both cases).

As in the infant literature, agreement was calculated in terms of the percent being within 0.5 or one octave of comparison acuity. However, the spatial frequency of the GT and the TACs were not identical across the range. There are no points showing exact agreement because there is no stimulus at which the spatial frequency is exactly equal in both the methods. This was because the step size of the spatial frequencies in GT did not have exact half octave steps, due to the pixel limitation of the screen. Therefore any point(s) close to the line for equality showed near to perfect agreement between the two methods. Similarly, there were some points that were very close to being one half or one octave in difference between these 2 methods. These were 2-7% greater than 0.5 octave or 5-12% greater than one octave. As they were close and because the next nearest lower spatial frequency would have been within 0.5 or one octave, these points were included in the calculation of the percent agreement within one half or one octave. The percent agreement within 0.5 or 1 octave is shown in Table 6-2 and the agreement within one octave can be seen in Figs 6-3 and 6-4.

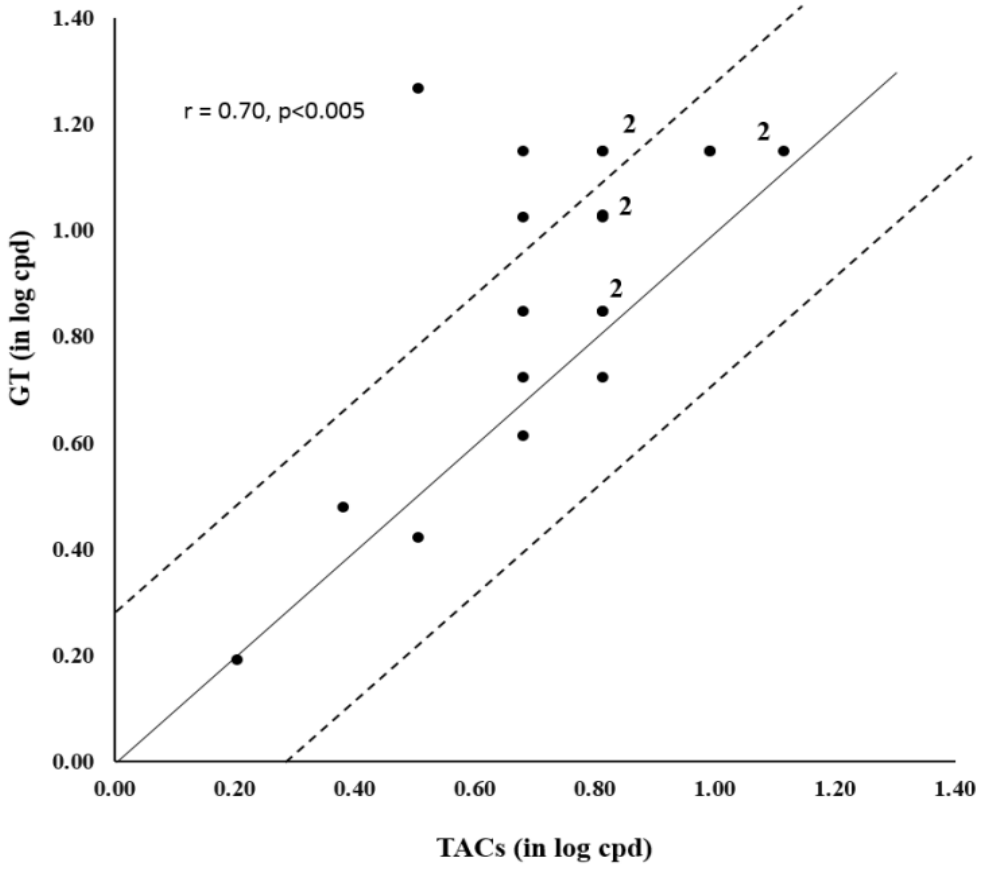


Figure 6-3. Scatterplots between GT and TACs – First visit

The solid line is the line for perfect agreement between 2 methods and the dotted lines indicate the one octave ranges. The digits indicate the data points where there is more than one data point in a given location.

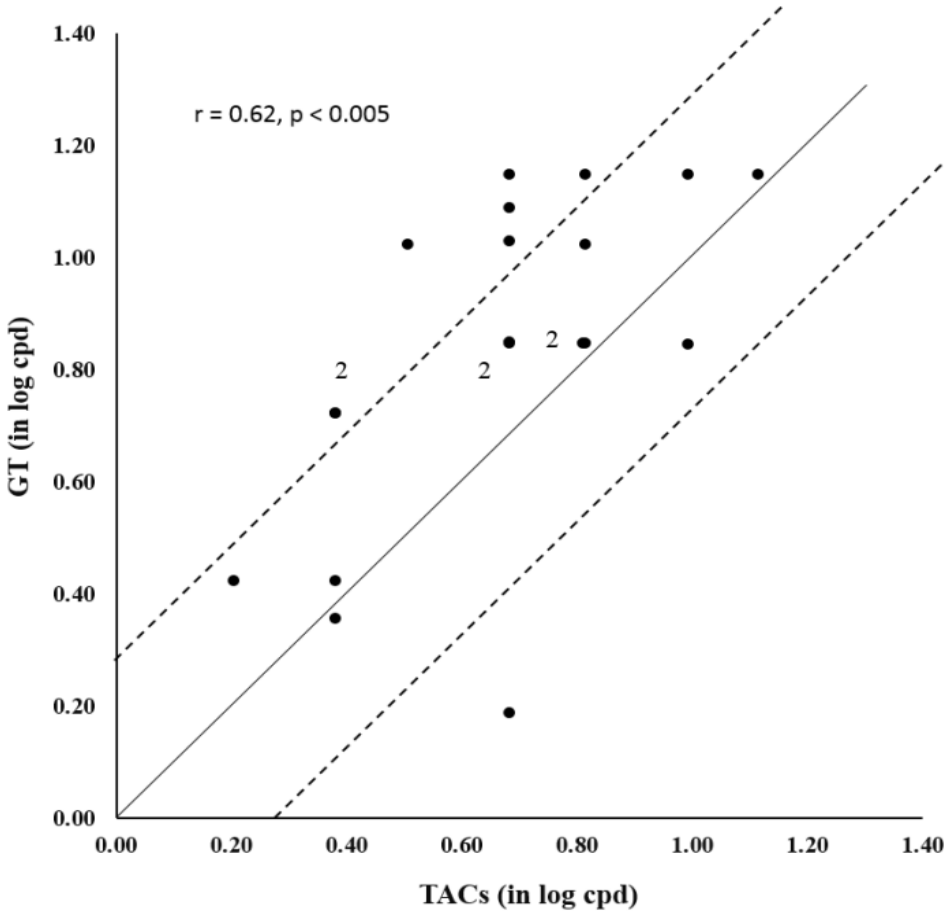


Figure 6-4. Scatterplots between GT and TACs – Second visit

The solid line is the line for perfect agreement between 2 methods and the dotted lines indicate the one octave ranges. The digits indicate the data points where there is more than one data point in a given location.

Table 6-2. Agreement between TACs and GT within half and one octaves

		Agreement between TACs and GT	
		Within the nearest 0.5	Within the nearest 1 octave
1	First Visit	63.2	89.5
2	Second visit	47	79

Since the spatial frequencies were not identical for both tests, they are taken to the nearest octave or half octave

Validity was also investigated by plotting visual acuity against age. There were significant correlations between objective GT acuities and age, as can be seen in Fig. 6-5 and 6-6. The correlations were significant at the $p<0.05$ level. The correlation coefficients are shown in Table 6-3.

Table 6-3. Correlation coefficients of GT and TAC acuities with age

	TACs 1st visit	TACs 2nd visit	GT 1st visit	GT 2nd visit
Correlation coefficient, r	0.58	0.45	0.80	0.73
p	0.009	0.053	<0.005	<0.005

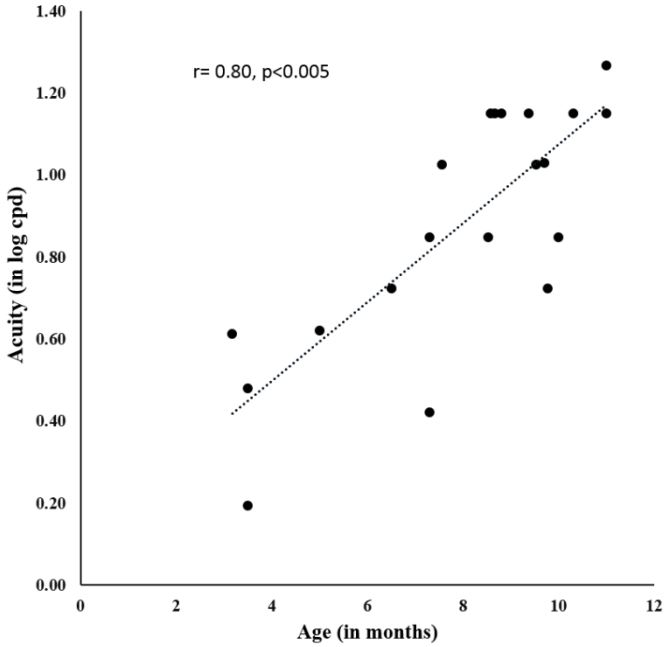


Figure 6-5. Scatterplot of VA against age – GT visit 1

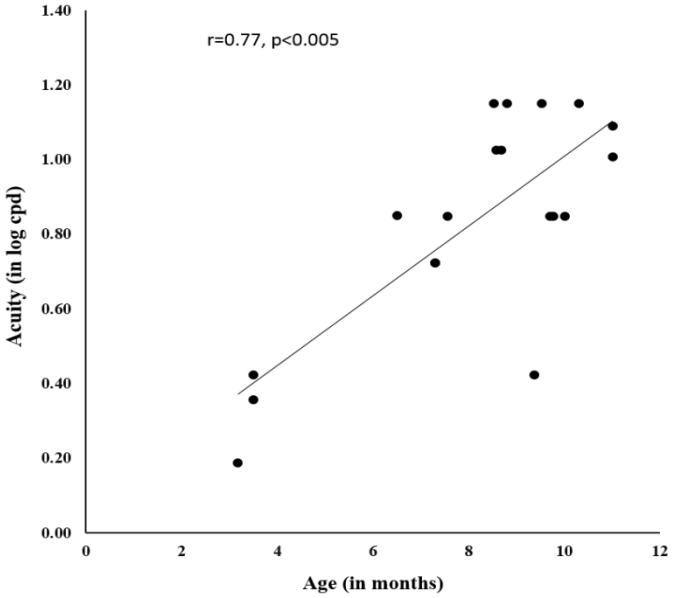


Figure 6-6. Scatterplot of VA against age – GT visit 2

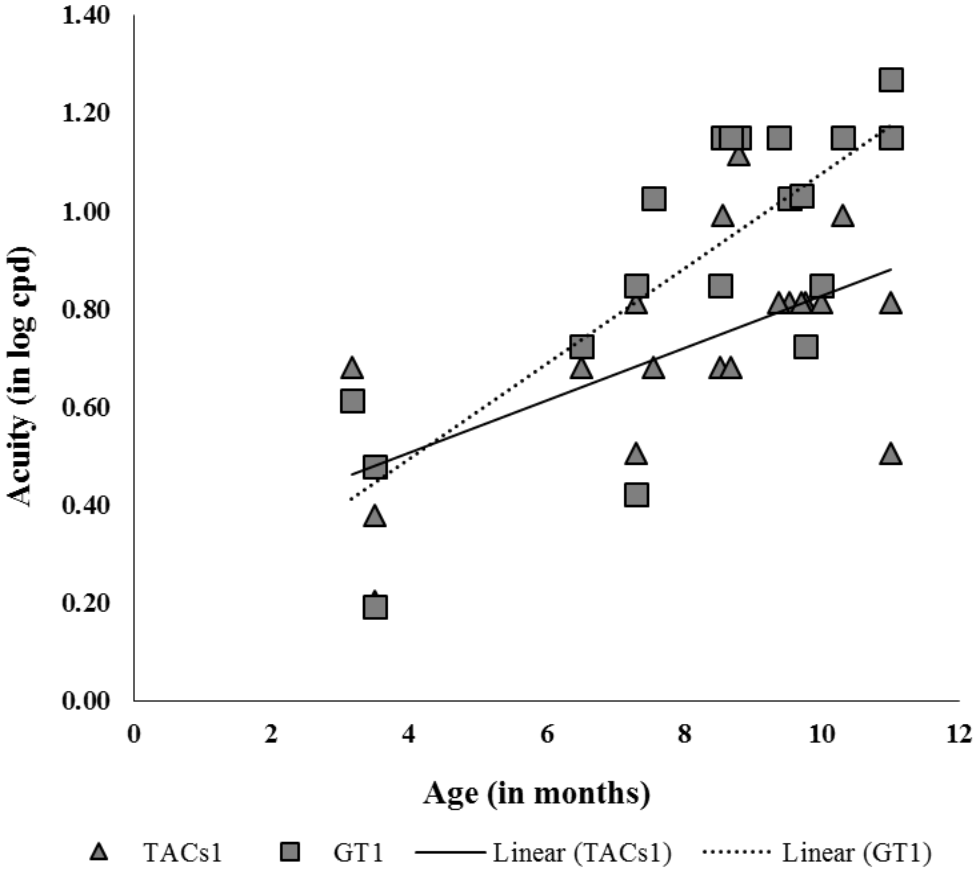


Figure 6-7. Scatterplots of GT and TACs against age – First visit.

The lines are the linear regression lines for TACs and GT

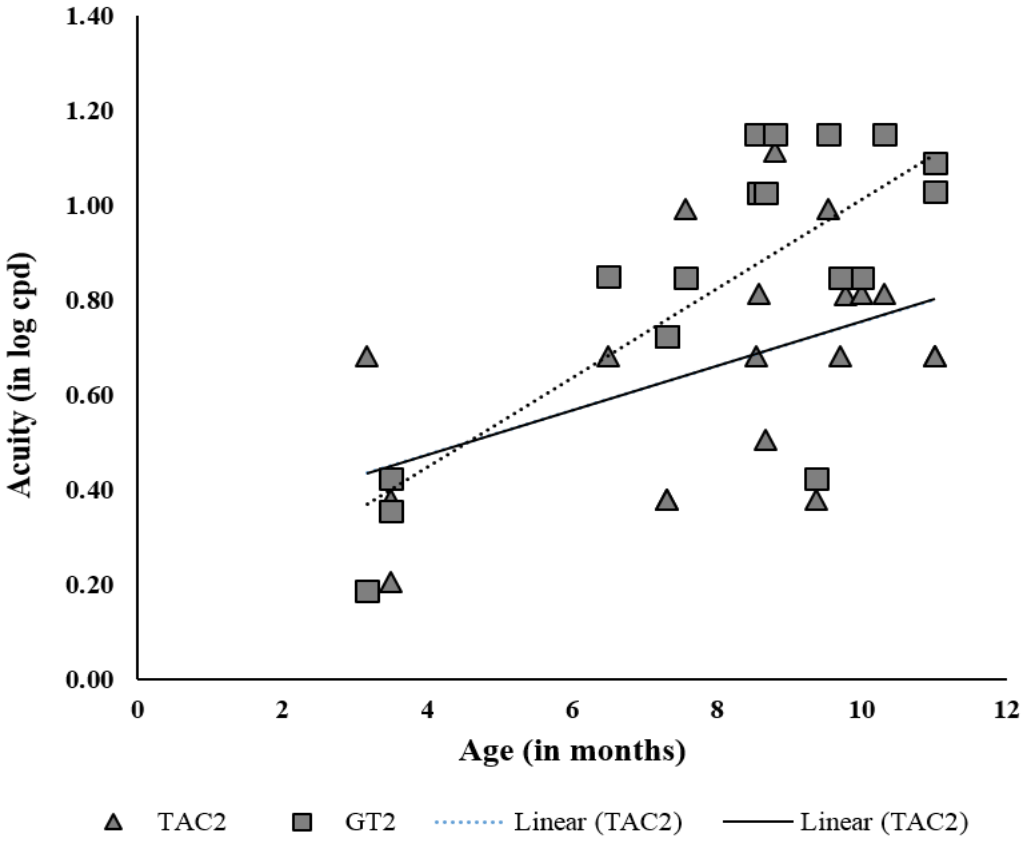


Figure 6-8. Scatterplots of GT and TAC against age – Second visit.

The lines are the linear regression lines for TACs and GT

Scatterplots between age and grating acuities for both TACs and GT are plotted for both the visits in Fig. 6-7 and Fig 6-8. It can be seen that, in both the visits, GT acuity increased more rapidly with age than TACs and also better correlation of acuity increase with age.

6.4.4 Repeatability

There was good correlation between the two visits for both the methods. The correlation coefficients for repeatability for both the methods were also very similar to each other.

The repeatability, as shown in the Fig.6-9 and 6-10, showed that there was agreement between visits in each of these methods. The agreement for repeatability within 1 octave was almost the same for both the visits, but there was slightly better agreement for TACs within half octave than GT as shown in Table 6-4.

Table 6-4. Repeatability of TACs and GT over 2 visits

	Repeatability Between Visits	
	Within 0.5 octave (%)	Within 1 octave (%)
TACs	84.2	89.5
GT	79	89.5

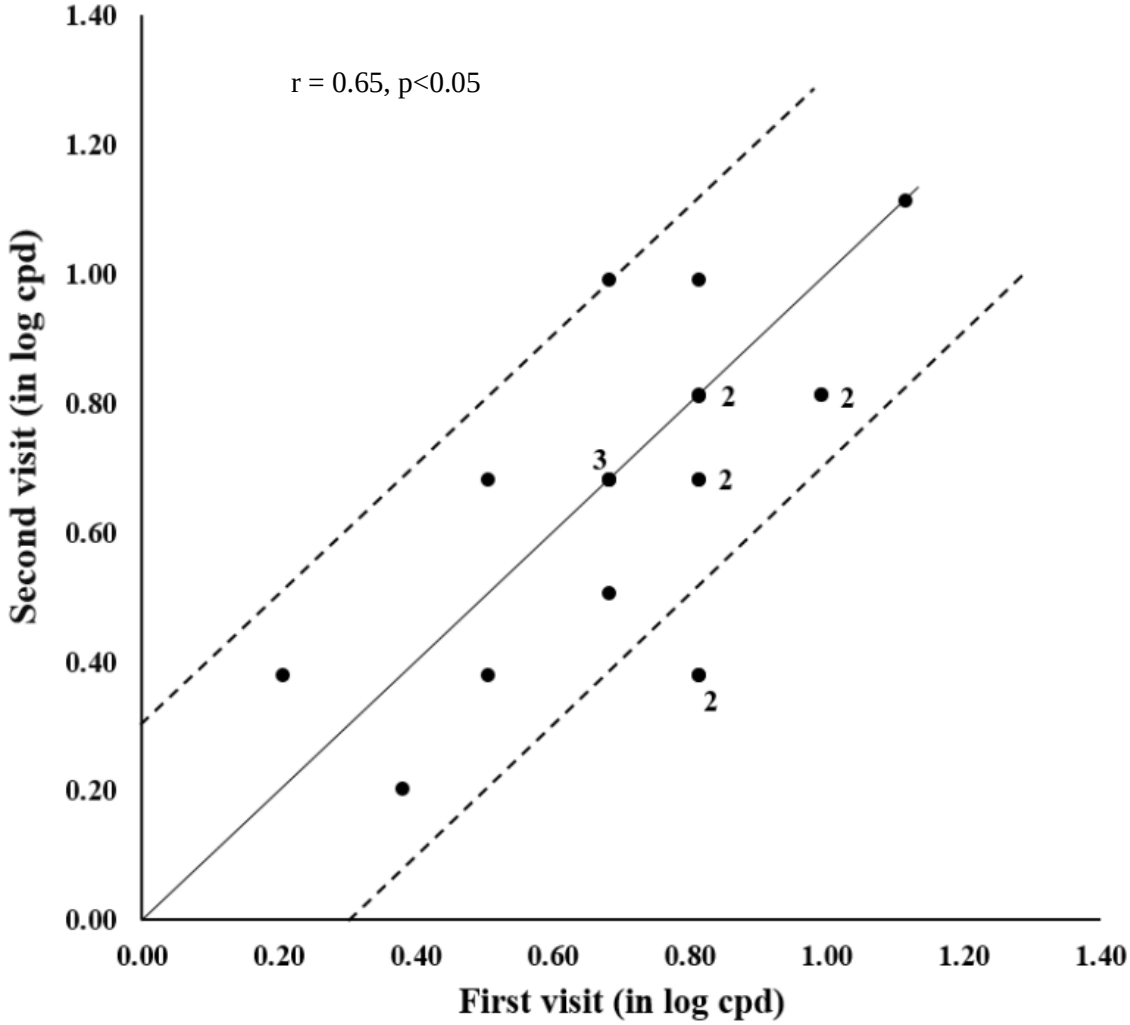


Figure 6-9. Repeatability of TAC acuities over 2 visits

The solid line is the line for perfect agreement between 2 methods and the dotted lines indicate the one octave ranges. The digits indicate the data points where there is more than one data point in a given location.

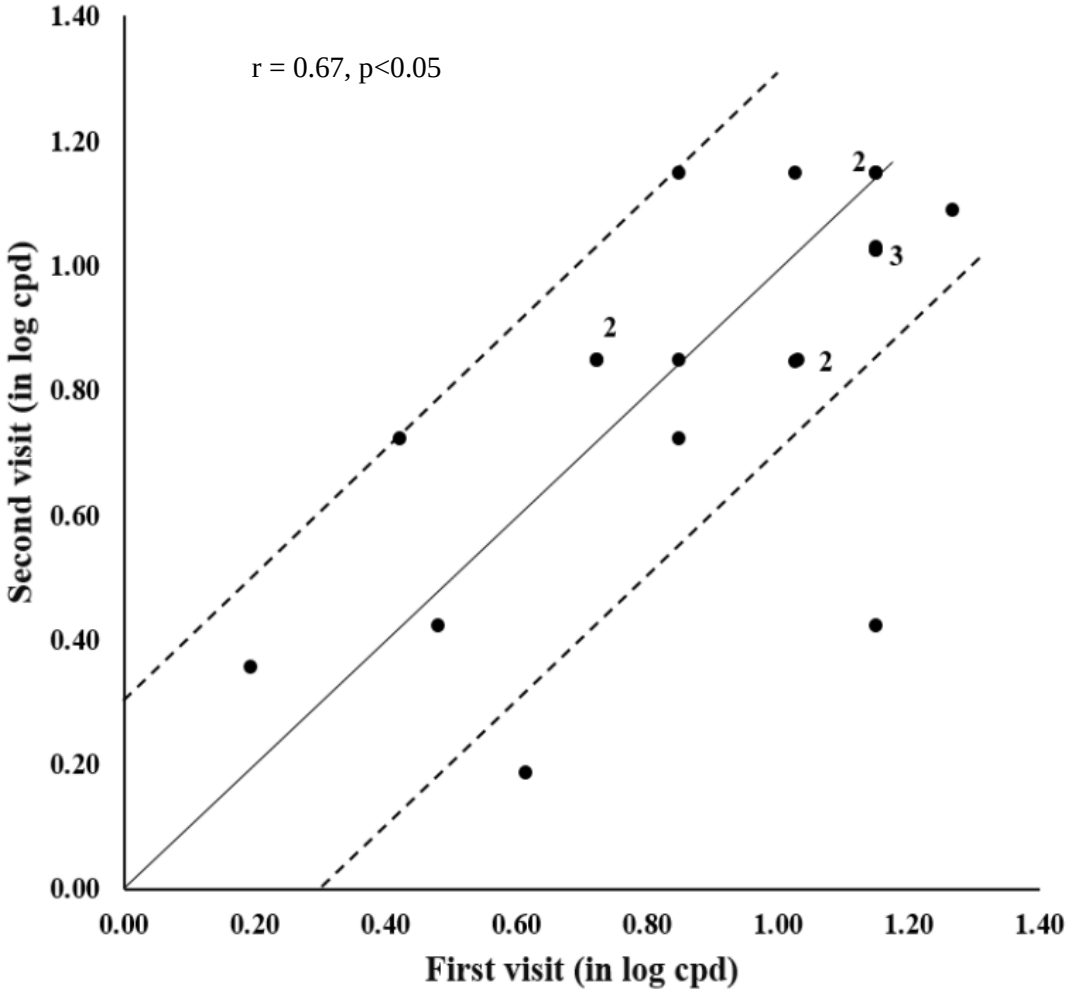


Figure 6-10. Repeatability of Gaze tracker over 2 visits

The solid line is the line for perfect agreement between 2 methods and the dotted lines indicate the one octave ranges. The digits indicate the data points where there is more than one data point in a given location.

6.4.5 Thresholds for GT and TACs

The mean thresholds from the objective GT protocols and the TACs from both the visits were compared and it was found that GT, on average, higher acuities than the Teller acuity cards. The GT estimates were on average, at least half octave higher than TACs as shown in the Fig.6-11.

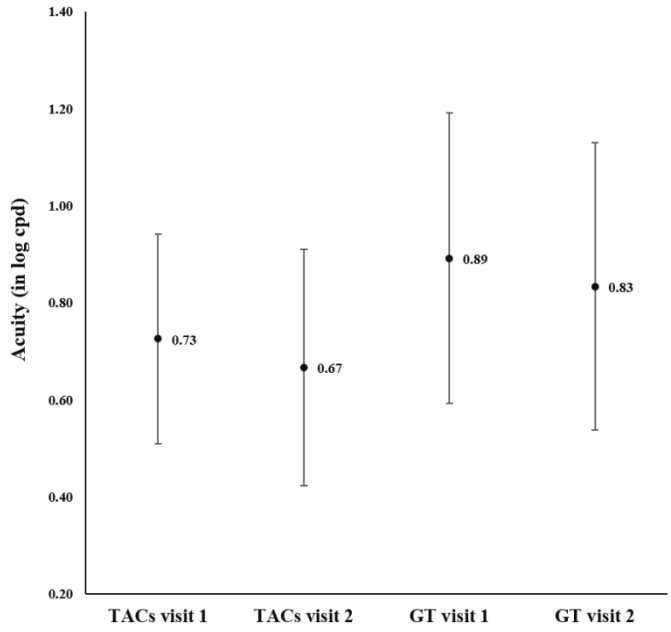


Figure 6-11.Error bars showing the mean time and ±1 SD for the acuities obtained

6.4.6 Time taken for protocols

A repeated measures ANOVA for the time taken (2 visits x 2 TAC/GT) showed a main effect of method ($F(1, 18) = 6.77$, $p = 0.02$), but no main effect of visit ($F(1, 18) = 1.3$, $p = 0.27$). There was no interaction between visit and method for the time taken ($F(1, 18) = 2.5$, $p = 0.13$).

The Teller acuity cards method was significantly faster than GT method as shown in the Fig.6-12. The second visit for TACs appears lower compared to the first visit but was not statistically significant, whereas the time taken for both visits for GT was similar.

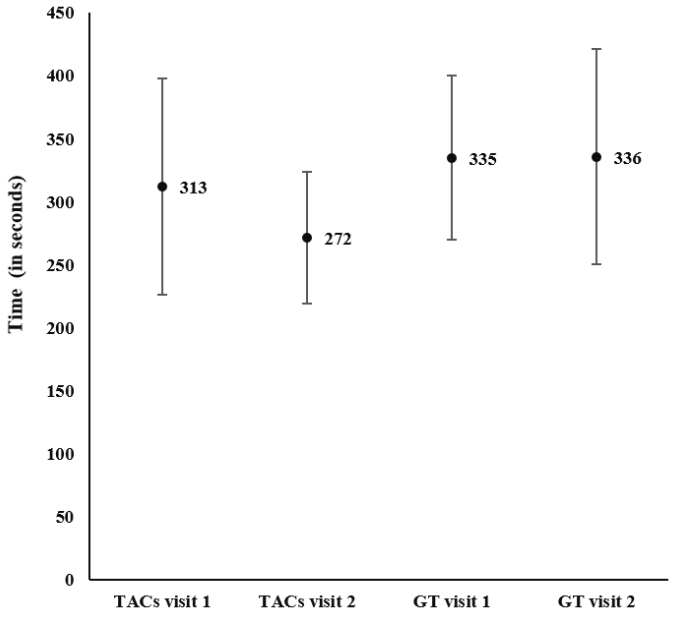


Figure 6-12. Error bars showing the mean time and ± 1 SD for both the methods

6.5 Discussion

The gaze tracker individual acuity values were compared with the age matched mean binocular VA norms (Salomao et al.[58]). The GT showed better agreement with norms than TACs and the gaze tracker gave significantly higher acuities compared to the TACs. The possible reasons for these differences will be discussed in the Chapter 7. The important finding in this chapter is that the gaze tracker showed good agreement and correlation with Teller acuity cards and that the GT had good repeatability as compared to TACs.

Table 6-5. Percent of agreement of GT and TACs with norms

	Comparison of our data with binocular VA norms[58]	
Methods	Within 0.5 octave	Within 1 octave
TACs	30.7	72
GT	61.5	87

6.6 Conclusions

GT on average gave higher acuities than TACs on both the visits. Therefore the hypothesis that there would be no difference in grating acuity by two methods namely Gaze tracker and TACs was not confirmed. Both GT and TACs gave equal repeatable measurements within 1 octave. Therefore the hypothesis that the GT measurement of VA will have no difference in

percentage of repeatable measurements as compared to TACs was confirmed. GT acuities showed improvement with increase in age, thus confirming the last hypothesis, although the improvement with age was faster with the GT than the TACs.

Chapter 7 - Discussion and Conclusions

The main aim of this thesis was validation of GT measurements of visual acuity in infants. The first step in the validation was to test adults and obtain valid results. This was an important step as that gave an essential indication about the feasibility and validity of the test for infants. If the test was not valid in adults, it would be unlikely to be successful in infant acuity testing.

7.1 Adult study

In chapter 4, I described the GT validation and testability in adults. The results showed that GT measurements agreed very well with Teller acuity cards, the standard test for clinical testing of VA in infants. We used the technique of observing the participant's actual fixations, using the software to objectively make the decision of seen/not seen for each grating. The participant was not given instructions except to view the screen. This concept was implemented to imitate the situation with infants for whom instructions cannot be given and for whom one would have to rely on naïve, natural eye movements. If the approach proved ineffective for adults (who have better attention levels) it would be deemed unlikely to be effective in infants.

The good agreement with TACs in adults showed that GT measurements are valid in principle for the measurement of grating acuity. The GT values also agreed well with the

subjective measurements of grating acuity, adding further validation to the GT measurements. There was no obvious difference in the mean GT and TAC values in the adult study. Thus the hypothesis, that there would be no difference in measured visual acuity obtained by the two methods namely, Gaze Tracker and TACs, was confirmed.

One of the most important purposes of the adult study was to test different protocols and to find the best and most efficient protocol for testing, which could then be used for obtaining VA in infants. The staircase and halving of range methods were found to be the most efficient methods in the adult study. Therefore, it was decided that these two methods would be used for measuring VA in infants. The correlation with uncorrected refractive error was an additional validation. Objective GT grating acuity correlated well with refractive error i.e., as uncorrected refractive error increased, visual acuity became poorer and there was a similar level of correlation with letter acuity. Thus the hypothesis that there will be a significant correlation between uncorrected refractive error and visual acuity in the adult population was confirmed. Although the correlations were similar, grating acuity was less susceptible to blur from uncorrected refractive error than letter acuity i.e. the loss in grating acuity with increasing refractive error was lower with grating acuity compared to letter acuity. The trend was similar to the results of the study by Thorn and Schwartz[106] in which they evaluated the effects of dioptric blur on grating and letter acuity (see Fig. 7-1). The gaze tracker grating acuity showed a mild decrease in grating acuity with respect to refractive error with a similar trend to the grating acuity in the Thorn and Schwartz[106] study (see Fig. 7-2), although the letter acuity in the current study did not show the same degree of reduction. There was

greater reduction in acuity compared to the magnitude of myopia in the Thorn et al.[106] study than in the current study. There are several possible reasons for the difference. Firstly, the letter chart used in their study was the Snellen chart as compared to the log MAR chart used in the current study. Secondly the sample size (n=7) was smaller in the Thorn et al. study. Lastly, the blur in their study was induced with plus lenses whereas in our study, uncorrected myopes were used. There is some evidence that myopes are better at interpreting blur than those who are not myopic[107]. Therefore myopes were less affected by the blur, and obtained slightly higher acuity than the induced myopes. This might explain why there was a greater reduction in acuity in the participants in the study by Thorn et al. study who were actually induced myopes.

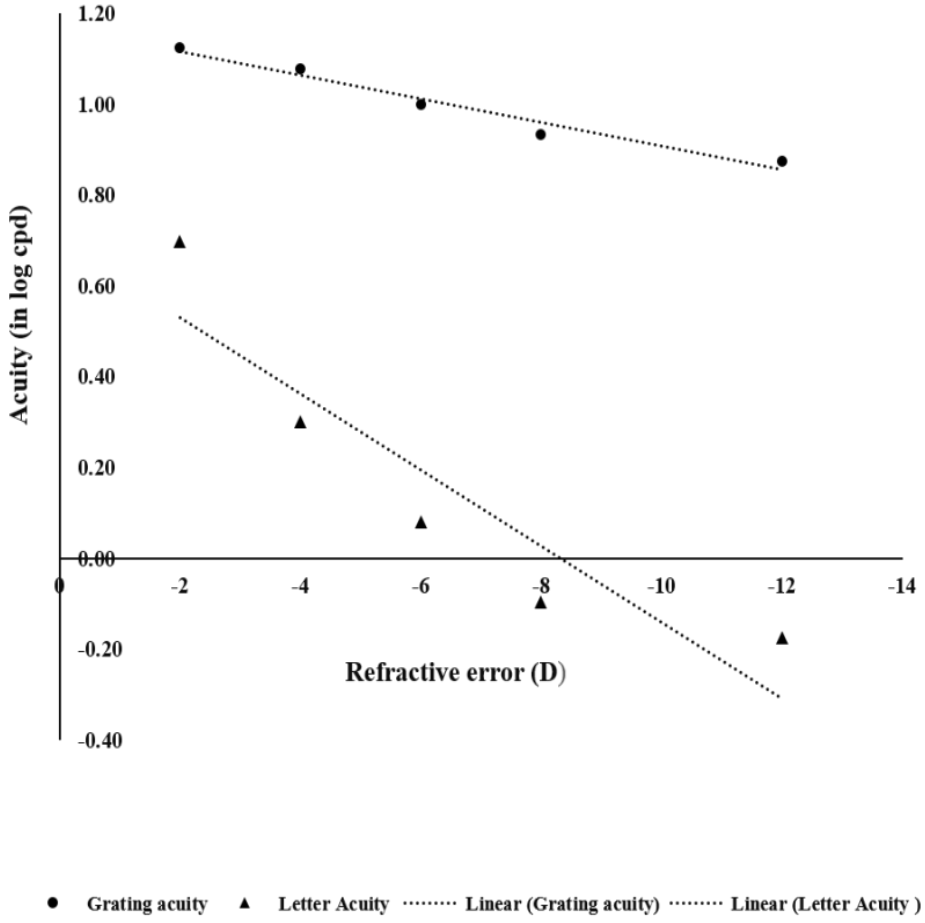


Figure 7-1. Letter and grating acuity against refractive error
(derived from Thorn & Schwartz, 1989)

The dotted lines represent least squares linear regression fits.

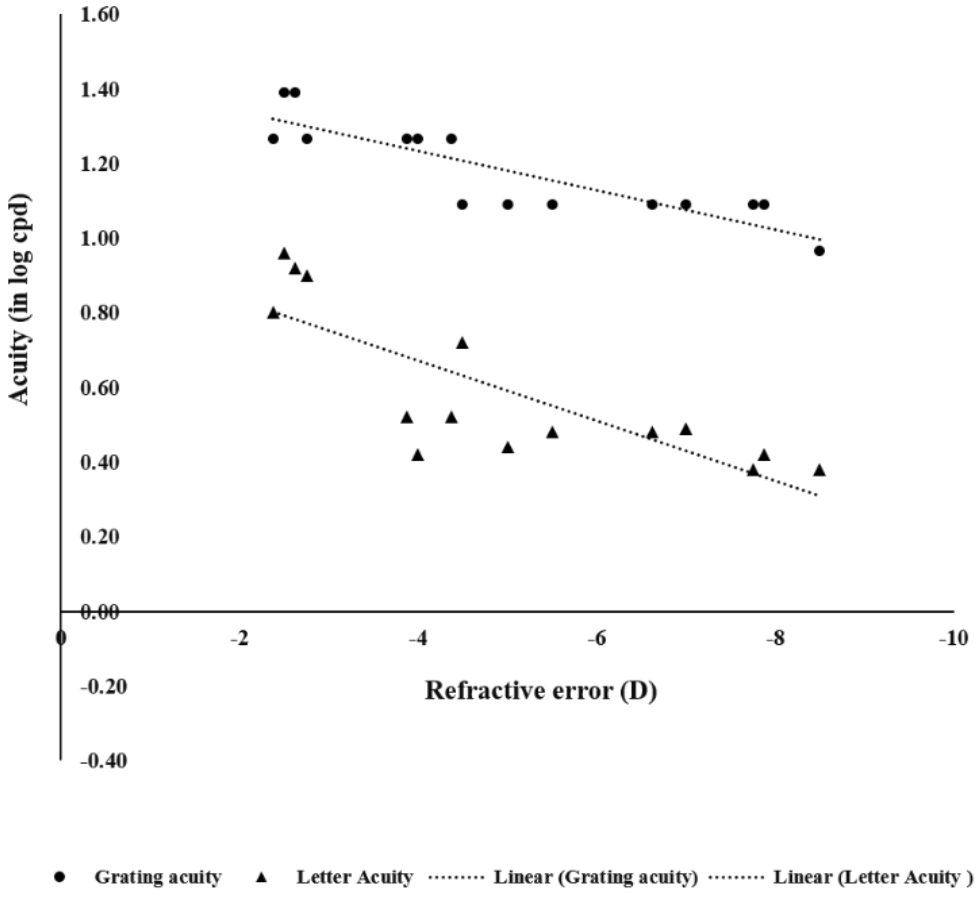


Figure 7-2. Letter and grating acuity against refractive error (current study)

The dotted lines represent least squares linear regression fits.

7.1.1 Possible limitations of the adult study

In the current study the experimenter knew the spatial frequencies he was presenting and also made the decision of whether each stimulus was seen or not seen. There is a possible bias as it was the same experimenter who performed both the TACs and the gaze tracker. Nevertheless, the good agreement of all the objective protocols with the subjective methods (both with the stimuli presented in the gaze tracker and TACs) helps to indicate that the agreements are genuine, as the experimenter would have little influence on the subjective results, which was based on a strict protocol.

7.2 Preliminary Infant trials

The next step was assessing the testability of the GT in infants. Chapter 5 described the feasibility and preliminary trial of VA testing using the GT in infants. This preliminary study in infants enabled testing for specific protocols and also the number of protocols that could be tested in a session. The results showed that there was a good testability of GT in infants and that the GT can give a VA measurement in infants. The results of the preliminary infant trials showed that infants get tired (habituated) to the stimuli when we had 4 different protocols, two each for GT and TACs. Therefore the result of the final protocol that was performed may not have been fully reliable - during that protocol there were many periods of loss of attention. Although, we used videos to keep the infant's attention, they still seemed to prefer to look at real faces and objects in real world. In fact, if a certain video was repeated, they also lost interest in that video. Therefore we had two different videos during set up and for use between stimulus presentations.

Calibration of the instrument had to be done before we actually tested the various spatial frequencies. The time taken for calibration depends on the infant's attention to fix and follow the target in the screen. If for some reason, infants were not interested in the targets for a while or calibration was not successful quickly, there was habituation to the calibration stimuli and the infant would lose attention and would get fussy. It became a situation of diminishing returns. If the investigators felt that the infant was too tired or fussy during calibration, we exited the test and tried again from the start of the experiment. If after the break, the infant was not happy and continue to be fussy, then the test was terminated. In some cases, the infant needed feeding or a diaper change, and after this, was able to continue testing. We discovered that it was important to ask the parent to let us know if they thought the baby needed either of these things, or a pacifier. The preliminary trials gave much insight in testing of infants using GT as well as TACs. It allowed us to optimize the procedure and protocols. The list of modifications that were made as a result of these preliminary trials can be seen in Chapter 6.

7.3 Main Infant study

The testability rates were very good and comparable to previously reported acuity card studies as shown in the Table 7-1.

Table 7-1 Testability of binocular acuity card studies

Study	Population	N	Age (in months)	Testability (%)
McDonald et al.[32]	Normally developing infants	32	1-6	100
McDonald et al.[49]	Normally developing infants	9	18	90
Sebris et al.[108]	Normal full term infants	168	6.6	97
Salomao et al[58]	Normal full term infants	726	0-36	99.3
GT (this study)	Normally developing children	20	3-12	95
TACs (this study)	Normally developing children	20	3-12	100

7.3.1 Validation assessment

7.3.1.1 Agreement between GT and TACs in current study

The final and most important step in the project was validation of VA measurements in infants using GT with the final protocol as described in Chapter 6. The validation was done by three methods namely, comparison with the Teller acuity cards, by comparing with infant binocular acuity norms and demonstrating an increase of GT VA with age. The acuities obtained by GT, on average were higher than TACs. Thus the first hypothesis, that there would be no difference in measured visual acuity obtained by the two methods namely, Gaze Tracker and TACs, was not confirmed. This is discussed in more detail below.

However, as in the infant literature, agreement was also calculated in terms of the percent being within 0.5 or one octave of comparison acuity. The agreement between GT and TACs was 79 – 90% within 1 octave and 47-63% within 0.5 octave. Since the spatial frequency of

the GT and the TACs were not identical across the range, there were some points that were very close to being one half or one octave in difference between these 2 methods. These were 2-7% greater than 0.5 octave or 5-12% greater than one octave. As they were close and because the next nearest lower spatial frequency would have been within 0.5 or one octave (which would have been the measure of VA if the higher frequency was not available), these points were included in the calculation of the percent agreement within one half or one octave. Identical TAC and GT acuities across the whole range of spatial frequencies cannot be obtained even at modification of testing distance. We compared our data with other infant studies which investigated agreement between two different techniques that were used for measuring acuity in infants. Our agreement measures are comparable with values obtained in the infant literature for binocular acuity card studies. The agreements were slightly lower between the GT and TACs in the second visit compared to the first visit (see Table 7-2). The possible reason could be that infants were habituated with the stimuli and had slightly lower attention levels.

Table 7-2 Agreement between ACP and PL binocular acuity card studies in the infant literature compared to agreement between GT and TAC in the current study

	Study	Age	% within 0.5 octave	% within 1 octave
1	McDonald et al.[32,49]	1-12 months and 18-36 months		95-100
2	Preston et al.[31]	2-8 months		75
3	Mohn et al.[51]	6-36 months		96
4	Lewis et al.[38]	15-30 months	44-67	66-75
5	Current study (between GT and TACs)	3-12 months	47 (2^{nd} visit) -63 (1^{st} visit)	79 (2^{nd} visit) - 90 (1^{st} visit)

* - percents were obtained by comparison with norms[58]

7.3.1.2 Agreement between GT and age-related norms from the infant literature

We also assessed the validity of the measurements by comparing the GT and TAC values with norms in the literature. The VA data of each of the 20 and 19 infants who completed the GT and TACs respectively were compared with the results of TACs for the similar age group in the infant literature[58]. The binocular data of Salomao et al.[58] were used for this purpose, as they measured binocular and monocular acuities in a large sample of 646 healthy children aged between 0 and 36 months of age. Both monocular and binocular VA showed a sharp improvement in acuity from birth and 6 months of age and slower growth thereafter. The half and one octave range of the Salomao et al.[58] data were calculated on either side of the mean acuity for each age group. The frequency of GT and TAC values in the current study that lie between those ranges was calculated as percent of agreement with the norms (Table 7-3).

Table 7-3. Agreement of GT and TACs with the binocular acuity norms

Current study	Age	Within 0.5 octave	Within 1 octave
GT	3-12 months	61.5	87.2
TACs	3-12 months	30	72.5

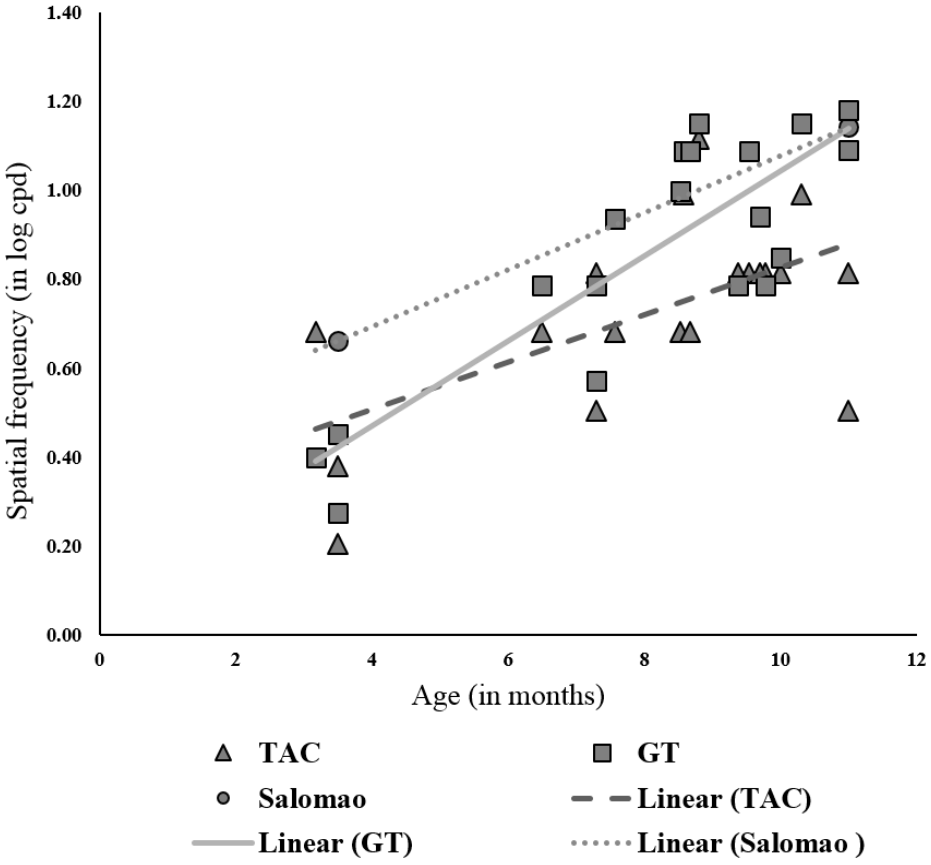


Figure 7-3. Scatter plots showing average acuities plotted against age

The lines are the linear regression lines for TACs, GT and Salamao et al.

The TAC values which did not agree with the published norms mostly fell below the lower limit of those norms as shown in the linear regression line in Figure 7-3. The regression lines between the published norms and the TACs values obtained in this study were close to parallel indicating there might be constant underestimation of VA of about 0.2 log units (approximately one level of the TACs) with the TACs in this study. The possible reason for TAC acuity tending to be lower than GT acuity might be that TACs were performed with the cards held in the vertical orientation. This scenario required infants to look at the top or the bottom of the cards. The observer noted that some infants had a strong tendency to prefer to look up rather than down or vice versa. This makes the judgment of the final acuity more difficult. In summary, it might be that horizontal orientation of the TACs yields better acuities than cards held vertically. This may explain the difference in the acuities in the vertically and horizontally held TACs in the preliminary infant trials. It was originally thought that this difference in the preliminary study was because of the lack of the stage for the vertically-held cards, but it may have been due to the difference in eye movements. It is not expected that there would be an actual difference in mean acuities obtained between horizontal and vertical gratings in infants[109].

The GT data in comparison with the norms and TACs (in this study) show a different pattern of VA development (see linear regression lines in the Fig 7-3,). At the younger age, the GT and current TACs data are in agreement, but the current and published TAC data are not in agreement. GT gave higher acuities for older age infants than the current TAC data but was in agreement with the published TAC data. One possibile explanation is that the requirement

for a vertical eye movement may cause this difference, if vertical eye movements were more difficult in the older infants. Vertical saccades have been shown to be less accurate in adults[110] The only evidence in the literature regarding vertical compared to horizontal saccades in infants is an article by Gredeback et al.[111] in which it was reported that vertical saccades had a shorter latency than horizontal saccades and that vertical saccades were made more frequently in 4-8 month olds. It is possible that horizontal saccades continue to develop later than vertical saccades, resulting in the difference found in adults. This might explain the lack of discrepancy at the younger age between the published norms and the present TAC results. As the infant gets older, the difference in saccades might develop, resulting in a difference between the published and present TAC results, but the GT results may be less affected, as smaller eye movements are required. Therefore the faster development of VA as measured with the GT may be a more accurate reflection of VA development.

7.3.1.3 Correlation with age

Another way of assessing the validity is to see that if the VAs correlated with age. The GT VA showed positive significant correlation with age (r=0.80 and 0.77 for the first and second visit respectively). All previous studies of VA in this age group showed that VA improves with age. The TAC results showed moderate but significant correlation with age (r=0.58 and r=0.45 for the first and second visit respectively). Thus the hypothesis that visual acuity, as measured by the GT, will improve with age in the infant group was confirmed.

7.3.2 Intraobserver reliability

The repeatability or the intraobserver reliability of GT was measured over two visits. The percent of GT values for intraobserver reliability within 1 octave were similar to acuity card studies in the infant literature. In other words, GT measurements had a good between-visits repeatability compared to a similar study in infant literature as shown in Table 7-4. McDonald et al.[32] found a correlation for intraobserver reliability for the acuity card procedure of 0.66 in infants aged between 1 to 6 months. Similarly values of correlation for intraobserver reliability for GT ($r=0.65$, $p<0.005$) and TACs ($r=0.67$, $p<0.005$) were obtained in this study. The interval between visits in McDonald et al.[32] studied was 1 day, whereas in our study it ranged from 7 to 10 days. The shorter time interval between the two visits in the study by McDonald et al. could be because the infants were younger in their study than the current study. The time interval between visits in our study was appropriate considering the visual development for the different infant age groups. Despite the longer interval, the correlation in the present study was similar to the study done by McDonald et al.

Table 7-4. Intraobserver reliability study in infant literature compared with the current study

S No.	Method	Age	% within 1 octave
1	Acuity card procedure[32]	1-6 months	87.5
2	GT (this study)	3-12 months	89.5
3	TACs (this study)	3-12 months	89.5

Both the gaze tracker and Teller acuity cards gave similar repeatability of about 89.5% within 1 octave between visits. Therefore, the hypothesis that the GT measurement of VA would have no difference in percentage of repeatable measurements as compared to TACs, was confirmed. This is an additional measure that is important as an indicator of the quality and usefulness of the measurement.

7.4 Limitations

One of the limitations of the study was that the spatial frequencies were not identical in the two tests. This did not matter for adults as their agreement was so good. This was because of the finite levels of spatial frequencies that are available in the GT because of the pixel nature of the screen. An unconventional testing distance for the TACs could have been used in the infant study to make the spatial frequencies more exact (although that would not totally have eliminated the differences in all spatial frequencies).

Another limitation was the fact that we used the TACs vertically, which is not the same as in the literature. We had a valid reason to use it that way; as we were measuring acuity in GT using horizontal gratings, it was important to have horizontal gratings as well in TAC testing. However it would have been useful in retrospect if we had infants return for a third visit for measurement of acuity using TACs in the conventional way (horizontal orientation). Practically, having infants come thrice during a short period would have been more difficult than the current scenario and might have decreased recruitment and increased drop-out rates

from the study. The alternate way to have avoided that problem would have been able to rotate the GT screen 90º to have vertical gratings, if it was technically feasible. This would have involved considerable software modifications, and was not feasable in the current study.

Another possible improvement that could have been done is to make the calibration faster. As the calibration, took longer time in some cases than the usual ones, infants lost attention and the performance of actual testing might have been afffected. It was possible to stop the calibration before completion, and run the experiment without calibration, but that makes judgement of eye tracking very difficult. Ideally, it would have been useful to be able to pause the calibration, if the infant lost interest, and to resume from we had left off. In addition, based on the observations made during this study, it would be better for testing infants to have the video cameras positioned slightly above the line of sight, rather than below it. This would enable tracking of infants' eyes even during bottle feeding, which is an ideal time to capture their attention. In the current situation, the bottle interrupts the camera view. This would optimise the use of the infants's limited attention span.

Another possible improvement would be to have two different observers for GT and TACs to keep it totally unbiased, although the method we have adopted in this study reduced most of the bias, as the observer did not know the spatial frequency threshold in either case. In addition, if the eye tracking information was displayed on a third monitor, so that the observer would not get to see the infant movements or direction of the gaze or the actual gratings, this would have removed any possibility of bias.

One source of error in measurement of both GT and TAC acuity might be a change in acuity values, if the testing distance is altered. The infant was not restrained physically to maintain the exact viewing distance. Measured acuity would apparently increase with a decrease in testing distance and vice versa. This effect would be greater with a smaller viewing distance and therefore a potentially greater problem with TACs than GT. With the TACs, a 10 cms change in viewing distance from the 55 cms distance used would result in a change in 0.04 log units. This change is less than 1/3rd octave (0.15 log units) and therefore significantly less than an increase or decrease to the next acuity card (0.5 octave). So this error is not considered substantial, as during the experiment the coordinator kept track of the distance during the test and corrected the distance, if the infant moved closer or farther.

7.5 Conclusions and Future work

This study gives evidence that GT has great potential for measuring VA objectively in infants. It showed 100 % testability for measuring VA in infants. In future, specific algorithms could be developed that fully automate VA testing in infants. The staircase method, established in this method could be used as the basis of an automated protocol. The raw data collected in this study, such as presentation times, time intervals between presentations and duration for first fixations can be used to develop these algorithms. This a wealth of detailed data is collected by the GT, but is not recorded with the TACs, which rely on the gross and observable looking behaviour. The feasability of such an automated method could be studied and may result in faster VA measurement times. This would eliminate all subjective bias that is encountered while using the GT and result in a fully objective measure of VA in infants.

The GT has most potential for pre-verbal infants and children up to the age when they can do a matching test of VA. The GT mainly relies on eye movement fixations to judge the VA and may be useful in the whole age group up to 3 years, who cannot give reliable subjective responses. The age range tested in this study is 3-12 months of age. Therefore, another step in GT validation is to study children aged > 1 year up to 3 years of age. GT validation would be undertaken in a similar experimental design to the present study.

Intraobserver reliability was studied in this study. In future, interobserver agreement should be studied and assessed for agreement. Another step in the validation process is with infants or children with ocular disorders. If poorer acuity is obtained in infants with ocular disorders, this would further validate the GT measurements. Now that we have obtained binocular acuities using GT, another step would be to demonstrate the feasability of GT for measuring monocular acuities in infants and establish the validity of this measurement also. Measuring monocular acuity is an important step in detecting ocular disorders and amblyopia.

Thus this study established the potential of the gaze tracking to measure VA in infants and indicates the value of further evaluations and complete automation in future.

Bibliography

1. Kniestedt C, Stamper RL. Visual acuity and its measurement. Ophthalmology Clinics of North America. 2003;16(2):155–70.

2. Leat SJ Westall CA. SRH. Assessing Children's Vision: A Handbook.; Butterworth-Heinemann Oxford, UK. 1999.

3. Hess RF, Dakin SC, Tewfik M, Brown B. Contour interaction in amblyopia: scale selection. Vision Research. 2001;41(17):2285–2296.

4. Tyler CW, Apkarian P, Levi DM, Nakayama K. Rapid assessment of visual function: An electronic sweep technique for the pattern visual evoked potential. Investigative Ophthalmology and Visual Science. 1979;18(7):703–713.

5. Almoqbel F, Leat SJ, Irving E. The technique, validity and clinical use of the sweep VEP. Ophthalmic and Physiological Optics. 2008;28(5):393–403.

6. Norcia AM, Tyler CW. Spatial frequency sweep VEP: Visual acuity during the first year of life. Vision Research. 1985;25(10):1399–1408.

7. Norcia A, Tyler C, Hamer R. Development of contrast sensitivity in the human infant. Vision Research. 1990;30(10):1475–1486.

8. Hamer RD, Norcia AM, Tyler CW, Hsu-Winges C. The development of monocular and binocular VEP acuity. Vision Research. 1989;29(4):397–408.

9. Rosenberg M, Jabbari B. Nystagmus and visual evoked potentials. Neuro-ophthalmology. 1987;7(3):133–138.

10. Hoffmann MB, Seufert PS, Bach M. Simulated nystagmus suppresses pattern-reversal but not pattern-onset visual evoked potentials. Clinical Neurophysiology. 2004;115(11):2659–2665.

11. Smith D. Controversies in pediatric ophthalmology - the clinical usefulness of the visually evoked potential. Journal of Pediatric Ophthalmology and Strabismus. 1984;21:235–236.

12. Naegele JR, Held R. The postnatal development of monocular optokinetic nystagmus in infants. Vision Research. 1982;22(3):341–346.

13. Catford G V, Oliver A. Development of visual acuity. Archives of Disease in Childhood. 1973;48(1):47–50.

14. Hopkisson B, Arnold P, Billingham B, McGarrigle M, Entwistle P. Visual assessment of infants: vernier targets for the Catford drum. The British Journal of Ophthalmology. 1991;75(5):280–3.

15. Van Hof-Van Duin J, Mohn G. Optokinetic and spontaneous nystagmus in children with neurological disorders. Behavioural Brain Research. 1983;10(1):163–175.

16. Fantz RL. Visual perception from birth as shown by pattern selectivity. Annals of the New York Academy of Sciences. 1965;118(21):793–814.

17. Fantz RL. Pattern vision in young infants. The Psychological Record,. 1958;8:43–47.

18. Fantz RL. Pattern vision in newborn infants. Science. 1963;140(3564):296–297.

19. Fantz RL, Ordy JM, Udelf MS. Maturation of pattern vision in infants during the first six months. Journal of Comparative and Physiological Psychology. 1962;55(6):907–917.

20. Fantz, R. L., Fagan, J. F., & Miranda SB. Early visual selectivity. In: L. B. Cohen & P. Salapatek (Eds.) Infant perception: From Sensation to Cognition,. New York: Academic Press; 1975:1, 249–346.

21. Teller DY. The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. Infant Behavior and Development. 1979;2:135–153.

22. Teller DY, Morse R, Borton R, Regal D. Visual acuity for vertical and diagonal gratings in human infants. Vision Research. 1974;14(12):1433–1439.

23. Dobson V, Fulton AB, Sebris SL. Cycloplegic refractions of infants and young children: the axis of astigmatism. Investigative Ophthalmology and Visual Science. 1984;25(1):83–87.

24. Howland H, Sayles N. Photorefractive measurements of astigmatism in infants and young children. Investigative Ophthalmology and Visual Science. 1984;25(1):93–102.

25. Howland HC, Atkinson J, Braddick O FJ. Infant astigmatism measured by photorefraction. Science. 1978;202:331–333.

26. Mohindra I, Held R, Gwiazda J, Brill S. Astigmatism in infants. Science. 1978;202(4365):329–331.

27. Gwiazda J, Scheiman M, Mohindra I, Held R. Astigmatism in children: changes in axis and amount from birth to six years. Investigative Ophthalmology and Visual Science. 1984;25(1):88–92.

28. Banks MS, Salapatek P. Contrast sensitivity function of the infant visual system. Vision Research. 1976;16(8):867–869.

29. Atkinson J, Braddick O, Moar K. Development of contrast sensitivity over the first 3 months of life in the human infant. Vision Research. 1977;17(9):1037–1044.

30. Dobson V, Schwartz TL, Sandstrom DJ, Michel L. Binocular visual acuity of neonates: the acuity card procedure. Developmental Medicine & Child Neurology. 1987;29(2):199–206.

31. Preston KL, McDonald M, Sebris SL. Validation of the acuity card procedure for assessment of infants with ocular disorders. Ophthalmology. 1987;94(6):644–653.

32. McDonald MA, Dobson V, Sebris SL, Baitch L, Varner D, Teller DY. The acuity card procedure: a rapid test of infant acuity. Investigative ophthalmology & visual science. 1985;26(8):1158–1162.

33. Clifford CE, Haynes BM, Dobson V. Are norms based on the original teller acuity cards appropriate for use with the new teller acuity cards II? Journal of AAPOS. 2005;9(5):475–479.

34. Clifford-Donaldson CE, Haynes BM, Dobson V. Teller Acuity Card norms with and without use of a testing stage. Journal of AAPOS. 2006;10(6):547–551.

35. Mayer DL, Dobson V. Assessment of vision in young children: a new operant approach yields estimates of acuity. Investigative Ophthalmology & Visual Science. 1980;19(5):566–570.

36. Wilson WR. Instrumentation for Operant Infant Auditory Assessment. Child Development and Mental Retardation Center. Seattle, WA. 1976.

37. Mayer DL, Dobson V. Visual acuity development in infants and young children, as assessed by operant preferential looking. Vision Research. 1982;22(9):1141–1151.

38. Lewis, T. L., Reed, M. J., Maurer, D., Wyngaarden, P. A., & Brent HP (1993). An evaluation of a acuity cards procedure. Clinical Vision Sciences. 1993;8(6):591 – 602.

39. Hertz BG. Acuity card testing of retarded children. Behavioural Brain Research. 1987;24(2):85–92.

40. Mohn G, Van Hof-van Duin J, Fetter WP, De Groot L, Hage M. Acuity assessment of non-verbal infants and children: clinical experience with the acuity card procedure. Developmental Medicine and Child Neurology. 1988;30(2):232–244.

41. Hertz BG, Rosenberg J, Sjö O, Warburg M. Acuity card testing of patients with cerebral visual impairment. Developmental Medicine & Child Neurology. 1988;30(5):632–637.

42. Hertz BG, Rosenberg J. Acuity card testing of spastic children: Preliminary results. Journal of Pediatric Ophthalmology and Strabismus. 1988;25(3):139–144.

43. Marx MS, Werner P, Cohen-Mansfield J, Hartmann EE. Visual acuity estimates in noncommunicative elderly persons. Investigative Ophthalmology & Visual Science. 1990;31(3):593–596.

44. Dobson V, Quinn GE, Biglan AW, Tung B, Flynn JT, Palmer EA. Acuity card assessment of visual function in the cryotherapy for retinopathy of prematurity trial. Investigative Ophthalmology and Visual Science. 1990;31(9):1702–1708.

45. Hertz BG, Rosenberg J. Effect of mental retardation and motor disability on testing with visual acuity cards. Developmental Medicine and Child Neurology. 1992;34(2):115–122.

46. Dobson V. Visual acuity testing in infants: from laboratory to clinic. Early Visual Development: Normal and Abnormal. New York, NY: Oxford University Press Inc,. 1993:318–334.

47. Van Hof-Van Duin J, Mohn G. The development of visual acuity in normal fullterm and preterm infants. Vision Research. 1986;26(6):909–916.

48. Mayer DL, Fulton AB, Hansen RM. Preferential looking acuity obtained with a staircase procedure in pediatric patients. Investigative ophthalmology & visual science. 1982;23(4):538.

49. McDonald M, Ankrum C, Preston K, Sebris SL, Dobson V. Monocular and binocular acuity estimation in 18- to 36-month-olds: acuity card results. American Journal of Optometry and Physiological Optics. 1986;63(3):181–186.

50. McDonald M, Sebris SL, Mohn G, Teller DY, Dobson V. Monocular acuity in normal infants: the acuity card procedure. American Journal of Optometry and Physiological Optics. 1986;63(2):127.

51. Mohn G, Van Hof-Van Duin J. Rapid assessment of visual acuity in infants and children in a clinical setting, using acuity cards. In: Detection and Measurement of Visual Impairment in Pre-Verbal Children.Vol 26. Springer Netherlands; 1986:363–371.

52. Brown AM, Yamamoto M. Visual acuity in newborn and preterm infants measured with grating acuity cards. American Journal of Ophthalmology. 1986;102(2):245–253.

53. Heersema DJ, Van Hof-Van Duin J. Age norms for visual acuity in toddlers using the acuity card procedure. Clinical Vision Sciences. 1990;5(2):167–174.

54. Jacobson SG, Mohindra I, Held R. Visual acuity of infants with ocular diseases. American Journal of Ophthalmology. 1982;93(2):198–209.

55. Mohindra I, Jacobson SG, Zwaan J, Held R. Psychophysical assessment of visual acuity in infants with visual disorders. Behavioural Brain Research. 1983;10(1):51–58.

56. Courage ML, Adams RJ. Visual acuity assessment from birth to three years using the acuity card procedure: cross-sectional and longitudinal samples. Optometry & Vision Science. 1990;67(9):713–718.

57. Mayer DL, Beiser AS, Warner AF, Pratt EM, Raye KN, Lang JM. Monocular acuity norms for the Teller Acuity Cards between ages one month and four years. Investigative Ophthalmology and Visual Science. 1995;36(3):671– 685.

58. Salomao SR, Ventura DF. Large sample population age norms for visual acuities obtained with Vistech-Teller Acuity Cards. Investigative Ophthalmology and Visual Science. 1995;36(3):657–670.

59. Luna B, Dobson V, Biglan AW. Development of grating acuity in infants with regressed stage 3 retinopathy of prematurity. Investigative Ophthalmology and Visual Science. 1990;31(10):2082–2087.

60. Birch EE, Spencer R. Visual outcome in infants with cicatricial retinopathy of prematurity. Investigative Ophthalmology and Visual Science. 1991;32(2):410–415.

61. Katsumi O, Mehta MC, Matsui Y, Tetsuka H, Hirose T. Development of vision in retinopathy of prematurity. Archives of Ophthalmology. 1991;109(10):1394–1398.

62. Drover JR, Wyatt LM, Stager DR, Birch EE. The Teller acuity cards are effective in detecting amblyopia. Optometry and Vision Science. 2009;86(6):755–759.

63. Paik HJ, Shin MK. The Clinical Interpretation of Teller Acuity Card Test in the Diagnosis of Amblyopia. Journal of the Korean Ophthalmological Society. 2001;42(7):1030–1036.

64. Lanzi G, Fazzi E, Uggetti C, et al. Cerebral visual impairment in periventricular leukomalacia. Neuropediatrics. 1998;29(3):145–150.

65. Kushner BJ, Lucchese NJ, Morton G V. Grating visual acuity with teller cards compared with Snellen visual acuity in literate patients. Archives of Ophthalmology. 1995;113(4):485–493.

66. Hertz BG. Use of the acuity card method to test retarded children in special schools. Child: Care, Health and Development. 1988;14(3):189–198.

67. Dobson V, Carpenter N. Interobserver agreement and observer bias in the acuity card procedure. Technical Digest on Noninvasive Assessment of the Visual system. 1991:20–23.

68. Mash C, Dobson V, Carpenter N. Interobserver agreement for measurement of grating acuity and interocular acuity differences with the teller acuity card procedure. Vision Research. 1995;35(2):303–312.

69. Getz LM, Dobson V, Luna B, Mash C. Interobserver reliability of the Teller Acuity Card procedure in pediatric patients. Investigative Ophthalmology and Visual Science. 1996;37(1):180–187.

70. Dobson, V., Carpenter, N. A., Bonvalot, K., & Bossler J. The acuity card procedure-interobserver agreement in infants with perinatal complications. Clinical Vision Sciences. 1990;6(1):39–48.

71. Atkinson J, Braddick O, Pimm-Smith E. “Preferential looking”for monocular and binocular acuity testing of infants. British Journal of Ophthalmology. 1982;66(4):264–268.

72. Maurer D, Lewis T, Brent H. The effects of deprivation on human visual development: Studies of children treated for cataracts. Applied Developmental Psychology. 1989;3:139–227.

73. Mash C, Dobson V. Long-term reliability and predictive validity of the Teller Acuity Card procedure. Vision Research. 1998;38(4):619–626.

74. Hall HL, Courage ML, Adams RJ. The predictive utility of the Teller acuity cards for assessing visual outcome in children with preterm birth and associated perinatal risks. Vision Research. 2000;40(15):2067–2076.

75. Mash C, Dobson V. Intraobserver reliability of the Teller Acuity Card procedure in infants with perinatal complications. Optometry & Vision Science. 2005;82(9):817–822.

76. Kohl P, Samek M. Refractive error and preferential looking visual acuity in infants 12-24 months of age: year 2 of a longitudinal study. Journal of the American Optometric Association. 1988;59(9):686–690.

77. Leat S, Yadav N, Irving E. Development of visual acuity and contrast sensitivity in children. Journal of Optometry. 2009;2(1):19–26.

78. Dobson V, Teller DY, Belgum J. Visual acuity in human infants: A review and comparison of behavioral and electrophysiological studies. Vision Research. 1978;18(11):1233–1238.

79. Gorman J, Cogan D, Gellis S. A device for testing visual acuity in infants. Sight-Saving Rev. 1959;29:80–84.

80. Marg Freeman, D. N., Peltzman, P., and Goldstein, P. J. E. Visual acuity development in human infants: evoked potential measurements. Investigative Ophthalmology and Visual Science. 1976;15:150–153.

81. Sokol S, Moskowitz A. Comparison of pattern VEPs and preferential-looking behavior in 3-month-old infants. Investigative Ophthalmology and Visual Science. 1985;26(3):359–65.

82. Riddell PM, Ladenheim B, Mast J, Catalano T, Nobile R, Hainline L. Comparison of measures of visual acuity in infants: Teller acuity cards and sweep visual evoked potentials. Optometry and Vision Science. 1997;74(9):702–707.

83. Tinelli F, Pei F, Guzzetta A, et al. The assessment of visual acuity in children with periventricular damage: a comparison of behavioural and electrophysiological techniques. Vision Research. 2008;48(10):1233–41.

84. Sokol S, Moskowitz A, McCormack G. Infant VEP and preferential looking acuity measured with phase alternating gratings. Investigative Ophthalmology and Visual Science. 1992;33(11):3156–3161.

85. Sokol S, Moskowitz A, Mccormack G, Augliere R. Infant grating acuity is temporally tuned. Vision Research. 1988;28(12):1357–66.

86. Sokol S, Moskowitz A, Hansen V. Electrophysiological evidence for the oblique effect in human infants. Investigative Ophthalmology and Visual Science. 1987;28(4):731–735.

87. Mohn G, Van Hof-van Duin J. Development of Spatial Vision. In: Vision and visual dysfunction. 10th ed. 1991:179–211.

88. Spinelli D, Pirchio M, Sandini G. Visual acuity in the young infant is highest in a small retinal area. Vision Research. 1983;23(10):1133–1136.

89. Mayer DL, Fulton AB, Hansen RM. Visual acuity of infants and children with retinal degenerations. Ophthalmic Paediatrics And Genetics. 1985;5(1-2):51–56.

90. Sokol S, Moskowitz A, Paul A. Evoked potential estimates of visual accommodation in infants. Vision Research. 1983;23(9):851–860.

91. Cassidy L, Taylor D, Harris C. Abnormal supranuclear eye movements in the child: A practical guide to examination and interpretation. Survey of Ophthalmology. 2000;44(6):479–506.

92. Irving EL, Steinbach MJ, Lillakas L, Babu RJ, Hutchings N. Horizontal saccade dynamics across the human life span. Investigative Ophthalmology & Visual Science. 2006;47(6):2478–2484.

93. Levene MI, Chervenak FA. The special senses. In: Fetal and Neonatal Neurology and Neurosurgery. Elsevier Health Sciences; 2009:744.

94. Guestrin ED, Eizenman M. Remote point-of-gaze estimation requiring a single-point calibration for applications with infants. In: Eye Tracking Research and Applications Symposium (ETRA).; 2008:267–274.

95. Mulvey, F., Villanueva, A., Sliney, D., Lange, R., Cotmore, S., & Donegan M. Exploration of safety issues in Eyetracking. Communication by Gaze Interaction (COGAIN). 2008;D5.

96. Cassel D. An Objective Methodology to Assess Visual Acuity Using Visual Scanning Parameters (masters thesis). Toronto, Ontario, Canada: University of Toronto. 2009.

97. Guestrin ED, Eizenman M. General theory of remote gaze estimation using the pupil center and corneal reflections. IEEE Transactions on Biomedical Engineering. 2006;53(6):1124–1133.

98. Sturm V, Cassel D, Eizenman M. Objective estimation of visual acuity with preferential looking. Investigative Ophthalmology and Visual Science. 2011;52(2):708–713.

99. Model D, Eizenman M. A probabilistic approach for the estimation of angle kappa in infants. Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '12. 2012:53.

100. Shih S, Wu Y. A calibration-free gaze tracking technique. Proceedings 15th International Conference on Pattern Recognition. ICPR-2000. 2000;4:201–204.

101. Nagamatsu T, Sugano R, Iwamoto Y, Kamahara J, Tanaka N. User-calibration-free gaze tracking with estimation of the horizontal angles between the visual and the optical axes of both eyes.

Proceedings of the 2010 Symposium on EyeTracking Research Applications ETRA 10. 2010;1(212):251–254.

102. Model D, Eizenman M. An automatic personal calibration procedure for advanced gaze estimation systems. IEEE transactions on bio-medical engineering. 2010;57(5):1031–1039.

103. Bailey IL, Lovie JE. New design principles for visual acuity letter charts. American journal of optometry and physiological optics. 1976;53(11):740–5.

104. Gwiazda J, Thorn F, Bauer J, Held R. Emmetropization and the progression of manifest refraction in children followed from infancy to puberty. Clinical Vision Sciences. 1993;8(4):337–344.

105. Gwiazda J, Grice K, Held R, McLellan J, Thorn F. Astigmatism and the development of myopia in children. Vision Research. 2000;40(8):1019–1026.

106. Thorn F, Schwartz F. Effects of dioptric blur on Snellen and grating acuity. Optometry and Vision Science. 1990;67(1):3–7.

107. Rosenfield M, Abraham-Cohen JA. Blur Sensitivity in Myopes. Optometry and Vision Science. 1999;76(5):303–307.

108. Sebris S, Dobson V, McDonald M, Teller D. Acuity cards for visual acuity assessment of infants and children in clinical settings. Clinical Vision Sciences. 1987;2(1):45–58.

109. Birch EE, Gwiazda J, Bauer JA, Naegele J, Held R. Visual acuity and its meridional variations in children aged 7-60 months. Vision research. 1983;23(10):1019–1024.

110. Collewijn H, Erkelens CJ, Steinman RM. Binocular co-ordination of human vertical saccadic eye movements. The Journal of physiology. 1988;404:183–97.

111. Gredebäck G, Ornkloo H, Von Hofsten C. The development of reactive saccade latencies. Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale. 2006;173(1):159–64.

112. Mayer DL. Cycloplegic Refractions in Healthy Children Aged 1 Through 48 Months. Archives of Ophthalmology. 2001;119(11):1625–1628.