INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

A Xerox Education Company

73-15,319

DENNIS, John Michael, 1942-
THE RELIABILITY OF SEVERAL RESPONSE INDICES
USED IN VESTIBULAR ASSESSMENT.

The University of Oklahoma, Ph.D., 1972
Health Sciences, general

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE RELIABILITY OF SEVERAL RESPONSE INDICES USED IN VESTIBULAR ASSESSMENT

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

JOHN MICHAEL DENNIS

Oklahoma City, Oklahoma

1972

THE RELIABILITY OF SEVERAL RESPONSE INDICES USED IN VESTIBULAR ASSESSMENT

APPROVED BY:

DISSERTATION COMMITTEE

PLEASE NOTE:

Some pages may have

indistinct print.

Filmed as received.

University Microfilms, A Xerox Education Company

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to the following:

To Dr. Gerald A. Studebaker for the advice, guidance and patience exercised during his direction of this experiment.

To the dissertation reading committee and final examination committee comprised of Doctors: Gerald Studebaker, Joseph Barry, Thomas Stokinger, Eugene Mencke and Willard Moran.

To the Department of Otolaryngology, University of Oklahoma Health Sciences Center, for providing the instruments and facilities for this study.

To the individuals who participated as subjects for this study.

To my wife, Judy, for her support and encouragement during my graduate studies.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

LIST OF ILLUSTRATIONS

THE RELIABILITY OF SEVERAL RESPONSE INDICES USED IN VESTIBULAR ASSESSMENT

CHAPTER I

INTRODUCTION

Caloric tests are used to assess the functional status of the vestibular system. These tests are customarily performed by introducing water of a temperature different from that of the human body into the external auditory canal. The temperature change in the vestibular apparatus produced by the water causes a flow of endolymph which, in turn, stimulates the vestibular receptors. This stimulation ultimately causes an involuntary movement of the eyes which is termed nystagmus. If the vestibular and the response systems are intact, the nystagmus is quite apparent. If the vestibular system is abnormal, the nystagmus under conditions of caloric stimulation may be diminished or absent.

The Fitzgerald-Hallpike test (24) is by far the most widely used caloric test procedure. As the test was originally described each ear is separately irrigated with two different temperatures of water, one seven degrees centigrade above body temperature and the other seven degrees centigrade below body temperature. The examiner observes the nystagmus visually and records the duration of the nystagmic reaction in seconds. The duration of the response is the parameter used to determine whether or not the reaction is normal.

Later, the technique of electronystagmography was applied to the assessment of vestibular function. A graphic display of the calorically induced nystagmus is obtained with this procedure by electro-mechanically recording the movement of the eyes. Electronystagmography allows the extraction of response parameters other than duration from the recorded data. These parameters include measurement of the various speed, frequency and amplitude measurements of the eye movements (2, 4, 37, 40, 41, 57, 66, 72, 79, 80).

Duration is the length of time in seconds that the induced nystagmus persists but provides little, if any, information concerning the magnitude of the ongoing reaction. However, the speed, frequency and amplitude measures of response available through electronystagmography do provide information about the magnitude of the reaction and examiners have used these measures to identify abnormal response in pathological cases where duration was within normal limits. Consequently, the response indices that reflect the magnitude of the ongoing caloric reaction are widely employed in electronystagmography. Each index is used to one extent or another depending upon individual preference and no common agreement has been reached as to which index is most suitable.

Empirical evidence concerning the reliability of the several parameters may provide information concerning which index or indices best fulfill clinical purposes. Several investigators have studied the effects of caloric testing on normal subjects and reported considerable inter-individual variability in the responses (4, 37, 44, 45, 72, 73, 75). Because of this variability across subjects, the responses evoked from one labyrinth of an individual are normally compared to those evoked from the opposite labyrinth of the same individual thus allowing the patient to serve as his own control.

Information concerning the reliability of the various response parameters when caloric testing is repeated on the same subject is limited (35, 37, 41). Clinically, many patients are retested in order to confirm the results of prior testing, to determine progress of treatment and to specify if the degree of induced nystagmus is commensurate with the current activity level of a disease. Therefore, information concerning the repeatability of the various response parameters in given individuals could have significant clinical implications.

The present study consists of an intrasubject replication of the Fitzgerald-Hallpike procedure to determine the repeatability of the several response indices of the induced nystagmic reaction. Normal human subjects were used.[1] Each subject was tested three times. The measures of nystagmus which were investigated were duration of the reaction, speed of the slow phase, total nystagmic beats during a specified time interval, culmination frequency (the greatest number of beats obtained in two sequential five-second intervals) and total amplitude (the sum in degrees of rotation of the amplitudes of all the slow phase components occurring during a specified interval of time). Except for duration these measurements were computed from within the temporal interval extending from 60 to 90 seconds from the onset of irrigation. A discussion of previous work most pertinent to this investigation is presented in the following chapter.

---

[1]The definition of normal for the purposes of this study is the absence of abnormal findings on a hearing screening test, neurological tests and a negative history of vestibular or chronic ear pathology. This is discussed in further detail in chapter iii.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

The term nystagmus is used to denote a regular, involuntary, conjugate movement of the eyes. Vestibular nystagmus evoked by thermal stimulation of the horizontal semicircular canal is the subject of study in this investigation. More specifically, an attempt was made to determine the reliability of several methods of expressing the magnitude of thermally induced nystagmus recorded electronystagmographically from normal human beings. This chaper is devoted to a discussion of vestibular testing. A review of the literature pertinent to the caloric technique and electronystagmographic recording is presented.

Neuronal discharge from the peripheral vestibular apparatus is carried via the vestibular portion of the eighth cranial nerve to the four vestibular nuclei located in the upper medulla and lower pons. From these structures impulses may be directed to connections in the cerebellum, the motor nuclei of the extra-ocular muscles and the reticular gray matter of the brain stem and to motor neurons at various levels of the spinal cord (7, 19, 28).

Caloric stimulation of the horizontal semicircular canal elicits nystagmus by changing the temperature in the region of the temporal bone which contains the vestibular structures. This temperature change has a cooling or a warming effect on the endolymph with respect to body temperature depending upon whether the water is below or above body temperature. During cold irrigation the portion of the endolymph in the lateral section of the horizontal canal is cooled. The cooling alters the specific gravity of the endolymph so that it becomes more dense and descends bringing the fluid above it into a position to be cooled. This process creates a movement of fluid within the horizontal semicircular canal. This fluid movement initiates neural action which causes the eyes to rotate in the direction of the movement or, in this case, toward the ear being irrigated. This vestibular eye movement is repeatedly interrupted by a rapid eye movement in the opposite direction. The former is termed the slow component of nystagmus and is of vestibular origin while the latter is termed the fast component of nystagmus and is considered to be of central origin (37, 40, 44, 47, 66). The direction of the induced nystagmus is designated by the direction of the fast component.

With warm irrigation the direction of nystagmus is reversed. Here, the warmed fluid becomes less dense than the surrounding fluid and rises. The slow component is then directed toward the side opposite irrigation and the fast component toward the irrigated ear. This type of vestibular activity is seen in the presence of a normal vestibular and ocular system. However, this system, like the other sensory systems, is subject to pathological disturbance at the peripheral or central level (1, 4, 8, 19, 24, 30, 36, 40, 46, 52, 72, 73, 74, 76).

Early Vestibular Testing

Schmiederkam discovered in 1868, that nystagmus could be produced by introducing a stream of water into the ear canal (70). This and other pioneering work in caloric testing prompted a substantial amount of investigation relating to tests that determine the functional activity or sensitivity of the vestibular system. In addition to thermic stimulation by water, other measures have been developed that utilize rotational stimulation (13, 17, 31, 53, 72, 73). These have not enjoyed a wide popularity in clinical use because of the difficulty in stimulating each labyrinth separately. The remainder of this discussion will be devoted principally to the thermal douching techniques.

The initial efforts in vestibular testing are reported in foreign journals beginning at the early part of the twentieth century. These journals are not conveniently available and require translation. Fortunately, more recent literature surveys this early era most adequately (24, 37, 40, 43, 44, 72, 78).

In 1906, Barany proposed a technique whereby a fairly large volume of water at 30°C or 20°C was injected into the external auditory meatus until a nystagmus was detected. He placed the patient in either a supine or sitting position. In both cases, the horizontal canal was placed in a vertical orientation.

Kobrak, in 1918, introduced the idea of minimal stimulation. He irrigated the ear with 5 to 10 cubic centimeters of water at 20°C or 27°C and 47°C until a nystagmus was seen. Grahe, in 1920, and de Kleyn, in 1928, used a similar procedure except that their lowest irrigating temperature was 17°C. Fischer and Wodak, in 1922, used minimal stimulation in a sitting position with the head tilted forward 30° so that the horizontal canal was in the horizontal plane. They irrigated for 10 seconds, waited one minute and then slowly moved the head backward taking 20 seconds until the horizontal canal was in the vertical plane. At this point, the temperature change was supposed to have its greatest effect and a nystagmus was easily observed. Veits, in 1928, used a similar technique syringing with 10 cubic centimeters of water in 10 seconds, waiting one minute but then moving the head to the optimum position in two seconds.

In 1942, Fitzgerald and Hallpike (24) developed a test of vestibular function that remains popular today. The technique separately elicited nystagmic reaction from both labyrinths by injecting first cold and then warm water into the external auditory meati. The patient was placed in a supine position with his head flexed 30° above the horizontal plane. This brought the horizontal semicircular canal into a vertical plane which is the desired position for maximum thermal excitation. Each ear was douched individually with water at 30°C and 44°C (seven degrees below and seven degrees above body temperature) for 40 seconds. Five minutes was allowed to elapse between douchings. During the test the patient was instructed to fixate his gaze on a mark on the ceiling and an assistant observed the duration of the nystagmic response visually and timed it with a stop watch.

Fitzgerald and Hallpike classified the results of the test as indicating normal response, a right or left canal paresis or a directional preponderance to the right or left. They illustrated a graphical method of depicting the results which is delineated in Figure 1. Each continuous line represented a total of three minutes, and was subdivided into one-minute, 20-second and 10-second intervals. The interrupted line denotes the duration of the nystagmic response measured from the application of the stimuli to the end of the visible nystagmus. Responses are numbered one through four. The sum of reactions one and three represent the response of the left labyrinth. The sum of reactions two and four represent the response of the right labyrinth. The sum of reactions one and four consist of nystagmus to the

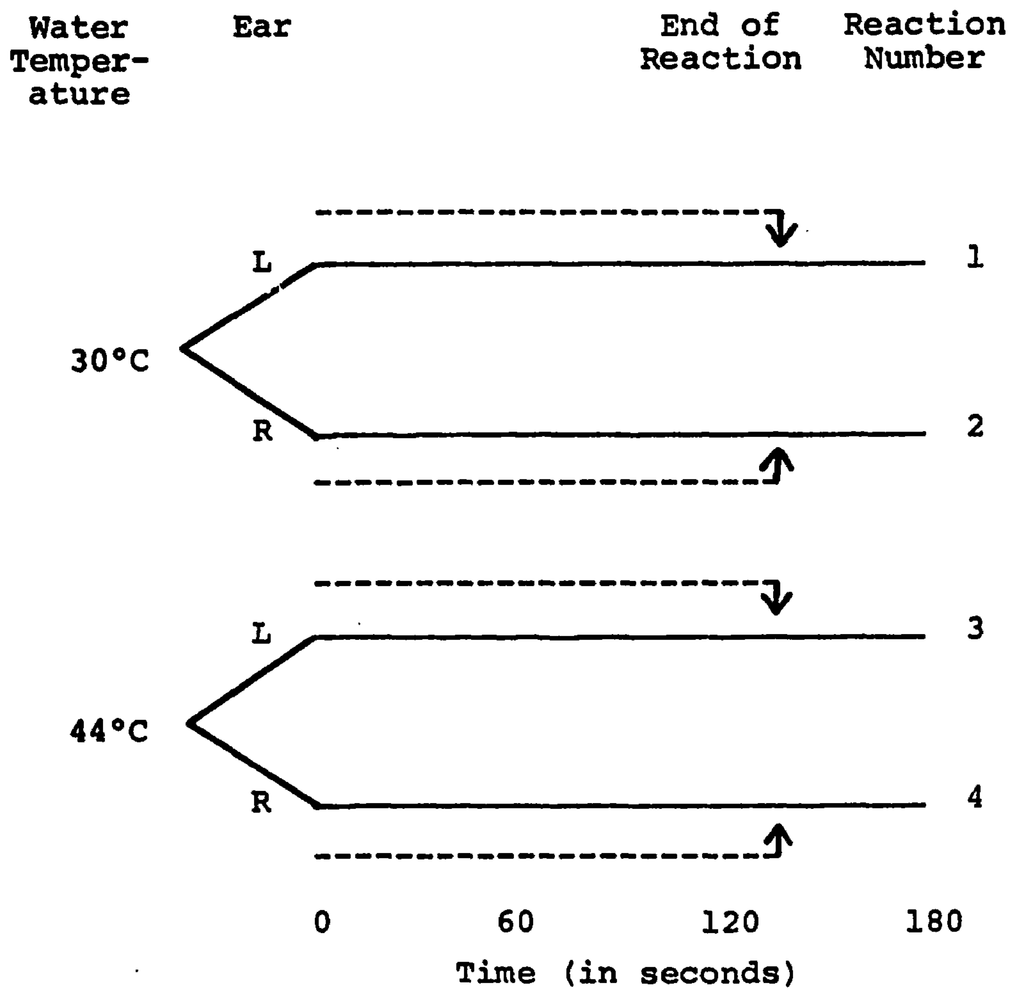


Fig. 1.--Graph Illustrating the Results of a Fitzgerald-Hallpike Test.

right (determined by direction of the fast component) while the sum of reactions two and three consist of nystagmus to the left.

Fitzgerald and Hallpike reported that durations between 120 and 240 seconds fell within the normal range of response. Because of this large intersubject variability they recommended a comparison of the reactions of the two sides. If the reactions of one canal were reduced with respect to the other the implication drawn was that a canal paresis existed. Thus, with a right canal paresis, responses two and four were reduced in comparison to responses one and three. The reverse was true with a left canal paresis. On the other hand, if the nystagmus response in one direction was facilitated while the nystagmic response in the opposite direction was slightly inhibited, the implication was that a directional preponderance existed. For example, with a directional preponderance to the right, responses one and four were enhanced while responses two and three were slightly reduced. The reverse was true for a directional preponderance to the left.

Fitzgerald and Hallpike (24) and Carmichael, Dix and Hallpike (11) stated that a canal paresis reflects lesions of the peripheral vestibular sensory elements and directional preponderance indicates an utricular or central lesion. However, the diagnostic significance of directional preponderance has been questioned in the literature (1, 36, 37, 45, 46, 67, 75).

Contemporary Vestibular Testing

The Fitzgerald-Hallpike technique has enjoyed wide popularity and is, at present, the principal procedure employed in vestibular testing. As originally proposed, the nystagmic response was monitored through direct observation by an observer or clinician with the patient wearing Frenzel's lenses which served to inhibit ocular fixation. The method of direct observation has been criticized on the following bases: nystagmus too weak to be easily visible may not be noted (73); details of the nystagmus response such as frequency, amplitude and speed of the eye deflection are not easily ascertained (4, 37, 40, 56, 72, 80); the difficulty that is encountered in differentiating atactical eye movements from true vestibular nystagmus (64, 73); and the fact that the duration of the response sometimes fails to demonstrate a vestibular system disturbance (4, 40, 56, 58, 59, 66, 73).

Electronystagmography

The above problems are circumvented by the use of a procedure that allows graphic recording of nystagmic activity. One such method, termed electronystagmography, is based upon recording changes of the corneoretinal potential sensed by fixed skin electrodes placed around the eyes.

Arnold, Giuliani and Stephens (2) report that the corneoretinal potential was first discovered by Dubois-Reymond in 1849, and was studied more recently by Mowrer in

1936 (61). The posterior part of the eye including the retina is electronegative relative to the anterior part of the eye including the cornea which is electropositive (2, 28, 43, 54, 61, 64, 73). The nystagmic movements of the eye that occur during caloric testing with electronystagmographic recording change the relative distance between the appropriately arranged electrodes and the electropositive and electronegative poles of the eyes. The changing distance between the electrodes and the poles of the eyes, which is caused by the nystagmic eye movements, varies the magnitude of the potential developed across the electrodes. The varying potential differences are electronically amplified and graphically recorded on a moving paper chart. With this procedure, it became possible to quantify the magnitude of the evoked vestibular reaction in terms of frequency or number of nystagmic beats per unit time, amplitude of nystagmic beats in degrees of eye movement and velocity of the slow phase of nystagmic beats in degrees per second.

Henriksson (40), in 1955, compared the speed of the slow phase and the duration of the response in calorically induced nystagmus. Utilizing the Fitzgerald-Hallpike test on normal subjects, he determined that the mean value for duration was 155 seconds with a standard deviation of 27 seconds. The overall mean value for speed of the slow phase was 29° per second with a standard deviation of 11° per second. Henriksson stated that with presentation of equal stimuli, i.e., water seven degrees above and below body temperature, duration appears to be a more constant measure of response than speed of the eye in the slow phase of nystagmus. However, when stimuli of increasing magnitude were applied, duration of response increased little, whereas the speed of the slow phase increased in a manner that closely reflected the increasing values of stimulation. These findings and several examples of pathological cases showing reduced eye speed but normal duration led Henriksson to conclude that the speed of the slow phase, not duration of the response, is the better expression of the sensitivity of the vestibular apparatus. He reports that this conclusion is in agreement with the work of Buys (10) who determined that the speed of the slow phase was proportional to the magnitude of the rate of the applied angular acceleration and with the work of Dohlman (20) who showed that the speed of the eye in the slow phase of nystagmus was proportional to the strength of thermal stimulus that is applied.

Aschan, Bergstedt and Stahle (4) reported a series of studies on normal individuals. In their manuscript the authors pointed out initially that the absolute values of the responses and the differences between hot and cold syringing are not considered clinically useful because of large individual variation. Clinically the important questions are whether symmetry exists between the response of the right and left labyrinths and/or between right and left beating nystagmus.

In Aschan, Bergstedt and Stahle's first two series of testing cold syringing only was done on 20 "national service" men aged 21-22 years and on one other group comprised of both sexes aged 12-67 years. The response was expressed in terms of duration in seconds which was specified as the time between the onset of irrigation to the last observable nystagmic beat. The findings were not treated statistically. However, the "national service" men showed a range of 0-40 seconds intrasubject difference between the labyrinths. This was determined by subtracting the response of one labyrinth from that of the other. The second group demonstrated greater variability with a range of 0-60 seconds intrasubject difference between the labyrinths. They attributed this larger variation in individual values to the inclusion of older subjects in the second group who might possibly have suffered from vestibular pathology.

A third series of normal individuals with an age range of 25 to 50 years was examined by Aschan, Bergstedt and Stahle using the Fitzgerald-Hallpike test. The technique was modified to the extent that syringing time was reduced to 30 seconds and nystagmus was recorded with the subject's eyes closed. In constrast to the Fitzgerald-Hallpike system illustrated in Figure 1, Aschan, Bergstedt and Stahle designate 30°C right as response one, 30°C left as response two, 44°C right as response three and 44°C left as response four. In order to determine the symmetry between the right and left labyrinths and between right beating and left beating nystagmus, the responses of this group were measured in seconds of duration, maximum speed of the slow phase in degrees per second and total amplitude.

The comparison of responses between labyrinths was a central issue to these investigators. They proceeded by subtracting the sum of the values for the left ear (2+4) from those for the right ear (1+3). This was termed a difference score. The same procedure was used for comparing right beating nystagmus (2+3) and left beating nystagmus (1+4).

In all cases, the differences between labyrinths were small and never approached statistical significance. Their derived norms are based on the assumption of a Gaussian distribution of difference scores in the normal population. The normal range was considered to be plus or minus two times the standard deviation of the difference scores. In the judgment of these and several other investigators in this area two times the standard deviation represents an "allowable difference."

Their results showed that a difference of 60 seconds[2] in duration of response can be assumed to constitute a normal

---

[2]It is conventional in the vestibular testing literature to refer to the differences noted as simply a difference of a certain magnitude rather than to state plus or minus but the meaning is the same as saying plus or minus a specified number of seconds or beats, etc.

range for both right/left sensitivity differences and right beating/left beating nystagmus differences. For maximum eye speed, the allowable difference between ears was 12° per second for both differences in ear sensitivity and differences in beating direction of the nystagmus. Total amplitude, which is seldom utilized clinically because of the time involved in computation, showed the greatest variation in difference scores. The allowable differences (two times the standard deviation) were 742° for the difference between ears and 1132° for the difference between the direction of nystagmus.

Later, Stahle (72) reported the results obtained from a series of 30 normal subjects. He used the same technique as Aschan, Bergstedt and Stahle (4) and nystagmus was assessed with regard to duration, total number of beats, total amplitude and maximum intensity. The latter is the mean speed of the slow phase during a 10-second period at the peak of the reaction. Stahle proposed a set of normal limits set at two times the standard deviation of the observed differences between ears. His data are reported in Table 1. These values reported by Stahle are in fairly close agreement with the previously reported values of Aschan, Bergstedt, and Stahle (4) and, with the exception of duration, the author reported that his normal limits are similar to those of Hamersma (37) whose work will be described later in this chapter.

TABLE 1

NORMAL LIMITS OF RESPONSE TO VESTIBULAR TESTING BY THE FITZGERALD-HALLPIKE METHOD AS PROPOSED BY STAHLE

	Right/Left Sensitivity (1+3) - (2+4)	Right Beating/Left Beating Nystagmus (2+3) - (1+4)
Duration:	60 seconds	60 seconds
Total number of beats:	100	175
Total amplitude:	1000°	1000°
Maximum intensity:	10° per second	10° per second

Stahle stated that, because all of the response parameters gave similar results, testing could be simplified by computing only one of these. He recommended speed of the slow phase or number of beats because they can be obtained in less time. He summarized by saying that the induced reaction should be assessed with regard to its duration and one of the other factors. He felt that this was necessary because in his pathological series, duration did not always reflect diminished labyrinthine function in peripheral lesions, whereas the diminished function was always revealed by the three other factors studied.

Torok (78) introduced culmination frequency as a means of expressing the magnitude of nystagmic reaction. He reported that after a rapid increase, the frequency of the thermally induced nystagmus beats reaches a ceiling value in a 10-second time interval between 25 and 35 seconds after the onset of the response. This frequency peak is then followed by a gradual decrease in rate. Torok's analysis was made by dividing the entire response into five-second intervals and counting the number of nystagmic beats in each five-second period. The culmination frequency was expressed as the highest number of beats reached in two successive or sequential intervals. He reported that the normal culmination frequency varies between 12 and 18 nystagmic beats per ten seconds for a thermic stimulation of 10 cubic centimeters of water at 68°F for five seconds.

Hamersma (37) investigated the responses of normal persons to thermic stimulation of the labyrinth utilizing electronystagmography. He compared the responses obtained by the Fitzgerald-Hallpike technique with those obtained by the Veits and the Kobrak methods. The records were analyzed for latency, duration, and total amplitude of the response and for the total number of beats and maximum speed of the slow phase in degrees per second. He also investigated the effect of various stimulus temperatures on the magnitude of the response as well as the influence of an obstruction (ear wax) of the external auditory meatus on the reaction.

He concluded that the Fitzgerald-Hallpike technique was the most suitable for vestibular testing. The Veits method, while yielding comparable results, proved to be difficult to employ in conjunction with electronystagmography because the movement of the head interfered with the recording. The position of the head called for by this procedure was also uncomfortable for the person being tested.

The Kobrak method possesses the disadvantage of smaller, more difficult to analyze recordings because of the weaker stimulus. With the exception of latency of response, all parameters of response were significantly smaller than those obtained with the Fitzgerald-Hallpike technique.

As the temperature of the water used for irrigation was lowered all parameters of response except duration showed a definite increase in magnitude. Duration also increased up to a point but leveled off with continued decreases in water temperature. Because of this Hamersma considered duration an indication of the interval of time required for the flow of blood through the temporal bone to restore the area to normal body temperature. He concluded that the other response parameters were better indicants of the vestibular response to stimuli of increasing strength.

Another finding of this study was that occlusion of the external auditory meatus by ear wax had no effect on the duration of response but that it did significantly reduce the total amplitude, total beats and speed of the slow phase. This result was attributed to the fact that the labyrinth is cooled to a greater extent when the wax is absent and that the extent of the temperature difference determines the magnitude of response as revealed by everything but duration. Duration may be of limited value as an indicator of strength of stimulation for the reason noted above.

Considering the results obtained by the Fitzgerald-Hallpike technique, Hamersma found no statistically significant differences between the responses of the left and right labyrinths for the response parameters used in this study. Fifty-five per cent of the subjects had a directional preponderance but this occurred an approximately equal number of times in each direction. The duration of response to cold water irrigations was significantly greater than to warm water irrigations. There were no differences due to irrigation temperatures measured by either total beats or total amplitude. The speed of the slow phase showed higher values with the warm stimulus than with the cold stimulus.

Hamersma retested four persons to determine if the results could be repeated within individuals. He states that the test-retest values were "fairly constant" but his data were not subjected to statistical analysis. His raw data which were compiled from a sample of only four subjects appear to show considerable variability for all aspects of the response except speed of the slow phase.

Finally, Hamersma derived normative values for right ear versus left ear sensitivity. He arbitrarily determined acceptable limits by doubling the standard deviation of the difference scores. These values which were obtained from the data collected on 47 subjects, are reported in Table 2. These limits, with the exception of those for duration, are similar to those proposed by Stahle.

Mehra (56) investigated the nystagmographic responses obtained from 31 normal adults who were given the Fitzgerald-Hallpike test. The results were analyzed for duration, latency, total beats, total amplitude and maximum speed of the slow phase. The differences between the mean values were evaluated statistically by using the *t* test for paired scores.

TABLE 2

NORMAL LIMITS OF RESPONSE FOR RIGHT/LEFT SENSITIVITY DIFFERENCES
OBTAINED BY THE FITZGERALD-HALLPIKE TECHNIQUE
AS PROPOSED BY HAMERSMA

Duration:	120 seconds
Total amplitude:	850°
Total number of beats:	130
Maximum speed of the slow phase:	13° per second

None of the indices of response revealed any significant differences between right and left ear sensitivity. Larger variations about the mean were seen in total beats and total amplitude than in the other measures. Two subjects showed directional preponderance to the left. In comparing the effect of cold irrigation to the effect of warm irrigation, the duration of response to the cold stimulus was significantly longer than the duration of response to the warm stimulus. However, the warm stimulus prompted a significantly greater reaction when the records were analyzed for total beats and maximum speed of the slow phase. A dysrhythmia of the nystagmus was seen for eight subjects in the form of a cessation of the beats for five to ten second periods at some point or points during the course of the reaction. Merha did not report normal limits of response for the indices studied.

The correlations between the different measures of post-caloric nystagmus were investigated by Maspetiol and Keravek (55). They examined the electronystagmographic records of 23 normal individuals who had been tested once by a procedure quite similar to the Fitzgerald-Hallpike. The measures studied were latency of response, duration of response, frequency, amplitude, total beats, and speed of the slow phase. Pearson product-moment correlation coefficients were computed for all combinations of response indices. The authors defined correlation coefficients of

.70 or greater as strong, .40-.69 as average and .20-.39 as weak.

Speed of the slow phase of the nystagmus, frequency and total beats showed correlations of each with the other of approximately .60. The authors interpreted this as "fairly strong." Amplitude correlated well (.60) only with speed of the slow phase. Duration correlated only with total beats while latency did not correlate with any of the other measures. Maspetoil and Keravek concluded that latency is a useless measurement and that duration of response should be corroborated by another measure of nystagmus.

Hinchcliffe (42) concluded that the maximum speed of the slow phase of nystagmus induced by the warm stimulus was the most valid measure of vestibular function. Using the conventional irrigating temperatures of 30°C and 44°C the duration of the response and the speed of the slow phase was compared to the auditory thresholds (250 through 8000 Hz) of 50 patients medically diagnosed as having Menieres disease by the Pearson product-moment technique. The greatest product-moment correlation coefficients were obtained for the speed of the slow phase induced by the hot stimulus. However, the highest correlation obtained was only slightly greater than .40.

Hinchcliffe postulated that, in a pathological disturbance such as Menieres disease which is common to both the auditory and vestibular portions of the labyrinth, the hypofunction of the vestibular labyrinth should to some extent be commensurate with impairment of function of the auditory labyrinth.

In a later paper Hinchcliffe (41) studied the number of nystagmic beats in the period from 60 to 90 seconds after the onset of irrigation as an index of caloric test response. He tested 20 otologically normal adults using water maintained at 44°C as the stimulus. Sixteen subjects were retested. The interval between tests was unspecified. The results were recorded electronystagmographically.

The results showed that the total number of beats gave a frequency distribution which was Gaussian in appearance. Hinchcliffe reports that less than 33 beats would be abnormal as would a difference between the results of the two ears of more than 20 beats. Pearson product-moment correlation coefficients were computed to determine the correlation between maximum speed of the slow phase and total beats. This was done for each ear separately on both testing occasions. The results show an overall correlation coefficient of .63 which was significantly different from no correlation at the .01 level of confidence. Hinchcliffe concluded that both measures of response are valid and recommended the use of total beats because of its ease of computation.

Torok (77) compared culmination frequency and speed of the slow phase in 13 normal subjects. Graded thermal stimuli of 100 cubic centimeters of water each were delivered alternately to each ear for 20 seconds at temperatures of 30, 25, 20, 15 and 10° centigrade. An interval of five minutes was allowed between irrigations. Graphs of the averaged responses to the various stimulus temperatures indicated that mean culmination frequency is more linearly related to stimulus strength than mean speed of the slow phase. In the individual responses no linear pattern could be seen on two out of 26 occasions for culmination frequency. No linear configuration was seen on seven out of 26 occasions for speed of the slow phase. However, the individual data also show that when there are gross deviations from linearity, the speed of the slow phase tends to show less deviation. The standard deviations for culmination frequency were more uniform throughout the various stimulus temperatures than were the standard deviations for speed of the slow phase. Pearson product-moment correlation coefficients between culmination frequency and speed of the slow phase at each stimulus strength were significant but declined in size at the stronger stimulation levels. Torok concluded that both culmination frequency and speed of the slow phase reflect stimulation magnitude because of the direct proportional relationship observed between stimulus strength and the means values of both response measures. Because culmination frequency is more linear with respect to stimulus strength and has more homogeneous standard deviations Torok proposed the utilization of culmination frequency as the more "reliable" parameter to express nystagmus magnitude.

A simplified test using water at 44°C as the stimulus was proposed by Halama and Hinchcliffe (35). Their procedure consisted of separately irrigating each ear with warm water for 30 seconds. The test was given to 20 normal people and it was repeated after a period of one month. The results were recorded electronystagmographically and analyzed for total beats during the interval between 60 to 90 seconds after the onset of irrigation. The data showed an approximate median value of 20 total beats for each ear on both tests. Reliability from test to test was estimated using product-moment correlation coefficients. The coefficients were .88 for the right ear and .77 for the left ear. However, a response from either side was not elicited from four subjects on the first test and three subjects on the second test. Because no response was elicited from some subjects, the authors concluded that the test is reliable but insensitive. The failure of some subjects to respond may have occurred because the subjects fixated their eyes on a ceiling marker during the course of the test. The effect that fixating the eyes on an object has on caloric test results will be explored in more detail later in this chapter.

Vestibular Screening Tests

Several investigators have proposed that the warm water stimulus be used as a screening test of vestibular function. Bernstein (8) contended that every type of vestibular abnormality is detected by the warm water stimulus. He reached this conclusion by reviewing many cases of vestibular pathology and noting that warm water testing alone always revealed the presence of dysfunction. In order to conserve time Bernstein recommended testing with warm water first and if the comparison between the ears was normal the examination was terminated. If the response was not normal then the complete Fitzgerald-Hallpike test was given.

Hart (39) discussed the possible effects of a unilateral peripheral pathological disturbance on the neurophysiology of the vestibular system. A canal paresis or a nonfunctioning labyrinth may possibly be accompanied by a directional preponderance toward the better labyrinth. This is because each labyrinth is connected to the muscles that move the eyes in a contralateral direction. Resting neural discharge from the intact labyrinth pulls the eyes toward the affected side (slow component). This occurs because there is little or no opposing action from the diseased labyrinth. A movement of the eyes in the opposite direction (fast component) is then centrally elicited to return the eyes to near the midline.

If directional preponderance is present, then cold water caloric testing which is widely used in the physician's office might not disclose any disparity between the ears. The diseased labyrinth may appear to be normally responsive because the directional preponderance elicits nystagmus beating in the same direction as induced cold water nystagmus. If the directional preponderance is of sufficient strength, it may appear that there is a canal paresis of the better side. The nystagmus caused by the directional preponderance opposes the cold water induced nystagmus of the better side because they are beating in opposite directions. This causes the better side to be reduced in responsivity. Because of this, Hart suggested the warm stimulus be used for screening purposes and if any abnormality is observed a full Fitzgerald-Hallpike test should be given. Greisen (30) states the same case in a recent publication.

The value of the warm water caloric test as a screening device was recently investigated by Barber, Wright and Demanuele (5). They initially described an unpublished study conducted by themselves on 114 normal subjects. Using speed of the slow phase as the response index, they correlated the right/left sensitivity differences obtained from Fitzgerald-Hallpike testing with the right/left sensitivity differences obtained from cold and from warm irrigations. The cold irrigations yielded a correlation coefficient of .62 while the warm irrigations yielded a correlation coefficient of .76. This result led the authors to study the value of the warm irrigations alone as a predictor of outcome on the Fitzgerald-Hallpike or bithermal tests.

Barber, Wright and Demanuele reviewed a series of 147 clinical patient files. On the basis of previous bithermal testing forty-two patients had demonstrated a hypofunction of one vestibular labyrinth as compared to the other. Forty of these forty-two patients would have been identified correctly by comparing the right/left sensitivity difference for the warm stimulus alone. All patients who showed normal bithermal results also showed normal results for the warm irrigations. These findings led the authors to conclude that the warm water caloric screening test could reliably predict the bithermal test results.

Graded Testing

Litton and McCabe (52) described the technique and the clinical application of Thermal Vestibulometry. Cold water irrigations at temperatures of 33, 29, 25, 21 and 17°C were given to a series of normal individuals. Each irrigation lasted for one minute and a five-minute rest was allowed between the termination of one reaction and the beginning of the next irrigation. One ear of each subject was irrigated. The choice of the test ear was determined by a counterbalanced schedule. Speed of the slow phase was selected as the response parameter for study. The mean slow phase speeds at each temperature were plotted as a function of stimulus temperature in graphic form. The graphs showed that speed of the slow phase increased in a linear fashion with linear decrease in stimulus temperature. The authors offered several examples of results obtained with various peripheral and central pathologies. Each type of pathology was examplified by a typical pattern of graded response differing from the linear pattern of normals.

A technique of continuous thermal vestibulometry was introduced by Steffen, Linthicum and Churchill (74). They designed an apparatus that continuously changed the temperature of the irrigating solution in a linear manner over time. The instrument provided a five-minute period of irrigation with greater than 250 cubic centimeters of water delivered per minute. The temperature continually decreased from 37°C to 19°C. Electronystagmographic recording began at the onset of irrigation and continued until 20 seconds after the cessation of the stimulus. Thirteen normal young adults were tested with this method. Speed of the slow phase was computed at 3°C stimulus temperature intervals. The increase in the speed of the slow phase from the onset of the nystagmus until it reached a maximum value was divided by the number of degrees of temperature change that produced this increase. This was termed the "slope of nystagmus." The normal range was 0.6 to 5.7 with a median of 1.8. Two pathological groups consisting of those with acoustic neuromas and those with Menieres disease showed a reduced slope with respect to normal but considerable overlap was evident between the pathological groups.

Torok administered a strong and a weak caloric stimulus to a large series of normal subjects (76). The strong stimulus consisted of 100 cubic centimeters of water at 20°C for 20 seconds. The weak stimulus consisted of 10 cubic centimeters of water at 20°C for five seconds. Both ears were tested and electronystagmographic recordings were analyzed for culmination frequency. The results were expressed as a ratio of the weak stimulus to the strong stimulus. The author considered normal to be any ratio between 1.3 and 3.5. Between 3.5 and 6.3 was considered potentially abnormal. When the ratio was higher than 6.3 the increment of response was thought to be disproportionately greater with respect to stimulus strengths than seen in normals. The author contended that this resembles the recruitment of loudness seen in audiology and cited results from patients with vestibular end organ lesions who presented this recruitment-like response. He also cited results obtained from patients with central nervous system lesions who do not show an increase in the magnitude of response with the strong stimulus. In some cases, a decrease in response magnitude to this stimulus was noted. This result was termed vestibular decruitment. A ratio of 1.3 or less was considered indicative of this phenomenon. Torok pointed out that these terms do not necessarily imply abnormal subjective sensation but rather an abnormal increase or decrease of neuromuscular reflex response.

Factors That Influence Thermally Evoked Nystagmus

Several variables affect the nystagmic response and all have been the subject of considerable study. Individual review of the pertinent research in these areas would render this chapter unnecessarily long and the contribution of a detailed review of these areas to the theme of the present investigation would be difficult to justify. However, these factors will be given brief explanation so that reference to any of them in the subsequent text will not be confusing.

Alcoholic and Drug Intake

The effects of drug and alcoholic ingestion on vestibular function known and much work has been accomplished in this area (3, 6, 9, 21, 22, 33, 51, 65, 68, 69, 73, 82). Alcohol even in small doses, causes a positional nystagmus which, depending upon the amount ingested, can be present for as long as 18 hours. Sleeping pills and sedatives, principally barbiturates, can cause a positional nystagmus. Any drugs that act as central nervous system depressants can also inhibit or weaken thermally induced nystagmus.

Visual Fixation

The influence of visual fixation has been extensively studied and its effects on thermally induced nystagmus are well described (14, 15, 18, 51, 54, 62, 71). Generally speaking, visual fixation inhibits and sometimes totally abolishes thermally induced nystagmus. Fixation on a distant object tends to decrease nystagmic response more than fixation on a near object. Because of the suppressing influence of fixation on nystagmus, most clinical testing is done with the patient's eyes lightly closed or with the patient in total darkness with the eyes open.

State of Mental Activity

The fact that state of mental activity can affect the nystagmic response was discovered early and considerable effort has been devoted to its study (12, 13, 14, 17, 18, 29, 31, 51, 54, 60, 63, 71). A stronger, more consistent nystagmus is observed when the individual being tested is occupied with a problem solving or decision making task. If the mental activity is not controlled and the individual is allowed to lapse into reverie, the nystagmus can be diminished, absent or dysrhythmic in nature. The influence of mental activity on nystagmus is well recognized and those involved with vestibular assessment should control this influence.

Habituation

Habituation refers to a process whereby response to a sensory input event diminishes or disappears when the stimulus is presented in a repetitive manner. This phenomenon has been examined extensively in vestibular research (15, 23, 27, 48, 49, 50). Many investigations using both human and animal subjects report a decline of response commonly called a response decline of nystagmus, to repetitious vestibular stimulation. Most of these studies employ extensive habituation trials, and to a certain extent, all parameters of the nystagmus response are affected. However, under the conditions of controlled mental activity and of testing in total darkness with the eyes open, Collins (15) demonstrated no significant habituation of nystagmus as measured by duration of response or by speed of the slow phase.

Justification for the Present Study

The quantification of thermally induced nystagmus provides a reasonable approach toward elucidating the functional condition of the peripheral and central vestibular systems. Although vestibular testing using the thermal stimulation provided by warm and cold water has been widely used since the early part of the twentieth century, there is little agreement as to how the magnitude of the nystagmus response should be quantified. Some investigators believe that any of the response parameters reflecting magnitude can be used and recommend using those that are least time consuming (4, 37, 55, 56, 58, 59, 72). Several authors contend that speed of the slow phase is the appropriate measure of nystagmus magnitude (5, 40, 43, 64, 74). Hinchcliffe (42) states that either speed of the slow phase or total beats is appropriate and advocates total beats because the relative ease of computing this index saves time. Torok (76, 77, 78, 79, 80), on the other hand, asserts that culmination frequency is the most suitable expression of response magnitude.

The difference of opinion concerning this issue is apparent. It would seem that information concerning the relative consistency or reliability of these various indices of response magnitude could provide valuable assistance in the resolution of this question. Hamersma (37) retested a portion of his subjects but his small sample size of four individuals coupled with the lack of statistical analysis of his data render his results inconclusive. Hinchcliffe (41) and Halama and Hinchcliffe (35) tested their subjects twice but only with the warm stimulus (44°C). In addition only two response parameters, total beats and speed of the slow phase were considered. Hamersma (37), Hinchcliffe (44) and Halama and Hinchcliffe (35) also considered only the test-retest stability of absolute scores and not reproducibility of difference scores between ears. The difference score is, of course, the score normally used in clinical evaluations.

The present study was designed to investigate the absolute and relative reliability of speed of the slow phase, total beats, culmination frequency, total amplitude and total duration. A test-retest sequence was utilized. A description of the experimental apparatus, the subject sample, and the procedures utilized are outlined in the following chapter.

CHAPTER III

PROCEDURE AND INSTRUMENTATION

Introduction

Various methods of expressing the magnitude of the calorically induced nystagmic reaction have been made available through the technique of electronystagmography (ENG). In order to investigate the reliability of several of these methods the Fitzgerald-Hallpike caloric test was administered to a sample of normal young adult males three times in succession with an interval of one-half hour between each administration. The response parameters investigated were speed of the slow phase, total nystagmic beats, culmination frequency, total amplitude and the duration of the reaction. With the exception of duration all response parameters were computed from within the interval between 60 and 90 seconds from the onset of the ear canal irrigation.

Subjects

Data were collected from 16 young adult males between the ages of 18 and 30. All subjects were required to demonstrate normal hearing sensitivity, a negative history of vestibular pathology or chronic ear pathology and a negative neurological history. A qualifying examination was administered to each subject. This procedure included a pure-tone air conduction screening test administered at 15 dB hearing level (ANSI-1969) at octave intervals from 250 to 8000 Hz inclusive and an otoneurological evaluation that included otoscopic examination and neurological testing. The pure-tone hearing screening test was administered by the investigator. The otoneurological examination was administered by first, second or third year Otorhinolaryngology residents in the Department of Otorhinolaryngology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma. In order for an individual to be included in the study all results of these examinations were required to be within normal limits.

All subjects were required to have refrained from alcohol and drug intake for 48 hours prior to an experimental session. No one with a history of recent experience with hallucinogenic drugs or marijuana was included in the study.

Apparatus

Figure 2 displays a simplified block diagram of the experimental apparatus. The electrical activity resulting from the varying corneoretinal potential difference developed between the recording electrodes placed approximately one centimeter lateral to the outer canthus of each eye was

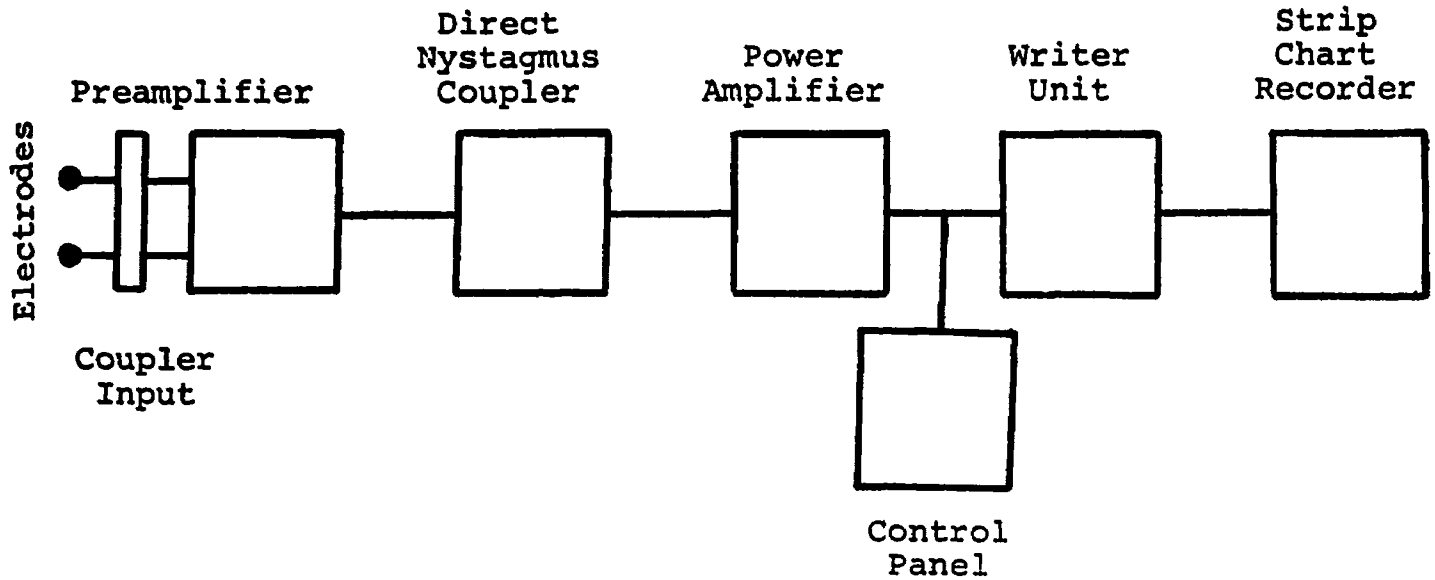


Fig. 2.--Simplified Block Diagram of Experimental Apparatus

conducted through the input of the direct nystagmus coupler (Beckman, Type 9859) to a preamplification stage (Beckman, 461B). The signal was then routed to the direct nystagmus coupler (Beckman, Type 9859). The coupler modified the frequency-response characteristics of the recording channel, regulated the zero baseline position of the stylus and provided for adjustment of the gain in the recording channel for the purposes of calibration.

The output of the Type 9859 coupler was directed to a power amplifier (Beckman, Type 462) which, with the preamplification stage, provided the overall gain of the recording channel. Further, the power amplifier unit provided control for additional frequency response manipulation, amplifier gain calibration and a control used in conjunction with the Type 9859 coupler to place the resting position of the pen (zero signal) on a selected chart line. The signal ultimately activated the stylus unit (Beckman 508) for graphic display. Selection of paper speed was provided by the Type A560 Control Panel. A chart speed of five millimeters per second was used. The recordings were displayed on curvilinear chart paper (Beckman Instruments, Inc.).

During the recording phase of the experiment the direct nystagmus coupler was set to provide a three-second time constant for the amplification system. This time constant is frequently used by clinicians and researchers because it is long enough to provide a relatively distortion free recording while short enough to prevent undue baseline shift caused by prolonged DC potential changes. The fast-slow mode switch which is used to determine the high frequency cutoff point of the coupler frequency response was positioned to allow the recording of both fast and slow components of the nystagmic signal. The power amplifier was adjusted to provide a high frequency rejection beginning at approximately 20 Hz. These settings together resulted in a frequency response specified by the manufacturer to be approximately 0.056 Hz to 20 Hz with the half-power point determining the beginning of low and high frequency rejection. The overall recorder sensitivity was determined by the preamplifier and the power amplifier controls.

Thermostatic Baths

The thermostatic bath apparatus consisted of two water baths (Grant Instruments, Type HPK) each complete with adjustable thermostatic controls, pumping units and dial thermometers for monitoring water temperature in degrees centigrade. Each pump was connected to flexible, insulated tubing which terminated in an irrigation nozzle. The rate of flow of the water was controlled by an adjustable stopcock on each delivery tube. Each bath contained approximately two gallons of water. One was held at 30°C and the other at 44°C. When the baths were activated but not in use, each irrigating nozzle was inserted in a holder on the inside of its appropriate bath making it possible for the water to circulate continuously throughout the system.

Calibration

Stimulus calibration. Constant temperature regulators in the two water baths controlled the temperature of the water used for the two stimuli at 30°C and 44°C. The temperature gauges were calibrated to the desired reading in degrees centigrade by means of a 0-50°C range mercury thermometer which is specified by the manufacturer to be accurate to 0.5°C. Prior to each irrigation the gauges were adjusted until the desired water temperature was read on the thermometer placed at the point of water emergence from the irrigating nozzle. This valve was set to deliver 6.25 cubic centimeters of water per second for a total of 187.5 cubic centimeters during the 30-second irrigation period. This flow rate is identical to the one used by Fitzgerald and Hallpike (24). To assure constancy of temperature throughout the irrigation system the water was allowed to circulate from the bath through the tubing and return to the bath for a period of not less than one-half hour before each experimental session.

Calibration of eye movements. The strength of the corneoretinal potential varies among individuals (38, 57). It follows, therefore, that for a given angular deflection of the eyes the amount of recorder-pen deflection will differ across subjects. For this reason, it is desirable to adjust the extent of pen deflection to a particular value in millimeters for a known angular deviation of the eyes. This calibration makes it possible to convert the graphic results into angular eye deviation in degrees. Calibration was accomplished by means of a nystagmic gonioscope. The particular gonioscope used in this investigation consisted of a row of three electric lights aligned horizontally. The distance between the center light and each outer light and the distance of the lights from the subject were such that the eyes had to be deviated by an angle of 10 degrees from their position when gazing at the center light in order to gaze directly at one of the outer lights. This produced a 20° deviation in the subject's eyes when he shifted his gaze from one outer light to the other outer light. The gonioscope was suspended from the ceiling in the subject's direct line of vision when he was maintaining a straight-ahead gaze with his head elevated 30° from the horizontal plane. The arrangement is illustrated in Figure 3.

After the electrodes were properly positioned (to be described in a subsequent section of this chapter) and affixed and the subject was placed in the test position, the calibration was begun. A standard calibration value of 1° of eye movement per 1 mm of pen deflection was used. The subject was instructed to gaze alternately left and right as each light was illuminated and to maintain his gaze upon the light for as long as the particular light was on. With the room light eliminated, the outer lights on the gonioscope

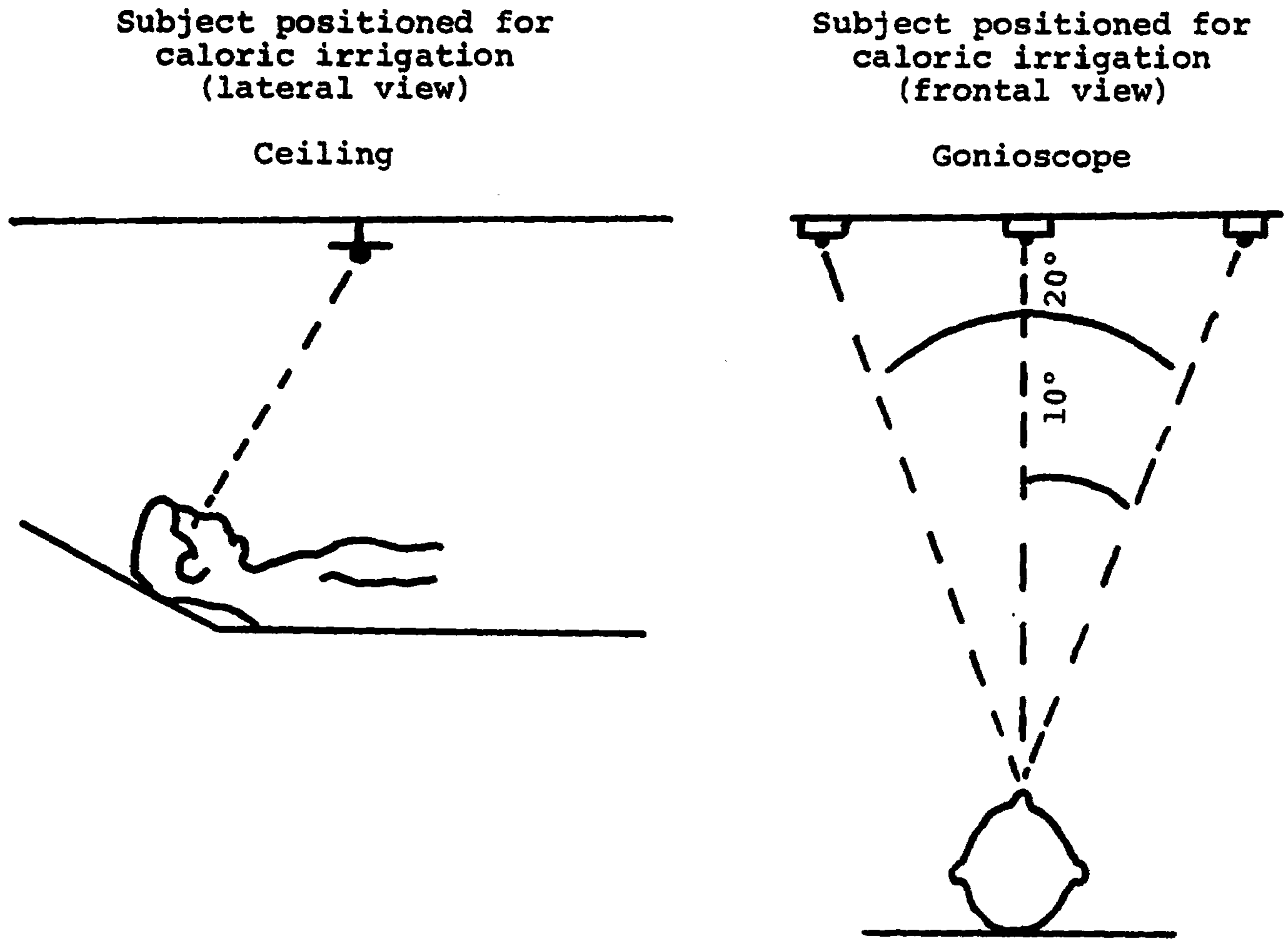


Fig. 3.--Position of Subject in Relation to Nystagmus Gonioscope During Calibration of Eye Movements

were flashed alternately at a rate of once per second. After the gain of the amplifier had been adjusted so that 20° of eye movement produced approximately a 20 mm pen deflection, ten pen deflections in each direction were recorded (twenty deflections). This produced a calibration recording for that subject on that occasion. A calibration recording is illustrated in Figure 4. An exact gain adjustment of 20 mm is not possible for each deflection. Consequently, to obtain a precise amplitude calibration the twenty deflections were averaged to give a mean pen deflection in millimeters. This procedure was done prior to the positional testing and prior to each irrigation. This precise calibration value was then used in the evaluation of the data.

Audiometer calibration. Prior to the testing of any subject, the acoustic output of each earphone (Telephonics, TDH-39, 10 ohm) activated by the audiometer (Belton 15C) was measured using an Allison Model 300 audiometer calibration unit. The audiometer calibration and the hearing screening test were completed in an acoustically-treated hearing test booth (Industrial Acoustic Corporation, Model 400A).

Experimental Procedure

The Fitzgerald-Hallpike caloric test was given three times to each of sixteen subjects. The eye movement responses of the subjects were recorded electronystagmographically on each occasion.

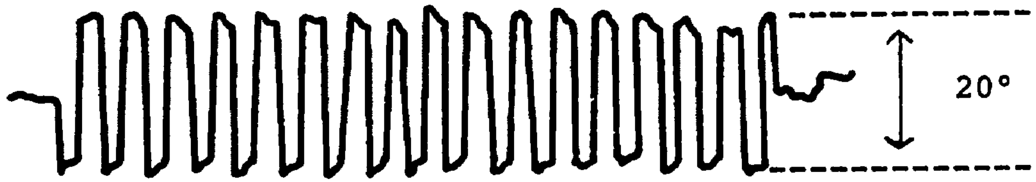


Fig. 4.--Typical Recording of Eye Movements During Calibration

Test Environment

The test room was located in the Otorhinolaryngology Clinic, Children's Memorial Hospital, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma. The room was maintained in a state of near total darkness to minimize visual influences. All data collection took place at night or when there was no other potentially disturbing activity in the immediate vicinity.

Preparation of the Subject

Each subject was situated on an examining table in supine position. Prior to electrode placement the surface area of the skin designated to receive the electrode was cleansed with alcohol. Each electrode was filled with a commercial electrode jelly (Sanborn Redux). One electrode was attached to the skin approximately one centimeter lateral to the temporal canthus of each eye by circular adhesive connectors. This particular arrangement provides for the recording of eye movements in the horizontal plane only. The indifferent electrode was affixed to the glabella.

Preliminary Testing

Before the caloric testing was begun care was taken to insure that a spontaneous or positional nystagmus was not present in the subject. That is, ENG recordings were made with the subject in the following positions and in the sequence listed:

1. Supine
2. Head rotated 90° to the right
3. Supine
4. Head rotated 90° to the left
5. Supine
6. Head hanging 30° below level of table
7. Sitting with head held in erect position

Recordings were obtained with the eyes closed and each position was maintained for one minute. If no spontaneous or positional nystagmus was observed, preparations were made for the Fitzgerald-Hallpike testing.

Fitzgerald-Hallpike Test

The subject was placed in a supine position on a padded examination table. The head was anteflexed to an angle of 30° with the horizontal plane by a fixed headrest designed for this purpose. This position places the horizontal semicircular canal in a vertical plane which is the optimum position for calorically induced endolymphatic flow. The electrodes were checked visually to insure that proper placement and contact were being maintained. A calibration of eye movements was made using the method previously described.

The Fitzgerald-Hallpike test was then begun. The sequences according to which the test temperatures and the test ears were counterbalanced are shown in Table 3. Inspection of the table shows that eight individuals received the cold water stimulus first and eight received the warm water stimulus first. In addition, the ear which was tested

TABLE 3

TEST ORDER FOR EARS AND TEMPERATURES

Irrigation Sequence			
1. Rc Lc Rw Lw	9. Rc Lc Rw Lw	2. Rw Lw Rc Lc	10. Rw Lw Rc Lc
3. Lc Rc Lw Rw	11. Lc Rc Lw Rw	4. Lw Rw Lc Rc	12. Lw Rw Lc Rc
5. Rc Lc Rw Lw	13. Rc Lc Rw Lw	6. Rw Lw Rc Lc	14. Rw Lw Rc Lc
7. Lc Rc Lw Rw	15. Lc Rc Lw Rw	8. Lw Rw Lc Rc	16. Lw Rw Lc Rc

1, 2, . . . = Subject Number
R = Right Ear
L = Left Ear
c = 30°C
w = 44°C

first was counterbalanced across subjects. The order in which the test temperatures and test ears were used remained the same for each subject throughout the three testing sessions.

The stimulus was introduced into the external auditory canal by means of a soft rubber nozzle placed adjacent to the canal wall. This allowed the water to flow down the canal wall to the tympanic membrane reducing the possibility of the formation of an insulating air bubble at that end of the ear canal. A small basin collected the irrigating water draining from the subject's ear. A complete test consisted of separate irrigation of each ear with both the warm and the cold water stimuli for a total of four irrigations. Each irrigation was timed to be precisely 30 seconds in duration (4, 14, 45, 72). All testing was done with the subject's eyes lightly closed and a 10-minute interval was allowed between successive irrigations. During this interval, the water temperature was monitored and, if necessary, adjustments were made to assure that the proper stimulus temperature was maintained at the beginning of the next irrigation. A re-calibration of pen deflection magnitude was also performed prior to each irrigation. Approximately one hour was required to complete a testing session.

During each irrigation and subsequent measurement period, the subject was required to perform serial arithmetic problems consisting of continuous subtraction. It has been previously discussed in chapter ii that tasks of this nature result in more stable and consistent responses. Verbal responses are required so that the examiner could be sure that the subject was attending to the task. It had been previously determined by the present investigator that the verbal responses did not introduce myogenic or other contamination into the recordings. This determination was accomplished by having several people fixate alternately on three adjacent wall targets while simultaneously counting out loud. Eye movement recordings were made with the individual remaining silent and compared to additional recordings made while the individual simultaneously counted out loud. In no cases were the recorded eye movements influenced by the muscular action produced by the counting.

Subsequent to the first Fitzgerald-Hallpike test the subject was tested a second and third time with a 30-minute interval between tests. The testing procedure remained the same throughout the three experimental sessions.

Measurement of Data

After the data had been accumulated and recorded the parameters of response were extracted from the graphic display. Excepting duration, all parameters of response were analyzed from the record during the interval of 60-90 seconds after the onset of irrigation. Previous investigations have shown that the magnitude of the reaction is greatest during this interval (4, 14, 15, 35, 37, 40, 41, 42, 58, 59, 72, 80).

Speed of Slow Phase

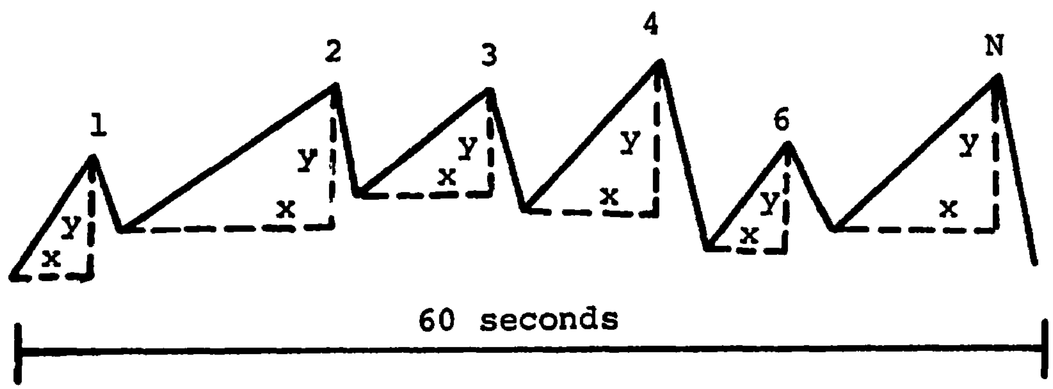
The speed of the slow phase is the average angular deviation of the eyes in degrees per second existing during the slow phase of nystagmus. Several methods of computing this index are available. The particular technique employed in the present study is illustrated in Figure 5. The average amplitude of each slow phase in millimeters and the average duration of each slow phase in millimeters over the 30-second measurement interval (i.e., the interval between 60 and 90 seconds from the onset of irrigation) were computed. The average amplitude was then divided by the average duration. The result was transformed into degrees per second by multiplying it by a conversion factor. The conversion factor was obtained by dividing the amplitude calibration (degrees deflection per millimeter) obtained prior to the particular irrigation being analyzed by the paper speed in seconds per millimeter.

Total Beats

Total beats is simply the number of the nystagmic beats occurring during the 30-second analysis period.

Culmination Frequency

Culmination frequency is the greatest number of beats obtained in any two successive five-second intervals within the analysis period. This value was calculated by dividing the 30-second analysis period into five-second



y = Amplitude in Millimeters

x = Duration in Millimeters

c = Conversion Factor

SSP = Speed of Slow Phase

$$SSP = \frac{\Sigma y/N}{\Sigma x/N} C$$

Fig. 5.--Speed of Slow Phase Computation

intervals and noting the highest number of beats reached in any two adjacent interval pairs.

Duration

Duration is the span of time in seconds that the recorded, induced nystagmus is no longer readily visable on the chart paper. The last beat of three consecutive nystagmus beats followed by ten seconds of no nystagmic activity specified the endpoint.

Total Amplitude

Total amplitude is the sum in angular degrees of the amplitudes of the slow phases of the beats occurring within the 30-second analysis period. This sum is said to represent the total eye rotation occurring in 30 seconds as though the slow phase were a continuous eye movement occurring without interruption by the fast phase return. The figure actually represents the eye rotation occurring in a period equal to 30 seconds minus the time occupied by the fast phase.

Analysis of the Data

Several statistical procedures were used to assist in the analysis and evaluation of the data. The principal statistic used was the reliability coefficient of intraclass correlation (R) which was utilized to estimate the reliability of the repeated measurements. Other procedures included

Student's t test for paired groups; Pearson's product-moment correlation ($\bar{r}$); Spearman's rank-difference correlation (r_s); the mean ($\bar{x}$), and the standard deviation (s). The .05 level of confidence was adopted as the cutoff point for rejecting the null hypothesis.

Intraclass Correlation Analysis

In measurement theory it is assumed that each obtained measurement is composed of two components: the true value and error. It follows then that the total variance of any set of measurements consists of the algebraic sum of the true variance and the error variance. In order to obtain an estimate of reliability one needs to determine the proportion of the total variance that is true variance. This may be accomplished by any of the statistical methods that indicate the degree of relationship or reliability in the form of a ratio of the true variance to the total variance. Presumably the error variance is computationally removed from the numerator while the denominator represents the sum of the true variance and the error variance.

It is well known that several sources of variation can be expressed as components of the total variance and that their relative magnitudes can be estimated by the analysis of variance technique. The method of intraclass correlation is based upon the partitioning of variances (34, 83). The method used in this experiment was a two-way anlaysis with one observation per cell. The formula used to compute the intraclass reliability coefficients was

$$R = \frac{\text{BCMS-RMS}}{\text{BCMS} + (k-1)\ \text{RMS}}.$$

BCMS is the between class variance. RMS is residual or error variance. K is the number of scores per class. Each subject is considered a class and repeated measurements on the same subject are scores within a class. The numerator of the formula represents an estimate of true variance and the denominator an estimate of the total variance. The ratio of the two provides an estimate of the proportion of the total variance that is true variance.

The method of intraclass correlation analysis was used in the analysis of the data for two principal reasons. First, three measurements were made on each subject. This necessitated the utilization of a statistic that estimates reliability when more than two measurements are made on the same subject. The measure of intraclass correlation satisfies this requirement because the reliability coefficient is estimated directly from the analysis of variance technique and, therefore, can be applied to any number of repeated measurements.

Second, the factors of fatigue, learning, memory and practice cannot be ignored in experiments where repeated measurements on the same subject are involved. These are variables that could produce a trend in a given direction in the scores derived from repeated trials. It is important to recognize trend effects and to remove their influence if, in the judgment of the investigator, they are irrelevant to the research question. Removing trend effects results in a smaller estimate of the error variance and a reliability coefficient which is not reduced by the trend effects. These effects can be detected and eliminated if desired by intraclass correlation analysis because the analysis of variance technique is utilized. Most other measures of reliability such as the Pearson product-moment correlation technique cannot do this because differentiation among error variances is not possible.

The data resulting from the experiment are presented and discussed in the following chapter.

CHAPTER IV

RESULTS AND DISCUSSION

Introduction

This experiment was undertaken to investigate the reliability of five indices used to express the magnitude of thermally induced vestibular nystagmus. Assessments of the nystagmus response to the Fitzgerald-Hallpike test were obtained on three occasions with equal intervals of time between occasions. Nystagmic responses were recorded from bitemporal electrodes positioned lateral to the outer canthus of each eye. The developed corneoretinal potential variations were amplified and recorded in graphic form on moving strip chart paper. Speed of the slow phase of nystagmus, total nystagmic beats, total amplitude of the slow phase of nystagmus, culmination frequency and duration of response were the indices of nystagmus magnitude studied.

Results

A result derived from one ear at one temperature and expressed by one of the five response parameters is called an absolute score. For example, analyzing for total beats, the value obtained from irrigating the right ear with water at 30°C is an absolute score. Similarly, the values obtained from irrigating the right ear with 44°C and the left ear with 30° and the left ear with 44°C are each absolute scores. A difference score (ds) is that difference between the sum of the two absolute scores obtained for one ear and the sum of the two absolute scores obtained from the other ear. A complete Fitzgerald-Hallpike test generates four absolute scores usually labeled 1, 2, 3 and 4 and one difference score [(1+3) - (2+4) = ds]. The absolute and difference score data obtained from all of the individual subjects under all conditions and trials are presented in Appendix A.

Analysis of Absolute Scores

Statistical analyses were performed on both the difference scores and on the individual absolute scores. Table 4 presents the means and standard deviations of the data obtained for each of the four absolute scores across all subjects. The five response parameters and the three test occasions are included in the table. Generally, the mean values for each index of response are fairly constant across temperature and ear conditions and across testing sessions. The reader is cautioned not to compare relative variability of the response parameters on the basis of the magnitude of their standard deviations as presented in Table 4 because the units of measurement are different in the case of each parameter.

TABLE 4

ABSOLUTE SCORE MEANS AND STANDARD DEVIATIONS
FOR EACH PARAMETER OF RESPONSE
ON EACH TEST OCCASION

	(1) 30°R			(3) 44°R			(2) 30°L			(4) 44°L		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
Speed of the slow phase												
$\bar{x}$	24.12	22.18	21.56	20.81	21.00	21.25	22.50	20.68	19.37	24.06	23.12	24.06
s	7.99	6.14	7.80	5.40	6.61	5.44	6.32	6.35	4.57	8.30	7.72	9.12
Total beats												
$\bar{x}$	60.18	60.68	61.31	63.62	64.18	64.18	61.50	61.00	63.68	63.56	67.06	66.37
s	12.14	10.53	10.07	14.60	15.38	14.46	17.21	16.44	15.36	14.50	14.50	15.03
Total amplitude												
$\bar{x}$	523.25	436.75	456.31	466.81	439.06	421.06	479.93	408.18	386.68	506.31	466.37	507.87
s	154.15	135.86	173.09	142.44	146.73	153.59	157.36	112.75	91.12	201.71	177.59	215.79
Culmination frequency												
$\bar{x}$	21.75	22.37	22.56	22.87	23.06	23.18	22.25	22.56	23.12	23.18	24.18	24.68
s	4.02	3.98	3.79	5.00	5.28	4.95	5.79	5.26	4.84	5.81	4.79	5.67
Duration												
$\bar{x}$	185.06	178.25	179.31	182.18	178.87	173.00	187.56	180.00	177.12	179.75	171.06	167.68
s	25.30	24.50	29.59	35.96	36.77	25.70	26.50	33.01	27.94	35.82	32.60	29.51

T_1 = Test 1

T_2 = Test 2

T_3 = Test 3

Comparison between ears. The hypothesis that there is no difference in sensitivity between the mean response of the two ears was tested by Student's *t* test for paired observations. The mean values for the right ear (1+3) were compared to those for the left ear (2+4). This was done for every response parameter on each of the three tests. The results of this analysis is contained in Table 5. The mean values representing right and left ears do not differ statistically from each other. This is true for all response indices on each test occasion. These findings are in agreement with those of previous investigators (4, 37, 56, 72).

Comparison between the effects of two test stimulus temperatures. This comparison between test stimulus temperatures was also done by utilizing Student's *t* statistic for paired observations. The mean response to cold stimulation (1+2) was compared to the mean response to warm stimulation (3+4) for each response parameter on each test occasion. Table 6 reports the results. For speed of the slow phase, the warm temperature stimulus was significantly stronger than the cold temperature stimulus on one out of three occasions. The effect of warm temperature stimulus was significantly stronger than the cold temperature stimulus on all three test occasions when the response was expressed by total beats. There was no significant difference between the effects of the stimuli when the response was determined by total amplitude. The same is true for duration of

TABLE 5

SUMMARY OF t SCORE ANALYSIS COMPARING RIGHT AND LEFT EAR RESPONSIVENESS ON EACH OF THREE TEST OCCASIONS

	Speed of the Slow Phase			Total Beats			Total Amplitude			Culmination Frequency			Duration		
	x_r	x_L	t value	x_r	x_L	t value	x_r	x_L	t value	x_r	x_L	t value	x_r	x_L	t value
Test 1	44.93	46.56	-0.8516	123.81	125.06	-0.3947	990.06	982.50	0.1511	44.62	45.43	-0.6539	367.25	367.31	-0.0069
Test 2	43.18	43.50	-0.1314	124.87	127.68	-1.1424	869.68	874.25	0.0266	45.12	46.75	-1.9436	357.12	351.06	0.8213
Test 3	42.81	43.43	-0.3071	125.00	130.06	-1.4530	877.37	894.75	-0.3907	45.75	47.81	-1.8542	352.31	344.81	0.9496

$\bar{x}_r$ = Mean value for right ear (1+3).

$\bar{x}_L$ = Mean value for left ear (2+4).

2.13 = t value needed to be significant at the .05 level of confidence.

TABLE 6

SUMMARY OF t SCORE ANALYSIS COMPARING THE RESPONSES PRODUCED BY WARM AND COLD WATER STIMULI ON THREE TEST OCCASIONS

	Speed of the Slow Phase			Total Beats			Total Amplitude			Culmination Frequency			Duration		
	x_c	x_w	t value	x_c	x_w	t value	x_c	x_w	t value	x_c	x_w	t value	x_c	x_w	t value
Test 1	46.62	44.87	0.8362	121.68	127.18	-2.3890[a]	981.50	973.12	0.1160	44.00	46.06	-2.2768[a]	372.62	361.93	1.2415
Test 2	42.87	43.81	-0.9215	121.68	131.25	-2.5192[a]	844.93	905.56	1.5295	44.93	47.25	-1.8061	358.25	349.93	0.8238
Test 3	40.93	45.31	-3.0019[a]	125.00	130.00	-2.1873[a]	843.18	928.93	-1.8105	46.31	47.87	-2.5807[a]	356.43	340.68	2.1243

$\bar{x}_c$ = Mean cold stimulus response (1+2).

$\bar{x}_w$ = Mean warm stimulus response (3+4).

2.13 = t value needed to be significant at the .05 level of confidence.

[a]Significant at the .05 level of confidence.

response. The effect of stimulus temperature when the response was expressed culmination frequency showed the warm temperature stimulus to be significantly stronger than the cold temperature stimulus on two out of three occasions.

The finding that the warm temperature stimulus is significantly stronger than the cold temperature stimulus when expressed by total beats and speed of the slow phase is in agreement with Mehra (56) and Hamersma (37). The finding of no difference in stimulus temperatures for total amplitude is also in agreement with these authors. The result of the stimulus temperature effect showing no difference when duration is the response index are in agreement with Aschan, Bergstedt and Stahle (4) and Henricksson (40). The result of no difference between cold and warm stimulus temperatures for duration is not in agreement with the findings of Hamersma (37), Hallpike (36) and Jongkees (44). Cold stimulus and warm stimulus temperature comparisons expressed by culmination frequency have not been previously reported.

The Occurrence of Trend Effects

The repeatability of the raw scores was estimated by intraclass reliability coefficients. A one-way analysis of variance technique may be used but the two-way analysis of variance technique (34, 83) computationally eliminates the "between columns" or trials sums of squares which are indicative of response decline, from the within classes sums of squares (see p. 57, chap. iii). This permits the use of the residual mean squares as an estimate of the intrasubject variability which is free of the influence of trend and further provides an indication of the significance of the trend effect, if present. The one-way analysis of variance technique does not provide this capability.

A significant trend effect was seen on four occasions out of the 20 analyses of variance performed on the absolute scores. These trend effects were apparently produced by a response decline to the stimuli over repeated trials. In particular, the absolute scores for conditions 30°C left and 30°C right, when expressed by speed of the slow phase and total amplitude, showed progressively smaller values with repeated testing. The progressive decline in the magnitude of the mean values for these conditions is evident in Table 4.

Tables 7 and 8, respectively, show a summary of the analyses of variance performed on the absolute scores for conditions 30°C right and 30°C left measured by speed of the slow phase and for conditions 30°C right and 30°C left measured by total amplitude. The between trials mean squares in both tables are significant at the .05 level of confidence. The statistical significance of the between trials mean square is the factor which indicates the existence of trend in the mean scores. The entire remainder of the data was also subjected to this type of analysis. However, there were no other instances of trend noted in the

TABLE 7

SUMMARY OF THE ANALYSES OF VARIANCE PERFORMED ON THE RAW SCORES FOR CONDITIONS 30°C RIGHT, 30°C LEFT, SPEED OF THE SLOW PHASE SHOWING SIGNIFICANT TREND

30°C Right					30°C Left				
Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Source	Degrees of Freedom	Sum of Squares	Mean Square	F
Classes (subjects)	15	2191.91	146.12	17.80[a]	Classes (subjects)	15	1352.64	90.17	16.24[a]
Trials	2	57.12	28.56	3.41[b]	Trials	2	78.79	39.39	7.09[b]
Residual	30	246.20	8.20		Residual	30	166.54	5.55	

[a]Significant at the .01 level of confidence. [a]Significant at the .01 level of confidence.

[b]Significant at the .05 level of confidence. [b]Significant at the .05 level of confidence.

2.01 = F value needed to be significant at the .01 level of confidence.

3.32 = F value needed to be significant at the .05 level of confidence.

TABLE 8

SUMMARY OF THE ANALYSES OF VARIANCE PERFORMED ON THE RAW SCORES FOR CONDITIONS 30°C RIGHT, 30°C LEFT, TOTAL AMPLITUDE SHOWING SIGNIFICANT TREND

30°C Right

Source	Degrees of Freedom	Sum of Squares	Mean Square	F
Classes (subjects)	15	966150.47	64410.03	16.56[a]
Trials	2	65843.04	32921.52	8.46[b]
Residual	30	116670.95	3889.03	

[a]Significant at the .01 level of confidence.

[b]Significant at the .05 level of confidence.

30°C Left

Source	Degrees of Freedom	Sum of Squares	Mean Square	F
Classes (subjects)	15	548701.00	36580.06	8.45[a]
Trials	2	69626.62	34813.18	8.04[b]
Residual	30	129809.62	4326.98	

[a]Significant at the .01 level of confidence.

[b]Significant at the .05 level of confidence.

2.01 = F value needed to be significant at the .01 level of confidence.

3.32 = F value needed to be significant at the .05 level of confidence.

absolute score analysis. A comparison of intraclass correlation coefficients derived from one-way and two-way analyses of variance performed on absolute scores is in Appendix B.

Reliability of Absolute Scores

The results of the intraclass correlations (intertest) performed on the absolute scores are displayed in Table 9 in the rows labeled R. It can be seen that the reliability coefficients generally fall within an interval between .80 and .90. Three exceptions to this are duration of response, 30°C right and 44°C left and total amplitude of response, 30°C left. The null hypothesis that the repeated measurements under consideration are not associated in the population and that the observed value of R differs from zero only by chance can be tested by the F-ratio computed from the same mean squares and degrees of freedom that were used to obtain R (34). The statistical significance of each R is identical to that of its associated F value (34). As Table 9 shows each value of R for every response parameter under all conditions is significantly different from .00 at the .01 level of confidence. The F value associated with each R value on the absolute score analysis is presented in Appendix C.

Table 9 also presents the confidence interval limits for each R at the .05 level. It is assumed that 90 per cent of such intervals will include the population R if a large number of samples is taken (34). With the .05 confidence

TABLE 9

THE RESULTS OF THE INTRACLASS CORRELATIONS PERFORMED ON THE RAW DATA AND THE .05 CONFIDENCE INTERVALS ASSOCIATED WITH EACH

		Speed of Slow Phase	Total Beats	Total Amplitude	Culmination Frequency	Duration
30°CR	R	.84[a]	.89[a]	.84[a]	.82[a]	.72[a]
	Confidence interval	.72-.92	.78-.94	.70-.91	.68-.90	.53-.85
44°CR	R	.80[a]	.87[a]	.82[a]	.87[a]	.83[a]
	Confidence interval	.64-.89	.76-.93	.67-.90	.76-.93	.70-.91
30°CL	R	.84[a]	.92[a]	.71[a]	.89[a]	.80[a]
	Confidence interval	.70-.91	.86-.96	.52-.84	.79-.94	.64-.89
44°CL	R	.84[a]	.83[a]	.83[a]	.84[a]	.78[a]
	Confidence interval	.71-.92	.70-.91	.70-.91	.71-.92	.62-.88

[a]Significant at the .01 level.

2.70 - level needed to be significant at the .01 level of significance.

level, the chances are one in 20 that the population R may be as high as or higher than the higher limit or as low as or lower than the lower limit. Considering only the lower values, thirteen of the 20 intervals reported in Table 9 have lower limits that are .70 or greater. Five of the lower limits are between .62 and .68 and the lowest two are .52 and .53.

The confidence intervals may also be used to estimate whether two R's differ significantly from each other (34). This is done by observing whether any two obtained R values being compared fall within the other's confidence interval. If they do not, the two R's may be said to differ with the specified level of confidence. The reliability coefficients for the individual response indices at the various temperature and ear conditions that are significantly different from each other are shown in Table 10. The R value for each of the 20 combinations of response parameters, temperatures and ears was compared to the other 19 combinations (see Table 9). On the basis of chance, one would expect the comparisons to appear to be significantly different once out of 20 times at the .05 level of confidence. Consequently, the occurrence of an R associated with any response parameter which is significantly larger or smaller than that associated with another parameter as frequently as once in 19 comparisons or less can be attributed to chance.

TABLE 10

A TALLY OF THE RESPONSE PARAMETERS WITH SIGNIFICANTLY DIFFERENT R's WITH THE DIRECTION OF THE DIFFERENCE INDICATED

RP	T	E		RP	T	E	RP	T	E		RP	T	E
TB	30°	R	>	DUR	30°	R	TB	30°	L	>	TA	44°	R
TB	30°	R	>	TA	30°	L	TB	30°	L	>	DUR	44°	R
SSP	30°	R	>	TA	30°	L	TB	30°	L	>	TB	44°	L
							TB	30°	L	>	TA	44°	L
TB	44°	R	>	DUR	30°	R	TB	30°	L	>	DUR	44°	L
TB	44°	R	>	TA	30°	L	TB	30°	L	>	SSP	30°	L
CF	44°	R	>	DUR	30°	R	TB	30°	L	>	TA	30°	L
							TB	30°	L	>	DUR	30°	L
TB	30°	L	>	TA	30°	R	CF	30°	L	>	TA	30°	L
TB	30°	L	>	CF	30°	R	CF	30°	L	>	DUR	30°	R
TB	30°	L	>	DUR	30°	R	CF	30°	L	>	DUR	44°	L
TB	30°	L	>	SSP	44°	R							

RP - Response Parameter
T - Temperature
E - Ear
TB - Total Beats
SSP - Speed of Slow Phase

CF - Culmination Frequency
TA - Total Amplitude
DUR - Duration
> - Statistically greater at .05 level of confidence.

The R for total beats for the left ear with 30°C stimulation is significantly greater than the R's for the other isolated response parameters on 12 out of 19 occasions or 63 per cent of the time. The R for culmination frequency 30°C left ear appeared significantly greater than the R's for the other parameters on three out of 19 occasions or 16 per cent of the time. The R's for total beats 30°C right ear and 44°C right ear appeared significantly greater than other parameters on two out of 19 occasions or 10 per cent of the time. These latter three response parameters have significantly greater R's on only two or three occasions. It seems probable that this slightly greater than chance occurrence rate in itself occurred by chance.

Duration of response with 30°C water delivered to the right ear was significantly less reliable than other conditions on six out of 19 occasions or 32 per cent of the time. Duration 44°C left ear was significantly less reliable on two out of 19 occasions or 10 per cent of the time. Total amplitude 30°C left ear produced significantly smaller reliability than the other parameters on five out of 19 occasions or 26 per cent of the time. Again, some of these lower rates in themselves could be chance occurrence rates.

Considered as isolated parameters of response, total beats with 30°C water delivered to the left ear appears to be more reliable than the other parameters most of the time while duration with 30°C water delivered to the right ear and total amplitude with 30°C water to the left ear, appear to be somewhat less reliable than the other isolated measures of response. In a comparison of the response indices, duration and total amplitude appear to be slightly less reliable than the others. Duration was significantly less reliable than the other response indices more often (9 occasions) than any other with total amplitude running a very close second (8 occasions). In no instances were the four reliability coefficients generated for duration significantly larger than those for any of the other indices under any condition.

The analysis of variance for absolute scores showed significant response decline on successive tests for speed of the slow phase 30°C right ear and 30°C left ear and total amplitude 30°C right ear and 30°C left ear.

As discussed earlier, R can be calculated with the effect of trends present in the data or with its effect removed. It was decided to utilize those R's obtained with the effect of trend removed for the following reasons. First, trend was a significant factor in only four out of twenty possible occasions. Therefore, the R's produced are significantly affected by a trend in the scores in only the four specific cases. Second, in clinical assessment, the usual interval of time between repeated tests of at least several days is sufficiently long so that a response decline of an absolute score will not occur. Third, trend does not appear in the difference score data of this study. The use of the procedures which remove the trend effect throughout permits comparison of the R's obtained from these two data groups to be made with fewer difficulties.

It is evident that some parameters of response magnitude under specific temperature and ear conditions are more or less reliable than are others. However, all of the absolute score response indices considered in this experiment are highly reliable. That is, repeated measurements on the same individual produce data with a high degree of homogeneity regardless of test temperature, test ear or response index utilized. Several observations support this conclusion. Every reliability coefficient is significantly different from zero at the .01 level of confidence. Seventeen of the 20 reliability coefficients are .80 or higher. The remaining three are .71 or greater. In addition, thirteen of the lower limits of the confidence intervals are greater than .70. Five are between .62 and .68 while the lowest two are .52 and .53.

The discussion up to this point has been concerned with the reliability of the response parameters based on the four absolute scores that result from the administration of the Fitzgerald-Hallpike caloric test. It will be remembered that in clinical application of this test these absolute scores are combined in order to generate a difference score. This score represents the difference between the algebraically added responses of the two ears to the warm and cold water stimuli. The reliability of the difference score is discussed in the following section.

Analysis of the Difference Scores

The difference score is obtained by subtracting the total score (the response to the warm water stimulus plus the response to the cold water stimulus) for one ear from the total score for the opposite ear. In the present investigation 30°C right ear was designated response number one; 30°C left ear was designated response number two; 44°C right ear was designated response number three and 44°C left ear was designated response number four. This is the most commonly used system. The total score for the right ear was obtained by adding responses one and three. The total score for the left ear was the sum of responses two and four. The difference score equals the difference between the total scores [(1+3) - (2+4)].

Each subject was tested three times. Hence, three sets of difference scores were derived for each response index. Table 11 presents the mean difference score and the standard deviation of the difference scores contributing to that mean score for each response parameter on each test occasion. The differences between the results obtained under these various conditions are the principal subjects of the analysis which follows in this chapter. Also, these values will be further discussed later in this chapter in the section devoted to the "normal" limits of response

TABLE 11

DIFFERENCE SCORE MEANS AND STANDARD DEVIATIONS FOR EACH PARAMETER OF RESPONSE ON EACH TEST OCCASION

	Speed of Slow Phase		Total Beats		Total Amplitude		Culmination Frequency		Duration	
	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S
T_1	-1.62	7.63	-1.25	10.25	37.81	196.48	-0.81	4.96	-0.06	36.10
T_2	-0.31	9.50	-2.81	9.84	1.37	206.06	-1.31	3.78	6.06	29.52
T_3	-0.87	8.14	-5.12	14.10	-17.37	177.86	-1.62	4.78	7.50	31.59

$\bar{X}$ - Mean Difference Score

S - Standard Deviation of Difference Scores

T_1 - Test 1

T_2 - Test 2

T_3 - Test 3

variability. Again the reader is cautioned not to attempt to compare the reliability of the indices on the basis of the relative numerical size of the standard deviations reported in the table. Further, the size of these standard deviations cannot be compared directly to those obtained on the absolute scores as will be explained subsequently.

In contrast to the results of the absolute score analysis there was no instance of "trend" or significant intertest differences obtained from the difference score analysis.

The intraclass reliability coefficients for the three sets of difference scores obtained on the three successive tests are displayed in Table 12. A comparison of intraclass correlation coefficients derived from one-way and two-way analyses of variance performed on the difference scores is presented in Appendix B. In comparing the difference score R values in Table 12 with those reported in Table 9 for absolute scores it appears that reliability of the difference score is generally poorer. Speed of the slow phase and total amplitude continue to exhibit high reliability coefficients although the magnitude of the R associated with these measures has declined slightly (by approximately .08) relative to the values associated with the absolute scores. On the absolute score analysis the total beats and culmination frequency data produced very high correlations (see Table 9). However, the analysis of the difference scores shows only a

TABLE 12

SUMMARY OF THE RESULTS OF INTRACLASS CORRELATION PERFORMED ON THE DIFFERENCE SCORES

Response Parameters	Speed of Slow Phase	Total Beats	Total Amplitude	Culmination Frequency	Duration
Reliability coefficient	.75[a]	.59[a]	.72[a]	.52[a]	.26[b]
Confidence interval	.56-.86	.36-.76	.54-.85	.28-.71	-.33-.50

[a]Significant at .01 level of confidence.

[b]Significant at .05 level of confidence.

moderate degree of correlation for these two response indices. Finally, duration of response presents a third picture. In the absolute score analysis reliability of the duration index was high, although generally lower than for any other response index. For difference scores, however, the reliability coefficient for duration is very much lower (.26). This value indicates no more than a slight, almost negligible, test-retest relationship. The obtained correlation was just barely significantly different from zero at the .05 level of confidence for this parameter. The F value associated with each R value is presented in Appendix C.

The absolute score and the difference score results appear to be inconsistent with one another. While the absolute data analysis showed a relatively high degree of reliability for all the response parameters, the difference score analysis revealed poorer reliability for all parameters. On the difference score analysis one R was much poorer than the rest, two were moderately reduced and two were slightly reduced with respect to the absolute score analysis.

In an effort to obtain a further understanding of this outcome, the results of the statistical analyses performed on the absolute score and difference score data were carefully examined. The average between class mean square and the average residual mean square obtained from the absolute score analyses were directly compared to the between class mean square and the residual mean square obtained from the difference score analysis.

As noted earlier, in the analyses of variance performed as a part of the intraclass correlation, the residual mean square was utilized as the error term. This mean square represents the within class (subject) variability. The intersubject variability is represented by the between class (subject) mean square.

In order to make a direct comparison, the mean squares from the difference score analysis were adjusted for the differences in magnitude between the absolute scores and the difference scores. The difference score is based on the sum of the responses for the right ear and the sum of the responses for the left ear. Consequently, the difference score based on total scores is on the average, twice as large as would be a difference score based on absolute scores. If a distribution of scores is formed by increasing the scores of a previously existing distribution by a factor of two, the variance (mean square) of the newly formed distribution will be increased by a factor of four over the variance of the previously existing distribution. Because the summing of the two absolute scores for each ear results in the difference scores being increased by a factor of two, the mean square resulting from the difference scores analysis must be reduced by a factor of four in order to properly compare the magnitude of the mean squares for absolute scores to the magnitude of the mean squares for difference scores.

Table 13 records the ratio of the residual mean squares and the between class mean squares from the intraclass correlation analyses performed on the absolute scores to the intraclass correlation analyses performed on the difference scores. The ratio was obtained by dividing each BCMS resulting from the difference score analyses by the mean BCMS resulting from the absolute score analyses and each RMS resulting from the difference score analysis by the mean RMS resulting from the absolute score analysis. This was done for each parameter of response.

Table 13 depicts that the BCMS is smaller on the difference score analysis for every response parameter. The RMS is also smaller on the difference score analysis for every response parameter except for duration where it is virtually equal in the two analyses.

The BCMS reduction reflects the fact that the range of difference score responses across subjects was sharply restricted relative to the intersubject range of absolute score responses. Like all parametric correlations, the intraclass correlation coefficient is strongly affected by the range of performance present in the sample. The narrower the range, the smaller will be the R value even while the error term remains unchanged. This relationship can be seen by examining the formula for intraclass correlation

TABLE 13

RATIOS DEPICTING THE CHANGES IN RESIDUAL MEAN SQUARES AND BETWEEN CLASSES MEAN SQUARES FROM THE ABSOLUTE SCORE ANALYSES TO THE DIFFERENCE SCORE ANALYSES FOR EACH RESPONSE PARAMETER

	Between Classes Mean Square	Residual Mean Square
Speed of slow phase	.35	.58
Total amplitude	.35	.56
Total beats	.13	.57
Culmination frequency	.16	.71
Duration	.17	1.00

$$R = \frac{BCMS - RMS}{BCMS - (k-1)\ RMS}.$$

The BCMS, or between subject variability, for speed of the slow phase and total amplitude was smaller on the difference score analysis than on the absolute score analysis by a factor of about .35. The BCMS was smaller in the difference score analysis by factors of .13 and .16, respectively, for total beats and culmination frequency. The BCMS for duration was smaller in the difference score analysis by a factor of .17. The reduced intersubject range in the difference score data is not surprising and is the result of each subject serving as his own control. However, the magnitude of the reduction was larger than had been expected. It is obvious that comparing the responsivity of one ear to the other ear greatly reduces the intersubject differences.

The residual mean squares are also reduced in the difference score analyses except for duration where there was no change. This reduction by itself would, of course, increase R. However, as can be seen by inspection of Table 13 the extent of the reduction of the RMS is always less and sometimes much less than the reduction in the BCMS. The net effect is the reduction in the R values presented earlier.

The difficulty of estimating reliability by the use and interpretation of correlation coefficients is apparent.

However, a study of the error term derived from the analysis of variance can be helpful. The absolute score reliability estimates for all parameters of response were high, although there was a tendency for the duration R values to be slightly lower than those for the rest of the indices. On the difference score analysis the error term was smaller by approximately equal amounts for speed of the slow phase, total amplitude and total beats. This outcome suggests greater reliability for the difference scores relative to the absolute scores in spite of the obtained R values. The reliability estimates declined because of an even greater reduction in the difference score differences across subjects. The R for total beats declined further than speed of the slow phase or total amplitude because the range of scores was even more restricted for this index than for the other two, e.g., BCMS ratio change of .13.

The BCMS for culmination frequency was greatly smaller on the difference score analysis than on the absolute score analysis. The RMS for the difference scores was also smaller than for the absolute scores suggesting higher intertest reliability for difference scores. However, the reduction in RMS for culmination frequency was smaller than that seen for total beats, total amplitude or speed of the slow phase. The result was a reduced reliability coefficient that is similar but slightly less than the one for toal beats.

The BCMS for duration decreased greatly on the difference score analysis but unlike the other indices the error term remained undiminished; that is, the within subjects variability was not decreased through the act of comparing the two ears of a given subject as was true for the other indices. For this reason, the reliability coefficient describing this response parameter is well below that associated with any other index.

The decrease in the RMS values on the difference score analysis for speed of the slow phase, total amplitude, total beats and culmination frequency is interpreted as suggesting that difference scores are more repeatable than absolute scores when these indices are used in spite of the observed smaller R values.

It is difficult to ascribe a hierarchy of reliability to the response indices. With the exception of duration, clear cut distinctions among the indices are not immediately evident. However, the following hierarchy is suggested by the writer. Total beats and culmination frequency appear to more reliable than speed of the slow phase and total amplitude. This conclusion is reached because higher reliability coefficients were generally obtained for total beats and culmination frequency on the absolute score analysis and because the reduced R values obtained from the difference score analysis are the result of a great reduction in the intersubject range of scores and not reduced repeatability.

In addition, a response decline on repeated measurements was not evident for these two response parameters in either analysis, although as previously mentioned, this is not a major point because the difference score is not significantly affected by a response decline on any parameter. Nevertheless, response decline is a factor which could have bearing in those clinical situations where a short term retesting is desired and only absolute scores are utilized to report the findings.

Speed of the slow phase is ranked ahead of total amplitude because it is highly reliable on both difference and absolute score analysis. Total amplitude, it will be recalled, had a high difference score R but a relatively low R on the absolute score analysis. Duration is the least reliable index last on the basis of a low difference score R and a high but, nevertheless, last place performance on the absolute score analysis.

Normative Values

The data from this experiment provide at least two types of information which can be used to establish norms. These include intersubject variability of both absolute and difference scores for each response parameter and also the allowable or "normal" test-retest variation in these scores.

It is customary in the literature dealing with the assessment of vestibular function by the Fitzgerald-Hallpike caloric test to define the normal limits of response as two times the standard deviation (intersubject) of the difference scores derived from a normal sample (4, 37, 41, 45, 72). For this reason, this standard is utilized in the discussion which follows. These limits will be referred to simply as the "normal limits." The two-sigma limits will also be used as the "normal" test-retest variability limits.

Table 14 presents the overall mean values and standard deviations for the absolute scores obtained for each of the response parameters. The measure of central tendency for each parameter was determined by averaging the absolute scores for all subjects obtained from both ears at both temperatures on all three occasions. The normal limits of the absolute scores are also given. These average values are not routinely used in reporting the outcome of Fitzgerald-Hallpike testing but they could be useful in the event that the two ears of a patient cannot be compared for some reason.

Table 15 depicts the mean difference scores, standard deviations and the normal limits for the difference scores (intersubject variability) for each parameter of response. The probability that a normal individual will produce a difference score outside these limits is approximately 4.56 per cent assuming a normal distribution of responses in the population.

TABLE 14

MEANS--INTERSUBJECT STANDARD DEVIATIONS AND RANGE OF NORMAL PERFORMANCE FOR THE ABSOLUTE SCORES FOR EACH RESPONSE PARAMETER

	Mean Difference (1+3) - (2+4)	Standard Deviation	Range of Normal Performance
Speed of slow phase	22°/second	6.9°/second	8.2 - 35.8
Total beats	63 beats/30 seconds	14 beats/30 seconds	35 - 91
Total amplitude	458 degrees	158 degrees	142 - 774
Culmination frequency	23 beats	5 beats	13 - 33
Duration	178 seconds	31 seconds	116 - 240

TABLE 15

MEANS--INTERSUBJECT STANDARD DEVIATIONS AND LIMITS OF NORMAL PERFORMANCE FOR THE DIFFERENCE SCORES FOR EACH RESPONSE PARAMETER

	Mean Difference (1+3) - (2+4)	Standard Deviation	Normal Limits
Speed of slow phase	-0.93°/second	7.48°/second	15°/second
Total beats	-3.06 beats/30 seconds	11.57 beats/30 seconds	23 beats/30 seconds
Total amplitude	7.27°	193.82°	388°
Culmination frequency	-1.27 beats/10 seconds	4.68 beats/10 seconds	9 beats/10 seconds
Duration	4.50 seconds	32.52 seconds	65 seconds

The standard method of reporting the results of the Fitzgerald-Hallpike caloric test is the difference score. The reason for this approach is the observation of relatively great intersubject variability in the absolute scores. Several investigators have reported that the absolute parameters of response are less useful because of their "great" interindividual variation (4, 24, 37, 72). When a comparison is made between the standard deviations of the absolute scores recorded in Table 14 and the standard deviations of the difference scores recorded in Table 15 for each parameter of response, they appear to be quite similar. It would seem that the intersubject variability of a particular distribution of absolute scores for a given response parameter is no greater than the intersubject variability of a distribution of difference scores for the same parameter of response. When the intersubject variability of the absolute score data is compared to the intersubject variability of the difference score data in Hamersma's (37) investigation, the same near equivalence of variability in the two types of data is also present.

However, a direct comparison of the intersubject variability values of the absolute scores and the difference scores is misleading and improper. This is because the difference score is not the difference between two absolute scores, but, rather, the difference between two sums made up of two absolute scores each, i.e., (1+3) - (2+4). Hence, the magnitude of these sums will be, on the average, twice as large as the absolute scores and the differences will be correspondingly larger. The intersubject variability values of the total scores for each ear (1+3) and 2+4) are much greater than the intersubject variability of the difference scores. This is shown in Table 16 which reports the standard deviations of the total scores. Generally speaking, it can be seen that the standard deviations reported in Table 16 are approximately twice as large as those reported in Table 15 for the difference scores.

Pearson product-moment correlation coefficients between total scores of the two ears of the individual subjects were calculated for every response parameter on each test occasion. These coefficients are also displayed in Table 16. The correlation coefficients between the total scores for the right ears and left ears are always high. This is true for each response parameter on all three tests. The high correlation coefficient between the total scores for the right ear and the total scores for the left ear and the decreased variability of the difference scores relative to the total scores both strongly support the common clinical practice of comparing the results for one ear of an individual with those for his own opposite ear; that is, of using difference scores.

Test-retest limits were also calculated for the difference score data. The mean change in difference scores

TABLE 16

STANDARD DEVIATIONS OF TOTAL SCORES AND PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS BETWEEN THE TOTAL SCORES FOR THE RIGHT EAR (1+3) AND THE TOTAL SCORES FOR THE LEFT EAR (2+4) ON EACH TEST OCCASION

		Speed of Slow Phase		Total Beats		Total Amplitude		Culmination Frequency		Duration	
		S	r	S	r	S	r	S	r	S	r
Test 1	R	11.01	.81	25.74	.94	267.33	.78	8.64	.90	52.46	.74
	L	13.18		30.32		321.21		11.09		46.67	
Test 2	R	11.73	.68	24.99	.93	251.20	.67	8.46	.93	47.33	.82
	L	12.11		27.05		257.58		9.04		52.26	
Test 3	R	12.32	.78	23.60	.87	287.16	.79	8.41	.89	35.33	.76
	L	12.46		29.02		257.25		9.90		48.96	

over testing sessions was computed and the normative values were determined by doubling the standard deviation associated with these data. Table 17 reports the normal limits of test-retest variability in difference scores when the Fitzgerald-Hallpike caloric test is given more than once to an individual. When the Fitzgerald-Hallpike caloric test is so repeated, approximately 95.44 per cent of the repeated test results done on normal individuals will not exceed the recorded values assuming a normal distribution of test-retest values in the population.

Comparison between the Magnitude of Absolute Scores and the Magnitude of Difference Scores

The possibility of any relationship between the size of an individual's absolute scores and the size of his difference score was explored. This was done to determine whether large difference scores are associated with large absolute scores (or vice versa). Should this be true, it might be necessary or desirable to specify different difference score limits depending upon the size of an individual's absolute scores. Spearman rank correlations were performed between the size of the absolute scores and the size of the difference scores across individual subjects. The test was applied to the results of each testing session. A nonparametric statistic was used because of uncertainties regarding the homogeneity of the variances of the absolute score-difference score distributions. The changes in the residual

TABLE 17

MEANS--STANDARD DEVIATIONS AND NORMAL TEST-RETEST LIMITS FOR DIFFERENCE SCORES FOR EACH RESPONSE PARAMETER

	Mean	Standard Deviation	Normal Limits
Speed of slow phase	-0.72	6.02	12
Total beats	1.25	11.27	22
Total amplitude	12.47	132.51	265
Culmination frequency	-0.16	4.96	10
Duration	-3.78	38.14	76

mean squares enhanced these uncertainties by showing differences in size of up to seven times even after adjustments for summation in the difference scores had been made (see Table 13). The ranked total absolute score responses, (1+2) + (3+4), were compared with the ranked difference scores. Table 18 depicts these results. The hypothesis that there is no association between the two measures in the population was rejected on only two of the 15 occasions. A significant correlation was seen in one out of three tests for duration and one out of three tests for culmination frequency. In both cases significance was barely attained at the .05 level of confidence.

The low incidence of significant correlation coefficients (two out of 15) and the fact that the two significant correlation coefficients which did occur barely attained significance at the .05 level suggest that there is little or no relationship between the magnitude of the absolute scores produced by an individual and the magnitude of the difference scores for the same individual. Therefore, the same limits of normal performance can be applied to the difference scores of anyone whose absolute scores fall within the range of absolute scores observed in this study.

Discussion

The principal goal of this experiment was to determine the reliability of several parameters of response used to express the magnitude of the thermally induced nystagmic

TABLE 18

SUMMARY OF SPEARMAN RANK CORRELATION BETWEEN THE MAGNITUDE OF THE RAW SCORES AND THE MAGNITUDE OF THE DIFFERENCE SCORES

	Speed of Slow Phase	Total Beats	Total Amplitude	Culmination Frequency	Duration
T_1	.30	.22	.23	.49[a]	.12
T_2	.03	.17	-.11	.16	.44[a]
T_3	-.27	.36	-.19	.41	.16

T_1 - Test 1

T_2 - Test 2

T_3 - Test 3

[a]Significant at the .05 level of confidence.

reaction. The analyses revealed several factors concerning the test-retest reliability of the indices of nystagmus response magnitude considered in this experiment.

The absolute scores for the four conditions that comprise a Fitzgerald-Hallpike test are highly reproducible upon repeated testing of the same individual for all parameters of response. Comparison of this result with the work of others must be limited because little has been reported in the area of repeated testing of the same individual. Hinchcliffe (41) and Halama and Hinchcliffe (35) tested their subject sample twice with the warm stimulus only. These authors concluded that absolute scores resulting from the warm stimulus and expressed by speed of the slow phase and total beats are reliable. Considering the warm stimulus, the results of the absolute score analysis in this investigation are similar to those of Hinchcliffe (41) and Halama and Hinchcliffe (35). However, the reliability coefficients for the absolute scores reported in this study are generally higher than those reported by these other authors.

The analysis of the reliability of the difference scores appeared to give different results than the analysis of the reliability of the absolute scores. The reliability coefficients for speed of the slow phase and total amplitude were slightly lower than those for the absolute scores but these two indices still exhibited relatively high reliability upon repeated measurements. The reliability coefficients for total beats and culmination frequency, however, indicated only moderate reliability on the difference score analysis. Finally, the reliability coefficient for duration of response was very low and suggested only a slight degree of test-retest reliability for this response index.

An analysis of the mean squares resulting from the analysis of variance suggested that speed of the slow phase, total amplitude, total beats and culmination frequency actually may be more reliable in expressing the relative magnitude of thermally induced nystagmus in the form of difference scores than in the form of absolute scores in spite of the apparent reduction in the R values on the difference score analysis. Duration, when reported as a difference score, is about as reliable as when reported as an absolute score but no better when conclusions are based on the magnitudes of the error terms. Taken together, the analyses suggested that total beats and culmination frequency were more reliable indices than speed of the slow phase which, in turn, was more reliable than total amplitude. However, all four of those indices are highly reliable and show greater reliability than duration on both difference score and absolute score analyses.

These results generally seem to support (if duration is excluded) those investigators who have recommended that any one of the response parameters that express the magnitude of the ongoing response may be used with equal confidence (4, 37, 55, 56, 66, 72). The data do not entirely support those who conclude that a particular response parameter is most reliable (5, 40, 41, 77, 78), except to clearly discourage the use of duration. Total beats and culmination frequency did appear somewhat more reliable than the others, however, a definite clear-cut delineation in favor of these indices is not present in the results.

With the exception of duration, the data of this investigation do not permit the recommendation for exclusive use of any one index over the other. The choice of which index to use clinically or for any other purpose may be based on other considerations such as convenience. In situations where time is a relevant factor, total beats or culmination frequency are desirable because of the relative ease of computing these indices. While the R values associated with these measures were relatively small, this result seemed to occur because of reduced intersubject range rather than a lack of repeatability as such. Duration, however, did appear less reliable than the other indices of response and if duration is used, it is recommended that one of the other four indices be reported in addition.

A trend effect in the form of a response decline was seen in four out of 20 instances in the absolute score data. This is a smaller incidence of response decline than some other sources have reported (49, 50). Collins (15), on the other hand, observed no significant response decline for speed of the slow phase, total beats or duration under the testing conditions of eyes open in total darkness with the subjects engaged in mental activity. The present investigation employed conditions of mental activity with the subject's eyes lightly closed in near dark surroundings. The conditions used in this study are similar to those used by Collins and may well account for the relatively small amount of response decline seen in this study.

The two sigma difference score limits are widely used in the electronystagmographical literature as arbitrarily defined normal intersubject limits. The values so obtained in this study are compared to similar data reported by other investigators in Table 19. It can be readily seen that the normative difference score limits for speed of the slow phase are in moderately good agreement across the investigators presented in the table. When total beats is the parameter considered, it can be seen that the normal intersubject limit values of Hinchcliffe and those of this investigation are similar. The values reported by Stahle and Hamersma appear at first inspection, to be divergent from the value derived in the present study, however, the time base used differs across these studies. In the present investigation and in Hinchcliffe's investigation, the total beats value was computed from within the 30-second interval between 60 and 90 seconds after the onset of irrigation.

TABLE 19

COMPARISON OF TWO SIGMA DIFFERENCE SCORE LIMITS FROM THIS INVESTIGATION WITH THOSE OF OTHER INVESTIGATORS

	Speed of Slow Phase	Total Beats	Total Amplitude	Duration
Present investigation	15°/second	23 beats/30 seconds	388°/30 seconds	65 seconds
Aschan, Bergstedt and Stahle	12°/second		742°/175 seconds (127°/30 seconds)	60 seconds
Stahle	10°/second	100 beats/177 seconds (17 beats/30 seconds)	1000°/177 seconds (169°/30 seconds)	60 seconds
Hamersma	13°/second	130 beats/212 seconds (18 beats/30 seconds)	850°/212 seconds (120°/30 seconds)	120 seconds
Hinchcliffe		20 beats/30 seconds		

Hence, the normative values were derived using the standard deviation of the number of total beats occurring in 30 seconds. In contrast, Stahle (72) and Hamersma (37) summed the beats from the entire course of the reaction. Consequently, their normative values were derived using the standard deviation of the total number of beats for 177- and 212-second intervals (the average response duration), respectively. When Stahle's and Hamersma's normative values were converted to beats per 30 seconds (recorded in the table), it can be seen that there is moderately good agreement across the investigations for this index of response as well.

The total amplitude data from Aschan, Bergstedt and Stahle (4), Stahle (72) and Hamersma (37) were also converted to average degrees per 30-second intervals. The response parameter of total amplitude so converted, exhibits considerably less intersubject variability in the work of these other investigators than in the present investigation. This outcome may possibly be explained by the fact that in later portions of the reaction, the amplitude of the nystagmus beats becomes more similar across subjects than it is during that portion of the record where the reaction is strongest. This was noted frequently in the records obtained during this investigation. Hence, the results averaged over the entire reaction will have less intersubject variation than when based on that part of the record where the amplitudes are greatest. Culmination frequency does not appear in the table because this response parameter was not included in the reports of other investigators. Finally, the values for duration are very similar across studies with the exception of the limit reported by Hamersma (37).

The allowable test-retest variation (two sigma limits) in difference scores can have very useful clinical application in the interpretation of changes observed in Fitzgerald-Hallpike test results. For example, considering the speed of the slow phase as the response parameter, an individual might achieve a difference score of +13° on the Fitzgerald-Hallpike caloric test and a difference score of -8° on a second or repeat test. Table 15 illustrates that the normative limit of difference between the ears for this parameter is 15°. Hence, both tests produced interear differences which were within normal limits. However, from the first test to the second test the difference score changed by 21°. Table 17 provides information which shows that such a change in difference score exceeds that which would be expected in the normal population.

CHAPTER V

SUMMARY AND CONCLUSIONS

For many years the inducement of nystagmus by irrigating the ear canal with a water stimulus with a temperature different from that of the normal body temperature has been a means of evaluating the responsivity of the vestibular system. The Fitzgerald-Hallpike caloric test is the most widely used method of vestibular assessment that employs the water-irrigation technique.

Originally, the duration of the thermally induced mystagmic reaction was the only means of quantifying the response to the Fitzgerald-Hallpike test. However, the development and use of electronystagmography has enabled the use of other parameters of response which can be used to determine the magnitude of the nystagmus reaction. These include various measurements of the amplitude, frequency and speed of the nystagmic eye movements. Although there has been extensive research in the area of vestibular assessment, little has been done relative to defining the reliability of the response parameters used in quantifying the magnitude of the thermally induced nystagmus.

This study investigated the reliability (repeatability) of several measures of the nystagmus response. Three complete Fitzgerald-Hallpike tests were administered to a sample of 16 neurologically normal adult males. There was an interval of one-half hour between successive test administrations. The induced nystagmus was recorded electronystagmographically and the records of each test were separately analyzed for speed of the slow phase of nystagmus, total beats of nystagmus, total amplitude of nystagmus, culmination frequency of nystagmus and duration of nystagmus. Estimates of the reliability of repeated measurements for each parameter of response were obtained principally by submitting the data to an analysis of intraclass correlation.

Findings

The results of the intraclass correlation analysis of absolute scores (the raw scores obtained for each ear at each temperature) showed that all of the parameters of response investigated are highly reliable when the caloric test is repeated on the same individual although duration and, to a lesser extent, total amplitude showed somewhat poorer reliability on the absolute score analysis than did the other indices.

The outcome of the Fitzgerald-Hallpike test is not routinely reported in the form of the actual scores (absolute scores) obtained but rather as difference scores which represent the difference in the response obtained when each of the two ears is stimulated. The analysis of the reliability of the difference scores showed different results than the analysis of the reliability of the absolute scores. The reliability coefficients of the difference scores for speed of the slow phase and total amplitude were high. However, the reliability of total beats and culmination frequency showed only moderate reliability on the difference score analysis. It was determined that the poorer reliability of total beats and culmination frequency occurred because of a restricted intersubject range of difference scores for these two parameters and not as a result of poorer repeatability as such.

If, instead, the small intersubject range had been like that seen for speed of the slow phase and total amplitude, the reliability coefficients for total beats and culmination frequency would have been similar to those for speed of the slow phase and total amplitude. The reliability coefficient for duration of response on difference score analysis was extremely low. The error term associated with duration indicated relatively poorer reproducibility on repeated measurements for this response index than for the others.

It was concluded that speed of the slow phase, total amplitude, total beats and culmination frequency are all highly reliable and can be used with confidence in the expression of nystagmus magnitude. Although substantial differences in reliability between speed of the slow phase, total amplitude, total beats and culmination frequency were not present in the data, the analyses suggested the following hierarchy of degree of reliability. Total beats and culmination frequency are more reliable than speed of the slow phase which is more reliable than total amplitude. All four of these indices are substantially more reliable than duration. Duration in the form of a difference score exhibits relatively low reliability and is not recommended for use as an index of response.

The data from this study were used to establish normal (two sigma) intersubject difference score limits as well as normal test-retest differences in difference scores. Additional findings of the investigation are summarized below:

1. A significant difference between ears was never seen for any response parameter on any of the three test occasions. A high correlation was seen between the responses produced by right ear stimulation and the response produced by left ear stimulation for all parameters of response on every test occasion.
2. Several significant differences between the two stimulus temperatures were observed. Warm water was a significantly stronger stimulus than cold water on one out of three test occasions for speed of the slow phase. When the response was expressed by total beats the response

to the warm stimulus was significantly stronger than the response to the cold temperature stimulus on all three test occasions. The warm temperature stimulus produced significantly greater response than the cold temperature stimulus on two out of three occasions for culmination frequency.

3. Response decline was observed in four out of 20 analyses on the absolute scores. These occurred for 30°C left ear, 30°C right ear speed of the slow phase and 30°C left ear and 30°C right ear total amplitude.

4. There was little, if any, significant correlation between the magnitude of the absolute scores and the magnitude of the difference scores.

BIBLIOGRAPHY

1. Anderson, H. Directional preponderance in some intracranial disorders. ACTA Oto-laryngologica, 1954, 44: 568-579.

2. Arnold, G., Giuliani, G., and Stephens, G. Electronystagmographic studies of vestibular function. Annals of Otology, Rhinology and Laryngology, 1959, 68: 129-144.

3. Aschan, G. Positional nystagmus in man during and after alcohol intoxication. Quarterly Journal in Study of Alcohol, 1956, 17: 381-405.

4. ________., Bergstedt, M., and Stahle, J. Nystagmography. ACTA Oto-laryngologica Supplement 129, 1956, 1-103.

5. Barber, H., Wright, G., and Demanuele, F. The hot caloric test as a clinical screening device. A.M.A. Archives of Otolaryngology, 1971, 94: 335-337.

6. Bergman, P., Nathanson, M., and Bender, M. Electrical recording of normal and abnormal eye movements modified by drugs. Archives of Neurology and Psychiatry, 1952, 67: 357.

7. Bernstein, L. The etiology and diagnosis of vertigo. A.M.A. Archives of Otolaryngology, 1962, 76: 329-337.

8. ________. Simplification of clinical caloric test. A.M.A. Archives of Otolaryngology, 1965, 81: 347-349.

9. Bochenek, S., and Omerod, F. The inhibitory action of certain substances on the response to vestibular stimulation. Journal of Laryngology, 1962, 76: 39-44.

10. Buys, E. Review d'oto-neuro-ocul, 1924, 2: 641. Cited in Henriksson, N. Speed of slow component and duration in caloric nystagmus. ACTA Oto-laryngol ogica Supplement 125, 1956, 3-29.

11. Carmichael, E., Dix, M., and Hallpike, C. Observations upon the neurological mechanism of directional preponderance of caloric nystagmus resulting from vascular lesions of the brain stem. Brain, 1965, 88: 51-74.

12. Collins, W. Effect of mental set upon vestibular nystagmus. Journal of Experimental Psychology, 1962, 63: 191-197.

13. ________. Manipulation of arousal and its effects upon human vestibular nystagmus induced by caloric irrigation and angular accelerations. Aerospace Medicine, 1963, 34: 124-125.

14. ________. Some methodological considerations in caloric tests of vestibular function. In Third Symposium on the Role of the Vestibular Organs in Space Exploration, chaired by Ashton Graybiel. Washington, D.C.: Office of Technology Utilization, National Aeronautics and Space Administration, 1968.

15. ________. Subjective responses and nystagmus following repeated unilateral caloric stimulation. Annals of Otology, Rhinology and Laryngology, 1965, 74: 1034-1054.

16. ________. Vestibular response from figure skaters. Aerospace Medicine, 1966, 37: 1098-1104.

17. ________., and Geudry, F. Arousal effects and nystagmus during prolonged constant angular acceleration. ACTA Oto-laryngologica, 1962.

18. ________., Guedry, F., and Posner, J. Control of caloric nystagmus by manipulating arousal and visual fixation distance. Annals of Otology, Rhinology and Laryngology, 1962, 71: 187-202.

19. Crosby, E., Humphrey, T., and Lauer, E. Correlative Anatomy of the Nervous System. New York: The Macmillan Company, 1962.

20. Dohlman, G. ACTA Oto-laryngologica Supplement 5, 1925. Cited in Henriksson, N. Speed of slow component and duration in caloric nystagmus. ACTA Oto-laryngologica Supplement 125, 1956, 3-29.

21. Evitar, A. Dizziness as related to menstrual cycle and hormonal contraceptives. A.M.A. Archives of Otolaryngology, 1969, 90: 301-306.

22. Faultz, S., and McCall, J. The laboratory evaluation of antivertigo drugs. Allied Therapeutics, 1967, 9: 843-845.

23. Fernandez, C., and Schmidt, R. Studies in habituation of vestibular reflexes: III. A revision. The Laryngoscope, 1962, 72: 939-953.

24. Fitzgerald, G., and Hallpike, C. Studies in human vestibular function: I. Observations on the directional preponderance of caloric nystagmus resulting from cerebral lesions. Brain, 1942, 65: 115-137.

25. Fluur, E. Influences of semicircular ducts on extraocular muscles. ACTA Oto-laryngologica Supplement 149, 1959.

26. ________. Vestibular compensation after labyrinthine destruction. ACTA Oto-laryngologica, 1960, 52: 367-375.

27. ________., and Mendel, E. Habituation, efference and vestibular interplay. I Monaural caloric habicuation. ACTA Oto-laryngologica, 1962, 55: 65-80.

28. Gacek, R. The innervation of the vestibular labyrinth. Annals of Otology, Rhinology and Laryngology, 1968, 77: 676-685.

29. Gillingham, K. Mental activity during ENG testing. Annals of Otology, Rhinology and Laryngology, 1969, 78: 575-586.

30. Greisen, O. Pseudocaloric nystagmus. ACTA Oto-laryngologica, 1972, 73: 341-343.

31. Guedry, F., and Lauver, L. Vestibular reactions during prolonged constant angular acceleration. Journal of Applied Physiology, 1961, 16: 215-220.

32. Guilford, J. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Company, 1956.

33. Gutner, L., Gould, N., and Batterman, R. Action of dimenhydrinated (Dramamine) and other drugs on vestibular function. A.M.A. Archives of Otolaryngology, 1951, 53: 308-315.

34. Haggard, E. Intraclass Correlation and the Analysis of Variance. New York: The Dryden Press, Inc., 1958.

35. Halama, A., and Hinchcliffe, R. Studies on a clinical caloric test. Journal of Laryngology and Otology, 1970, 84: 149-153.

36. Hallpike, C. The caloric test: a review of its principles and practice with especial reference to the phenomenon of directional preponderance. In The Vestibular System and Its Diseases, edited by Robert J. Wolfson, Philadelphia University of Pennsylvania Press, 1966.

37. Hamersma, H. The caloric test. A nystagmographical study. Doctoral dissertation, University of Amsterdam, Amsterdam, Holland, 1957.

38. Hart, C. Corneo-retinal potential variation and the bithermal caloric test. Annals of Otology, Rhinology and Laryngology, 1969, 78: 181-186.

39. ________. The value of the hot caloric test. The Laryngoscope, 1965, 75: 302-315.

40. Henriksson, N. Speed of slow component and duration in caloric nystagmus. ACTA Oto-laryngologica Supplement 125, 1956, 3-29.

41. Hinchcliffe, R. Nystagmus rate as an index of caloric test response. ACTA Oto-laryngologica, 1968, 65: 311-315.

42. ________. Validity of measures of caloric test response. ACTA Oto-laryngologica, 1967, 63: 69-73.

43. ________., and Voots, R. An electronystagmographic technique for the examination of vestibular function. Neurology, 1962, 12: 686-697.

44. Jongkees, L. The evaluation of the vestibular caloric test. In The Vestibular System and Its Diseases, edited by Robert J. Wolfson, Philadelphia, University of Pennsylvania Press, 1966.

45. ________. Value of the caloric test of the labyrinth. A.M.A. Archives of Otolaryngology, 1948, 48: 402-417.

46. Kirstein, L., and Preber, L. Directional preponderance of caloric nystagmus in patients with organic brain disease. ACTA Oto-laryngologica, 1954, 44: 265-273.

47. Koike, Y. An observation of the eye speed of nystagmus. ACTA Oto-laryngologica, 1959, 50: 377-390.

48. Lidvall, H. F. Mechanisms of motion sickness as reflected in the vertigo and nystagmus responses to repeated caloric stimuli. ACTA Oto-laryngologica, 1962, 55: 527-536.

49. ________. Vertigo and nystagmus responses to caloric stimuli repeated at short intervals. ACTA Oto-laryngologica, 1961, 53: 33-44.

50. ________. Vertigo and nystagmus responses to caloric stimuli repeated at short and long intervals. ACTA Oto-laryngologica, 1961, 53: 507-518.

51. Litton, W., and McCabe, B. Controllable variables in vestibulometry. A.M.A. Archives of Otolaryngology, 1967, 86: 111-114.

52. ________. Thermal Vestibulometry: Technique and Clinical Aspects. In Sensorineural Hearing Processes and Disorders, edited by A Bruce Graham. Boston: Little, Brown and Company, 1967.

53. McLay, K., Madigan, M., and Omerod, F. Anomolies in the recorded movements of the eye during optokinetic, rotary and caloric nystagmus. Annals of Otology, Rhinology and Laryngology, 1957, 66: 473-486.

54. Mahoney, J., Harlan, W., and Bickford, R. Visual and other factors influencing caloric nystagmus in normal subjects. A.M.A. Archives of Otolaryngology, 1957, 66: 46-53.

55. Maspetiol, R., and Keravec, J. Etude du coefficient de correlation entre les differentes mesures du nystagmus postcalorique. Annals Oto-laryngologica (Paris), 1962, 79: 909-913.

56. Mehra, Y. Electronystagmography: a study of caloric tests in normal subjects. Journal of Laryngology and Otology, 1964, 78: 520-529.

57. Milojevic, B. Electronystagmography. The Laryngoscope, 1965, 75: 243-258.

58. Miskolczy-Fodor, F. Electronystagmography: its perspective, advantages and limitations in routine vestibular testing. In The Vestibular System and Its Diseases, edited by Robert J. Wolfson, Philadelphia, University of Pennsylvania Press, 1966.

59. ________., and Arnold, G. The vestibulogram: a graphic record for the evaluation of vestibular nystagmus. Transactions of the American Academy of Ophthamology and Otolaryngology, 1960, 64: 168-181.

60. Mowrer, O. Influence of "excitement" on the deviation of postrotational nystagmus. A.M.A. Archives of Otolaryngology, 1934, 19: 46-54.

61. ________., Ruch, T., and Miller, N. Corneo-retinal potential difference as the basis of the galvanometric method of recording eye movements. American Journal of Physiology, 1936, 114: 423-428.

62. Naits, T. The effect of eye closure upon nystagmus. ACTA Oto-laryngologica Supplement 179, 1963, 72-85.

63. Pearcy, J., and Hayden, D. Reflexes from the gastrointestinal tract to the labyrinth. American Journal of Physiology, 1928, 87: 196-199.

64. Philipszoon, A. Electronystagmography in daily ENT practice. A.M.A. Archives of Otolaryngology, 1967, 86: 107-110.

65. Rashbass, C., and Russell, G. Action of a barbituate drug (amylobarbitone sodium) on the vestibulo-ocular reflex. Brain, 1961, 84: 329-335.

66. Rubin, W. Nystagmography. A.M.A. Archives of Otolaryngology, 1968, 87: 266-271.

67. Rundle, F., and McCabe, B. A clinical evaluation of caloric testing by the Fitzgerald-Hallpike-Cawthorne method. The Laryngoscope, 1961, 71: 1186-1195.

68. Schroeder, D. Some effects of alcohol on nystagmus and vertigo during caloric and optokinetic stimulation. Annals of Otology, Rhinology and Laryngology, 1972, 81: 218-228.

69. Sekitani, T., McCabe, B., and Ryer, J. Drug effects on the medial vestibular nucleus. A.M.A. Archives of Otolaryngology, 1971, 93: 581-589

70. Smiederkan. Cited in Hinchcliffe, R., and Voots, R. An electronystagmographic technique for the examination of vestibular function. Neurology, 1962, 12: 686-697.

71. Sokolovski, A. The influence of mental activity and visual fixation upon caloric-induced nystagmus in normal subjects. ACTA Oto-laryngologica, 1966, 61: 209-220.

72. Stahle, J. Electronystagmography in the caloric and rotary tests. ACTA Oto-laryngologica Supplement 137, 1958, 8-83.

73. ________. Electronystagmography: its value as a diagnostic tool. In The Vestibular System and Its Diseases, edited by Robert J. Wolfson, Philadelphia, University of Pennsylvania Press, 1966.

74. Steffen, T., Linthicum, F., and Churchill, D. Continuous thermal vestibulometry: a new technique of caloric examination. Annals of Otology, Rhinology and Laryngology, 1970, 79: 619-632.

75. Thomsen, K. The caloric test in a normal material. ACTA Oto-laryngologica Supplement 109, 1953, 189-196.

76. Torok, N. A new parameter of vestibular sensitivity. Annals of Otology, Rhinology and Laryngology, 1970, 79: 808-817.

77. ________. Nystagmus frequency versus slow phase velocity in rotary and caloric nystagmus. Annals of Otology, Rhinology and Laryngology, 1969, 78: 625-639.

78. ________. Significance of the frequency in caloric nystagmus. ACTA Oto-laryngologica, 1948, 36: 38-50.

79. ________. Some observations on culmination and directional preponderance of the poststimulatory nystagmus. The Laryngoscope, 1962, 72: 79-103.

80. ________. The culmination phenomenon and frequency pattern of thermic nystagmus. ACTA Oto-laryngologica, 1957, 48: 530-535.

81. ________., Guillemin, V., and Barnothy, J. Photoelectric nystagmography. Annals of Otology, Rhinology and Laryngology, 1951, 60: 917-926.

82. Vancil, M., Hemenway, W., Spindler, J., and Black, F. Suppression of alcohol induced nystagmus by Innovar. A.M.A. Archives of Otolaryngology, 1969, 90: 104-110.

83. Winer, B. *Statistical Principles in Experimental Design*. New York: McGraw-Hill Company, 1962.

A P P E N D I C E S

APPENDIX A

INDIVIDUAL SUBJECT DATA

TABLE 20

RAW SCORE DATA (1,2,3,4) FOR EACH SUBJECT
ON ALL CONDITIONS, TRIALS
AND RESPONSE PARAMETERS

30°C Right (1)

Subject Number	SSP T_1	SSP T_2	SSP T_3	TB T_1	TB T_2	TB T_3	TA T_1	TA T_2	TA T_3	CF T_1	CF T_2	CF T_3	DUR T_1	DUR T_2	DUR T_3
1	39	33	27	58	52	58	760	590	658	21	20	20	183	197	210
2	22	20	17	89	86	87	431	364	353	31	30	31	167	153	182
3	26	21	14	58	52	58	625	405	257	21	19	22	228	190	160
4	22	19	19	65	70	73	463	392	384	23	25	26	154	169	162
5	19	21	18	71	72	69	388	402	385	25	29	29	181	192	181
6	16	12	16	51	49	55	318	278	308	20	18	22	183	176	164
7	15	16	15	51	55	54	314	320	304	19	20	22	179	193	173
8	24	22	21	67	58	60	560	500	521	24	21	21	153	131	140
9	45	36	45	62	61	60	783	696	797	22	22	23	209	220	229
10	25	28	31	44	47	50	785	741	786	16	17	18	190	172	173
11	27	26	24	48	66	56	591	348	599	18	25	20	171	150	163
12	25	25	23	68	69	69	542	493	425	24	26	24	188	160	192
13	24	19	20	61	59	62	478	325	332	21	22	23	189	175	168
14	15	19	16	47	51	46	368	390	350	17	19	16	148	159	130
15	21	20	23	48	55	55	515	432	529	18	19	20	198	199	202
16	21	18	16	75	69	69	451	312	313	28	26	24	240	216	240

30°C Left (2)

Subject Number	SSP T_1	SSP T_2	SSP T_3	TB T_1	TB T_2	TB T_3	TA T_1	TA T_2	TA T_3	CF T_1	CF T_2	CF T_3	DUR T_1	DUR T_2	DUR T_3
1	26	23	22	51	43	46	710	443	507	18	17	17	206	187	185
2	23	24	24	109	108	112	447	476	489	38	38	38	193	161	161
3	28	25	20	69	66	67	650	482	310	26	24	24	226	242	213
4	16	14	16	73	79	68	380	309	376	26	28	24	170	144	140
5	17	13	15	66	64	61	294	236	299	25	25	22	172	191	160
6	25	23	20	64	62	64	512	486	445	22	21	24	209	209	198
7	14	16	12	49	51	58	250	301	228	18	20	20	162	176	186
8	25	20	23	55	53	50	496	419	472	20	24	22	178	162	150
9	33	33	30	60	54	67	783	652	557	21	20	23	177	164	180
10	20	17	18	47	46	62	346	309	298	20	17	23	146	136	128
11	23	16	14	40	47	57	429	331	317	14	20	20	168	150	173
12	35	35	24	83	78	80	653	605	423	29	27	29	186	175	170
13	26	21	21	55	53	60	550	369	370	20	20	23	240	243	214
14	17	16	18	49	47	52	382	350	391	19	18	19	155	142	163
15	16	19	17	45	54	51	387	413	399	16	20	19	216	188	233
16	16	16	16	69	65	64	350	350	309	24	22	23	197	210	180

SSP = Speed of the Slow Phase
TB = Total Beats
TA = Total Amplitude
CF = Culmination Frequency
DUR = Duration

T_1 = Test 1

T_2 = Test 2

T_3 = Test 3

TABLE 20--Continued

44°C Right (3)

Subject Number	SSP T_1	SSP T_2	SSP T_3	TB T_1	TB T_2	TB T_3	TA T_1	TA T_2	TA T_3	CF T_1	CF T_2	CF T_3	DUR T_1	DUR T_2	DUR T_3
1	18	23	21	53	45	48	462	471	429	18	17	17	150	133	152
2	18	14	17	101	98	101	345	306	346	35	34	37	201	210	175
3	16	15	12	59	66	59	376	317	214	22	24	21	243	251	214
4	20	20	26	75	81	83	593	640	743	27	28	29	145	139	143
5	20	22	15	73	81	66	329	381	198	26	30	25	165	180	150
6	17	15	20	61	49	56	341	312	388	22	18	21	200	210	209
7	15	17	16	52	59	44	352	321	285	21	24	20	177	164	165
8	24	22	22	58	53	60	512	391	420	22	20	23	180	151	172
9	24	30	31	67	57	70	730	709	715	25	21	24	186	176	165
10	27	27	25	51	60	56	552	549	421	18	21	21	141	138	133
11	33	33	30	57	65	54	687	567	596	21	24	21	133	163	152
12	30	32	26	86	82	78	681	703	450	30	28	27	180	168	196
13	20	22	24	58	70	69	424	389	406	20	25	25	270	251	220
14	18	13	20	55	50	60	348	315	372	19	16	18	165	150	179
15	19	14	19	43	43	50	436	309	440	15	15	18	185	198	179
16	14	17	16	69	68	64	301	347	314	25	24	24	194	180	164

44°C Left (4)

Subject Number	SSP T_1	SSP T_2	SSP T_3	TB T_1	TB T_2	TB T_3	TA T_1	TA T_2	TA T_3	CF T_1	CF T_2	CF T_3	DUR T_1	DUR T_2	DUR T_3
1	28	31	24	52	66	57	664	689	516	19	23	20	161	145	193
2	24	26	23	103	93	108	432	482	539	39	32	40	198	165	151
3	23	27	24	74	72	80	523	514	428	28	27	31	178	178	192
4	16	14	21	63	66	70	371	298	401	21	24	26	145	126	124
5	16	17	20	75	74	83	238	261	349	30	28	31	178	175	168
6	14	12	12	48	47	59	300	320	335	19	17	21	150	150	130
7	11	13	12	61	65	62	234	250	244	22	27	24	148	140	145
8	23	25	22	60	55	55	388	412	361	21	23	20	162	149	152
9	45	38	50	63	67	61	1014	715	992	22	24	21	226	198	231
10	25	29	38	43	50	54	769	816	1019	15	17	19	200	184	150
11	29	26	23	59	81	58	518	333	579	20	25	23	160	169	168
12	26	25	26	73	77	71	530	684	559	26	28	27	140	150	159
13	35	26	26	72	78	77	594	525	448	25	27	28	164	170	161
14	19	15	17	54	49	50	395	341	368	21	18	20	170	164	150
15	25	20	24	48	49	57	611	460	551	17	17	20	226	213	203
16	25	21	23	69	84	66	520	362	437	26	30	24	270	261	206

SSP = Speed of the Slow Phase
TB = Total Beats
TA = Total Amplitude
CF = Culmination Frequency
DUR = Duration

T_1 = Test 1

T_2 = Test 2

T_3 = Test 3

TABLE 21

DIFFERENCE SCORE DATA [(1+3)-(2+4)] FOR EACH SUBJECT ON EACH RESPONSE PARAMETER

Subject Number	Speed of the Slow Phase			Total Beats			Total Amplitude			Culmination Frequency			Duration		
	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
1	3	2	2	8	-12	3	-152	-71	64	2	-3	0	-34	-2	-16
2	-8	-16	-13	-22	-17	-32	-103	-288	-329	-11	-6	-10	-23	37	45
3	-9	-16	-18	-26	-20	-30	-172	-274	-267	-11	-8	-12	67	21	-31
4	10	11	8	4	6	18	305	425	350	3	1	5	-16	38	41
5	6	13	-2	3	15	-9	185	286	-65	-4	6	1	-4	6	3
6	-6	-8	4	0	-11	-12	-153	-216	-84	1	-2	5	24	27	45
7	5	4	7	-7	-2	-22	182	90	117	0	-3	-2	46	41	7
8	0	-1	-2	10	3	15	188	60	108	5	-6	2	-7	-29	10
9	-9	-5	-4	6	-3	2	-284	38	-37	4	-1	3	-8	34	-17
10	7	9	0	5	11	-10	222	165	-110	-1	4	-3	-15	-10	28
11	8	17	17	6	3	-5	331	251	299	5	4	-2	-24	-6	-26
12	-6	-3	-1	-2	-4	-4	40	-93	-107	-1	-1	-5	42	3	59
13	-17	-6	-3	-8	-2	-6	242	-180	-80	-4	0	-3	55	13	13
14	-3	1	-3	-1	5	4	-61	14	-37	-4	-1	-5	-12	3	-4
15	-1	-5	1	-2	-5	3	-47	-132	19	0	-3	-1	-59	-4	-55
16	-6	-2	-7	6	-12	3	-118	-53	-119	3	-2	1	-33	-75	18

T_1 = Test 1

T_2 = Test 2

T_3 = Test 3

APPENDIX B

INTRACLASS CORRELATION COEFFICIENTS DERIVED FROM ONE-WAY AND TWO-WAY ANALYSES OF VARIANCE

TABLE 22

INTRACLASS CORRELATION COEFFICIENTS DERIVED FROM ONE-WAY AND TWO-WAY ANALYSES OF VARIANCE PERFORMED ON THE ABSOLUTE SCORES

		Speed of Slow Phase	Total Beats	Total Amplitude	Culmination Frequency	Duration
30°C Right	2 Way R	*.84	.89	*.84	.82	.72
	1 Way R	.83	.89	.78	.82	.72
44°C Right	2 Way R	.80	.87	.82	.87	.83
	1 Way R	.80	.88	.81	.88	.82
30°C Left	2 Way R	*.84	.92	*.71	.89	.80
	1 Way R	.78	.92	.62	.89	.78
44°C Left	2 Way R	.84	.83	.83	.84	.78
	1 Way R	.84	.82	.82	.83	.76

* Denotes Trend Effect

TABLE 23

INTRACLASS CORRELATION COEFFICIENTS DERIVED FROM ONE-WAY AND TWO-WAY ANALYSES OF VARIANCE PERFORMED ON THE DIFFERENCE SCORES

	Speed of Slow Phase	Total Beats	Total Amplitude	Culmination Frequency	Duration
2 Way R	.75	.59	.72	.52	.26
1 Way R	.75	.58	.73	.53	.27

APPENDIX C

THE F VALUE ASSOCIATED WITH EACH RELIABILITY COEFFICIENT ON THE ABSOLUTE AND DIFFERENCE SCORE ANALYSES

TABLE 24

THE F VALUE ASSOCIATED WITH EACH RELIABILITY COEFFICIENT ON THE ABSOLUTE SCORE ANALYSIS

		Speed of Slow Phase	Total Beats	Total Amplitude	Culmination Frequency	Duration
30°CR	Reliability coefficient	.84[a]	.89[a]	.84[a]	.82[a]	.72[a]
	F Value	17.81	24.31	16.56	14.82	8.95
44°CR	Reliability coefficient	.80[a]	.87[a]	.82[a]	87[a]	83[a]
	F Value	12.72	21.43	15.19	21.11	16.16
30°CL	Reliability coefficient	84[a]	.92[a]	.71[a]	.89[a]	.80[a]
	F Value	16.24	38.62	8.45	25.32	12.94
44°CL	Reliability coefficient	.84[a]	.83[a]	.83[a]	.84	.78[a]
	F Value	16.68	13.27	15.93	14.82	11.73

[a]Significant at the .01 level.

2.70 - F value needed to be significant at the .01 level of significance.

TABLE 25

THE F VALUE ASSOCIATED WITH EACH RELIABILITY COEFFICIENT ON THE DIFFERENCE SCORE ANALYSIS

	Speed of Slow Phase	Total Beats	Total Amplitude	Culmination Frequency	Duration
Reliability coefficient	.75[a]	.59[a]	.72[a]	.52[a]	.26[b]
F Value	9.76	5.32	8.87	4.28	2.03

[a]Significant at .01 level.

[b]Significant at .05 level.

2.70 - F value needed to be significant at the .01 level of significance.

2.03 - F value needed to be significant at the .05 level of significance.