Reliability of a computer-based system for measuring visual performance skills

Graham B. Erickson, O.D., a Karl Citek, O.D., Ph.D., a Michelle Cove, O.D., b Jennifer Wilczek, O.D., c Carolyn Linster, a Brendon Bjarnason, a and Nathan Langemo a

a Pacific University College of Optometry, Forest Grove, Oregon; b Luxottica Group, Buffalo, New York; and c Parker Optometry, Prince George, British Columbia.

KEYWORDS
Visual performance; Sports vision; Visual acuity; Dynamic visual acuity; Contrast sensitivity; Accommodative facility; Stereopsis; Perception span; Eye-hand coordination; Visual-motor reaction time

Abstract
BACKGROUND: Athletes have demonstrated better visual abilities than nonathletes. A vision assessment for an athlete should include methods to evaluate the quality of visual performance skills in the most appropriate, accurate, and repeatable manner. This study determines the reliability of the visual performance measures assessed with a computer-based system, known as the Nike Sensory Station.

METHODS: One hundred twenty-five subjects (56 men, 69 women), age 18 to 30, completed Phase I of the study. Subjects attended 2 sessions, separated by at least 1 week, in which identical protocols were followed. Subjects completed the following assessments: Visual Clarity, Contrast Sensitivity, Depth Perception, Near–Far Quickness, Target Capture, Perception Span, Eye–Hand Coordination, Go/No Go, and Reaction Time. An additional 36 subjects (20 men, 16 women), age 22 to 35, completed Phase II of the study involving modifications to the equipment, instructions, and protocols from Phase I.

RESULTS: Results show no significant change in performance over time on assessments of Visual Clarity, Contrast Sensitivity, Depth Perception, Target Capture, Perception Span, and Reaction Time. Performance did improve over time for Near-Far Quickness, Eye-Hand Coordination, and Go/No Go.

CONCLUSIONS: The results of this study show that many of the Nike Sensory Station assessments show repeatability and no learning effect over time. The measures that did improve across sessions show an expected learning effect caused by the motor response characteristics being measured.

Optometry 2011;82:528-542

Researchers and clinicians have sought to discover the specific vision skills that correlate to success in sports. The vision and visual perceptual skills identified as important include static and dynamic visual acuities, contrast sensitivity, distance stereopsis, accommodative-vergence facility, span of perception, central eye-hand reaction and response speeds, and peripheral eye-hand response speed. 1,2 Two extensive review articles cite numerous studies to conclude that athletes have demonstrated better visual abilities than nonathletes, and that top athletes—those who are most successful—often have visual abilities that are superior to lower-level or less successful athletes. 3,4 Some aspects of these skills commonly are assessed as part of a routine vision examination, but many are not evaluated for various reasons. For some vision skills, there is little or no standardization of assessment techniques, and some instrumentation may be outdated, if available at all.

Disclosure: The authors thank Nike, Inc., for its generous support of this project and John R. Hayes, Ph.D., for assistance with statistical analyses. Nike provided the Nike Sensory Station and subject compensation to facilitate this study. Drs. Erickson and Citek frequently provide consultation services to Nike but have no financial relationship involving the products tested or discussed in this article.

* Corresponding author: Graham B. Erickson, O.D., Pacific University College of Optometry, 2043 College Way, Forest Grove, OR 97116.
E-mail: ericksog@pacificu.edu

©2011 American Optometric Association. All rights reserved.
doi:10.1016/j.optm.2011.01.012
To provide specialized vision care to an athlete, the eye care practitioner should identify the vision factors essential to successful performance of the sport tasks. The vision assessment should then include methods to evaluate the quality of those skills in the most appropriate, accurate, and repeatable manner. A significant body of research compares performance on various measures of visual function in athlete populations to guide the clinician in selecting appropriate visual measures, as described below. The battery of tests comprising the Nike Sensory Station (Nike, Inc., Beaverton, Oregon) attempts to address this clinical need.

Static visual acuity

Assessment of visual performance skills routinely begins with measurement of static visual acuity (SVA). Compromised SVA can negatively affect other areas of visual performance. Previous research has found mixed results regarding SVA in athlete populations; when SVA is assessed using chart systems with 20/20 as the best acuity measurable, there is no statistically significant difference in the visual ability of athletes compared with that of nonathletes. Even when a best acuity demand of 20/15 is presented, Laby et al. found that 81% of professional baseball players could achieve that level. Laby et al. subsequently modified their assessment method to achieve acuity demands down to 20/7.5, reporting overall mean SVAs of approximately 20/13 with several athletes attaining SVAs of 20/9.2 or better. However, several studies in which visual acuity was degraded with plus addition lenses did not find a detrimental effect of defocus. Review of the study protocols reveals that subjects were assessed on predictable, repetitive motor tasks. One of the goals of the current study is to determine which assessment tasks, if any, are subject to a simple motor learning effect with repeated measurement.

Dynamic visual acuity

Dynamic visual acuity (DVA) generally is defined as the ability of the visual system to resolve detail when there is relative movement between the target and the observer. Many sports involve extensive movement of an object (e.g., ball, puck), competitors, teammates, the athlete, or all simultaneously. Often at elite levels of sport, the velocity of movement between the athlete and the target is tremendously high; therefore, it is essential for athletes to be able to accurately perceive and identify critical target features during dynamic situations. Stine et al. in a review of the literature, found that athletes show superior DVA abilities compared with nonathletes and that elite athletes have better DVA than do amateur or nonelite athletes. These findings suggest that there is an important link between elite athletes and DVA ability. On the other hand, Ward and Williams reported no significant differences in performance on a DVA test between elite and subelite youth soccer players. However, their use of a predictable rotator device to measure this function may not have been environmentally appropriate to simulate the visual task demands of a large-field, dynamic sport, such as soccer. Although many researchers agree about the importance of DVA in sport, this visual skill is often not assessed in clinical practice because of limitations in commercial instruments available to measure it.

DVA decreases with increasing target velocity. To have a high DVA requires the ability to resolve targets at a higher velocity than average. Brown concluded that the relationship between DVA and target angular velocity is approximately linear for targets moving at velocities up to 90° per second. DVA is affected by both the target parameters (e.g., target luminance, target/observer velocity, and exposure time of target) and the physiologic abilities of the observer (e.g., resolving power of the eye, oculomotor abilities, peripheral awareness, and psychological abilities to interpret what is seen). Enhancement of either the target parameters or physiologic abilities of the subject can improve DVA abilities.

Contrast sensitivity

Contrast sensitivity measures the visual system’s ability to process spatial or temporal information about objects and their backgrounds under varying lighting conditions. Measuring an athlete’s contrast sensitivity is important because most sports involve interpreting visual information at contrast levels below what is measured with a typical visual acuity chart. Performance of athletes on contrast sensitivity testing is significantly better than that of nonathletes across all spatial frequencies evaluated.

Ginsburg showed that contrast sensitivity decreases as the velocity of a target increases and that, as target size decreases, contrast sensitivity demand increases. Ginsburg also showed that higher spatial frequencies will be affected before lower spatial frequencies by changes in target distance, movement direction, and low illumination. Contrast sensitivity also may be degraded in contact lens wearers if lens fit or water content are not optimal for the patient, resulting in corneal edema even when SVA seems acceptable. Finally, contrast sensitivity can increase or decrease after refractive surgery.

Many commercial systems are available to measure contrast sensitivity. Several use linear grating patterns that vary in spatial frequency, contrast level, and, possibly, orientation. Others use letters or numbers of varying contrast levels or size. Contrast sensitivity measurements usually involve determination of a threshold contrast level at specific spatial frequencies.

Stereopsis

Determining distance and spatial localization of an object is a necessity for athletes in many sports. These judgments can be made using monocular depth cues, but it is suggested that superior binocular depth perception (stereopsis) can be advantageous for the athlete. The research
results on assessment of stereopsis have produced mixed results: some studies found no difference between athlete and nonathlete populations, whereas other studies found better performance in athletes. It has been suggested that the difference in findings may be because of the lack of standardized testing procedures, the use of simulated depth targets, and the limitations of the instruments to measure threshold stereocuity. Previous studies used near stereo tests or testing at far with the American Optical Vectorscope Projector Slide (Southbridge, Massachusetts) or a Howard-Dolman apparatus. As many sports are dynamic, athletes possibly would perform better with a dynamic stereopsis assessment because static testing may not reveal any difference between athletes and nonathletes.

Accommodative-vergence facility

Competitive sports rarely occur at one distance. Most athletes need to look between far, intermediate, and near distances quickly, requiring rapid accommodative-vergence responses. This visual skill can be assessed with the Haynes distance rock test. A study using this test presented normative data for a population of elite athletes but did not compare performance with that of nonathletes.

Perception span

Perception span, or central visual recognition accuracy, uses tachistoscopic presentation to measure the speed and span of recognition. Several studies have investigated speed of recognition abilities in athletes. Most studies have found that experienced athletes can evaluate sport-relevant information more rapidly than inexperienced observers; sport situations studied include cricket, volleyball, tennis, and “fast ball” sports. Other studies have investigated both speed and span of recognition by evaluating the ability to recall a sequence of numbers presented tachistoscopically for 1/50 of a second and found no difference in athletes compared with nonathletes. However, Melcher and Lund did find significant differences in performance both for span and speed of recognition, which were also present when distraction factors were added to the task to simulate competition conditions. In consideration of these differences in research results, the authors conclude that it is the use of numerical stimuli confounding the assessment of speed of recognition in athletes; the use of target parameters that more closely simulate the visual information processed in sport situations can yield better discrimination of perception span abilities that correlate with sports performance.

Central eye-hand reaction and response time

Visual-motor reaction and response speeds are critical to performance. Reaction time is the elapsed time between the onset of a visual stimulus and the initiation of a motor response. Response time is the total time required by the visual system to process a stimulus plus the time needed to complete the motor response. Several studies report that athletes in various sports have faster reaction times compared with nonathletes and that reaction time is a discriminator between expertise levels. However, other studies have not found this difference. A gender bias also has been reported, with men achieving faster times than women on average. A previously available instrument for measuring this performance skill was the Reaction Plus timer (WR Medical Electronics, Stillwater, Minnesota). Following brief training regimens, eye-hand reaction time could be improved, making this a potentially valuable assessment and goal for the athlete.

Peripheral eye-hand response

Overall ability to process and respond to visual stimuli strongly enhances an athlete’s eye-hand coordination. The typical instrumentation used for evaluating eye-hand coordination has been a 2-dimensional panel with an array of lights mounted on a wall, such as the Wayne Saccadic Fixator (Wayne Engineering, Skokie, Illinois). The athlete is required to press a randomly lit button as rapidly as possible with one hand, then another button is lit in a random position on the instrument, and the reaction time reflex cycle is repeated for the selected test time period. The panel is set at the athlete’s arm length and is larger than the central visual field, thus assessing a peripheral eye-hand response. The instrument typically is programmed to test in 2 primary modes: visual proaction time, a self-paced mode for a set time period in which each light stays lit until the button is pressed, then the next random light is lit; and visual reaction time, an instrument-paced stimulus presentation in which each light stays lit for a preset amount of time (typically 0.75 seconds) before automatically switching to another light regardless of whether the button is pressed. One study found better visual proaction times in youth athletes than nonathletes, whereas another study found no such difference between adult athletes and nonathletes. Visual reaction time has been compared in athletes and nonathletes in only 1 study, with athletes performing better than nonathletes. The level of ambient room lighting affects test outcomes: performance improves significantly as room illumination is decreased.

The Nike Sensory Station is designed to test vision skills that previously have been identified as important for sports, including static and dynamic visual acuities, contrast sensitivity, distance stereopsis, accommodative-vergence facility, span of perception, central eye-hand reaction and response speeds, and peripheral eye-hand response speed. It is designed to provide a visual performance profile that graphically represents the athlete’s visual strengths and weaknesses. The purpose of this study is to determine the test-retest reliability of the visual performance measures assessed with the Nike Sensory Station.
Methods

Phase I

Subjects. An Institutional Review Board proposal for the use of human subjects in research was submitted and approved. One hundred thirty-four subjects were recruited from the Pacific University student body and surrounding community to participate in Phase I of this study. All subjects signed an informed consent form at the time of the initial screening, and all were required to pass a vision screening with a minimum visual acuity (VA) of 20/50 in each eye. One hundred twenty-five subjects (56 men, 69 women), age 18 to 30, completed the study. Subjects who completed the study were compensated with passes to the Nike employee store located in Beaverton, Oregon.

Test protocol. The Nike Sensory Station consists of a single computer controlling 2 high-resolution liquid crystal display monitors (both 0.28 mm dot pitch): one 22-inch diagonal display and one 42-inch diagonal touch-sensitive display. A handheld Apple iPod touch® (Apple Corporation, Cupertino, California), connected via wireless input to the computer, is used in several assessments as described below. Custom software controls the displays, input acquisition, and test procedures based on subject responses. Prerecorded instructions are automatically played at the start of each assessment (see Appendix); if a subject has questions concerning a test procedure, the prerecorded instructions are repeated with no additional coaching from the researcher.

Subjects attended 2 sessions in which identical protocols were followed. At the end of session 1, subjects were asked to return in 1 to 2 weeks to repeat the assessments. This time interval was established in an attempt to minimize the potential learning effects for the assessments. Most subjects returned within 10 to 14 days, with the longest duration between tests being 30 days. Monocular and binocular VAs were measured at the start of each session under normal room illumination (~250 lux) with a printed logMAR chart at 6 m. Subjects then were assessed on each section of the Sensory Station in order as described below. Ambient illumination at the 22-inch display was ~80 lux; at the 42-inch display, ~140 lux; and at the iPod touch, ~100 lux.

Subjects completed the Sensory Station assessments adhering to a standardized protocol: Visual Clarity, Contrast Sensitivity, Depth Perception, Near-Far Quickness, and Target Capture at 16 feet (4.9 m) from the respective displays. Subjects were then moved to within arm’s length of the 42-inch display and performed the Perception Span, Eye–Hand Coordination, Go/No Go, and Reaction Time assessments. Each session was completed in about 20 minutes. For clarification, when referring to a specific assessment on the Nike Sensory Station, the proper name the manufacturer has given the assessment will be used and capitalized.

Nike Sensory Station assessments. Visual Clarity (visual acuity). SVA was measured at 16 feet (4.9 m) with the 22-inch display. Black Landolt rings, with gaps at the top, bottom, left, and right, were presented on a white background in random order at preset acuity demands. Subjects were instructed to swipe the screen of the iPod touch in the direction of the gap in the ring as soon as it was identified. Animation examples were shown, followed by 3 practice trials. If the gap direction was not easily discriminated, the subject was encouraged to guess, per the instructions. Final threshold acuity was measured between the demands of 20/8 and 20/99 using a staircase reversal algorithm. The algorithm begins with a large (20/50 equivalent) stimulus, decreasing in size until the subject does not correctly identify the stimulus. When this occurs, the stimulus increases in size until it is identified correctly. This procedure continues until several reversal points are achieved; the exact number of reversal points for the algorithm is proprietary and not available for publication. SVA is based on the number of correct responses, with consideration for guessing. An occluder was held by the examiner in front of the subject’s nontested eye during the monocular measurements. The sequence of the SVA assessment was always right eye (O.D.), left eye (O.S.), and both eyes (OU).

Contrast Sensitivity. Four black circles, each of which subtended 0.82°, were presented on a light gray background in a diamond configuration covering 2.35° on the 22-inch display at 16 feet (4.9 m) (see Figure 1). One circle at random contained a pattern of concentric rings that varied sinusoidally in brightness from the center to the edge. Subjects were instructed to swipe the screen of the iPod touch in the direction of the circle with the pattern. Animation examples were shown, followed by 3 practice trials. If the patterned circle was not easily discriminated, the subject was encouraged to guess, per the instructions. Contrast sensitivity was measured binocularly at 2 spatial frequencies, 6 and 18 cycles per degree (cpd), using a staircase reversal algorithm similar to that described previously. Final threshold contrast sensitivity was measured between 10% and 1.0%.

Figure 1 Concentric ring target for Contrast Sensitivity assessment.
(1.0 to 2.0 log units) contrast at 6 cpd, and between 32% and 2.5% (0.5 to 1.6 log units) contrast at 18 cpd.

Depth Perception (stereopsis). The subject wore a pair of liquid crystal goggles (NVIDIA® 3D Vision™, Santa Clara, California), connected via wireless link to the computer, while viewing the 22-inch display at 16 feet (4.9 m). The liquid crystal shutter system created simulated depth in 1 of 4 black rings presented on a white background, such that the ring should appear to float in front of the screen (see Figure 2). The sizes and arrangement of the rings were identical to those of the circles used in Contrast Sensitivity. The width of the lines defining each ring was 12 mm, subtending 0.14°. Subjects were instructed to swipe the screen of the iPod touch in the direction of the floating ring and were encouraged to respond as quickly as possible. Animation examples were shown, followed by 3 practice trials. If the floating ring was not easily discriminated, the subject was encouraged to guess, per the instructions. Threshold stereopsis was measured between 237 and 12 arc seconds using a staircase reversal algorithm similar to that described previously. In addition, response time for the first 2 stimulus presentations at the subject’s threshold was recorded, and an average response time for the testing was calculated.

Near-Far Quickness (accommodative-vergence facility). The subject remained at 16 feet (4.9 m) aligned with the 22-inch display. The subject also held the iPod touch at 16 inches (40 cm) from the eyes, with the top edge positioned just below the bottom of the far screen (see Figure 3). Positioning and instructions were presented with an animation example; if needed, the examiner helped the subject with the positioning adjustment. In alternating style, a 20/80-equivalent black Landolt ring was presented in a box on the handheld screen, and a black Landolt ring 0.1 log unit above the threshold determined with the Visual Clarity assessment was presented on the far screen. The subject was instructed to swipe the screen of the iPod touch in the perceived direction of the gap in the ring presented on each display; incorrect responses would not change the target presentation. The assessment began with 3 practice trials. The first Landolt ring was always presented on the far screen. After the correct response was recorded, the Landolt ring appeared on the handheld screen. The subject then continually switched focus between far and near for 30 seconds, trying to correctly identify as many rings as possible. The number of correct responses determined the score.

Target Capture (DVA). The subject now was aligned with the 42-inch display at 16 feet (4.9 m) and instructed to fixate a central white dot until a yellow-green Landolt ring (dominant wavelength about 555 nm at maximum saturation possible on the display) appeared briefly in 1 of the 4 corners of the screen (see Figure 4). The size of the Landolt ring was 0.1 log unit above the threshold determined with the Visual Clarity assessment, and the angular distance along the diagonal from the fixation dot to the center of the Landolt ring was approximately 6.1°. Because of the reduction in VA away from the fovea, individuals with VA of 20/50 or better would need to saccade from the fixation dot to the Landolt ring to correctly discriminate the direction of the gap. This is a method of assessing DVA. As before, subjects indicated the perceived direction of the gap by swiping the screen of the iPod touch. Animation examples were shown followed by 3 practice trials. If the subject could not see the orientation of the gap in the ring, guessing...
was encouraged. The duration of the Landolt ring presenta-
tion started at 500 milliseconds and was progressively
shortened after correct responses. The threshold stimulus
exposure duration was determined using a staircase reversal
algorithm.

Perception Span. The subject was positioned within arm’s-
length of the 42-inch touch-sensitive display, with the
center of the screen at about eye level. Automated instruc-
tions directed the subject to focus on a shrinking white dot
in the center of a grid pattern composed of up to 30 circles
(see Figure 5). When the dot disappeared, a pattern of
yellow-green dots (same color parameters as above) flashed
simultaneously for 100 milliseconds within the grid. The
subject then touched the screen to recreate the pattern of
dots. If the subject achieved a passing score (≥ 75% cor-
rect), the grid pattern increased in size with an increasing
number of dots. The first 2 levels had 6 circles in the grid
pattern with 2 and 3 dots, the next 5 levels had 18 circles
with 3 to 7 dots, and the last 4 levels had 30 circles with
7 to 10 dots. Each circle was 19 mm in diameter, and the
largest grid pattern was 18 cm in diameter. The grids and
dot patterns were preset to maintain standardization. The
dot patterns at each level were pseudorandomized to main-
tain equivalent spatial distribution of the dots for each pre-
sentation and to eliminate “clustering” of dots and easily
recognizable patterns or shapes. Animation examples
were shown, followed by 2 practice trials. The overall score
for this assessment was based on the cumulative number of
correct responses; missed responses and extra guesses were
subtracted from the cumulative score. If the subject did not
achieve a passing score on a level, that level was repeated.
If the subject again failed to pass that level, the assessment
was terminated. If the subject achieved a passing score on
the second attempt, only the higher score was used for
the overall score, and testing continued. The maximum
score possible on this assessment was 64.

Eye-Hand Coordination (peripheral eye-hand response).
For this assessment, subjects held their arms at shoulder
height within easy reach of a grid of circles presented on
the 42-inch touch-sensitive display (see Figure 6). The grid
consisted of 8 columns (68.6 cm) and 6 rows (44.5 cm) of
equally spaced circles, with each circle 48 mm in diameter.
During the assessment, a yellow-green dot (same color pa-
rameters as above) appeared within 1 circle of the grid. Au-
tomated instructions directed the subject to touch the dot as
quickly as possible using either hand. As soon as the dot
was touched, a subsequent dot would be presented. A se-
quence of 96 dots was pseudorandomized to maintain
equivalent spatial distribution within each presentation
and to eliminate “clustering” of dots and easily recog-
nizable patterns. Animation examples were shown, followed
by 1 full practice trial. The score was the total time to touch
all 96 dots.

Go/No Go. The positioning of the subject and the grid
pattern on the display for this assessment were identical to
those for Eye-Hand Coordination. However, the dot stim-
ulus could be either yellow-green (same color parameters
as above) or red (dominant wavelength about 620 nm at
maximum saturation possible on the display). Although
these colors could be confused by some color-deficient
individuals, the difference in apparent brightnesses of the
dots is sufficient to allow easy discrimination. If the dot was
yellow-green, the subject was directed to touch it as before.
But if the dot was red, the subject was instructed not to
touch it. Both the red and yellow-green dots appeared at
random locations for only 450 milliseconds, with no time
gap between dot presentations. If a yellow-green dot was
not touched within this time, no point was awarded for that
presentation; if a red dot was touched, a point was
subtracted from the overall score. Again, subjects were
encouraged to touch as many yellow-green dots as possible.
Automated instructions and animation examples were
shown, but there was no practice trial for this assessment.
Ninety-six total dots (64 yellow-green, 32 red) were
presented in a pseudorandomized sequence to maintain
equivalent spatial distribution within each presentation and

Figure 5 Thirty-circle grid for Perception Span.

Figure 6 Grid for Eye-Hand Coordination and Go/No Go.
to eliminate “clustering” of dots and easily recognizable patterns. The overall score was the cumulative number of yellow-green dots touched minus any red dots touched. **Reaction Time** (central eye-hand reaction and response time). For the final assessment, subjects remained at arm’s length from the 42-inch touch-sensitive display. Two annular patterns appeared on the screen with centers 30.5 cm apart; each annulus consisted of 2 concentric circles, 11.4 cm and 3.2 cm in diameter (see Figure 7). Automated instructions directed the subject to place the fingertips of the dominant hand on the inner circle of the annulus on that side of the screen, with no portion of the hand extending across the boundary line marked on the screen. If the hand was aligned correctly, this control annulus would change color to yellow-green (same color parameters as above). The subject then was instructed to center the body in front of the opposite (test) annulus and focus attention on the center of that annulus. After a randomized delay of 2, 3, or 4 seconds, the test annulus turned yellow-green, and the subject moved the hand to touch its inner circle as quickly as possible. Animation examples were shown, followed by 2 practice trials. Five trials were conducted per subject to calculate average reaction and response times. Reaction Time was measured as the elapsed time between onset of the test annulus and release of the control annulus. Response time was measured as the elapsed time between onset and touching of the test annulus. After 5 trials, the computer calculated the averages and standard deviations for the reaction and response times. If any single measure differed from the mean by more than 2 standard deviations in either direction, another trial was conducted to replace the outlying measure for that trial. The software was programmed such that no more than 2 trials could be repeated for any subject.

**Phase II**

Subsequent to subject feedback and analysis of Phase I data, modifications were made to the equipment, instructions, and protocols in an effort to improve the reliability of the assessments. Ambient illumination was identical to that in Phase I. Likewise, data collection was conducted in the same manner as in Phase I.

**Subjects.** Thirty-six subjects (20 men, 16 women), age 22 to 35, were recruited from the Pacific University student body and surrounding community to participate in Phase II of this study; none of these subjects participated in Phase I. As in Phase I, all subjects signed an Informed Consent Form at the time of the initial screening, and all were required to pass a vision screening with a minimum VA of 20/50 in each eye.

Subjects attended 2 sessions in which identical protocols were followed. At the end of session 1, subjects were asked to return within about 1 week to repeat the assessments, with the shortest interval being 3 days and the longest being 10 days. Subjects were compensated with passes to the Nike employee store located in Beaverton, Oregon.

**Test protocol modifications.** During Phase I, researcher observation and subject feedback consistently indicated that multiple swipes on the iPod touch occasionally were necessary to register a response. Although this may have been a minor annoyance and of no consequence for the outcome on most assessments, it could have impacted the results on Depth Perception response time and Near-Far Quickness. Thus, for Phase II the iPod touch had a plastic sleeve placed on it to improve the subject’s accuracy when swiping a response.

Because of observed ceiling effects, Contrast Sensitivity upper limits were increased to 2.4 log units (0.4% contrast) for 6 cpd and 2.0 log units (1.0%) contrast for 18 cpd. The instructions for Near-Far Quickness and Target Capture were improved to minimize confusion, and Go/No Go instructions reflected the addition of a practice trial (see Appendix). The practice trial protocols were altered for Near-Far Quickness (increased the practice trials to 6) and Eye-Hand Coordination (reduced the length of the practice trial by one half: 48 targets instead of 96). The “touch zones” for Reaction Time were increased to include any portion of the control and test annuli. Finally, programming changes were made to the protocols of some of the assessments. Analysis of individual subject data sets found a minor error in the staircase reversal algorithm used for scoring Visual Clarity, Contrast Sensitivity, Depth Perception, Target Capture, and Perception Span. The software programming for the staircase reversal algorithms was not consistently applied as intended, necessitating this second phase of data collection to assess repeatability. A separate software error in Perception Span resulted in an unexpected bimodal distribution of the scores during Phase I.

**Statistical analyses**

Results were analyzed primarily using repeated measures analysis of variance (ANOVA) and Student’s-paired t tests, as appropriate. Difference-versus-means plots and 95% limits of agreement (LOA) are reported to demonstrate test-retest reliability. Subsequent to the changes to the instrumentation and protocol described above, results for
Contrast Sensitivity, Depth Perception, Target Capture, Perception Span, and Reaction Time are reported only for Phase II data. Changes to the remaining assessments did not result in significantly different distributions or variances for data measured in Phase II with respect to Phase I. Therefore, results for Visual Clarity, Near-Far Quickness, Eye-Hand Coordination, and Go/No Go are pooled across phases and reported together for all subjects.

**Results**

**Visual Clarity**

Repeated-measures ANOVA found significant differences in binocular versus monocular acuity assessment, \((F_{2,320} = 31.93, P < 0.001)\). However, there are no significant differences between sessions 1 and 2, \((F_{1,160} = 0.32, P = 0.57)\), and no interaction effect of session and eye used, \((F_{2,320} = 0.50, P = 0.61)\). On average, binocular acuity is better than either monocular acuity: mean (SD) [Snellen equivalent] OU –0.205 (0.146) [20/12.5] versus O.D. –0.117 (0.167) [20/15.3] and O.S. –0.132 (0.171) [20/14.8].

Figure 8 plots the difference in binocular acuities between sessions versus mean binocular acuity for all subjects. Because multiple results have the same difference-mean combination, some of the symbols on the figure are “jiggled” in either or both directions to show the actual number of datapoints. Monocular difference-mean plots are not shown for the sake of brevity but are similar to those of the binocular results. The average differences are OU 0.006, O.D. –0.005, and O.S. 0.016. The average differences are not significantly different from zero, OU \((t_{160} = 0.40, P = 0.69)\), O.D. \((t_{160} = 0.32, P = 0.75)\), and O.S. \((t_{160} = 1.02, P = 0.31)\). However, the slopes of the linear regressions are significantly different from zero, OU 0.328, \((t_{160} = 20.78, P < 0.001)\); O.D. 0.159, \((t_{160} = 9.51, P < 0.001)\); and O.S. –0.152, \((t_{160} = 9.43, P < 0.001)\).

![Figure 8](image)

**Figure 8** Difference-versus-mean plot of binocular (OU) acuity for Visual Clarity. Average difference (solid black line), 95% limits of agreement (dashed lines), linear regression (solid red line). Some of the symbols are “jiggled” in either or both directions to show the actual number of datapoints.

**Contrast Sensitivity**

Repeated measures ANOVA found a significant difference in sensitivities between spatial frequencies, \((F_{1,35} = 298.08, P < 0.001)\). However, there are no significant differences in sessions, \((F_{1,35} = 1.11, P = 0.30)\), nor the interaction effect of spatial frequency and session, \((F_{1,35} = 0.21, P = 0.65)\).

Figure 9 plots the difference in log sensitivities between sessions versus mean log sensitivity for both spatial frequencies assessed. Because multiple results have the same difference-mean combination, some of the symbols on each figure are “jiggled” in either or both directions to show the actual number of datapoints. The average difference is –0.02 for 6 cpd and –0.04 for 18 cpd. Neither of the average differences are significantly different from zero, \((t_{35} = 0.70, P = 0.49)\) for 6 cpd and \((t_{35} = 0.96, P = 0.35)\) for 18 cpd. However, the slopes of the linear regressions, 0.240 and 0.159, are significantly different from zero \((t_{35} = 7.59, P < 0.001)\) and \((t_{35} = 3.42, P = 0.002)\), respectively.

![Figure 9](image)

**Figure 9** Difference-versus-mean plot of log Contrast Sensitivity for 6 cpd: data (blue +’s), average difference (solid black line), 95% limits of agreement (solid double lines), linear regression (solid red line); and for 18 cpd: data (blue squares), average difference (short-dashed black line), 95% limits of agreement (long-dashed black lines), linear regression (dashed red line). Some of the symbols are “jiggled” in either or both directions to show the actual number of datapoints.

**Depth Perception**

The Student’s \(t\)-test shows no significant difference between sessions for either threshold \((t_{35} = 0.83, P = 0.413)\), or mean response time \((t_{35} = 1.32, P = 0.20)\). Figure 10A plots the difference in thresholds between sessions versus mean threshold. Because multiple results have the same difference-mean combination, some of the symbols on the figure are “jiggled” in either or both directions to show the actual number of datapoints. The average difference (LOA) is \(-5.58 (-84.89, 79.31)\) arc seconds. The average difference is not significantly different from zero (same statistic as above). The slope of the linear regression, –0.489, is not significantly different from zero \((t_{35} = 0.07, P = 0.94)\).
Figure 10B plots the difference in mean response times between sessions versus average mean response time. The average difference (LOA) is $-159.82 \, (-1584.75, 1424.93)$ milliseconds. The average difference is not significantly different from zero (same statistic as above). The slope of the linear regression, $-0.275$, is not significantly different from zero, ($t_{35} = 0.002, P = 1$).

**Near-Far Quickness**

The score for 1 subject was not recorded by the software in Phase I, session 1; therefore, that subject’s score for session 2 also is not included for analysis. There is an overall significant improvement in score from session 1, mean (SD), 23.1 (5.30), to session 2, 25.7 (5.74), ($t_{159} = 6.20, P < 0.001$). Figure 11 plots the difference in correct responses between sessions versus mean correct response. The average difference (LOA) is $2.68 \, (-8.04, 13.40)$. The average difference is significantly different from zero (same statistic as above). The slope of the linear regression, 0.104, is not significantly different from zero, ($t_{159} = 0.24, P = 0.81$).

**Target Capture**

The Student’s $t$-test found no significant difference in threshold between sessions, ($t_{35} = 0.70, P = 0.49$). Figure 12 plots the difference in thresholds between sessions versus mean threshold. The average difference (LOA) is $18.75 \, (-295.63, 333.31)$. The average difference is not significantly different from zero (same statistic as above). The slope of the linear regression, $-0.545$, is not significantly different from zero, ($t_{35} = 0.02, P = 0.98$).

**Perception Span**

The Student’s $t$-test found no significant difference in scores between sessions, ($t_{35} = 1.68, P = 0.10$).

**Eye-Hand Coordination**

The score for 1 subject was not recorded by the software in Phase I, session 1; therefore, that subject’s score for session 2 also is not included for analysis. There is an overall significant improvement (reduction) in total time from session 1, mean (SD) 49.7 (4.66), to session 2, 47.0 (4.58), ($t_{159} = 11.76, P < 0.001$). Figure 14 plots the difference in total times between sessions versus mean total time. The average difference (LOA) is $-2.70 \, (-8.04, 2.99)$ seconds. The average difference is significantly different from zero (same statistic as above). The slope of the linear regression, $0.020$, is not significantly different from zero, ($t_{159} = 0, P = 1$).

Figure 10  A, Difference-versus-mean plot of Depth Perception thresholds, in arc seconds. Average difference (solid black line), 95% limits of agreement (dashed lines), linear regression (solid red line). Some of the symbols are “jiggled” in either or both directions to show the actual number of datapoints. B, Difference-versus-mean plot of Depth Perception mean response times, in milliseconds. Average difference (solid black line), 95% limits of agreement (dashed lines), linear regression (solid red line).

Figure 11  Difference-versus-mean plot of Near-Far Quickness correct responses. Average difference (solid black line), 95% limits of agreement (dashed lines), linear regression (solid red line).
Go/No Go

There is an overall significant improvement in score from session 1, mean (SD) 26.6 (12.85), to session 2, 34.1 (11.27), \((t_{160} = 9.90, P < 0.001)\). Figure 15 plots the difference in scores between sessions versus mean score. The average difference (LOA) is 7.58 (211.46, 26.61). The average difference is significantly different from zero (same statistic as above). The slope of the linear regression, \(-0.155\), is not significantly different from zero, \((t_{160} = 0.20, P = 0.84)\).

Reaction Time

Response time for 1 subject in session 1 was longer than 1.1 seconds, more than 5 standard deviations greater than the mean response time. Consequently, all data for this subject were removed for analysis.

For Reaction Time, the Student’s \(t\)-test found no significant difference between sessions, \((t_{34} = 1.11, P = 0.27)\). Figure 16A plots the difference in reaction times between sessions versus mean reaction time. The average difference (LOA) is \(-4.69 (-53.69, 44.30)\). The average difference is not significantly different from zero (same statistic as above). The slope of the linear regression, 0.302, is not significantly different from zero, \((t_{34} = 0.07, P = 0.94)\).

For response time, the Student’s \(t\)-test found no significant difference between sessions, \((t_{34} = 0.52, P = 0.61)\). Figure 16B plots the difference in response times between sessions versus mean response time. The average difference (LOA) is 4.47 (295.15, 104.09). The average difference is not significantly different from zero (same statistic as above). The slope of the linear regression, 0.353, is not significantly different from zero, \((t_{34} = 0.04, P = 0.97)\).

Discussion

The purpose of this study is to determine the test-retest reliability of the visual performance measures assessed with the Nike Sensory Station. To establish repeatability, the assessments were conducted over 2 sessions separated by about 1 week. Results show no significant change in performance over time on assessments of Visual Clarity, Contrast Sensitivity, Depth Perception, Target Capture, Perception Span, and Reaction Time. This demonstrates no learning effect over time and establishes repeatability of these measures.
Performance did improve from session 1 to session 2 for Near-Far Quickness, Eye-Hand Coordination, and Go/No Go. These measures that improved across sessions show an expected learning effect caused by the motor response characteristics being measured.

As expected, binocular measures on Visual Clarity are about 0.08 log unit, or about 4 letters, better than monocular measures. Because the sequence of this assessment was always O.D., O.S., and OU, the results could have been influenced by a practice effect. Nonetheless, in absolute terms, our result differences may not be of practical significance. Interestingly, performance also varied slightly across subjects between sessions for Visual Clarity and Contrast Sensitivity based on mean threshold levels, as evidenced by the shallow but significant linear regression slopes. For Visual Clarity, binocular acuity decreases by about 0.03 log unit, or about 1.5 letters, for each 1-line decrease in mean binocular acuity. Similarly, right eye acuity decreases by about 0.016 log unit, or less than 1 letter, for each 1-line decrease in mean O.D. acuity, but left eye acuity increases by about 0.015 log unit, or less than 1 letter, for each 1-line decrease in mean O.S. acuity. For Contrast Sensitivity, threshold improves for either spatial frequency by about 0.02 log unit for each 0.1-log unit improvement in mean contrast sensitivity. Presumably, these effects result from an increased likelihood of response variability when initial measurements are less than maximum values. However, additional study investigating these phenomena, as well as validation against similar established instruments, is warranted.

In addition, for Contrast Sensitivity results, the ceiling effects evident in Phase I data are eliminated for 18 cpd measures but remain for 6 cpd measures in Phase II. Contrast of 2.4 log units (0.4%) is the minimum capable of being displayed on the monitor used, thus limiting the ability to assess true threshold at 6 cpd. This ceiling effect may mask performance differences between testing sessions. Nonetheless, contrast sensitivity beyond 2.0 log units at any spatial frequency may be only of academic interest, because the application of such high sensitivity has yet to be established.

Addition of the plastic sleeve on the iPod touch reduced the occurrence of the necessity for multiple swipes, as observed and reported by the individual researchers. This is one possible reason for the difference in Depth Perception results between Phase I and Phase II. Depth Perception results also show an apparent ceiling effect at the resolution limit of the instrument (12 arc seconds). This is surprisingly good stereoacuity, especially for a far measure. The results suggest that a study to determine validity of the assessment is needed.

For Near-Far Quickness, on average, subjects completed about 1 additional cycle of target presentations during session 2, which equates to about a 10% improvement. The iPod touch multiple-swipe problem did not have a significant effect on test performance, as evidenced by the similar distributions and variances of the data for the 2 study phases. For Eye-Hand Coordination, the reduction in practice trial time instituted in Phase II did not significantly decrease the motor learning effect. For Go/No Go, on average, scores improved by about 28% between sessions. However, the addition of a practice trial in Phase II did not result in improved test performance. Further study is needed to determine if these improvements measured on repeated testing are clinically relevant.
A factor that may have affected the results is screen glare from the glossy glass surface of the 42-inch display. Because of the reflective nature of the display surface, direct light sources facing the screen create reflections that can be distracting to the user. Measures were taken to minimize the impact of avoidable sources of reflected glare from varying light conditions during the data collection: the windows in the research room were draped with heavy black cloth, and the overhead lighting was adjusted to minimize reflectance on the monitors. The remaining unavoidable reflections on the display primarily arose from individual subject appearance and the clothing they may have been wearing (see Figure 17). These reflections could have affected performance on Target Capture, Perception Span, Eye-Hand Coordination, Go/No Go, and Reaction Time. However, because we did not record or control what subjects were wearing, we have no method of determining whether this actually was a significant factor in subject performance.

This study was limited to the assessments on the Nike Sensory Station. Although some of the assessments are similar to those made with other previously or currently available instruments, the measurements made in this study are not intended to be normative. Likewise, the subjects in this study were not assessed with any other instrumentation. Therefore, a direct comparison of the current study data with results from other instruments is not appropriate. Future studies could conduct such comparisons to determine the validity of the individual assessments, although standardized and validated assessment tools are not available for all of the assessment areas measured with the Nike Sensory Station. It also remains to be determined whether these assessments measure visual performance skills that reflect the requisite skills for athletes or differentiate athletes by skill level. Future studies also could investigate athlete populations as subjects and develop normative performance data for specific ages, skill levels, and sports applications.

Conclusions

The results of this study show that many of the Nike Sensory Station assessments demonstrate repeatability and no learning effect over time. The measures that did improve across sessions demonstrate an expected learning effect. However, because we did not record or control what subjects were wearing, we have no method of determining whether this actually was a significant factor in subject performance.

References


Appendix

Nike Sensory Station Instruction Set (Phase I with strikethrough, Phase II with underline)

Visual Clarity

Let’s begin with visual clarity, how clearly you see a stationary object. Stand facing the screen with your toes on the line. For this assessment you will need to focus on the 4 gray circles. One of these circles contains a pattern comprised of light and dark shaded rings while the other 3 are solid gray. Swipe the screen as soon as you identify the correct circle. If you cannot see the gap, take your best guess. Now let’s begin with a practice run. If you have any questions now is a good time to ask your trainer. Remember to swipe the screen as soon as you see the gap.

Contrast Sensitivity

Our next assessment will be contrast sensitivity, how well you see subtle differences in brightness. For this assessment you will need to focus on the 4 gray circles. If the gap is at the left of the ring, swipe the screen from right to left. If the gap is at the top of the ring, swipe the screen from bottom to top. The ring will change in size after each response. If you cannot see the gap, take your best guess. Let’s begin with a practice run. If you have any questions now is a good time to ask your trainer now.

Depth Perception

Our next assessment is depth perception, how quickly and accurately you judge target distances. Your trainer will hand you the glasses. Remember, respond as quickly as you can.

Near Far Quickness

Next is near far quickness, how quickly and accurately your eyes shift focus between near and far targets. For this assessment the trainer will show you how to properly hold and respond to the hand held display. You will hold the display 16 inches away from your face, with the top edge just below the far screen. While looking at the hand held screen, a ring with a gap will appear. As before, quickly swipe in the direction of the gap. As soon as you have responded correctly to the near ring, a ring will appear inside the box on the far screen. Again swipe the screen in the direction of the gap. Continue this back and forth, focusing between each target as fast as you can. This assessment is timed, you have 30 seconds to correctly identify as many rings as you can. Before we start the assessment, let’s do a practice run to get you warmed up. Remember this is timed, so push yourself to change focus as quickly as you can.

Target Capture

Moving now to target capture, how quickly you identify a peripheral target. Keep your eyes on the white dot in the center of the screen. A ring with a gap will appear for a split second in 1 of the 4 corners of the display screen. Quickly move your eyes to focus on the ring. Identify the gap direction in the ring and swipe the screen in that direction of the gap. We will start with 6 practice trials. Remember keep your eyes on the dot until the ring appears, then quickly move your eyes to focus and determine the direction of the gap on the ring.

Perception Span

Next is perception span - how much information you process in a split second. Keep your eyes on the shrinking dot in the center of the grid. When the circle completely shrinks and dot disappears, green dots will flash inside some of the surrounding circles. Remember the pattern of green dots, and using the touch screen in front of you, tap on the matching circles to recreate the same pattern. When you have recreated the pattern as best as you can, tap enter. As you progress, the number of green dots and the size of the grid may increase. We will start with 2 practice runs. Work at your own pace, and remember, accuracy is important, speed is not.

Eye-Hand Coordination

Next is eye-hand coordination - how accurately and quickly your hands move to a visual target. For this assessment you will stand in front of a large grid of circles. Position yourself directly in front of the screen so that with your arms extended your fingertips just touch the screen. Once you are in correct position, take your hands away from the screen and hold at shoulder height to begin testing. A single light will appear in a circle at random locations on the grid. Using either hand hit the lights as quickly as you can with your fingertips. Once hit another light will immediately appear.
appear somewhere else on the grid. Continue as fast as you can until you hear the whistle. You will have 1 practice round. Remember, push yourself to hit as many lights as you can.

Go/No Go

Next is go/no go - how quickly you make decisions. This assessment is similar to the last, but now there are 2 types of lights, green and red. If the light is green, use the fingertips of either hand to hit the light as quickly as you can. If the light is red, do not hit the light. You will be penalized for hitting any red light. Both the red and green lights will only appear for only a short period, so you must react quickly. Continue as fast as you can until you hear the whistle. For this test you will not have a practice round. Remember to hit as many green lights as quickly as you can.

Reaction Time

Moving to reaction time - a measure of how quick you are. On the screen in front of you are 2 buttons. When we begin testing, if you are right handed you will rest the fingertips of your right hand on the start button on the right, if you are left handed you will rest the fingertips of your left hand on the start button on the left. If you are ambidextrous, pick the hand you are most comfortable with. No portion of your hand should extend across the boundary line marked on the screen. Now stand directly in front of the opposite button, with your eyes fixed on its center. When ready, place your fingertips on the start button. After a random period of time, the light directly in front of you will turn on. Move your hand over and hit the light as quickly as you can. You will practice twice before starting the assessment. Remember, this assessment is all about quickness.