



Reliability of Online MemTrax Digital Photo Recognition, Word Recognition and H-Scan Screening Tests



Rolf Martin, Ph.D.^{1,2}; Terence L. Rose, Ph.D.¹; Susan Kaib, M.D.³; Joseph Vogelmann, D.E.E.⁴; Anthony Wetherell, Ph.D.⁵ & J. Wesson Ashford, M.D., Ph.D.⁶
¹MMT Corporation, Sherman, CT 06784; ²Blueberry Study Team, Sherman, New Fairfield, Danbury, Bethel and Mansfield, CT and Cold Spring-on-Hudson, NY;
³Kronos Science Laboratory, Phoenix, AZ 85016; ⁴Orentreich Foundation for the Adv of Sci, Cold-Spring-on-Hudson, NY 10516;
⁵Durrington, United Kingdom; ⁶Stanford/VA Aging Clinical Rsch Center, VA Palo Alto Health Care System, Palo Alto, CA 94304

Abstract

BACKGROUND: To determine which offline and online performance measurements have highest reliability and statistical power and are therefore most suitable for screening and long-term performance monitoring, the Kronos Science Laboratory, MemTrax and Blueberry Study (BBS/MMT) groups directly compared their offline and online measurements by measuring test-retest reliability during a two-week period of repeated daily measurement.

METHODS: For detailed descriptions of methods and the online measurement sites themselves, please visit Memtrax.net, Blueberrystudy.com/KSL and Blueberrystudy.com/Measurements.

RESULTS: For offline visual accommodation (VA), offline highest audible pitch (HAP), online HAP, online face recognition (FR), online MemTrax and online BBS word recall measurements, test-retest and split half reliability (correlation, r) ranged from poor ($r < 0.7$) to relatively high (MemTrax up to 0.87; BBS word recall up to 0.96; offline HAP 0.98) depending on how many repeated measurement measurements were averaged prior to test-retest reliability calculation. Correlations with age were: HAP offline, -0.610; HAP online, -0.787; MemTrax (beta), 0.610; online FR (reaction time), -0.27; and offline VA, -0.646. MemTrax and FR correlated more closely with HAP than chronological age (e.g., MemTrax(beta)-age $r = 0.610$ while MemTrax(beta)-HAP $r = 0.767$) suggesting that HAP may represent biological rather than chronological age. Correlations between offline and online HAP measurements were 0.927, online MemTrax and online FR, 0.752, and offline VA and online HAP, 0.712.

CONCLUSIONS: Among measurements that were evaluated, MemTrax photo recognition, and BBS word recall and H-Scan highest audible pitch had highest test-retest reliability values (0.87-0.98) when repeated measurements were performed, indicating that relatively precise memory and hearing measurements can be obtained for either screening or performance monitoring.

Introduction

Since beginning continuous word memory monitoring in 2003, the Blueberry Study (BBS) group has reported each year at annual American Aging Association meetings that test-retest reliability values obtained online are as high as 0.97, and statistical power derived in major part from large numbers of repeated measurements is sufficient to enable individual participants to measure year to year changes in their performance as small as 1 percent. Each participant can therefore measure whether use of medications, nutrient supplements and healthful diets are associated with more or less rapid age-associated memory decline.

A key question raised about these BBS results is whether online measurement pages constructed for the study provide more power than other measurements such as MemTrax or H-SCAN that are used to screen and monitor age-associated change, or whether the appearance of power results solely from large numbers of repeated measurements.

Drs. Wes Ashford, Susan Kaib, Mitch Harman (Kronos Longevity Research Institute) and Rolf Martin agreed to conduct a direct comparison of their measurement methods at the 2008 annual American Aging Association meeting. Results of this comparison are presented here together with data conducted by each group separately.

Methods

Repeated mid-morning offline and online measurement sessions were completed on weekdays only by 12 participants at the Kronos Science

Laboratory in Phoenix, AZ. Each session lasted approximately 20 minutes and included 2 to 4 minutes for each measurement. Measurements were offline and online highest audible pitch (HAP), offline vibrotactile sensation, visual accommodation (VA), MemTrax continuous photographic image recognition memory and BBS face recognition memory, with faces presented in groups of four rather than continuously, following the protocol for BBS word recognition measurements since 2002. Offline H-SCAN (Hochschild, R., Corona del Mar, CA) and offline and online MemTrax and online BBS measurements had previously been performed for 11, 11 and 10 years, respectively, in Arizona, California and Connecticut/NY, providing large, long-term data sets that also were compared.

For detailed descriptions of methods and online measurement sites please visit Memtrax.net, Blueberrystudy.com/KSL and Blueberrystudy.com/Measurements.

Results

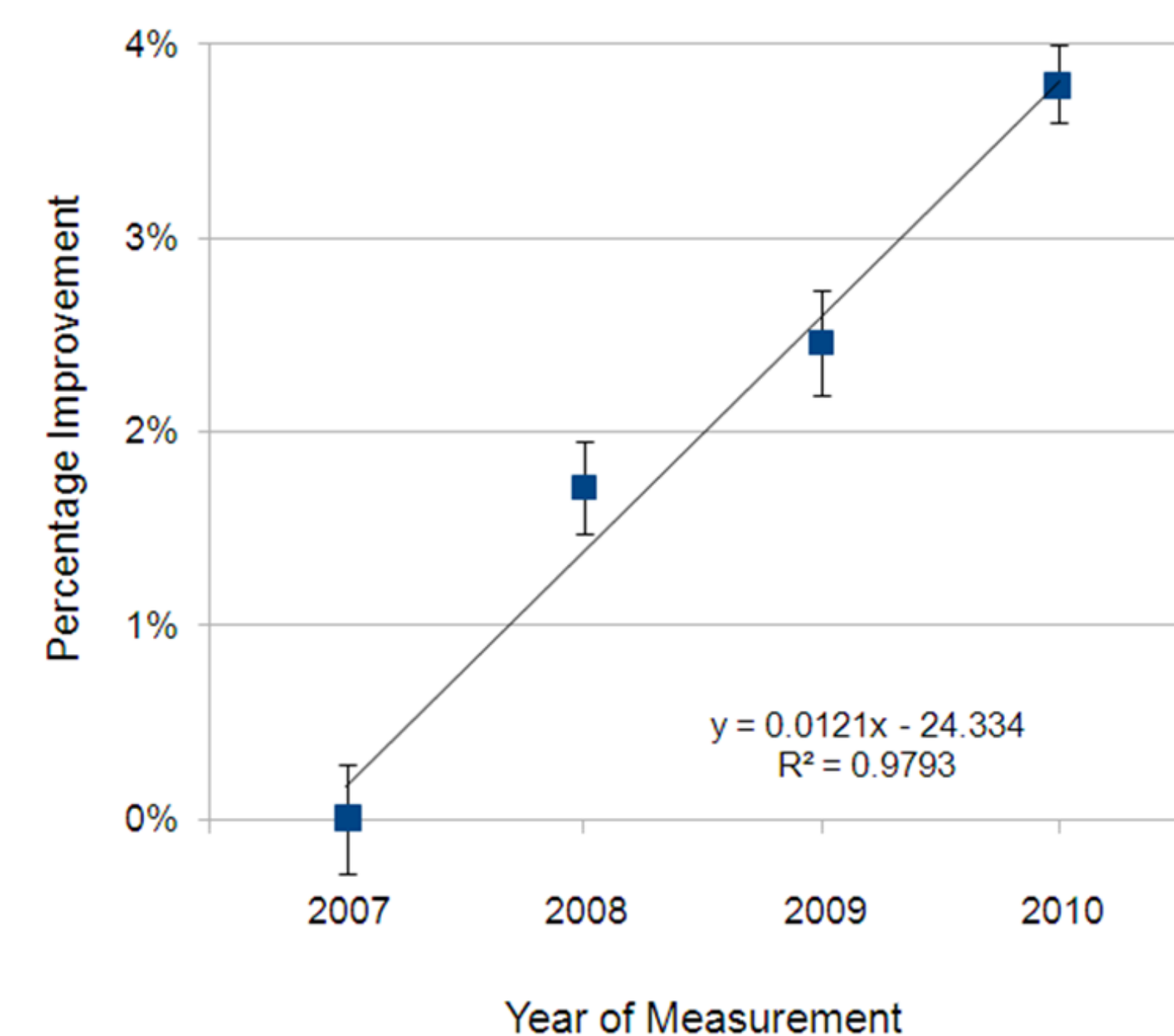
The high precision of online BB study measurement pages enables individuals to track year-to-year changes in their personal performance by conducting 5-10 minute online measurements every 2 or 3 days. The four year data set shown in Figure 1 (below) was obtained by one study participant. Scores in this person's data set increased significantly each year compared to the previous year at 95% confidence (2-tailed t test $p = 0.030$ for 2009 vs. 2008 and less for other years).

To date, two of 400 participants have completed over 100 measurements per year during consecutive years. Both reached the 95% confidence threshold for year-to-year changes below 1%. These dedicated individuals have evidently demonstrated that anyone familiar with the web can begin high-confidence personal health monitoring and obtain statistically powerful evidence concerning factors that may help them resist Alzheimer's disease, multiple sclerosis and other age-associated chronic illnesses that affect cognitive performance.

Low online within-person standard deviations yield strong correlations with age. Figure 2 (top right) shows HAP and MemTrax correlations with age. This figure compares correlations based on different numbers of values so should be interpreted cautiously.

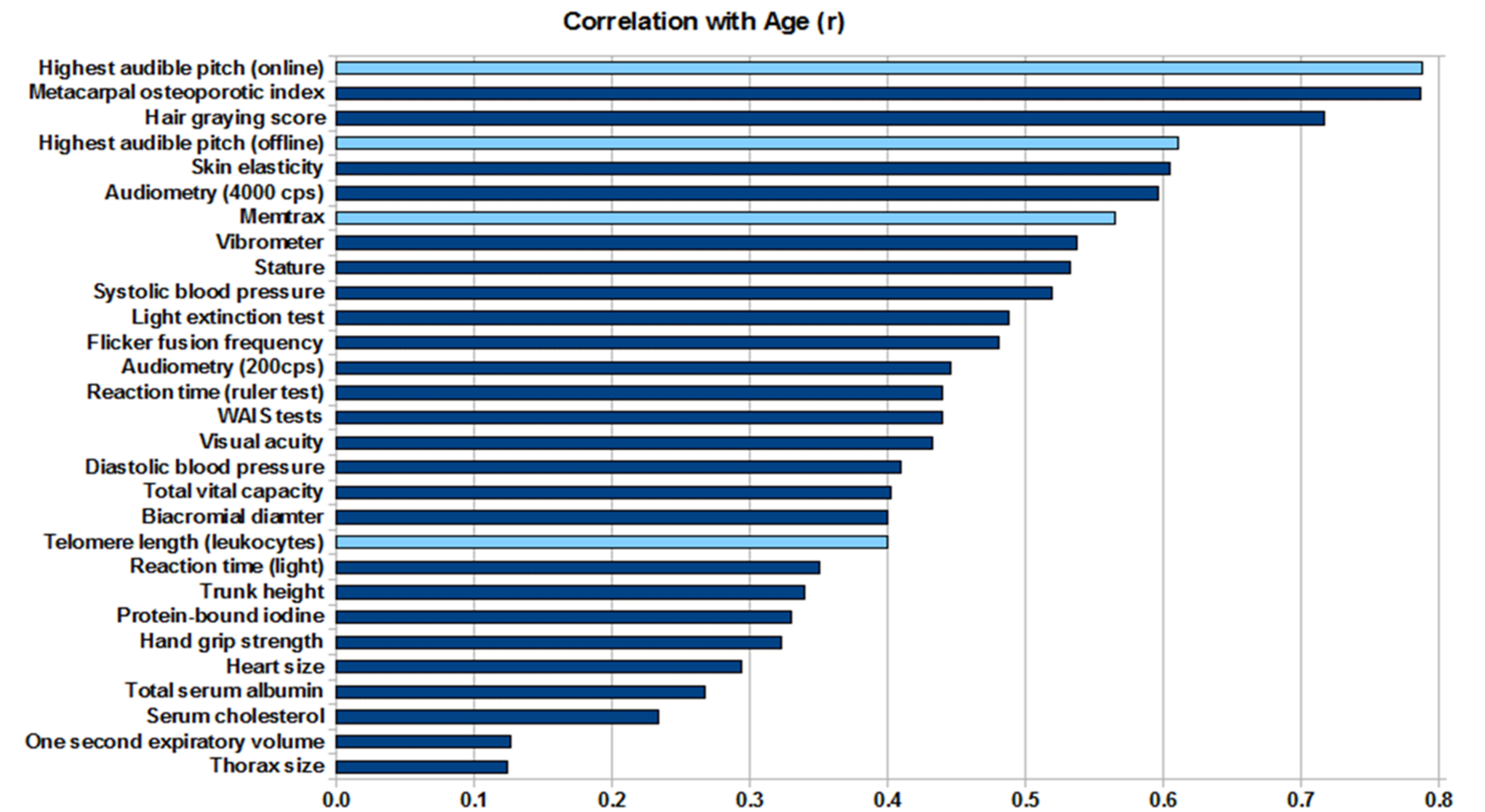
Four year data set of Participant DB

Each value shown represents one year of 100 or more measurements including recognition memory for common nouns, uncommon nouns, verbs and proper names. Scores from each year to the next improved significantly for this person at the 95% confidence level. Error bars display standard error of the mean.



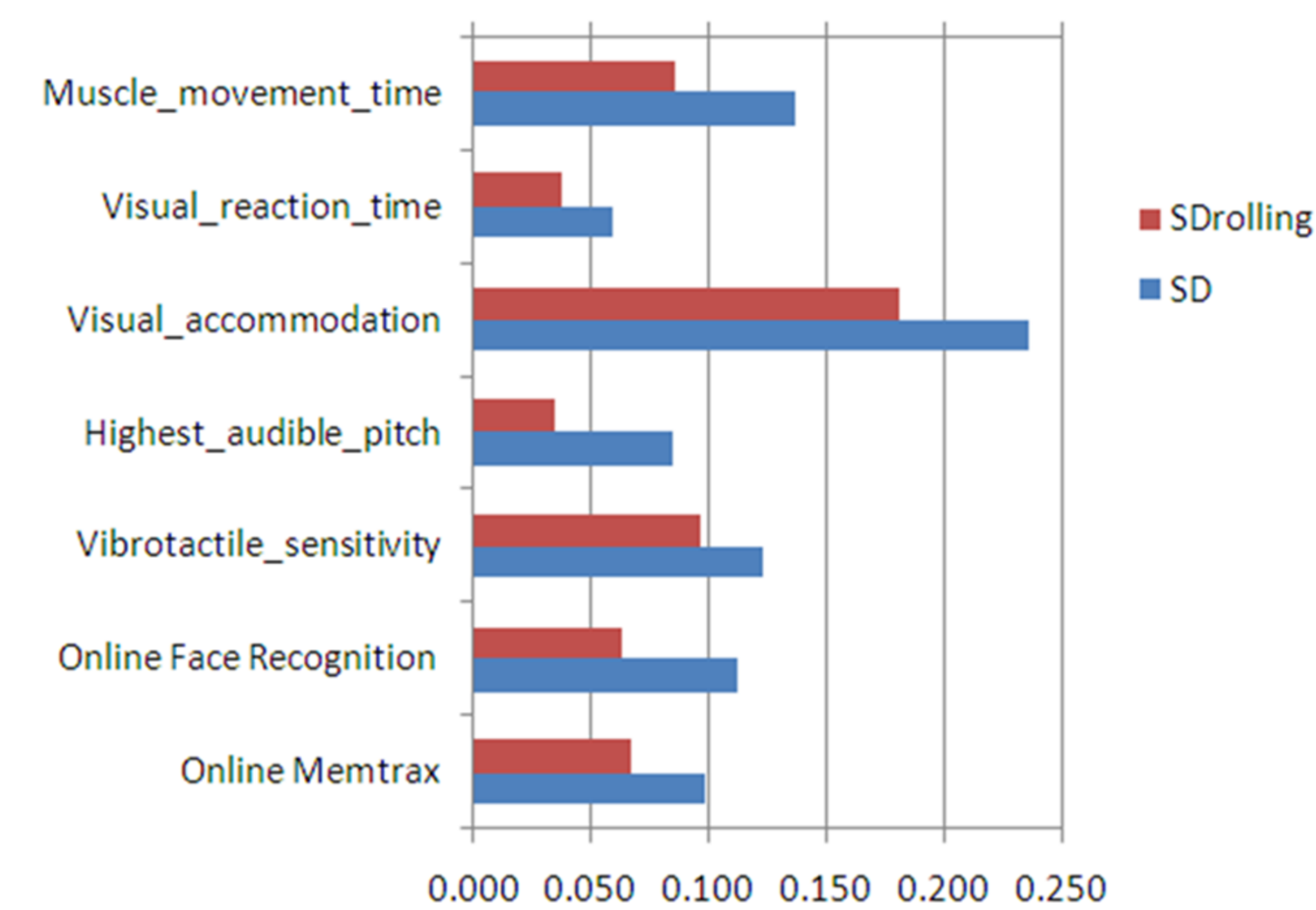
Highest Audible Pitch and MemTrax Correlated Strongly With Age Compared to Telomeres and Comfort's Test Battery

From Alex Comfort, The Biology of Senescence, Elsevier, NY, Third Edition, 1979.



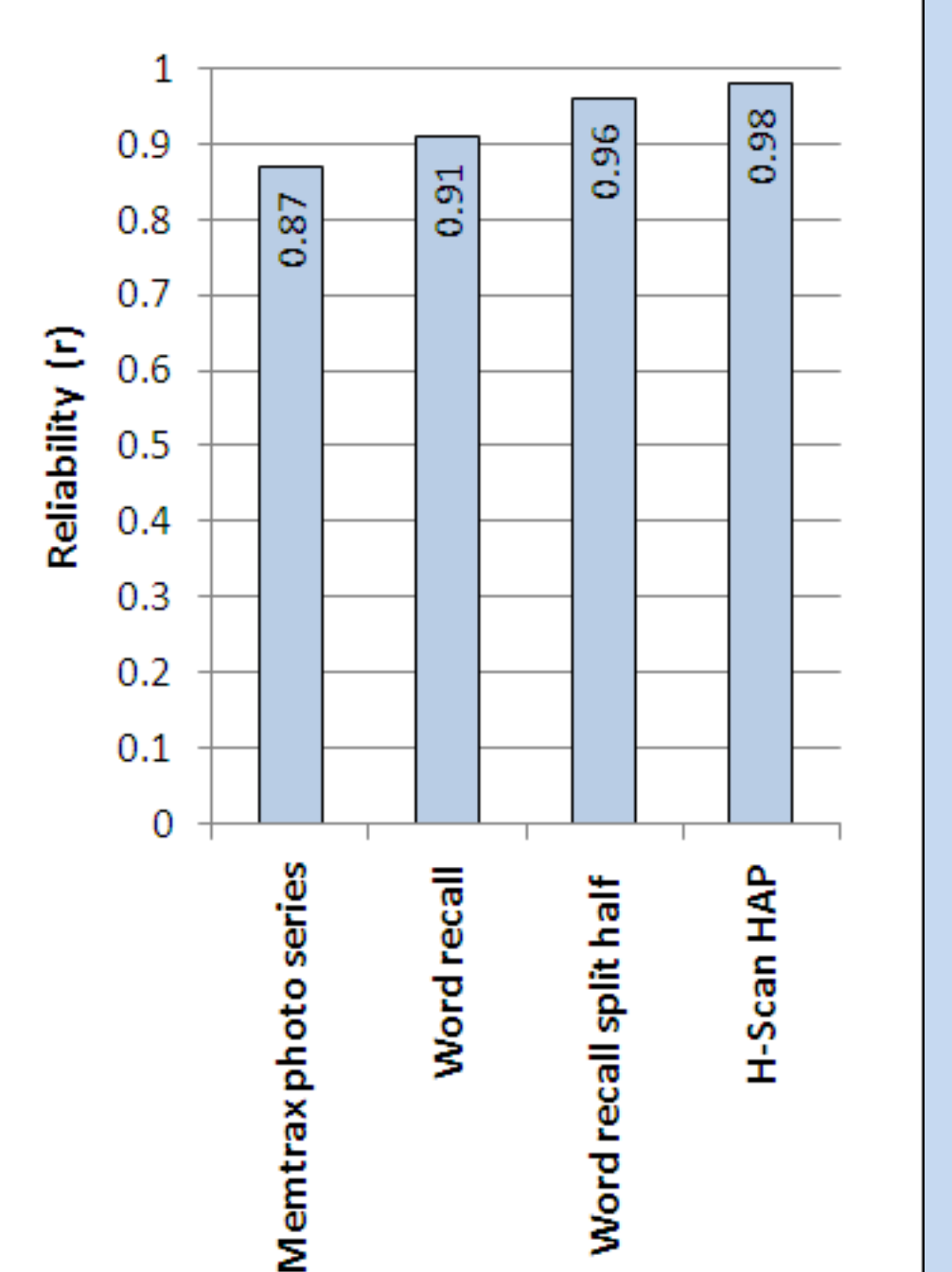
Relative Percent Standard Deviations of Offline H-SCAN and Online Memtrax and BBS Assessments Are Similar

Rolling standard deviations similar to rolling averages were used to remove relatively long-term interference such as colds and practice effects.



Image, Word and H-Scan HAP Test-Retest Reliability Values

(Obtained from 5-10 measurements per week over two weeks)



Conclusions

Very high precision obtained by some Blueberry study participants is based primarily on large numbers of repeated measurements. MemTrax, H-Scan HAP and other performance measurements can provide similar precision if motivated individuals "self screen" and then monitor progression toward or protection from Alzheimer's and other chronic age-associated illnesses.