International Journal of Research in Exercise Physiology

Original Research Article

Comparing Reaction Time Testing using Traditional Methods versus a Portable Interactive Light System

Madison E. Bunton[1], David S. Senchina[1]*

[1]Biology Department, Drake University, Des Moines, IA, USA; * Faculty mentor

Abstract

Introduction: Neuromuscular training systems for athletes, such as portable interactive light systems (PILS), are used for both training and testing, though recent research has emphasized the former. **Purpose:** This study compared a reaction time (RT) test from two different configurations of a PILS system (with subjects front- or side-facing) to the traditional ruler drop test and a color change-based app test. The main hypothesis was that performance times would correlate between all four RT tests, though the PILS times would be greater due to the gross motor movements involved. **Methods:** Forty-one subjects participated, split equally between female/male and athlete/non-athlete. Each subject completed five trials of all four tests, with tests administered in counterbalanced order. Subjects subjectively evaluated their performance and RT test difficulty after each test and after completing the test battery. **Results:** The main hypothesis was only partially supported because all RT tests correlated with each other (Pearson's $R \geq 0.44$, $p \leq 0.004$) except for the ruler drop test. Average reaction times for all four tests were: PILS-Front 0.61 ± 0.12 s, PILS-Side 0.64 ± 0.12 s, MA 0.44 ± 0.06 s, and RD 0.12 ± 0.05 s. There were sex differences only for PILS tests (females 0.68 ± 0.12 s *vs*. males 0.57 ± 0.1 s) but no other tests. The app-based test was the lone test where athletes were faster. When asked to report which trials were the fastest or slowest, subjects were accurate only half the time. **Conclusion:** These results suggest that PILS-based RT test results may correlate to other electronic color-based visual stimulus RT tests despite their different motor demands.

Key words: portable interactive light system, reaction test, reaction time, response time.

Introduction

Reaction time (RT) is crucial in both athletic and daily living activities[1]. RT is the amount of time that passes between the start of a stimulus and the first detectable response[2]. Several factors may influence RT, including sex, age, biomechanics, nutritional status, fitness status, and (in the case of testing) former training or practice[3–5]. The ability to react swiftly to stimuli can significantly impact success in various sports, where quick decision-making and coordination are essential. In experimental settings, athletes typically have significantly faster RT's compared to non-athletes, likely due to the cognitive demands of athletics where quick responses are crucial[6], though RT can vary by sex[7,8] and sport[9,10] and is affected by both the presence, absence, and type of warm-up performed[11].

Contemporary standard RT tests are typically computer-based and in response to changes in light or shape signals. An older traditional method, where the subject catches a dropped ruler[12], is still occasionally used. Both tests involve finer motor movements than athletes typically perform in their sport. Portable interactive light systems (PILS) designed for neuromuscular training may offer a more athletic-specific approach to RT testing that involves gross muscle movement. These systems consist of wirelessly controlled discs that light up in various colors and patterns, which can be arranged in countless configurations at either hand or foot level. The versatility of PILS may make them particularly effective for evaluating athletes' RT in sport-specific setups. There is no consensus in the psychophysics literature about what might differentiate a "pure reaction time" test from something involving gross motor movement (e.g., a "response time" test involving both the RT and movement time); thus, for this manuscript, all reactions/responses are referred to as RT regardless of whether or how much gross motor movement is involved.

Most research on PILS has focused on their efficacy as a training tool for sports such as basketball, handball, martial arts, tennis, soccer, and volleyball, and emphasized long-term RT training [13–21]. Only one system's RT test has been formally investigated in terms of test-retest reliability[22]. None have been formally compared to more traditional RT tests.

This study aimed to compare standard/more commonplace RT tests with PILS RT tests. Additionally, this study sought to determine if there were any notable differences in RT between athletes and non-athletes, or females and males. Four hypotheses were formulated: (1) PILS RT test times would be greater/longer than those obtained from traditional methods because of the gross body movements required, (2) performance scores would positively correlate across all four tests, (3) males would have faster RT than females across all tests, and (4) athletes would have faster RT than non-athletes across all tests.

Methods

Approval to perform the experiment was granted by the Drake University Institutional Review Board (2023-24051). Forty-one subjects consented to participate. Based on their sex and sport participation, subjects were partitioned into four groups: female athletes, female non-athletes, male athletes, and male non-athletes. "Athlete" was defined as anyone currently participating in an NCAA sport (whether in-season or out) or participating in club sports or squads of comparable training commitment.

Participation involved a single visit lasting approximately 30 minutes. After consent was obtained, subjects were screened for any confounding injuries, reported their athletic background and past PILS experiences, and had their height and weight recorded.

Subjects performed four different RT tests in a counterbalanced order. Given four tests, there were 24 different possible test sequences. All unique sequences were used before repeating any; thus, across 41 subjects, each test sequence was utilized only once or twice. By these permutations each test was first of four 25% of the time, second of four 25% of the time, and so forth.

Two of the four tests were standardized RT tests. One was the classic ruler drop test (designated as RD[12]). For the RD, RT was calculated using a standardized equation based on the distance dropped. Another was a tap-based mobile device test on a tablet (Reaction Speed-Test by Vicki Patridge v. 3; designated as MA) where the color of the screen changes as a cue to initiate movement. Subjects rested their palms on a table surface with their third finger touching the bottom of the tablet. An investigator tapped the screen to start the test. The screen color changed at variable intervals. For the mobile app, RT was the time between screen color change and the subjects' tapping the screen as measured by the app.

The other two RT tests were performed with a PILS system (FitLight; FitLight Corp., Miami, FL) with a starter disc at waist height and six other discs at chest height on a magnetic board, arranged in two rows of three discs each. The starter disc was on a tripod and its distance from the magnetic board was adjusted for each subject based on their arm length. The two PILS RT tests differed in how the subject was positioned relative to the discs (Figure 1). In PILS-Front, the subject stood face-on to the board, with everything in front of the subject and within their visual field, requiring no lower body movement and minimal torso movement. In PILS-Side, the subject stood perpendicular to the board and the start disc was 90°off to the side; thus, subjects had to rotate their torso/hips 90° after hitting the start disc (some used foot pivoting and compensatory leg movements). For both PILS tests, RT was the time from when the subjects engaged the start disc to when they engaged the target disc as measured by the PILS' app.

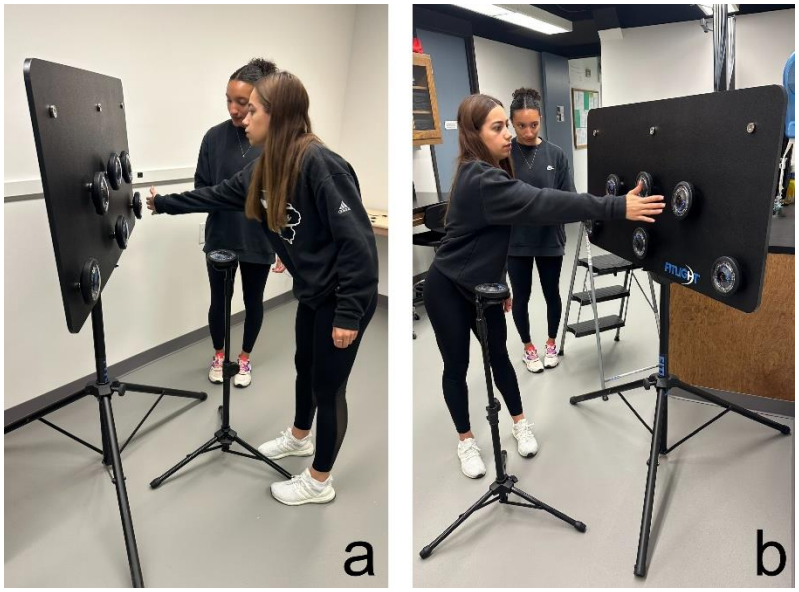

Figure 1. Configuration of the portable interactive light system (PILS) for PILS-Front (a) and PILS-Side (b).

Each of the four RT tests (RD, MA, PILS-Front, and PILS-Side) was administered five consecutive times, for a total of 20 RT tests across the whole visit. The five performances of a given RT were averaged prior to analysis; clearly aberrant performances were excluded prior to calculating averages. After completing all five trials of a given RT test, subjects were presented with two 10-cm visual analogue scales (VAS). The "Difficulty VAS" asked them to evaluate the RT in terms of complexity on a continuum of "easy" to "hard", and the "Performance VAS" asked them how well they felt the RT captured their actual abilities on a continuum of "well" to "poorly". Subjects indicated their perceptions by drawing a vertical line on the scale, and the distance from the left anchor to their mark was measured. They were also asked to recall which of the five trials was their fastest and which was their slowest. After they'd completed all four RT tests, subjects were asked to rank the tests from easiest ("1") to most difficult ("4").

Statistics were performed in JAMOVI 2.3.28. Typically, RT data is not normally distributed. When the RT data were checked for normality, the PILS-Front and PILS-Side data were normally distributed, but the MA and RD data were not. Erring on the side of minimizing data manipulation, only raw data results were reported. Depending on the outcome, either ANOVA or two-way ANOVA were used for comparisons. Whenever a statistical difference was found, follow-up *post hoc* tests utilizing Bonferroni corrections were employed as appropriate. Pearson correlations were used to compare outcomes between tests. A paired samples t-test was used for determining if there was

a difference in recall accuracy when subjects were asked to remember which trial within a set of five was the fastest or slowest. McNemar's tests were used for comparing the same responses across RT tests. Significance was defined as $p \leq 0.05$. Statistical trends ($0.05 < p < 0.1$) were reported out of thoroughness, but no follow-up tests were ever performed on trends.

Results

Subject Characteristics

Anthropometric data are shown in Table 1. Expectedly, t-tests indicated males were significantly taller and heavier than females (both $p<0.001$). Within females, there was no difference in height or weight between non-athletes vs. athletes ($p \geq 0.49$). Within males, there was no difference in weight ($p=0.549$), however, athlete males were significantly taller than non-athlete males ($p=0.008$). Female athletes reported being members of track and field ($n=6$), softball (2), cross country (1), and cheer (1). Male athletes reported being members of track and field ($n=5$), football (3), baseball (1), or long-distance sprint/cross country (1).

Table 1. Anthropometric characteristics of the 41 subjects included in the study.

	Non-Athlete	**Athlete**
Female	*n*=11 167.2 ± 5.4 cm 59.4 ± 5.4 kg	*n*=10 168.8 ± 5.0 cm 61.3 ± 9.0 kg
Male	*n*=10 173.1 ± 6.3 cm 81.0 ± 18.0 kg	*n*=10 182.6 ± 7.3 cm 86.8 ± 23.0 kg

Reaction Times

Average reaction times for all four tests were: PILS-Front 0.61 ± 0.12 s, PILS-Side 0.64 ± 0.12 s, MA 0.44 ± 0.06 s, and RD 0.12 ± 0.05 s (Figure 2). ANOVA revealed a significant difference by test ($p<0.001$). *Post-hoc* comparisons using Bonferonni corrections indicated no difference between PILS-Front and PILS-Side, but significant differences for all other comparisons (all $p \leq 0.001$) such that RT were longest for the PILS tests and shortest for the RD. Pearson correlations on RT by test revealed that performance on three of the tests (PILS-Front, PILS-Side, and MA) positively and moderately correlated with each other, whereas RD did not correlate with any of those three (Table 2).

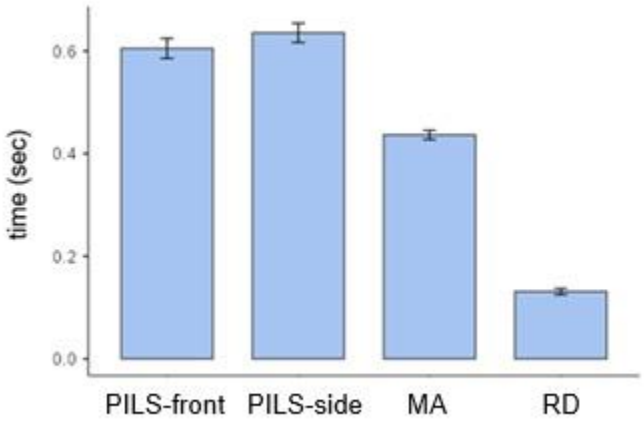

Figure 2. Average RT for the four tests expressed as time in seconds. Abbreviations: PILS = portable interactive light systems, MA = mobile app, RD = ruler drop.

Table 2. Correlation matrix for performance times across the four reaction tests.

	PILS-Front	**PILS-Side**	**MA**	**RD**
PILS-Front	--	--	--	--
PILS-Side	*r*=0.44, *p*=0.004	--	--	--
MA	*r*=0.51, *p*<0.001	*r*=0.49, *p*<0.001	--	--
RD	*r*=0.07, *p*=0.691	*r*=-0.14, *p*=0.391	*r*=0.19, *p*=0.25	--

Abbreviations: PILS = portable interactive light systems, MA = mobile app, RD = ruler drop.

Figure 3 displays RT data for each of the four tests broken down by subject groups (female vs. male, non-athlete vs. athlete). For the PILS-Front results, two-way ANOVA indicated a significant difference by sex ($p<0.001$) such that males performed faster than females, but no significant differences by athlete status ($p=0.162$). For PILS-Side results, there was a trend for a significant difference by sex ($p=0.066$) such that males performed faster than females, but no significant difference by athlete status ($p=0.507$). For MA, there was no significant difference by sex ($p=0.106$) but a significant difference by athlete status ($p=0.048$) such that athletes performed faster than non-athletes. For RD, there were no significant differences by sex or athlete status ($p\geq0.653$).

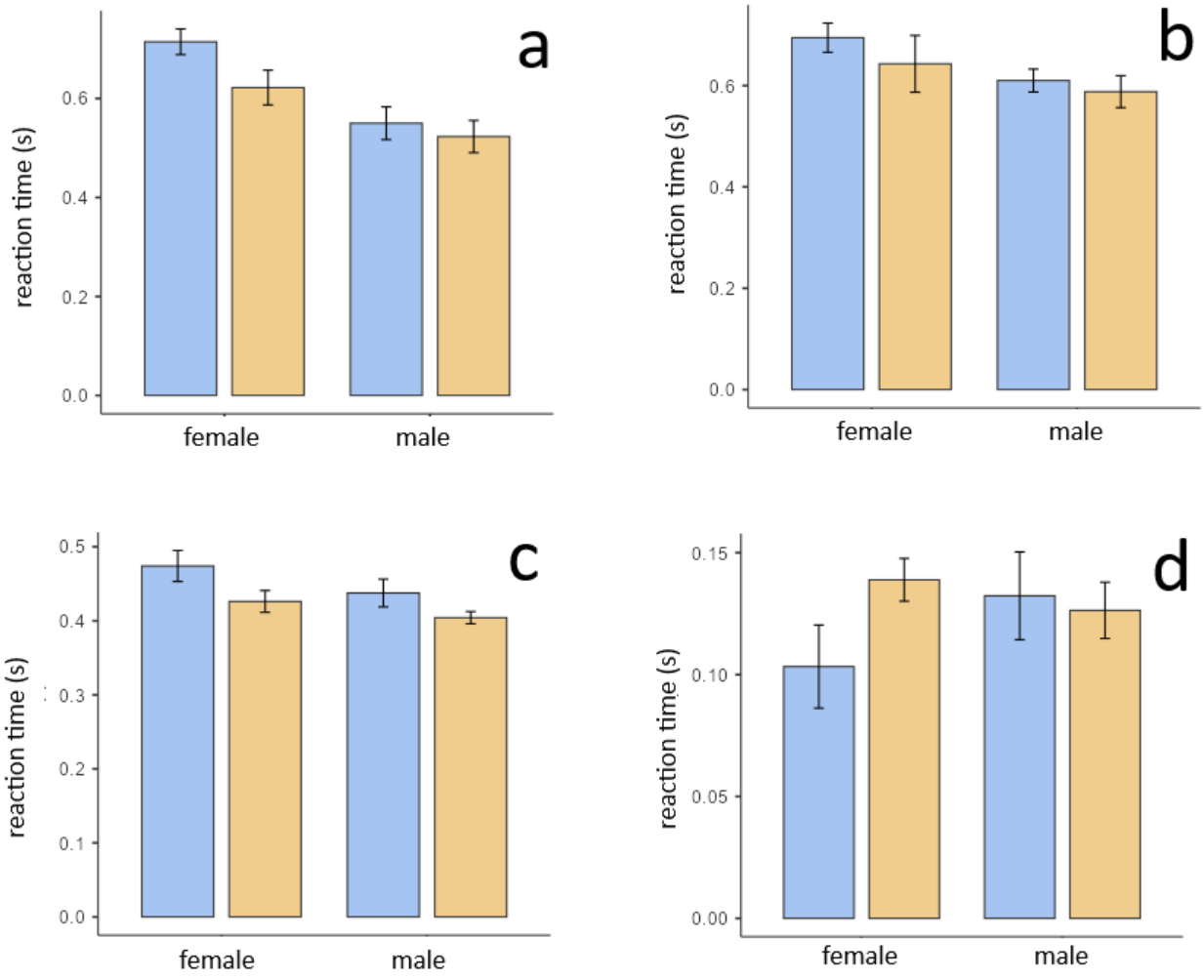

Figure 3. Reaction time data (expressed as averages +/- standard deviation, in seconds) for the four RT tests: (a) PILS-Front, (b) PILS-Side, (c) mobile app, AND (d) ruler drop. Athlete status is indicated by color: non-athletes (blue), athletes (orange).

Acute Perceptual Scores

Average perceptual scores for both the Difficulty VAS and Performance VAS across all four RT tests are shown in Figure 4. Separate ANOVAs indicated only trends for significant differences in scores by test (Difficulty VAS p=0.065, Performance VAS p=0.084) and thus were not plumbed further. Two-way ANOVAs for sex or athlete status revealed only two trends across all comparisons (sex differences in both VAS scales for RD, with Difficulty VAS p=0.066 and Performance VAS p=0.087).

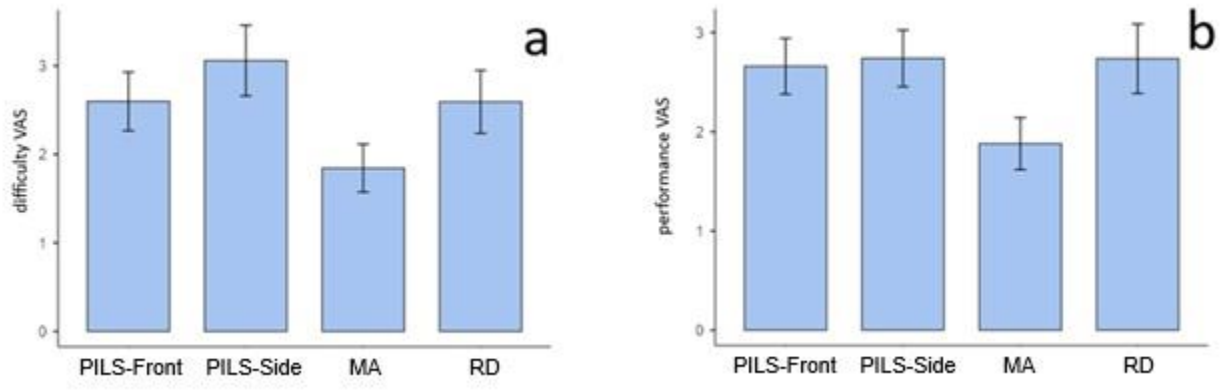

Figure 4. Subjective measures of (a) difficulty and (b) self-performance as assessed by 10-cm visual analogue scales (VAS). Values are expressed in cm as averages +/- standard deviation. Abbreviations: PILS = portable interactive light systems, MA = mobile app, RD = ruler drop.

When subjects were asked to recall which of the five trials within an RT test was their fastest and which was their slowest, response accuracy varied greatly across subjects. Subjects were 56.7 ± 31.1% accurate in recalling the fastest trial, and 62.0 ± 28.5% accurate in recalling the slowest trial; a t-test revealed no significant difference between the two recall accuracy rates. By McNemar's test and after adjusting the alpha value for the number of comparisons, there was never a significant difference in the recall accuracy between subjects' perceived fastest and slowest trials within each of the four RT; likewise, there were no significant differences between subjects' accuracy rate in perceptions of fastest trials for any pair of RT tests, nor for slowest trials.

Overall Perceptual Scores

Average rank scores for relative difficulty across all four RT tests are shown in Figure 5. There was a significant difference by test (p=0.001) such that PILS-Side was perceived as significantly more difficult than the other three RT tests by Tukey *post hoc* tests (all p<0.001); however, the other three tests were not significantly different from each other.

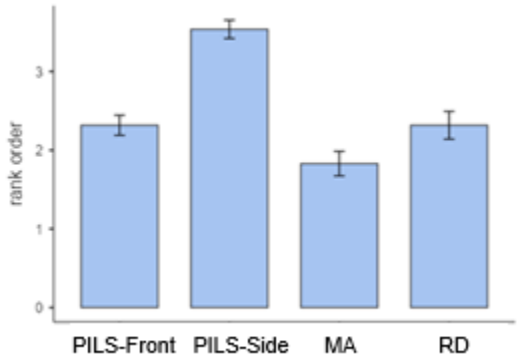

Figure 5. Rank-order scores for all four RT tests, expressed as average score +/- standard deviation based on a four-point scale. Abbreviations: PILS = portable interactive light systems, MA = mobile app, RD = ruler drop.

Discussion

The results of this study provide insight into how different RT tests compare to each other. Supporting hypothesis #1, performance times were significantly greater for the two PILS tests versus MA or RD (Figure 2). This is consistent with the PILS tests requiring more gross body movements than either the MA or RD, which require only hand or short limb movements.

Correlations between tests showed that PILS-Front, PILS-Side, and MA positively and moderately correlated together (Table 2), indicating that performance on one test tended to be similar to performance on the other test. However, RD did not correlate with any of the other three tests. These data mostly, yet only partially, support Hypothesis #2. Both the PILS tests and the MA test measure RT in response to a change in color, whereas RD measures RT to a dropped physical object, which may explain why RD performance did not correlate with other tests. Since the PILS RT significantly and positively correlated with MA, these results may suggest that tests such as the MA may function as less expensive and simpler alternatives to more expensive PILS devices for simple RT measurement. However, since the tests differ in their motor demands, that conclusion should be treated cautiously. The PILS tests require substantial upper body and trunk movement, getting at the issue of "pure reaction time" versus "response time" alluded to earlier in the Introduction. Alternatively still, if PILS systems are available, and since PILS systems can be configured into different sport-specific arrangements, then PILS-based RT tests may better mirror the motor demands of athletic situations and/or be more motivating to athletes, though our study did not formally address either of those possibilities.

The analysis of sex and athlete status revealed some interesting comparisons

(Figure 3). For PILS-Front males performed significantly faster than females. A similar though non-significant trend was observed for PILS-Side. Both PILS tests involve greater gross muscle whole-body movements than either the MA or RD. The greater average percentage of upper body muscle mass and strength in males[23,24] compared to females may explain these differences. At best, the results may only half support Hypothesis #3 and only for the PILS RT tests. Athletes performed significantly better than non-athletes in the MA test, but not other tests (Figure 3), contrary to what was predicted in Hypothesis #4. This finding is surprising given that one might expect athletes to have faster RT compared to non-athletes given their sport training[6,9,10]. There are three possible explanations for the general lack of difference between athletes and non-athletes in this study. One is that the MA test utilized the same daily body movements that both athletes and non-athletes regularly use; since both groups have the same amount of "practice" with those motor patterns, the app would not differentiate them. Another is the diversity of sports played by athletes in this study. A less likely explanation is that RT has a large genetic component[3,25], so it is plausible that these specific RT outcomes reflect genetics more so than training. No other sex or athlete status comparisons were significant.

When assessed using VAS administered acutely after each test, participants generally perceived all tests to be similar in difficulty, regardless of test type (Figure 4). However, when asked to rank-order the four tests in terms of relative difficulty after trying all four, subjects gauged PILS-Side as being more difficult than the other tests (Figure 5). The rank-order assessment forced subjects to compare the four tests against each other, whereas the Difficulty VAS assessment asked them to consider the difficulty on a scale from "easy" to "hard" and not in comparison to each other, which may explain the discrepancy.

Subjects were generally inaccurate when asked to report which trials were the fastest or slowest, being accurate only about half of the time. Several factors may explain this. With some subjects, their performances across the five trials within a specific RT test were very similar, sometimes differing by only a quarter-of-a-second, which is likely below most people's abilities to perceive under these circumstances[26]. Alternatively, subjects could have been focusing on an upcoming performance more than trying to remember past performances or could have been inattentive or flippant to the recall reporting.

This study had several limitations. As noted in the second paragraph of the discussion, the RT tests differed by their involvement of gross motor movements; thus, some incorporate an additional "response time" along with the "pure reaction time." Another limitation is the small sample size per sex-by-athlete group (approximately 10 subjects for each of the 4 groups); a larger sample size may yield different results. This

study's athlete population was diverse within itself in terms of sport type and training, so future research could narrow in on athletes of specific sports to see if sport-specific differences emerge. As this study tested only healthy college students, it does not represent the full human lifespan. This study did not take into consideration other factors that may influence performance on reaction time, such as age, sleep, stress, diet (e.g., amounts of caffeine or similar stimulants consumed), recent physical activity, or within-experiment cognitive-based fatigue. This study cannot speak towards whether the specific PILS used here or similar neuromuscular training systems have the capacity to enhance neuromuscular skills, as the experiment was limited to a single session. We did not explore test-retest validity.

Conclusions

PILS-based RT tests used in this study were comparable to an app-based RT test but not the traditional ruler drop test (possibly because the former three are color change-based whereas the ruler drop test is not). Males exhibited faster RT for tests involving more gross motor movement. There were no differences in test performance for athletes compared to non-athletes. PILS-based RT testing is more expensive than other forms of RT testing, however, PILS-based systems allow more sport-specific testing of RT. When asked to subjectively evaluate their performance, subjects struggled to differentiate trials within a test, possibly due to the low variation in performance times between trials.

Acknowledgments

This research was partially supported by the Cliff G. Rea endowment for undergraduate research. MEB completed this study in partial fulfillment of biology degree thesis research requirements and with the assistance of first-year students in Bio 031 Kinesiology Orientation. Christopher Kliethereems provided advice on some statistical analyses. Both co-authors contributed to manuscript writing.

Address for Correspondence

Faculty advisor: David S. Senchina (PhD), Kinesiology Program, Biology Department, Drake University, Des Moines, IA, USA. Tel. 1 515 271 2956, Fax 1 515 271 3702, E-mail dssenchina@drake.edu.

References

1. Paul, D. J., Gabbett, T. J. & Nassis, G. P. (2016). Agility in Team Sports: Testing, Training and Factors Affecting Performance. *Sports Med.* 46, 421–442.
2. Collet, C. (1999). Strategic aspects of reaction time in world-class sprinters. *Percept. Mot. Skills* 88, 65–75.
3. Davies, G. *et al.* (2016). Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Mol. Psychiatry* 21, 758–767.
4. Reigal, R. E. *et al.* (2019). Relationships Between Reaction Time, Selective Attention, Physical Activity, and Physical Fitness in Children. *Front. Psychol.* 10, 2278,.
5. Wong, A. L., Goldsmith, J., Forrence, A. D., Haith, A. M. & Krakauer, J. W. (2017). Reaction times can reflect habits rather than computations. *eLife* 6, e28075.
6. Tomczyk, C. P., Mormile, M., Wittenberg, M. S., Langdon, J. L. & Hunt, T. N. (2018). An Examination of Adolescent Athletes and Nonathletes on Baseline Neuropsychological Test Scores. *J. Athl. Train.* 53, 404–409.

7. Bleeker, M., Bolla-Wilson, K., Agnew, J. & Meyers, D. (1987). Simple visual reaction time: Sex and age differences. *Dev. Neuropsychol.* 3, 165–172.
8. Der, G. & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychol. Aging* 21, 62–73.
9. Meng, K. Y., Zuhairi, N., Abd Manan, F., Feizal, V. & Omar, R. (2018). Visual reaction time and visual anticipation time between athletes and non-athletes. *Malays. J. Public Health Med.* S1, 135–141.
10. Caccese, J. B. *et al.* (2020). Clinical Reaction-Time Performance Factors in Healthy Collegiate Athletes. *J. Athl. Train.* 55, 601–607.
11. McGowan, C. J., Pyne, D. B., Thompson, K. G. & Rattray, B. (2015). Warm-Up Strategies for Sport and Exercise: Mechanisms and Applications. *Sports Med.* 45, 1523–1546.
12. Johnson, B. L. & Nelson, J. K. (1969). *Practical Measurements for Evaluation in Physical Education*..
13. Badau, D. & Badau, A. (2022). Optimizing Reaction Time in Relation to Manual and Foot Laterality in Children Using the Fitlight Technological Systems. *Sensors* 22, 8785.
14. Badau, D. *et al.* (2023). The Impact of Peripheral Vision on Manual Reaction Time Using Fitlight Technology for Handball, Basketball and Volleyball Players. *Bioengineering* 10, 697.
15. Forni, F., Farinini, E., Leardi, R. & Rinaldo, A. (2022). Effects of visual training on motor performance in young tennis players using FitLight trainer. *J. Sports Med. Phys. Fitness* 62, 585–592.
16. Hassan, A. K., Alhumaid, M. M. & Hamad, B. E. (2022). The Effect of Using Reactive Agility Exercises with the FITLIGHT Training System on the Speed of Visual Reaction Time and Dribbling Skill of Basketball Players. *Sports* 10, 176.
17. Silvestri, F. *et al.* (2023). Acute Effects of Fitlight Training on Cognitive-Motor Processes in Young Basketball Players. *Int. J. Environ. Res. Public. Health* 20, 817.
18. Stoica, M., Sorin, C., Vișan, R. & Dreve, A. (2024). AI, BlazePod Sensors, and Head Vests Implemented in Assessments on Reaction Time and Gaze Training Program in U10 Football Game. *Appl. Sci.* 14, 6323.
19. Polechoński, J. *et al.* (2025). Assessment of the reliability and validity of simple and complex reaction speed tests in mixed martial arts athletes using the BlazePod system. *Balt. J. Health Phys. Act.* 17, 2,.
20. Potisaen, J., Potisaen, T. & Srimuang, T. (2023). Effect of BlazePod Training Program Optimization on Agilities and Leg Muscle Strength among Male Volleyball Players. *J. Exerc. Physiol. Online* 26, 79–88.
21. Gradinaru, L., Petracovschi, S., Bota, E., Merghes, P. & Oravitan, M. (2023). The effect of Blazepod Flash Reflex Training program on vertical jump in U15 female volleyball players - ProQuest. *Timisoara Phys. Educ. Rehabil. J.* 16, 31–37.
22. de-Oliveira, L. A., Matos, M. V., Fernandes, I. G. S., Nascimento, D. A. & da Silva-Grigoletto, M. E. (2021). Test-Retest Reliability of a Visual-Cognitive Technology (BlazePod™) to Measure Response Time. *J. Sports Sci. Med.* 20, 179–180.
23. Janssen, I., Heymsfield, S. B., Wang, Z. M. & Ross, R. (2000). Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr. *J. Appl. Physiol.* 89, 81–88.
24. Perez, M. A., Urrejola-Contreras, G. P., Hernández, J., Silva, P. & Torres-Banduc, M. (2024). Sex differences in upper and lower strength and their association with body composition among university students. *Phys. Act. Nutr.* 28, 64–71.
25. McGue, M., Bouchard, T. J., Lykken, D. T. & Feuer, D. (1984). Information processing abilities in twins reared apart. *Intelligence* 8, 239–258.
26. Haigh, A., Apthorp, D. & Bizo, L. A. (2021). The role of Weber's law in human time perception. *Atten. Percept. Psychophys.* 83, 435–447.