Journal of Athletic Training 2014;49(3):356–359
doi: 10.4085/1062-6059-49.2.04

www.natajournals.org

original research

# Practice Effects Associated With Repeated Assessment of a Clinical Test of Reaction Time

**Gianluca Del Rossi, PhD, ATC; Alfonso Malaguti, BSc, ATC; Samanta Del Rossi, BSc**

University of South Florida, Tampa



***Context:*** Researchers have confirmed that the ruler-drop test could be included as part of a multifaceted concussion-assessment battery and potentially as a way to track recovery from head injury. However, it is unclear if this clinical test of reaction time would be characterized by inconsistent performance because of practice effects.

***Objective:*** To determine if the ruler-drop test is susceptible to practice effects after serial administration.

***Design:*** Descriptive laboratory study.

***Setting:*** Sports medicine research laboratory.

***Patients or Other Participants:*** Forty-three persons (age = 21.8 ± 2.6 years).

***Intervention(s):*** Ten sessions were completed over 5 weeks. Participants completed 10 trials of the ruler-drop test during each session.

***Main Outcome Measure(s):*** The mean reaction times calculated for all participants from each test session were analyzed to determine if there was any meaningful change (ie, improvement) in reaction time over the course of the investigation.

***Results:*** Simple reaction time improved (ie, decreased) after repeated administration of the ruler-drop test, and the most pronounced improvement occurred between the first 2 test sessions. Between the first and second test sessions, reaction time decreased by almost 7 milliseconds, and there was an overall improvement of almost 13 milliseconds between the first and tenth sessions. Although the pairwise comparisons between the first and second and the first and third sessions were not significant, the change in mean reaction time between the first session and most of the other sessions was significant. We noted no differences when successive sessions were compared.

***Conclusions:*** To prevent practice-related improvements in reaction time, practitioners should allow at least 1 practice session before recording baseline results on the ruler-drop test.

***Key Words:*** yardstick test, concussion-assessment battery, head injuries

**Key Points**

- If the ruler-drop test is to be used as part of a multifaceted concussion-assessment battery or as a way to track recovery from head injury, practice-related improvements in reaction time must be accounted for before a person's baseline measure is established.
- Allowing people to repeat the ruler-drop test until their performance level has stabilized makes it possible to account for and thereby limit the effects of practice on performance.
- Practice-related improvements in reaction time on the ruler-drop test may be lessened by administering at least 1 practice session before recording baseline values.

After a sport-related concussion, patients typically demonstrate measureable delays in reaction time.[1] As these injuries resolve, reaction times eventually return to normal.[2,3] This transient departure from normal status after injury and eventual return to baseline as recovery progresses enables clinicians to use reaction time as an indicator of concussion and as a way to gauge recovery from injury.

In concussion patients, *reaction time*, defined as the time it takes to initiate a behavioral response after the presentation of a sensory stimulus, is typically assessed using computerized neuropsychological testing software. Although numerous forms of these test batteries are available to practitioners, the cost and the specific equipment and personnel requirements associated with test administration often limit how frequently these measurements are actually used for concussion assessment and management. A simple, less expensive clinical measure of reaction time that can be used in the absence of a more formal computer-generated assessment is the traditional ruler-drop test. Although the ruler-drop test actually measures reaction time plus movement time, this well-established test continues to be an acceptable means of approximating simple reaction time in a clinical setting.

Often used in elementary and high schools as part of a battery of sport-performance evaluation tests, the ruler-drop test requires the athlete to catch a measuring stick that has been dropped; the clinician then measures the length or distance the ruler has traveled before being grasped to provide a measure of simple clinical reaction time. The clinical usefulness of the ruler-drop test has been previously studied.[4–7] In fact, researchers[4–7] have confirmed that this clinical test of reaction time could be a part of a multifaceted concussion-assessment battery and potentially a way to track recovery from head injury. The ruler-drop test has acceptable test-retest reliability that compares favorably with computerized measures of reaction time.[5–7] The problem, however, is that using such a test to evaluate recovery from concussion requires serial administration of the test; whether repeating such a test over time (ie, multiple days and weeks) would be characterized by inconsistent performance due to practice effects is unknown.

A clinical test that is susceptible to practice effects poses a potential problem because improved reaction time that is merely the result of repeated assessments can obscure the true extent of a person's recovery from concussion. Unfortunately, the results of previous research studies (although not specifically designed to investigate this question) appear to suggest that a person's performance on clinical reaction time tests may indeed be affected by practice effects.[5,6] Therefore, our objective was to determine if serial administration of the ruler-drop test over 10 sessions would lead to improvements in reaction time. The ruler-drop test is a simple and straightforward test of reaction time, but for someone who has not previously performed such a test, it represents a novel task that is likely to be performed suboptimally at first and better with more exposure. Thus, we hypothesized that we would observe practice-related improvements in reaction time.

## METHODS

### Participants

Forty-three participants (21 men and 22 women; age = 21.8 ± 2.6 years; 42 right-hand dominant, 1 left-hand dominant) were recruited for this repeated-measures prospective investigation. We obtained ethical approval for the study from the university institutional review board and performed all testing in a controlled setting (research laboratory). All study procedures were fully explained to participants, who then voluntarily gave consent before testing. Participants with any existing neurologic or musculoskeletal condition that would affect their ability to complete the study were excluded.

### Apparatus and Procedure

Simple reaction time was evaluated using a measuring stick that was 60-cm long and marked in 1-mm increments. The participant sat in a chair and rested the forearm and lateral aspect of the palm of the dominant hand on the armrest with the fingers suspended off the edge (Figure). The measuring stick was hidden from view by suspending it vertically within a polyvinyl chloride (PVC) pipe that was 6.0 cm in diameter and open at the top and bottom. The ruler was positioned within the pipe so that the zero point was level with the lower open edge of the PVC pipe (Figure). The participant positioned the open hand against the open lower edge of the pipe so that the zero point of the ruler was directly above the thumb and index finger. Also, to standardize the starting position of the thumb and fingers during all trials, we asked all participants to keep their fingers lined up with the PVC pipe (and approximately 4 cm apart). This was a critical step because the reaction time assessed in our study actually consisted of reaction time plus movement time. That is, we did not just measure the time that transpired between the presentation of the stimulus and the initiation of the muscular response to the ruler dropping (true reaction time) but rather we measured the time required to initiate the grasping motion as well as the time it took the fingers to come together and clutch the ruler (movement time). Therefore, it was important for the distance between the thumb and fingers to remain constant.

**Figure. Apparatus used to keep the measuring stick out of sight of the participant. The ruler was initially positioned within a polyvinyl chloride pipe so that the zero point of the ruler was level with the lower open edge of the pipe. The participant was required to catch the ruler as soon as possible after seeing it.**

Once the participant was ready for testing, the researcher dropped the measuring stick from inside the pipe at self-generated random intervals (between 1 and 5 seconds) to prevent the participant from anticipating the time of release. Although it happened infrequently, if at any point the participant anticipated the release, the trial was repeated. An anticipatory grasp was determined to have occurred if movement of the thumb and fingers was noted just before the ruler was released. Once the ruler was released from inside the PVC pipe and visualized, the participant was required to catch it as quickly as possible with minimal movement of the hand from the starting position. The distance or length the ruler had fallen before it was grasped was measured at the most superior aspect of the participant's thumb and was converted into a reaction time (in milliseconds) using the formula for a falling body ($d = \frac{1}{2}gt^2$), where d is distance, $g$ is acceleration due to gravity, and t is time. Given the smooth surface of the measuring stick, it was possible for the ruler to slip between the fingers when the participant attempted to grasp it, thus potentially increasing the reaction time for that trial. Therefore, any

**Table 1. Mean Reaction Time for the Ruler-Drop Test by Session, Mean ± SD**

| Session | Reaction Time, ms |
|---|---|
| 1 | 264.9 ± 17.1 |
| 2 | 257.7 ± 18.2 |
| 3 | 258.2 ± 16.2 |
| 4 | 256.1 ± 18.1[a] |
| 5 | 253.0 ± 20.2[a] |
| 6 | 254.4 ± 16.4[a] |
| 7 | 254.2 ± 19.7 |
| 8 | 252.6 ± 17.5[a] |
| 9 | 254.4 ± 17.2[a] |
| 10 | 252.0 ± 15.7[a] |

[a] Indicates different from session 1 ($P \leq .05$).

trial in which we noted slippage was discarded and repeated.

## Data Analysis

All participants were required to complete 10 trials on 10 different days (sessions), with 2 sessions typically occurring within a single week (eg, Monday and Wednesday or Tuesday and Thursday). Each session was completed in less than 10 minutes. The 3 fastest and 3 slowest times during each test session were eliminated, and the middle 4 times were averaged[8] to eliminate potential outliers. We calculated the mean reaction time from each of the 10 sessions and used those values in all statistical tests. To determine if any practice effects were associated with serial administration of the reaction time, we performed a repeated-measures analysis of variance. Post hoc pairwise comparisons with Bonferroni adjustments were conducted if necessary. The Bonferroni correction is an adjustment made to reduce type 1 errors when multiple pairwise comparisons are performed. The correction requires that the a priori critical $P$ value ($\alpha$) be divided by the number of comparisons being made. All statistical analyses were performed using SPSS statistical software (version 20.0; IBM Inc, Armonk, NY) with the level of significance set a priori at $\alpha \leq .05$.

## RESULTS

Statistical tests performed on the mean reaction time values from the 10 test sessions (Table 1) revealed a main effect, although the Mauchly sphericity test was significant (Mauchly W = 0.185, $P = .021$), so we applied the Huynh-Feldt adjustment for degrees of freedom ($F_{8,337} = 6.25$, $P < .0001$). A total of 17 pairwise comparisons were performed. Each session was compared with the preceding and successive sessions as well as with the baseline (first) session (Table 2). The new critical $P$ value calculated by applying the Bonferroni correction was $P = .0029$ (0.05/17). Post hoc tests using the Bonferroni adjustment for multiple pairwise comparisons revealed differences only between the first session and the fourth, fifth, sixth, eighth, ninth, and tenth sessions (Table 2).

## DISCUSSION

Simple reaction time decreased (ie, improved) after repeated assessments, and the most pronounced improvement occurred between the first 2 test sessions. Between the first and second test sessions, reaction time decreased by almost 7 milliseconds; overall improvement between the first and tenth sessions was almost 13 milliseconds. To some degree, these data parallel those reported in a previous study[6]: that simple reaction time improved by about 11 milliseconds between the original and subsequent administrations of the modified ruler-drop test. Unfortunately, with data from only 2 test sessions, the researchers were unable to determine if reaction time would have continued to improve with additional exposure.[6]

**Table 2. Pairwise Session Comparisons for Reaction Times on the Ruler-Drop Test**

| Session Comparison | Mean Difference, ms | *P* Value |
|---|---|---|
| 1 versus 2 | 7.2 | .057 |
| 1 versus 3 | 6.8 | .055 |
| 1 versus 4 | 8.9 | .002[a] |
| 1 versus 5 | 11.9 | < .001[a] |
| 1 versus 6 | 10.5 | < .001[a] |
| 1 versus 7 | 10.7 | .011 |
| 1 versus 8 | 12.3 | .001[a] |
| 1 versus 9 | 10.5 | .001[a] |
| 1 versus 10 | 12.9 | < .001[a] |
| 2 versus 3 | –0.45 | 1.0 |
| 3 versus 4 | 2.1 | 1.0 |
| 4 versus 5 | 3.1 | 1.0 |
| 5 versus 6 | 1.4 | 1.0 |
| 6 versus 7 | 0.23 | 1.0 |
| 7 versus 8 | 1.6 | 1.0 |
| 8 versus 9 | –1.8 | 1.0 |
| 9 versus 10 | 2.4 | 1.0 |

[a] Indicates difference (Bonferroni-adjusted significance level of $P < .0029$).

If the ruler-drop test is to be used serially to monitor recovery from concussion, it is important to identify ways of minimizing the practice effects or eliminating them altogether. We believe the strength of our study is that test sessions were repeated 10 times. Conducting more than 2 test sessions allowed us to better assess (ie, longitudinally) the extent of these effects. Enabling people to repeat test sessions until their performance levels have stabilized makes it possible to account for and thereby limit practice effects on performance.[9] Based on the results of our study, reaction time appeared to plateau after the first test session (ie, starting with the second test session) because we noted no differences in reaction time between any of the last 9 test sessions. Therefore, we believe that in order to protect against practice-related improvements in reaction time, clinicians should allow at least 1 practice session before recording baseline measurements on the ruler-drop test. We are the first to demonstrate that the ruler-drop test—and tests similar to it—appear to be susceptible to performance improvement with repeated exposure.

A comparison of our data with those from previous studies[5–7] reveals a notable difference in reported mean reaction times. Using the ruler-drop test, we detailed an average reaction time that was 50 to 60 milliseconds slower than that for a modified version of the test.[5–7] Although the ruler-drop test and the previously studied clinical test of reaction time are conceptually similar, methodologic differences between the tests may explain the discrepancy among studies. Perhaps the most obvious and likely the most significant difference between the reaction time tests is related to visualization. With the ruler-drop test, the ruler was kept out of sight from the participant, whereas in previous studies, it was not.[5–7] Our reason for hiding the ruler from view was to limit or control the participant's urge to anticipate the release of the measuring stick. We assume that because the participant's eye level was above the height of the PVC pipe opening, the ruler would likely have fallen for a short period of time before it was noticed and then grasped by the participant, thus increasing reaction time. Additional minor differences existed between tests (eg, in previous studies, the ruler was coated or wrapped with high-friction tape, whereas our ruler was not), but we do not believe these methodologic differences had an adverse effect on reaction time. The slower reaction times we report do not in any way diminish the implication of our findings: that a clinical reaction time test that involves catching a measuring stick appears to be susceptible to practice effects.

We theorize that the improved reaction times in each subsequent test session, although minimal, were likely the result of visual feedback derived from completing earlier test trials. More specifically, we believe visual information related to how the task was completed at the outset provided valuable information to the participant, which subsequently guided him or her to make the necessary or appropriate adjustments during future test sessions.[10] In fact, it has been suggested that the practice effects associated with reaction time occur because participants learn to distinguish between their faster and slower responses, which then allows them to avoid slow responses in the future.[11]

All research investigations have limitations that affect the generalizability of the results because of the methods used. The various methodologic limitations of our investigation include the fact that participants were not questioned at all during the study regarding their motivation levels. It has been reported that the ruler-drop test is an intrinsically motivating task,[12] and our assumption was that motivation levels remained constant throughout the study. Additionally, we did not control for or monitor the participants' diet (including alcohol consumption), fatigue level, or amount of sleep, all of which are known to potentially affect reaction time.[13–16] Also, we studied a convenience sample, which consisted mostly of college-aged students; therefore, the data obtained from this group may not represent the general population. Finally, although we found that reaction time data began to plateau after the first test session, it remains unclear for how long the practice effects persist. Therefore, future researchers should attempt to determine if there is a limit to how long practice effects are retained.

## CONCLUSIONS

In summary, the ruler-drop test is susceptible to practice effects; therefore, practitioners using this test for diagnostic purposes or to track recovery from concussions should consider administering at least 1 practice session before establishing a patient's baseline value.

## REFERENCES


1. Warden DL, Bleiberg J, Cameron KL, et al. Persistent prolongation of simple reaction time in sports concussion. *Neurology*. 2001;57(3):524–526.
2. Covassin T, Elbin RJ, Nakayama Y. Tracking neurocognitive performance following concussion in high school athletes. *Phys Sportsmed*. 2010;38(4):87–93.
3. Iverson GL, Brooks BL, Collins MW, Lovell MR. Tracking neuropsychological recovery following concussion in sport. *Brain Inj*. 2006;20(3):245–252.
4. Eckner JT, Whitacre RD, Kirsch N, Richardson JK. Evaluating a clinical measure of reaction time: an observational study. *Percept Mot Skills*. 2009;108(3):717–720.
5. Eckner JT, Kutcher JS, Richardson JK. Pilot evaluation of a novel clinical test of reaction time in National Collegiate Athletic Association Division I football players. *J Athl Train*. 2010;45(4):327–332.
6. Eckner JT, Kutcher JS, Richardson JK. Between-seasons test-retest reliability of clinically measured reaction time in National Collegiate Athletic Association Division I athletes. *J Athl Train*. 2011;46(4):409–414.
7. Eckner JT, Kutcher JS, Richardson JK. Effect of concussion on clinically measured reaction time in 9 NCAA Division I collegiate athletes: a preliminary study. *PM R* 2011;3(3):212–218.
8. Udermann BE, Murray SR, Mayer JM, Sagendorf K. Influence of cup stacking on hand-eye coordination and reaction time of second-grade students. *Percept Mot Skills*. 2004;98(2):409–414.
9. Duff K, Westervelt HJ, McCaffrey RJ, Haase RF. Practice effects, test-retest stability, and dual baseline assessments with the California Verbal Learning Test in an HIV sample. *Arch Clin Neuropsychol*. 2001;16(5):461–476.
10. Abernethy B, Hanrahan SJ, Kippers V, MacKinnon LT, Pandy MG. *The Biophysical Foundations of Human Movement*. 2nd ed. Champaign, IL: Human Kinetics; 2005:227–231.
11. Maylor EA, Rabbitt PM, James GH, Kerr SA. Effects of alcohol, practice, and task complexity on reaction time distributions. *Q J Exp Psychol A*. 1992;44(1):119–139.
12. Eckner JT, Chandran S, Richardson JK. Investigating the role of feedback and motivation in clinical reaction time assessment. *PM R*. 2011;3(12):1092–1097.
13. Cote KA, Milner CE, Smith BA, et al. CNS arousal and neurobehavioral performance in a short-term sleep restriction paradigm. *J Sleep Res*. 2009;18(3):291–303.
14. Hernández OH, Vogel-Sprott M, Ke-Aznar VI. Alcohol impairs the cognitive component of reaction time to an omitted stimulus: a replication and an extension. *J Stud Alcohol Drugs*. 2007;68(2):276–281.
15. Jauch-Chara K, Hallschmid M, Schmid SM, Bandorf N, Born J, Schultes B. Sleep loss does not aggravate the deteriorating effect of hypoglycemia on neurocognitive function in healthy men. *Psychoneuroendocrinology*. 2010;35(4):624–628.
16. Van den Berg J, Neely G. Performance on a simple reaction time task while sleep deprived. *Percept Mot Skills*. 2006;102(2):589–599.


*Address correspondence to Gianluca Del Rossi, PhD, ATC, University of South Florida, 13220 USF Laurel Drive, MDF 5117, Mail code MDC 106, Tampa, FL 33612. Address e-mail to gdelross@health.usf.edu.*