J Forensic Sci, March 2008, Vol. 53, No. 2
doi: 10.1111/j.1556-4029.2008.00670.x
Available online at: www.blackwell-synergy.com

Everett B. Lohman III,[1] *D.P.T.Sc, P.T. O.C.S.; Donna L. Thorpe,*[1] *Dr.P.H.; Marilyn Prior,*[1] *M.P.T., P.T.; James George,*[1] *D.P.T., P.T.; and Jimmy P. Kim,*[1] *D.P.T., P.T.*

Can APB 2000 Be Used to Discern Sincerity of Effort in Unimpaired Subjects from Maximal Performance in Subjects with Shoulder Pain?

ABSTRACT: The automated pegboard (APB 2000), which has been found to objectively quantify motor performance, was used to differentiate maximal motor performance among subjects with shoulder pain, healthy unimpaired subjects performing normally and also while feigning shoulder pain. Six participants with shoulder pain and 15 healthy unimpaired individuals participated. Individuals with shoulder pain were tested on the APB 2000 using their affected upper extremity. Unimpaired participants were instructed to perform normally on the test with randomly selected upper extremity and to feign shoulder pain with the other upper extremity. The two tests for the unimpaired participants were conducted 1 week apart. There were significant differences in mean performance time for normal, patient, and feigned performance, with 80, 111, and 149 sec for the three groups respectively ($p < 0.0005$). There was also considerable overlap in the three distributions of performance times. These preliminary findings suggest that the APB 2000 is able to distinguish performance time between these three groups. Whether it can be used to distinguish between maximal performance and submaximal performance in individuals suspected of submaximal performance requires further study.

KEYWORDS: forensic science, shoulder pain, feigned effort, motor performance, automated pegboard

Suspect effort, or malingering can be broadly defined as all forms of fraud relating to matters of health (1). More specifically, it is the simulation or exaggeration of a physical or mental disease or defect (1). Malingering, exaggeration of pain symptoms, feigning weakness, or submaximal effort has been a continued problem in the medical legal arena (2–4). Malingering for secondary gains costs millions of dollars a year in compensation. In efforts to control fraudulent claims, multiple studies and tests are used by medical practitioners in attempts to identify the malingering patient (2,3).

In a study reviewing tests currently being used to identify malingering, none had been found to consistently succeed in distinguishing cases from noncases among patients with symptoms of chronic pain (3). There are a significant number of studies on malingering using neuropsychological tests to detect intentional dysperformance. Fishbain et al. (3) reviewed studies using questionnaires to detect malingering of pain and determined that "Overall, this research evidence indicates that a pain malingering profile cannot be identified by questionnaire."

The Jamar (dynamometer handgrip) has been used in studies attempting to determine malingering. Loss of grip strength is a frequent problem encountered in patients with chronic pain. "Loss of grip strength is a ratable item in the determination of permanent impairment; therefore, it is important to determine whether such loss of strength is malingered" (3). After reviewing six studies utilizing the Jamar, Fishbain et al. (3) concluded that not all studies were successful in utilizing the Jamar for detection of malingering.

Isokinetic testing has also been used in efforts to determine whether the patient with chronic pain has produced a submaximum rather than maximum effort (3). Isokinetic testing, which measures the amount of resistance according to the subject's muscular contraction, offers an advantage over the isometric testing (Jamar) in determining submaximal effort. The computerized isokinetic dynamometer used in isokinetic testing can be a powerful tool in measuring maximal or feigned muscle effort (4).

Fishbain et al. criticizes the use of the concept of submaximal effort to detect malingering in chronic pain patients (CPPs) (3). The researchers opine that the following factors may result in legitimate submaximal effort in CPPs: muscle fatigue, pain, or psychological factors such as anxiety, depression, or fear of pain (3). These factors would make it difficult to determine the occurrence of simulated pain. After reviewing multiple studies on identifying malingering, Fishbain et al. (3) concluded that there are currently no reliable methods to identify malingering patients with chronic pain.

The automated pegboard (APB 2000) has been shown to be a reliable tool in objectively measuring the temporal components of motor performance, or the time taken to complete a motor task as specified by the computer in the pegboard (5). The total time required for each stage of the motor task, which includes the time to comprehend instructions, as well as executing, and completing the instructions (perception, cognition, and action), are objectively measured, tracked, and reported statistically and graphically. The system also measures hand, arm, and upper body dexterity.

The purpose of this study was to use the APB 2000 to compare data on performance patterns of simulated malingering as compared to nonmalingering performance by healthy subjects and patients with shoulder pain. As shoulder pain complaints are common among working age adults (6), this study proposed to use the APB 2000 to determine if there are systematic differences in test time between (1) genuine shoulder pain patients and subjects who simulate pain and (2) simulated malingering and maximal performance by healthy controls. Because malingerers tend to exhibit submaximal effort during performance tests (2–4,7), it is our hypothesis that

[1]Department of Physical Therapy, Loma Linda University, School of Allied Health Professions, Loma Linda, CA 92350.

Received 20 Jan. 2007; and in revised form 18 Aug. 2007; accepted 19 Aug. 2007.

subjects simulating malingering on the APB 2000 pegboard will require more time to complete a given motor task compared to the time taken by a person with real shoulder pain.

Methods

Participants

Two groups of participants, patients with shoulder pain, and healthy unimpaired individuals, aged 21–60 years, participated in this study. Six patients undergoing physical therapy for genuine shoulder pain with mild impairments and functional limitations were recruited from two local outpatient clinics by their physical therapist. A visual analog scale for pain was not collected. These individuals were able to perform normal activities of daily living with at least fair muscle grade for abduction, flexion, and horizontal adduction using manual muscle testing. They were also required to have shoulder range of motion of at least 0–140° of flexion and abduction, 0–120° of horizontal adduction, and an overall grip strength of 5 lbs of force using a Jamar hand dynamometer. Fifteen physical therapy students served as healthy unimpaired participants. Individuals were excluded from this study if they had complaints of neck pain radiating to the shoulder or if they were unable to read and execute the commands on the pegboard. Age for each subject was collected; however, height and weight were not because the APB 2000 is adjustable to accommodate each subject's stature. All study participants signed an informed consent approved by the Institutional Review Board.

Instruments

The Reflex Development & Testing's Automated Peg Board System 2000 (APB 2000) consists of a computer-driven panel containing 21 colored and shaped pegs in precisely located sockets (Fig. 1). It measures the display time (the time taken by the subject to read the instructions given by the computer and then remove the designated peg from its place), distance time (the time to remove the single peg and place it in the designated vacant slot), and the return time (the time to hit the home switch after placing the peg in its slot) for each of the 21 moves. The computer also measures the total trip time (the time taken by the subject to complete all the 21 moves) during a single trial.

The APB 2000 had been found to be a highly reliable tool to measure performance time in both pathological and nonpathological populations (5). The intraclass correlation coefficient for the mean of repeated measurements, used to establish reliability, exceeded 0.90.

FIG. 1—*Automated peg board system (APB) 2000.*

Procedures

After informed consent was obtained, screening procedures for eligibility were conducted. The subjects were screened for any cognitive deficits using the Mini-Mental State Examination. The subjects were required to score a minimum of 24 out of the total 30 points to be included in the study. They were also screened for visual deficits using the Titmus Vision Test and asked to read a short sentence within 8 sec that was in the same font size as that of the computer screen display and at a distance of 3 feet. Upper extremity range of motion (ROM) of the involved shoulder in subjects with shoulder pain was measured using a goniometer by a single examiner, a physical therapist with 20 years of experience, to determine eligibility according to the inclusion and exclusion criteria. Intra-rater reliability was not performed.

Healthy individuals and individuals with shoulder pain completed six consecutive 21-move trials on the APB 2000. Following maximal performance, the healthy participants were asked to simulate malingering or suspect performance by falsifying the performance of an individual with legitimate shoulder pain during a second set of six trials using the remaining untested upper extremity. The operational definition of a malingering or suspect performance in this article is a feigned maximal performance with an intention to exaggerate performance impairments from a feigned shoulder injury. All subjects were given two practice trials before they performed the six test trials.

A random assignment by block was employed to assign the testing order of the faked or genuine maximal performance to the healthy participants. Both upper extremities were tested for the faked or genuine maximal performance to the healthy participants. A second random assignment determined which upper extremity was to be used for the first set of six trials (dominant or nondominant). The second set of six trials was performed 1 week later.

Data Analysis

The mean performance time (in seconds) per trial for the six trials for maximal patient performance and simulated malingering were compared using the independent *t*-test. Mean performance time for healthy maximal performance was compared to feigned maximal performance using the matched pairs *t*-test. Mean time for the first three trials compared to the second three trials was also made using the matched pairs *t*-test.

Simple linear regression analyses were run to determine the tolerance limits with 95% prediction intervals, for both healthy normal and feigned maximal performance. Tolerance limits are boundary values for a variable expected with 95% of a target population. In our study, the values for the two target populations were patients performing maximally and malingerers. A previously collected dataset of pegboard performance time by healthy unimpaired subjects between the ages of 20 and 60 years was used as a surrogate for the patient population to estimate tolerance limits for performance time, as well as maximal performance time by participants in this study.

Finally, receiver operator curve (ROC) was examined for potential cut-off values for performance time to define feigned performance with desired specificity assumed to be at least 95%. Data were analyzed using the SPSS version 12.0 program (8).

Results

Fifteen individuals were recruited from the population of physical therapy students and faculty to perform as healthy participants and again as malingerers. Six patients who were being treated for shoulder pain were recruited from a local physical therapy clinic. The healthy individuals, comprising seven males and eight females, were between the ages of 21and 60 with a mean age of 36.6 years (Table 1). There were three males and three females with a mean age of 47.8 among patients with shoulder pain. The mix of males to females was similar in the control and patient groups ($p = 0.82$). There was no significant difference in age between healthy individuals and shoulder pain patients ($p = 0.07$). Three patients had had recent surgery for rotator cuff repair and three had diagnoses that included rotator cuff tendonitis and/or adhesive capsulitis. All had pain with reaching in daily activities.

The mean performance time per trial over six trials for controls, patients, and feigners was 80, 111, and 149 sec, respectively (Fig. 2). There were significant differences between the mean performance time for maximal healthy versus feigned effort ($p < 0.0005$), maximal healthy versus maximal patient effort ($p = 0.007$), and feigned versus maximal patient effort ($p = 0.008$).

Performance time for each of the six trials is displayed in Fig. 3. The maximal healthy performance, maximal patient, and feigned effort exhibit consistent differences over all six trials. The feigned performance time is markedly higher compared to the maximal healthy and maximal patient effort. In all three comparison conditions, there is a significant improvement in mean performance time over trials 4, 5, and 6 when compared to mean performance time over trials 1–3: maximal healthy controls ($p < 0.0005$), patients ($p < 0.001$), and submaximal or feigned effort ($p < 0.0005$).

TABLE 1—*Comparison of upper extremity motor performance on the pegboard test by subjects with shoulder pain, healthy subjects feigning a painful shoulder and performing normally.*

	Control ($n = 15$)	Feigners ($n = 14$)	Patients ($n = 6$)	p-value
Gender n, [%]				
Male	7 (46.6%)	6 (42.9%)	3 (50%)	0.82*
Female	8 (53.3%)	8 (57.1%)	3 (50%)	
Age (years) mean, [SD]	36.6 (12.6)		47.8 (10.9)	0.06†
Total time (in sec), mean, [SD]	79.9 (15.8)	146.0 (32.2)	111.1 (18.0)	

*Chi-square test.
†Independent t-test.

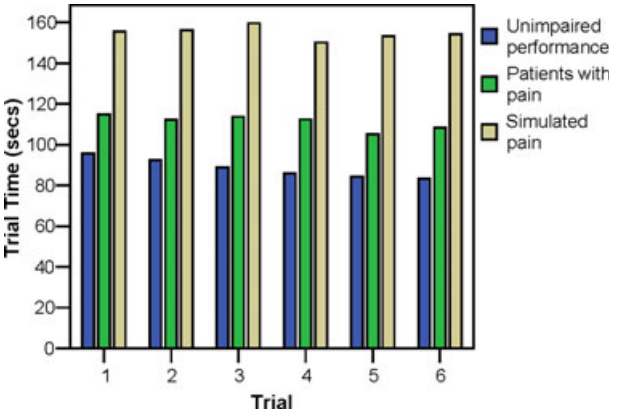

FIG. 3—*The performance time for each of the six trials for each comparison. In all three comparison conditions, there is a significant improvement in mean performance time over trials 4, 5, and 6 when compared to mean performance time over trials 1–3.*

The standard deviation (SD) for the distribution of consecutive trial scores for the first three trials, the first four, the first five, and finally all six trials were calculated for each participant. Figure 4 displays the SD for the four distributions for all three comparisons. The healthy normal performance shows minimal variability over trials 3–6. The patients begin to show an increase in variability with the additions of trials 5 and 6. The submaximal performances consistently showed marked variability over all the trials, but began to lessen by the cumulative sixth trial.

To establish a test to detect feigned performance it is essential to describe the range of normal performance, and to compare it to that of feigned performance (9). Table 2 shows that although there is a large difference in mean performance time between normal and feigned performance, there is a large overlap of at least 30 sec (88.5–121.9 sec) between the two distributions. There is an even greater overlap between the range of maximal patient performance

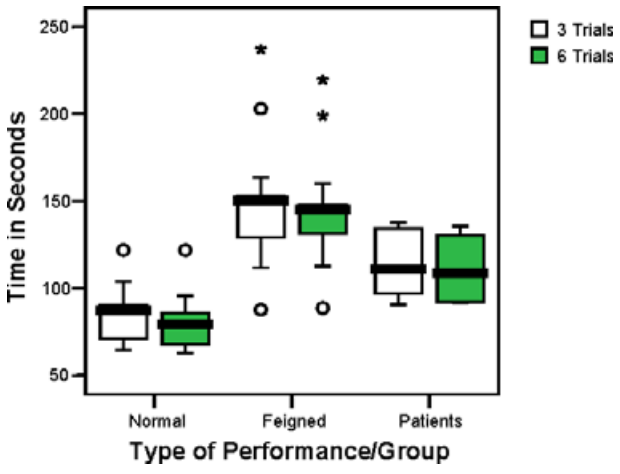

FIG. 2—*Comparison of time per trial over three and six trials for normal and feigned performance for healthy subjects and normal performance for patients.*

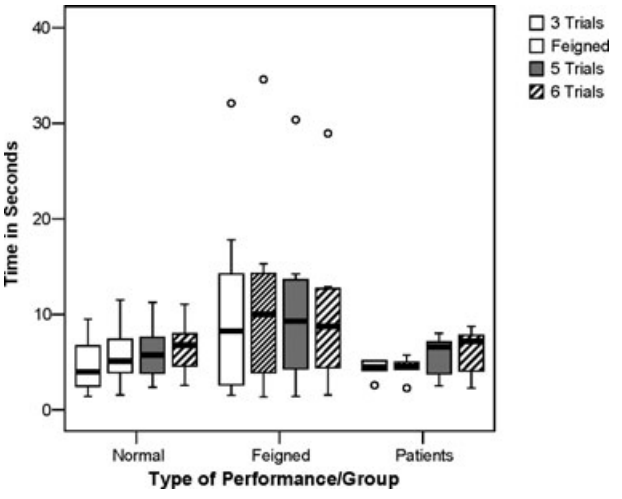

FIG. 4—*Comparison of standard deviations for performance time over three to six trials for normal and feigned performance by healthy subjects and normal performance by patients.*

TABLE 2—*Comparison of statistics for performance time/trial for three trials of maximum honest effort for normal subjects and patients, and for feigned effort compared to estimated tolerance limits (TL) for the target population.*

					95% TL for comparison groups		95% TL for feigned performance	
	n	Mean	Min	Max	Lower	Upper	Lower	Upper
Feigned performance	14	148.77	88.5	219.6				
Patient performance	6	113.60	91.5	135.6	39.4	187.8	77.6	219.9
Maximal performance	15	83.55	62.7	121.9	24.6	142.5	89.7	207.9

TABLE 3—*Comparison of the coefficient of variation (CV) for performance time/trial for three trials of maximum honest effort for normal subjects and patients and for feigned effort.*

					95% TL for comparison populations		95% TL for feigned performance	
	n	Mean %	Min	Max				
Feigned performance	14	6.25	1.37	13.53				
Patient performance	6	3.98	1.93	5.30	0	12.14	0	14.06
Maximal performance	15	5.70	2.11	12.41	0	13.63	0	14.20

TL, tolerance limit.

TABLE 4—*Maximal performance time data over three trials previously collected on unimpaired reference groups by age compared to estimated tolerance limits (TL) for the target population obtained on the same data.*

							Upper TL for comparison groups		Lower TL for feigned performance
	n	Mean	Min	Max	95th percentile	99th percentile	95%	99%	95%
40–59	40	96.64	70.2	121.6	121.2	124.6	141.9	156.7	103.1
20–39	30	83.18	58.9	115.7	107.8	117.9	129.6	146.0	99.4

and the range of values for feigned performance time (88.5–135.6 sec). Because these values were based on small sample sizes, tolerance limits estimating the boundaries for performance time for 95% of the target population are also presented. This widens the overlap to roughly 40 sec between normal and feigned performance. When tolerance limits for feigned performance time is compared to patient performance, the overlap more than doubles. However, this large an overlap is likely to be an overestimation strongly influenced by the very small sample of patients.

To examine consistency of performance across repetitions the coefficient of variation (CV) was calculated for each subject under the three conditions (Table 3). Mean CVs for the three distributions varied little, from 3.98 to 6.25%. Not surprisingly, the range of values for the CVs of the three conditions almost totally overlapped.

Performance time for the three conditions were also compared to data previously collected on 70 physically and cognitively unimpaired subjects, aged 20–59, during maximal performance (Table 4). A comparison of maximal performance for the 15 subjects in this study reveals that mean performance by the subjects is typical of the 20–39 age group. However, the maximum time for the study participants is similar to the maximum for the 40–59 year age group. Both of these observations are consistent with the fact that the 15 subjects were largely from the younger age group, with a few in the older group. The 95% upper tolerance limit for normal maximal performance based on this larger dataset of 40–59 age group (141.9) is very close to the upper limit (142.5) for the normal performance of study participants.

The ROC analysis was conducted using feigned performance time as positive outcomes and normal performance by patients, study participants, and participants from a previous study, as negative outcomes. With specificity at least 95% chosen as the criteria for a legitimate test of feigning, the analysis yielded a score of 121.6 sec as the smallest cut-off score and a sensitivity of 86%. The area under the curve was 0.94 (95% CI 0.86, 1.02). The prevalence of feigning in the sample was 13%.

Discussion

Two parameters, consistency of effort as measured by the CV for performance time and performance time per trial, were investigated as potential tests to distinguish between feigned and honest effort. According to Guides for the Evaluation of Permanent Impairment, one of the most applied criteria for honest effort has been consistency of effort (10). In the physical performance domain, the CV has successfully been used to discriminate feigned and normal behavior for active cervical ROM (10), but not for grip strength (9). From the almost complete overlap of CV values for timed performance on the pegboard by all three performances in this study, the CV cannot be useful for distinguishing feigned from normal unimpaired or patient performance.

On the other hand, the difference in the distribution between maximal performance time and feigned performance time by the same individuals is significant, suggesting that performance time may be a legitimate parameter for distinguishing between honest and feigned effort. However, in the medico-legal arena, a test of honest effort must distinguish between patients with genuine impairments of the upper extremity, which may affect performance time. From the small number of patients with shoulder pain, it is apparent that feigned performance may not be easily distinguished from honest patient effort.

For the exploratory ROC analysis for potential cut-off values of performance time for a proposed test of honest effort, we assumed that a specificity of at least 95% was desirable. The sample used for the analysis consisted of the 15 feigned performances, the six patients, and the 40 unimpaired individuals who were 40 and older. The results showed that the minimum cut-off value would be 126.9 sec for such a test, which would then have a sensitivity of 79%. This would mean that for every disability claim determined to be genuine, 5% or less would test positive for feigning and 79% of feigners would be positively identified. These results should be considered very preliminary, as this is based on several key assumptions which cannot be validated at this point.

Appropriate ROC analysis requires the selection of a valid sample that represents the target population and is adequate in size (11). Because there were only six patients, maximal performance by unimpaired 40–59-year-old individuals was included in the sample. Whether or not performance time of these unimpaired individuals is reasonably similar to patients to be considered appropriate surrogates remains to be seen. The 95% tolerance limits from the regression analysis suggests that the overlap in performance time for patients and those who would fake maximal performance in the target population could be considerably greater that seen in the sample used. Furthermore, it is also not known how similarly the performance of participants who are asked to feign would be to those who would actually do so. In studies of feigning cognitive impairments, it appears that feigning by students in psychology is more difficult to detect than by nonstudent adults (11).

Finally, a valid estimate of the prevalence of feigning among those who would have an opportunity to do so is also needed, as an adequate sample size depends on the prevalence of a positive test (12,13). For example, if the prevalence of feigning is 1 per 100, a sample size of 1000 is required (12). Thus, further research is required to verify the findings of this study.

Conclusion

This is the first study to examine the possibility of using a functional test in the physical domain, rather than a test of impairment such as strength or ROM, to identify submaximal performance. The difference in mean performance time among maximal normal effort, maximal patient effort, and feigned maximal effort by healthy participants, was significant. Yet, there is also considerable overlap in the distribution of these three sets of performances. Preliminary evidence suggesting that the APB 2000 pegboard may be able to detect submaximal performance in a population submitting claims of disability appears encouraging, but will require further study.

References

1. Gorman WF. Defining malingering. J Forensic Sci 1982;27(2):401–7.
2. Resnick PJ. The detection of malingered mental illness. Behav Sci Law 1984;2(1):21–38.
3. Fishbain DA, Cutler R, Rosomoff HL, Rosomoff RS. Chronic pain disability exaggeration/malingering and submaximal effort research. Clin J Pain 1999;15(4):244–74.
4. Dvir Z. Reproducibility of performance and certainty of judgment in maximal vs. feigned muscular effort. Percept Mot Skill 1999;88(3 Pt 3):1078–80.
5. Lohman EB, Johnson EG, Miguel AM, Cordett TK, Kang D. Automated pegboard system. Reliability and validity of a new tool. J App Res 2003;3(3):262–9.
6. Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, et al. Multidisciplinary biopsychosocial rehabilitation for neck and shoulder pain among working age adults. Cochrane Database Syst Rev 2003;2:CD002193.
7. Mittenberg W, Payton C, Canyock CM, Condit DC. Base rates of malingering and symptom exaggeration. J Clin Exp Neuropsychol 2002;24(8):1094–102.
8. SPS Inc. SPSS user's guide, Version 10.0 for Windows. Chicago, IL: SPS Inc., 1999.
9. Dvir Z. Coefficient of variation in maximal and feigned static and dynamic grip efforts. Am J Phys Med Rehabil 1999;78(3):216–21.
10. Dvir Z. Maximal versus feigned active cervical motion in healthy patients. Spine 2001;26(15):1680–8.
11. Haines ME, Norris MP. Comparing student and patient simulated malingerers' performance on standard neuropsychological measures to detect feigned cognitive deficits. Clin Neuropsychol 2001;15(2):171–82.
12. Westin LK. Receiver operating characteristic (ROC) analysis: evaluating discriminance effects among decision support systems. 2001; http://www.cs.umu.se/research/reports/show.cgi?year=2001&nr=018, accessed August 31, 2006.
13. Rosenfeld B, Sands SA, Van Gorp WG. Have we forgotten the base rate problem? Methodological issues in the detection of distortion Arch Clin Neurophychol 2000;15(4):349–59.

Additional information and reprint requests:
Dr. Everett B. Lohman III,
Associate Professor
Department of Physical Therapy
Loma Linda University
Loma Linda, CA 92350
E-mail: elohman@llu.edu