Using Grip Strength Force Curves to Detect Simulation: A Preliminary Investigation

Jeremy J. Davis¹,†, Jacqueline R. Wall¹,*, Crystal K. Ramos¹, Kriscinda A. Whitney²,³, Mark T. Barisa⁴

¹School of Psychological Sciences, University of Indianapolis, Indianapolis, IN, USA
²Richard L. Roudebush Veterans Affairs Medical Center, Indianapolis, IN, USA
³Indiana University School of Medicine, Indianapolis, IN, USA
⁴Baylor Institute for Rehabilitation, Dallas, TX, USA

*Corresponding author at: University of Indianapolis, 1400 East Hanna Avenue, Indianapolis, IN 46227, USA. Tel.: +1-317-788-6142; fax: +1-317-788-2120.
E-mail address: jwall@uindy.edu (J.R. Wall).

Accepted 8 January 2010

Abstract

Analysis of grip strength force curves has successfully detected suboptimal effort in industrial rehabilitation research. This study examined force curve analysis as an effort measure when grip strength was administered according to standardized neuropsychological procedures in a sample without reported neurologic and upper extremity injury. Eighty-two undergraduates were randomized to control (n = 26), naïve simulator (n = 28), and coached simulator (n = 28) conditions. Outcome measures included grip strength in kilograms, variables calculated from grip strength force curves, and the Word Memory Test (WMT). While average force in kilograms was not significantly different between groups, significant differences were found on the average to peak force ratio as calculated from grip strength force curves. The classification accuracy of average to peak force ratio was lower than the WMT, but comparable to other effort measures. Force curve analysis may warrant further study in a clinical sample.

Keywords: Grip strength; Effort assessment; Symptom validity testing; Simulated malingering

Introduction

Current practice guidelines recommend a multimethod approach to effort assessment that may include clinical observations, freestanding symptom validity tests, and embedded effort measures (American Academy of Clinical Neuropsychology, 2007; Bush et al., 2005). Given the need for clinical efficiency, it is important to maximize diagnostic yield while minimizing test administration time. Since embedded effort measures are derived from other neuropsychological instruments, they increase the data from which inferences regarding effort can be drawn without adding length to the test battery. Much of the research on effort assessment with standard neuropsychological instruments has focused on tests of intellectual ability and memory (Iverson & Binder, 2002; Nies & Sweet, 1994), but other components of a full neuropsychological evaluation have shown promise as effort measures including tests of motor function (Arnold & Boone, 2007; Greiffenstein, 2007).

Motor measures were included in a seminal study in the effort assessment literature that compared clinical judgment and discriminant function analysis in the detection of simulated malingering (Heaton, Smith, Lehman, & Vogt, 1978). Grip strength and finger tapping were among the tests that significantly contributed to the discriminant function, which was more accurate than clinical judgment at detecting simulators. More recent research has examined the classification accuracy of cut scores for finger tapping (Arnold et al., 2005) and the impact of coaching on grooved pegboard performance (Johnson & Lesniak-Karpiak, 1997). Concurrent with studies of single motor measures, other research has explored the configuration of...
scores on multiple motor measures including the profile of grip strength, finger tapping, and grooved pegboard (Greiffenstein, Baker, & Gola, 1996) and variability across trials of grip strength and finger tapping (Wall & Millis, 1998).

Greiffenstein and colleagues (1996) reported a motor profile among patients with postconcussive syndrome that departed from a hierarchical theory of motor function (Haaland & Yeo, 1989). Greiffenstein and colleagues found that patients with a history of traumatic brain injury (TBI) demonstrated reduced performance in the order of increasing task complexity: Grip strength performance was better than finger tapping, which was better than grooved pegboard. Patients with postconcussive syndrome, however, showed a different pattern with worse performance on grip strength than on finger tapping and grooved pegboard, which the authors described as a “nonphysiologic and functional” (p. 483) profile.

In an effort to increase understanding of both neurologically normal and nonphysiologic motor profiles, Rapport, Farchione, Coleman, and Axelrod (1998) examined the motor profiles of nonclinical participants randomized to control, and naive and coached simulated malingering conditions. Control participants demonstrated significantly higher scores than both groups of simulators on all motor tests. Naive simulators performed significantly worse on grooved pegboard than on grip strength and finger tapping, which were not significantly different. Coached simulators performed worse on finger tapping than grooved pegboard, but neither measure was different from grip strength. Although motor profiles did not closely follow the expected patterns, the large effect sizes on motor measures (e.g., average Cohen’s $d = 1.19$ for grip strength) suggested that, given the right method of analysis, these tasks might be useful in differentiating effort.

Additional support for the use of motor measures in effort assessment emerged in an examination of grip strength and finger tapping in a sample of brain injury simulators (Wall & Millis, 1998). Simulator and control groups were significantly different on grip strength and finger tapping. The authors calculated standard deviations of grip strength and finger tapping trials of individual participants. Control participants showed significantly greater variability across finger tapping trials, bilaterally. Injury simulators showed significantly greater variability on dominant-hand grip strength trials; differences in variability on nondominant-hand grip strength trials were not significant. These findings suggest that consideration of performance during motor tasks in addition to the total score might improve the differentiation of legitimate deficits from suboptimal performance. Although grip strength total scores, for example, might be low in both cases, additional variables describing the grip strength trial might differ in detectable ways.

**Effort Assessment in Industrial Rehabilitation**

Obtaining best effort performance on grip strength is especially important to evaluations conducted in the context of industrial rehabilitation and associated disciplines like occupational and physical therapy (Innes, 1999). Effort assessment in these areas has been studied with several approaches including comparing grip strength across dynamometer handle positions (Niebuhr & Marion, 1987), considering cut scores based on the coefficient of variation of grip strength trials (Shechtman, 2001a, 2001b), and using variables calculated from grip strength force curves (Shechtman, Sindhu, & Davenport, 2007). Force curves of grip strength may provide a more accurate characterization of the examinee at work than the static representation given by the maximum grip in kilograms.

In one of the earliest studies using grip strength force curves, Gilbert and Knowlton (1983) assigned nonclinical volunteers to sincere and insincere groups with respective instructions to perform to the best of their ability or to give a high submaximal effort with the appearance of best effort. Participants were instructed to squeeze the dynamometer to maximal effort and maintain it for 5 s. Several force curve variables were examined including slope to peak force, average to peak force ratio, and peak force to body weight ratio. Slope to peak force was calculated using the standard formula for the slope of a line with time and force plotted on the $x$- and the $y$-axes, respectively. Average to peak force ratio was calculated by dividing average force during the middle 3 s of the trial by peak force. Peak force to body weight ratio was calculated by dividing peak force by body weight. Separate stepwise discriminant function analyses were performed for female and male participants with the three force curve variables as predictors. For women, the average to peak force ratio was the only significant predictor, accurately classifying 94% of participants. For men, slope to peak force and average to peak force ratio were significant predictors, accurately classifying 85% of participants.

Other studies have also shown the utility of force curve variables in grip strength effort assessment. Smith, Nelson, Sadoff, and Sadoff (1989) administered two sets of six trials (three per hand) with instructions to produce best effort with one hand and to feign best effort (50% of maximum) with the other hand in each trial. Five force curve variables were calculated: Coefficient of variation, average to peak ratio, ratio difference, peak-average difference, and peak-average root difference. The coefficient of variation was calculated by dividing the standard deviation of force by the average force during the interval from the time at 90% of peak force through the end of the trial. Average to peak ratio was calculated as described by Gilbert and Knowlton (1983). The rationale for and derivation of the latter three variables are beyond the scope of this review and are covered in detail by Smith and colleagues (1989). Cut scores based on these variables correctly identified 52.2%—87.0% of women on
feigning trials and 92.5%–100% of men on feigning trials. Composite cut scores based on multiple discriminator variables increased the classification rates of women on feigning trials to 84.8%–93.5%. Chengalur, Smith, Nelson, and Sadoff (1990) extended the findings of Smith and colleagues (1989) in a clinical sample with unilateral hand injury. The discriminator variables, calculated as above, identified 28.3%–79.2% of feigning trials for injured hand in women and 30.0%–85.0% of feigning trials for injured hand in men.

Although industrial rehabilitation research has evidenced the utility of force curve analysis in grip strength effort assessment, the method of administering grip strength was often different from the standardized instructions common to neuropsychological assessment (Reitan & Wolfson, 1993). Some studies specified the duration of the grip strength trial (Gilbert & Knowlton, 1983), and others varied the handle position of the dynamometer (Niebuhr and Marion, 1987). Differences in administration methods might affect examinee performance and limit the generalizability of the findings to evaluations using a different administration method.

Research Aims and Hypotheses

The goal of the current study was to extend neuropsychological research in effort assessment using grip strength (Rapport et al., 1998; Wall & Millis, 1998) by incorporating the methodology of force curve analysis from industrial rehabilitation (Gilbert & Knowlton, 1983; Smith et al., 1989). Given the promising results of a pilot study (Davis, Killilea, & Wall, 2007), force curve analysis was compared with standard grip strength assessment (Reitan & Wolfson, 1993) and a widely used symptom validity test. Consideration of findings from the industrial rehabilitation literature and pilot study suggested five force curve variables that were likely to classify effort: Trial length, average force, peak force, slope to peak force, and average to peak force ratio.

It was hypothesized that control and simulator participants would perform differently on standard grip strength assessment with control participants showing greater grip strength than both naïve and coached simulators. Force curve variables were hypothesized to differentiate control participants from naïve and coached simulator participants. Specifically, control participants were expected to show shorter trial length, increased average and peak force, greater slope to peak force, and higher ratio of average to peak force compared with simulators. Finally, it was hypothesized that force curve analysis would distinguish control and simulator participants at a level of accuracy comparable to that achieved by other methods of determining examinee effort.

Materials and Methods

Participants

Participants were solicited from introductory psychology courses and the general student population at a Midwestern university. Institutional review board approval was obtained before initiating recruitment, and participants gave informed consent prior to beginning the study. Volunteers currently enrolled in introductory psychology courses received research participation credit, and all volunteers who completed the assessment procedures were compensated with a $10 gift card to a local store. Inclusion criteria included age of at least 18 years, current student status, and English fluency. Exclusion criteria included hypertension; history of injury or motor impairment in the arm or hand; and history of neurological condition.

Ninety-one individuals volunteered for the project; nine were excluded due to the history of neurological condition (n = 3), broken wrist (n = 2), hand ligament injury (n = 1), and finger injury (n = 3). The final sample consisted of 82 adults randomly assigned to control (n = 26), naïve simulator (n = 28), and coached simulator (n = 28) groups. Seventy-two participants were woman (88%), and the average age of participants was 20.6 years (SD = 4.7). In terms of ethnicity, the sample was 84% European American, 15% African American, and 1% Latino/Hispanic. The average educational level was 12.9 years (SD = 1.1), and 90% of participants were right-handed.

Measures

Smedley hand dynamometer. Grip strength in kilograms was measured using the Smedley hand dynamometer (Reitan & Wolfson, 1993). Average raw scores in kilograms from two trials (or the most consistent two trials of three, if the first and second trials differed by more than 5 kg) per hand were averaged across dominant and nondominant hands.

Digitally modified dynamometer. Grip strength was also assessed with a digitally modified Jamar dynamometer (Asimov Engineering, n.d.). The Jamar dynamometer is a hydraulic device, and the digital modification involved replacing the
analog dial, which normally provides a reading in kilograms, with a pressure to voltage transducer (Ashcroft model K 17M0215F220). The transducer was connected to a digital voltage meter (Voltcraft model DMM M-3890D USB Multimeter) that provided a continuous reading of the voltage. In this way, the voltage readings provided a measure of grip strength, albeit converted from the usual metric. Voltage readings were captured at 0.1-s intervals by the voltage meter and were transmitted via universal serial bus port to a computer for storage and analysis. Force curves were plotted using a program distributed by the manufacturer of the meter (DMM-Profilab) and were imported into Microsoft Excel for analysis. Force curve variables included trial length (s), average force (mV; defined as the sum of force readings over the entire trial divided by the length of the trial in tenths of seconds), peak force (mV; maximum force over the entire trial), slope to peak force (derived from the general equation for the slope of a line, this variable was defined as peak force divided by time at peak force), and average to peak force ratio (defined as average force divided by peak force).

Reading quiz. A nine-item, multiple-choice measure was administered to assess comprehension of information in brief reading passages that were distributed prior to the instructions. The quiz was used as a manipulation check in previous research (Rapport et al., 1998). Participants in control and naïve simulator groups received a passage and quiz on spinal cord injury; participants in the coached simulator group received a passage and quiz on minor TBI.

Adherence questionnaire. The adherence questionnaire is a combination of Likert items designed to assess participants’ adherence to instructions and perceived success at completing the tasks according to the instructions. All participants rated the extent to which they followed instructions and believed their efforts were successful on a scale from 1 (“not at all”) to 5 (“very much so”). Participants in simulators groups also rated the extent to which they believed that their test performance was convincing to the test administrator.

Word Memory Test. The Word Memory Test (WMT; Green, 2002) is a computer-administered test of learning and effort. The effort portion includes three relatively easy subtests comprised of forced-choice recognition items; the test also has four subtests that are comparable to other list learning tasks. The present examination focused on the effort subtests: Immediate Recognition, Delayed Recognition, and Consistency. Interpretation of these subtests is detailed in the test manual and is not discussed here to protect test security.

Procedures

Informed consent was obtained from participants after verifying that they met inclusion criteria and did not have a history of upper extremity injury or neurologic condition. Next, participants were given a reading passage and quiz that varied according to condition (Rapport et al., 1998). Participants randomly assigned to control and naïve simulator groups received a passage and quiz on spinal cord injury; participants in the coached simulator condition received a passage and quiz on brain injury. Following completion of the quiz, participants were provided with instructions that varied according to condition. Naïve simulators were instructed to feign symptoms due to an accident but were not given specific instructions on how to do that. Coached simulators were instructed to feign symptoms due to an accident and were given some guidelines on how to better feign impairment without detection. The instructions were modified from those used in other simulation studies (Martin, Hayes, & Gouvier, 1996) and are available from the authors on written request.

After participants read the instructions, clinical psychology doctoral students, who were blind to participant condition, administered the WMT and both grip strength measures. Grip strength measures were counterbalanced to reduce potential order effects. As this project was part of a larger research effort examining both embedded effort measures and separate symptom validity tests, additional neuropsychological instruments were administered next in standardized order. Following test administration, participants completed the adherence questionnaire. Research assistants debriefed participants regarding the aims of the project and distributed remuneration.

Statistical Analyses

Groups were compared in terms of gender, ethnicity, handedness, and responses to the adherence questionnaire using \( \chi^2 \) analyses. Univariate analysis of variance (ANOVA) was used to compare the groups on age, education, and performance on the reading quiz. Preliminary data exploration revealed skewed distributions and heterogeneity of variance on some force curve and WMT variables, so nonparametric analyses were used to examine the research hypotheses. The Kruskal–Wallis test was used with group as the independent variable and force curve variables, grip strength in kilograms, and WMT scores as dependent variables. Given the number of analyses, \( \alpha \) was set at 0.01 to reduce the likelihood of a Type I
error. Post hoc analyses (Mann–Whitney) were conducted to explore further the relationships demonstrated by the Kruskal–Wallis test. Effect sizes of Mann–Whitney post hoc tests were estimated using \( r \), which was calculated from the \( z \)-score corresponding to the test statistic using the formula \( r = \frac{z}{\sqrt{N}} \) (Rosenthal, 1991). Cohen’s (1988) guidelines were used to classify the resulting effect sizes as small (\( r = .10 \)), medium (\( r = .30 \)), and large (\( r = .50 \)). Receiver operating characteristic (ROC) analysis was used to identify optimal cut scores on force curve variables for differentiating control and simulator participants. The areas under the ROC curves were compared using by Stata 9 software (StataCorp, 2005), which implements the approach described by DeLong, DeLong, and Clarke-Pearson (1988).

### Results

Control, naïve simulator, and coached simulator groups were not significantly different with regard to gender, \( \chi^2(2, n = 82) = 0.36, p = .84 \), ethnicity, \( \chi^2(4, n = 82) = 6.52, p = .16 \), or handedness, \( \chi^2(2, n = 82) = 3.43, p = .18 \). The groups were also comparable in terms of age, \( F(2,79) = 0.017, p = .98 \), education, \( F(2, 9) = 0.52, p = .59 \), and in their scores on the quiz assessing comprehension of the reading passages, \( F(2,79) = 1.32, p = .27 \).

All participants reported moderate instruction adherence or better (i.e., three or higher on the 5-point scale). The groups were comparable in the percentage of participants endorsing high or very high adherence to instructions, \( \chi^2(2, n = 82) = 3.49, p = .17 \). The majority of participants (67.1%) rated their success at producing the results asked of them as four or higher on the 5-point scale. The groups were comparable in the percentage of participants endorsing moderate success or better, \( \chi^2(2, n = 82) = 1.95, p = .38 \). The naïve and coached simulator groups were comparable in the percentage of participants reporting a moderately convincing performance or better, \( \chi^2(1, n = 82) = 1.90, p = .17 \).

Descriptive statistics across groups on grip strength and WMT variables are presented in Table 1. With regard to force curve variables, significant differences were found on average to peak ratio, \( H(2) = 17.41, p < .001 \). Post hoc tests revealed differences in average to peak ratio between control participants and naïve simulators, \( U = 203, p = .005, r = .38 \); control participants and coached simulators, \( U = 123, p < .001, r = .57 \); but not between naïve and coached simulators, \( U = 328, p = .29, r = .14 \). Length of trial, average force, peak force, slope to peak force, and grip strength in kilograms were not significantly different across groups.

Group differences on the WMT were significant for Immediate Recognition, \( H(2) = 32.78, p < .001 \), Delayed Recognition, \( H(2) = 37.47, p < .001 \), and Consistency, \( H(2) = 39.36, p < .001 \). Post hoc comparisons revealed differences between control participants and both naïve and coached simulators on Immediate Recognition, Delayed Recognition, and Consistency. Naïve and coached simulators were different on Delayed Recognition, \( U = 224, p < .001, r = .37 \), but not on Immediate Recognition and Consistency.

In order to examine the classification accuracy of force curve variables, ROC curves of average to peak ratio were compared with those of the WMT in both groups of simulators. Among naïve simulators, the area under the curve (AUC) of average to peak ratio was 0.72 (95% confidence interval [CI] = 0.59–0.86), which was not different from Immediate Recognition (AUC = 0.84; CI = 0.73–0.95), \( \chi^2(1, n = 54) = 1.85, p = .17 \). The AUC of average to peak ratio was different from Delayed Recognition (AUC = 0.88; CI = 0.78–0.98), \( \chi^2(1, n = 54) = 4.78, p = .03 \), and Consistency (AUC = 0.89; CI = 0.78–0.98).

### Table 1. Grip strength and WMT performance

<table>
<thead>
<tr>
<th></th>
<th>Control Mdn (25, 75%ile)</th>
<th>Naïve simulator Mdn (25, 75%ile)</th>
<th>Coached simulator Mdn (25, 75%ile)</th>
<th>( H )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Force curve</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time (s)</td>
<td>2.78 (1.87, 4.03)</td>
<td>3.31 (2.24, 4.99)</td>
<td>2.44 (1.71, 4.51)</td>
<td>2.26</td>
</tr>
<tr>
<td>Average force (mV)</td>
<td>138 (111, 192)</td>
<td>115 (83, 167)</td>
<td>96 (73, 138)</td>
<td>9.05</td>
</tr>
<tr>
<td>Peak force (mV)</td>
<td>180 (139, 242)</td>
<td>151 (128, 233)</td>
<td>138 (102, 208)</td>
<td>4.36</td>
</tr>
<tr>
<td>Slope to peak force</td>
<td>0.23 (0.12, 0.50)</td>
<td>0.15 (0.08, 0.25)</td>
<td>0.18 (0.11, 0.26)</td>
<td>3.89</td>
</tr>
<tr>
<td>Average peak ratio</td>
<td>0.79 (0.75, 0.85)</td>
<td>0.71 (0.65, 0.81)</td>
<td>0.70 (0.65, 0.75)</td>
<td>17.41*</td>
</tr>
<tr>
<td>Grip strength (kg)</td>
<td>25.9 (20.8, 33.8)</td>
<td>26.9 (21.9, 29.8)</td>
<td>23.2 (17.8, 30.1)</td>
<td>2.69</td>
</tr>
<tr>
<td><strong>WMT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR</td>
<td>100 (99.4, 100)*</td>
<td>52.5 (40.6, 99.4)*</td>
<td>77.5 (67.5, 94.4)*</td>
<td>32.78*</td>
</tr>
<tr>
<td>DR</td>
<td>100 (99.4, 100)*</td>
<td>47.5 (31.3, 91.9)*</td>
<td>75.0 (67.5, 95.0)*</td>
<td>37.47*</td>
</tr>
<tr>
<td>Consistency</td>
<td>100 (97.5, 100)*</td>
<td>57.5 (50.6, 86.9)*</td>
<td>73.8 (62.5, 89.4)*</td>
<td>39.36*</td>
</tr>
</tbody>
</table>

Notes: WMT = Word Memory Test; Mdn = median; %ile = percentile; IR = immediate recognition; DR = Delayed Recognition. Distribution descriptive statistics and Kruskal–Wallis test of grip strength and WMT. Cells with different superscript letters are significantly different on post hoc comparison (Mann–Whitney).

\( ^* p < .001 \).
CI = 0.79–0.99), $\chi^2(1, n = 54) = 4.38, p = .04$. Among coached simulators, the AUC of average to peak ratio was 0.83 ($CI = 0.72–0.94$). It was not different from Immediate Recognition (AUC = 0.91; $CI = 0.82–0.99$), $\chi^2(1, n = 54) = 1.16, p = .28$; Delayed Recognition (AUC = 0.90; $CI = 0.82–0.99$), $\chi^2(1, n = 54) = 1.01, p = .31$; or Consistency (AUC = 0.95; $CI = 0.00–1.00$), $\chi^2(1, n = 54) = 3.51, p = .06$. The classification accuracy of the WMT and average to peak ratio are presented in Table 2.

Discussion

This study examined the diagnostic utility of variables calculated from grip strength force curves in a simulated malingering design. University students randomly assigned to control, naïve simulator, and coached simulator conditions were assessed with two measures of grip strength (grip strength in kilograms using the Smedley dynamometer and force curves produced by a digitally modified Jamar dynamometer), the WMT, and a posttest to determine adherence to instructions. Five variables were calculated from grip strength force curves: Trial length, average force, peak force, slope to peak force, and average to peak force ratio.

The first hypothesis that control and simulator participants would demonstrate different levels of performance on both measures of grip strength was inconsistently supported. Control participants did not demonstrate greater grip strength in kilograms than either group of simulator participants. The similarity across groups in maximal grip strength (in kg) was unexpected, given the effect sizes observed in a pilot study (dominant hand, $d = 0.85$; nondominant hand, $d = 0.91$; Davis et al., 2007). It is inconsistent with other research in neuropsychology (Rapport et al., 1998; Wall & Millis, 1998) and industrial rehabilitation (Smith et al., 1989). Participants were excluded from the study if they had a history of upper extremity injury or neurological disorder. Random assignment to condition was utilized, and the experimental groups were comparable in gender ratio. The expected group differences were observed on the WMT, and on a posttest adherence questionnaire, participants in all groups reported attempting to follow the instructions. Therefore, it is not likely that this lack of difference arose from an issue with participant selection or with the experimental manipulation.

When grip strength was assessed using the digitally modified dynamometer, the force curves of all three groups were similar in terms of trial length, average force, peak force, and slope to peak force. The average to peak force ratio, however, differentiated control and simulator participants. The expected distinction between naïve and coached malingerers (i.e., that coached simulators would perform at a lower level than control participants but at a higher level than naïve simulators) was not supported. It is possible that the sample, which was sufficient to detect a large effect, was not large enough to detect smaller differences between simulator groups. Other research has inconsistently demonstrated the influence of coaching on motor and cognitive measures (Johnson & Lesniak-Karpiak, 1997; Rapport et al., 1998).

Average to peak force ratio appeared promising in that it differentiated control participants from both simulator groups. The statistical comparison of ROC curves suggested that the average to peak force ratio was comparable to the WMT in terms of the AUC for coached simulators. The results were inconsistent for naïve simulators as the AUC of average to peak ratio was similar to one WMT variable, Immediate Recognition, but not the other two. Examination of achieved sensitivity and specificity revealed absolute differences in classification accuracy between the average to peak force ratio and WMT for both groups.
of simulators. These differences might not have been detected in the ROC analyses due to the reduced sample size ($n = 54$) in the AUC comparisons. With no false positives (i.e., specificity = 1.00), the sensitivity of the WMT was 0.79 for naïve simulators and 0.71 for coached simulators. With specificity of 0.92, the best sensitivity of average to peak force ratio was 0.54 for naïve simulators and 0.57 for coached simulators. With more restrictive specificity of 0.96, the sensitivity of average to peak force ratio was 0.25 for naïve simulators and 0.22 for coached simulators. Thus, despite some similarities in terms of the AUC, the observed classification accuracy of average to peak force ratio in this sample was lower than that of the WMT. It was in the range of reported findings from other research using force curve analysis (e.g., Chengalur et al., 1990).

The present results are promising, but not without limitation. The use of a simulated malingering design is primary among the limitations of this study. Common criticisms of simulation research include reduced external validity arising from the lower age, higher education, and reduced incentive to under-perform among participants instructed to simulate impairment in comparison to litigants or disability claimants. Another argument against simulated malingering designs is the simulation paradox: Simulators are asked to do their best to feign impairment in order to develop methods of detecting individuals who feign impairment when asked to try their best (Rogers & Cavanaugh, 1983). Despite the limitations of simulation research, it serves an important function in the early stages of the development of an effort measure (Demakis, 2004). The integration of findings from simulation and known-groups designs has been suggested as an appropriate method of balancing the greater control over internal validity offered by simulation designs with the greater external validity provided by known-groups research (Rogers, Harrell, & Liff, 1993). Other limitations are related to the sample that was one of convenience with deviations from national demographic data. Notably, the modal participant in this sample was a 19-year-old, Caucasian, female with 13 years of education.

Overall, findings of the present study suggest that, while grip strength is a simple task that can be summarized in terms of the maximum force in kilograms, examination of variables reflecting change over time during the task may provide greater descriptive value and diagnostic utility. The approach used in this study has implications for clinical practice in that force curve variables could be considered in addition to maximum grip strength as a means of assessing examinee effort. Limitations notwithstanding, the results suggest future research. An important first step would be the design and implementation of another digitally modified dynamometer that would be portable. Another possibility would be to examine the feasibility of generating force curves from the output of commercially available digital dynamometers. Either instrumentation option would facilitate data collection in different settings and allow an experimental protocol to be conducted in a hospital or outpatient facility. Data from a clinical sample would potentially extend these initial findings.

Funding

This work was supported by a Faculty Summer Research Grant from the University of Indianapolis to JRW.

Conflict of Interest

None declared.

Acknowledgements

The authors would like to thank David Wall, PhD, for assistance with design and implementation of the digitally modified dynamometer; Carolyn Sherer, MA, for assistance with data collection and poster preparation; and Michelle Leslie, MA, for assistance with data collection. The authors are also grateful to Scott R. Millis, PhD, ABPP (CN,CL,RP), CStat, for assistance with Stata.

References


