Work 55 (2016) 873–882
DOI:10.3233/WOR-162450
IOS Press

Accuracy of distraction based lifting criteria for the identification of insincere effort utilizing the under loading method

Robert Townsend[a,*], Shane Bell[a] and John Harry[b]
[a] *University of Memphis, Memphis, TN, USA*
[b] *University of Nevada Las Vegas, Las Vegas, NV, USA*

Received 27 May 2015
Accepted 23 March 2016

Abstract.
BACKGROUND: Distraction based material handling tests are used to measure effort during functional capacity evaluations.
OBJECTIVE: To identify validity of effort using the under loading method with the XRTS Lever Arm. Classification of effort between the two sessions (100% effort and estimated 50% effort) was compared with current validity criteria.
METHODS: Fifty healthy and asymptomatic subjects were tested under two conditions (100% effort and an estimated 50% effort). Comparisons were made between percent changes from crate lifts to lever arm lifts from three starting heights (0.25 m, 0.38 m and 0.051 m).
RESULTS: During the 100% effort sessions, no subject had a mean percent change >20%, a majority of lift comparison >25% or a single lift comparison >30%. The specificity of the current validity criteria is 100%. The under loading method of using the XRTS lever arm was 20% sensitive in identifying an equivocal or invalid test result when subjects gave an estimated 50% effort.
CONCLUSIONS: The results indicate the under loading method of distraction based testing is very specific but lacks sensitivity in identifying feigned weakness with asymptomatic, non-injured individuals. The high specificity of the test should eliminate concern of having a false positive for insincere effort during functional testing.

Keywords: Functional capacity evaluation, cross-reference testing, percent change

1. Introduction

Measuring the material handling capabilities of an individual is typically assessed during Functional Capacity Evaluations (FCE's). Research concerning the effectiveness of FCE's has been conducted in numerous countries around the world [1]. The usefulness of measured data contained within a FCE report is dependent on two factors: reliability of testing methodology, and the willingness by the claimant to exert a maximum safe effort during the assessment [2]. The accuracy of FCE results is crucial in determining whether an individual is ready to return to work by meeting physical demands similar to normal job duties, or if restrictions need to be placed on a work injury claimant. Secondary gain by remaining in the patient role can give the injury claimant incentive for sub-maximal effort and magnification of symptoms [2–5].

Recently, interpretations and clinical applicability of many validity/sincerity of effort testing have been questioned [6–10]. Lifting tasks are defined

*Address for correspondence: Robert Townsend, MS, P.O. Box 381093, Germantown, TN 38183, USA. Tel.: +1 901 484 5247; Fax: +1 901 328 6382; E-mail: rjtwnsn1@memphis.edu.

as normal occupational material handling according to the U.S. Department of Labor. Traditionally two approaches have been used for determining when safe maximal workloads during a lifting assessment have been performed: the kinesiophysical and psychophysical approaches. A kinesiophysical approach allows a trained evaluator to determine the safe maximal workload for a claimant during material handling testing. This approach lacks accuracy (less than 50% correct classification) of interpretation during visual estimation of individual contribution [7]. The psychophysical approach allows the individual being tested to determine their maximal / terminal workload during lifting tasks. Clinical interpretations of effort through measuring physiologic variables such as heart rate have also come into question regarding accuracy [7, 10]. Accurate testing regarding methodology and results for identification of claimant contribution allows the outcomes to be utilized for returning to work decisions and/or making physical restriction recommendations. Validity of effort testing also may be required to prevent evaluator bias towards the claimant [10]. Previous studies have suggested inaccurate methods of assessment of an individual's effort during material handling be discontinued [9].

Distraction based testing has been used to identify patients who present with exaggerated pain complaints by comparing physical responses between two similar, benign tests [8, 11]. Waddell, et al. [11] first introduced the concept of distraction based testing with comparison of straight leg raise capability in a supine to seated position commonly known as the "Flip Test". To the authors' knowledge, the 2010 patient based study by St. James. et al. [9] is the only study examining distraction based protocol for assessing validity of effort during a lifting task. The cross-referencing of demonstrated lifting capabilities between workloads of two different physical appearances (crate and lever arm lifts) provides an element of distraction as defined by Waddell [11]. This distraction of comparing workloads of different physical appearances is intended to prevent work injury claimants from controlling test results while feigning weakness. Utilizing distraction based testing comparing weights in a crate and cross-referencing the results with the XRTS lever arm has shown increased accuracy in validating dynamic lift testing using the psychophysical approach [8].

Lifting on the XRTS lever arm is compared to the lifting of unmarked steel bars in a crate. The XRTS lever arm is a non-computerized patented second-class lever in which a handle plate, held by the subject, is attached with an adjustable chain. The handle plate is the same structural dimensions for hand placement by the subjects as lifting unmarked steel bars in a crate (30.05 cm $\times$ 30.05 cm). This allows subjects to have the same starting height from the floor and the same width of the hands for both the lever arm and crate lifts. Unmarked barbell plates can be loaded onto a steel bar attached to a moveable carriage. Workloads can have different configurations (varied assortments of unmarked plates) to change the visual appearance of the same weight on the carriage. This carriage can be maneuvered along the length of the lever arm. Equally spaced holes along the length of the lever arm allow for the weighted carriage to be secured into place as shown in Fig. 1. The workload for the lever arm can be determined by predictable linear increases in weight loaded on the carriage in one location (position) along the length of the lever arm. Workloads can also be determined by maneuvering the carriage with the same weight to a different location on the length of the lever arm.

The study by St. James. et al. [9] determined cross-reference validity criteria using the lever arm device. The criteria used to determine validity of effort measured the percent difference between reported maximal lifts of the XRTS lever arm and the unmarked weights in a crate. The percent difference in comparative lifts correlated with failed criteria in hand and pinch grip testing with 99% accuracy of effort classification. As a result, distraction based lift testing validity criteria may be useful in identifying individuals who show exaggerated complaints of pain [11]. The increased accuracy presented by distraction based testing may have consequences on medical case closure and allocation of health care dollars. To this end, the purpose of this investigation was to examine the accuracy of the current existing validity criteria between dynamic lifting of weights in a crate and lifting on the XRTS lever arm device using the under loading method. At the time of this study, there were no investigations available using the XRTS lever arm current existing validity criteria with a control group of non-injured asymptomatic subjects.

1.1. Research hypothesis

The research hypothesis is that asymptomatic, non-injured subjects will be accurately classified for effort using the under loading method of cross-reference testing on the XRTS lever arm. Current validity

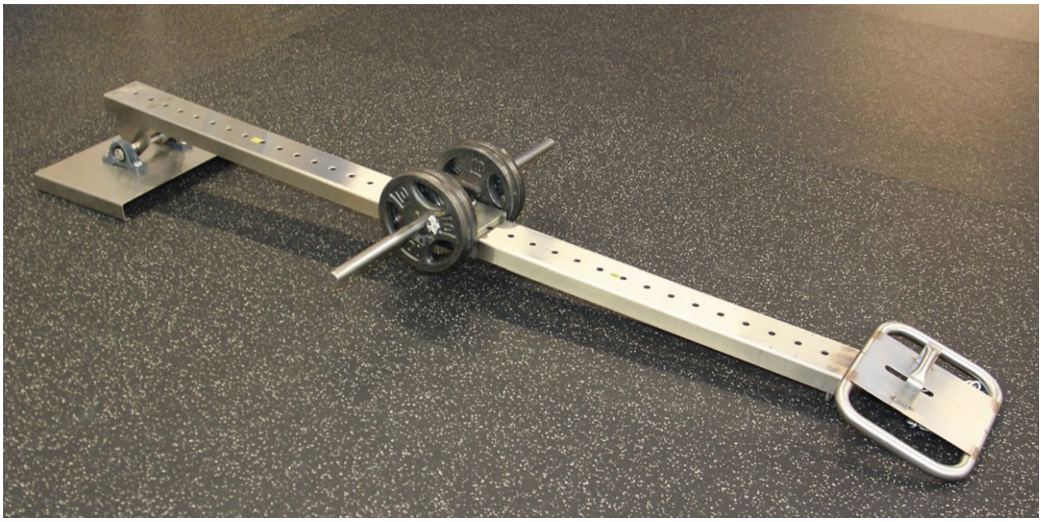

Fig. 1. The XRTS Lever Arm.

criteria, as outlined by the manual for the lever arm device as seen in Table 1, were used to classify test results as valid or invalid efforts.

2. Methods

2.1. Subjects

A convenience sample of 50 volunteers (17 women and 33 men) with a mean age of 28.0 years (SD = 9.9) was used in this analysis. A power analysis using G* Power was run for investigating the possibility of Type II Error. In the case of dependent *t* tests with proposed alpha of 0.05, power (1-beta) of 0.95, and an effect size of 0.5, the power analysis suggested 45 participants. The population of this study was 50 which exceeds the suggested population of 45 subjects. This sample was recruited by word of mouth and posted flyers. Secondary to examining percent differences between crate and lever arm lifts opposed to absolute strength measures, Male and female subjects were assumed to respond similarly. No separate validity criterion between genders exists in the clinical setting using XRTS protocol. All subjects were asymptomatic individuals who were not receiving medical or rehabilitative treatment in their arms, back or legs and did not have any known lifting restrictions prescribed by a medical physician. No compensation was given to the subjects for this study. Participation was strictly voluntary. This experiment was approved by the local Institutional Review Board.

2.2. Procedure

Subjects participated in two laboratory sessions. Subjects were randomly assigned which session, 100% effort or 50% effort was performed first. During one session, subjects were instructed to give a full safe effort in dynamic lifting of unmarked steel bars in a crate and the X-RTS lever arm. A full safe effort maximum was defined as a workload that could be safely lifted 1 or 2 times per hour for an entire eight hour work day without assistance from another person. Crate and lever arm lifting was performed with starting heights from the floor at 0.25 m, 0.38 m, and 0.51 m. The lift was initiated by standing upright with the crate and bringing the distal metacarpals to waist level. Subjects returned the crate or lever arm handles to the starting height to complete the lift. Each reported baseline lift of unmarked weights in a crate was cross-referenced using the XRTS lever arm using the under loading method. The subjects were blinded to both the absolute value of workload and the percentage of the baseline crate lifts when performing the lever arm lifts.

The under loading method is the process for selecting an initial weight for lifting the XRTS lever arm to cross-reference crate lifts. The under loading method consists of selecting a lever arm workload that is approximately equal to 30% of the previous crate lift for the same starting height. For example, as crate lift of 40 kg would have beginning workload of 12 kg on the XRTS lever arm. The weight configuration of weight loaded on the lever arm was within ±2.27 kg

Table 1
Validity criteria for effort classification using the XRTS Lever Arm

Effort classification	Average % difference of 3 lifts	At least half of comparative lifts	Single set comparative data sets	Two or more sets of comparative data sets
Valid	<20%	<25%	None ≥ 30%	--
Invalid	≥ 25%	≥ 25%	Any ≥ 40%	≥ 30%
Equivocal	≥ 20% but < 25%	<25%	None ≥ 30%	--

of the reported safe maximum with the corresponding crate lift. If an identical or similar reported safe maximum existed between two or more baseline lifts of 0.25 m, 0.38 m and 0.51 m, a different configuration of an additional 2.27 kg or 2.27 kg less than the original baseline lift was used. The purpose of the configuration change was to ensure the subjects never had the same visual configuration of weights on the lever arm for two comparative lifts. Once the configuration of weight loaded on the lever arm was determined, the carriage was placed at a specific positioned on the length of the lever arm. The carriage position was selected so the force required to lift the lever arm was approximately equal to 30% of the previous crate lift for the same starting height.

Once the initial workload for cross-reference testing was established, a progressive lift test using the lever arm was performed. Subjects performed a single repetition lifts increasing in weight until they reported achieving their safe maximal effort. Increases in workload by approximately 15% of the baseline lift were made for each attempt. This process was repeated until a reported safe maximum on the lever arm lift was achieved. Immediately following the dynamic lifting cross-reference test on the lever arm, a repeat maximal baseline lift was performed using the weights in the crate to rule out any potential fatigue effects. During the alternate session the same procedures applied, only the subject was instructed to give an estimated 50% effort during dynamic lifting of both the unmarked steel bars and the lever arm lift.

2.3. Data analysis

Both conditions, 100% safe maximal lifting capactiy and estimated 50% of maximal safe lifting capactiy were compared for accuracy in classification of effort. Curent validiy criteria were used as defined by the protocol in the lever arm manual. The classifications of effort, valid, invalid and equivocal were initially used to classify each subjects results for both conditions.

The criteria used for classifying subject effort were documented previously [8]. The classifications of effort, Valid, Invalid and Equivocal as defined by the lever arm manual were used. Equivocal or atypical results do not qualify as valid test results. Equivocal and atypical results were grouped with invalid test results in a category considered “non-valid”. The category of non-valid results was used to calculate parametric statistics and frequencies of test result classification. The criteria for each classification were defined as outlined in the lever arm manual and are indicated in Table 1. Data are presented as mean ± SD. Standard checks for assumptions of parametric statistics were performed. Crate and lever arm lifts were compared via dependent *t*-tests, with standardized mean differences corrected for repeated measures. Statistical analyses were performed with SPSS v22.0. Diagnostic test evaluation statistics were performed with MedCalc statistical software, version 14.120.

3. Results

There were no false positives for an invalid or equivocal effort during the condition in which subjets were instructed to lift 100% of their safe maximal lifting capability. The under loading method of distraction based testing was measured 100% specific in identifying non-valid tests. The 95% Confidence Interval for specificity 92.9 to 100%. The sensitivity in identification of a non-valid test was 22% with a 95% Confidence Interval of 11.54 to 35.97%. The Positive Predictive Value was 100% defined as the probability that feigned weakness is present when the test criteria are positive for non-valid classifications (equivocal or invalid). Table 8 has the number of tests that correctly identified the 50% effort and 100% effort conditions. The sensitivity and specificity results can be see in Table 9.

3.1. 100% effort condition

The condition of 100% safe maximal lifting capacity had a valid classifications for all 50 subjects. The criteria defined by XRTS protocol uses a percent

Table 2
Asymptomatic subject under loading study 100% effort condition

Bilateral lift height	Crate lift absolute load in Kg (Mean, SD)	Lever Arm lift absolute load in Kg (Mean, SD)	Percent difference of comparative lifts
0.25 m ($n = 50$)	44.15 ± 10.53	48.56 ± 10.57	6.97 ± 5.66
0.38 m ($n = 50$)	44.60 ± 10.55	44.46 ± 10.68	7.84 ± 5.44
0.51 m ($n = 50$)	45.19 ± 11.26	44.76 ± 11.10	7.87 ± 5.16

Table 3
Asymptomatic under loading study 50% effort condition

Bilateral lift height	Crate lift absolute load in Kg (Mean, SD)	Lever Arm lift absolute load in Kg (Mean, SD)	Percent difference of comparative lifts
0.25 m ($n = 50$)	25.3 ± 6.25	25.43 ± 6.66	12.1 ± 9.08
0.38 m ($n = 50$)	26.89 ± 7.22	26.05 ± 7.59	13.50 ± 10.33
0.51 m ($n = 50$)	27.61 ± 8.31	26.74 ± 9.22	14.95 ± 11.72

difference between crate and lever arm lifts from the same starting height. Comparing each criterion for classification of a valid effort during the safe maximal lifting condition indicated no false positives in any criteria. The comparison of mean loads and percent variances between lever arm and crate lifts can be seen in Table 2. The average variability between three comparative lifts from subjects giving 100% safe maximal lifting was 7.56% (±4.41). The criteria requiring at least half of all comparative lifts have variability <25% would require 2 or 3 lifts for each subject. During the maximal safe effort condition there were no comparative lifts that had variability ≥25%. The final criteria of not having any single set of comparative lift variability ≥30% was met by all 50 subjects.

3.2. *50% effort condition*

The condition of having subjects lift an estimated 50% of their maximal safe capabilities had subjects fitting all three categories of effort classification, valid, invalid and equivocal. The invalid classification occurred in 5 of 50 subjects. The values for the mean workloads and percent change during the 50% condition can be seen in Table 3. With the exception of one subject, the remaining subjects were classified as giving a valid effort during the estimated 50% of their maximal safe effort condition. The subject that did not fit a category classification had an average percent difference of 19.91%, peak percent difference of 30.39%. Additionally, a majority of comparative lifts for this individual were less than 25%.

Each of the subjects classified with equivocal test results during the condition in which an estimated 50% effort was given had an average variability >20% but <25%. The criterion of having a single set of comparative lifts ≥30% was not present in any of the subjects with the equivocal classification. All of the subjects who had an equivocal classification had a majority of their lift comparisons less than 25%.

Data collection order of lifting went from the 0.25 m, 0.38 m then 0.51 m. This occurred during both baseline crate lifting and cross-reference lever arm lifts. The data collected suggest there was no learned effect or potential counting of repetitions to cause a decrease in percent change from crate to lever arm lifts. The data indicates no reduction in percent change when a subject progressed from the 0.25 m to the 0.51 m lifts. This data can be seen in Table 2.

All data met assumptions of parametric statistics except for the 50% trial on the lever arm from the 0.51 m height. There were no significant differences between any of the pairs, with p-values ranging from 0.111 to 0.943 (Table 4). For the 50% trial at the 0.51 m height, a Wilcoxon signed rank test was performed, and was not significant ($p = 0.077$). All standardized mean differences in comparative lifts were trivial-to-small according to Hopkins (1999), and ranged from 0.01 to 0.31 as seen in Table 5.

4. Discussion

The results of this study had mean percent changes between crate and lever arm lifts significantly lower than the criteria for valid test classification. This occurred during both the 100% and 50% effort conditions. The validity criteria were outlined in a previous investigation using a clinical population of 200 consecutive work injury claimants [8]. Within that study distraction based lifting results were organized between individuals who have passed validity criteria in hand and pinch grip testing along with clinical

Table 4
Paired samples test

Paired samples	95% Confidence interval	t	df	Sig. (2-tailed)
100% 0.51 m Crate to 0.51 m LA	3.836	0.917	54	0.363
100% 0.38 m Crate to 0.38 m LA	2.473	0.072	54	0.943
100% 0.25 m Crate to 0.25 m LA	1.007	–1.056	54	0.296
50% 0.51 m Crate to 0.51 m LA	4.580	1.487	54	0.143
50% 0.38 m Crate to 0.38 m LA	4.201	1.622	54	0.111
50% 0.25 m Crate to 0.25 m LA	1.839	–0.123	54	0.903

Table 5
Standardized mean differences between comparative lifts

Paired samples	Cohen's d
100% 0.51 m Crate to 0.51 m LA	0.18
100% 0.38 m Crate to 0.38 m LA	0.01
100% 0.25 m Crate to 0.25 m LA	0.20
50% 0.51 m Crate to 0.51 m LA	0.29
50% 0.38 m Crate to 0.38 m LA	0.31
50% 0.25 m Crate to 0.25 m LA	0.02

observation of abnormal pain behaviors. When comparing the results of this study of 50 asymptomatic, non-injured subjects with the under loading method to the clinical population, smaller percent differences were seen. The 100% effort condition used in the asymptomatic group (Table 2) can be compared to injury claimants who passed all the criteria for hand strength assessment in Table 6. The comparison of the 50 subjects in the asymptomatic group when giving an intentional 50% effort with injury claimants who failed validity of effort testing with a hand strength assessment can be seen in Table 7. The differences in comparative lifts that exist between asymptomatic subjects using the under loading method and injury claimants may be due to multiple variables. These variables include the protocol in the control group opposed to injury claimant group along with physical and psychological differences in the populations of these two groups.

In the clinical setting, the under loading method has specific recommendations for use in distraction based testing during functional capacity evaluations. An individual's absolute values during baseline lifts in relation to normal job demands and testing behaviors will determine the procedures for lever arm cross-reference testing. Therefore, the under loading procedure used during this study may not reflect the procedures used during all functional capacity evaluations in a clinical setting. This may account for some of the differences seen in percent change values being higher in the original clinical based study by St. James et al. [9] when compared to this under loading study.

The procedure of the under loading method for the asymptomatic subjects began all lever arm loads with the same 30 percent of baseline lifts. Despite different configurations of weight being used for the 0.25 m, 0.38 m and 0.51 m lifts, all initial workloads were 30% of baseline crate lifts. After each lever arm lift, the workload was increased at an interval of 15% for each following lift. This process was repeated until the subject reported reaching a safe maximal lift during the 100% effort condition. During the 50% effort condition, the same procedure was used until the subject reached a perceived 50% of safe maximum. Using the same starting point and same interval increase with the lever arm lifts, allows for similar number of attempts to be performed before reaching a terminal load. This allows for an individual to potentially predict when a workload similar to the baseline crate lift occurs by counting attempts. However, the percent changes between the first and last cross-reference lift did not indicate any learned effect. These results can be seen in Table 3.

4.1. *Physical comparison of patient and control populations*

The population used for this study had the physical requirements of being asymptomatic and not receiving medical or rehabilitative treatment. The population of the original clinical based study all had applied for benefits in connection to work-related injuries or for long term disability [8]. The physical differences between the two populations can be potentially defined by the work-related injured

Table 6
Injury claimants who passed all criteria for hand strength assessment from St. James, Schapmire, et al. [8]

Bilateral lift height	Crate lift absolute load in Kg (Mean, SD)	Lever Arm lift absolute load in Kg (Mean, SD)	Percent difference
0.25 m ($n = 31$)	17.37 ± 6.4	19.23 ± 6.58	22.3 ± 25.4
0.38 m ($n = 37$)	18.46 ± 0.93	20.94 ± 8.07	20.7 ± 24.95
0.51 m ($n = 38$)	19.05 ± 7.93	20.77 ± 8.07	19.6 ± 11.26

Table 7
Injury Claimants who failed two or more criteria for hand and strength assessment (St. James, Schapmire, et al.) [8]

Bilateral lift height	Crate lift absolute load in Kg (Mean, SD)	Lever Arm lift absolute load in Kg (Mean, SD)	Percent difference
0.25 m ($n = 48$)	8.8 ± 5.08	13.02 ± 6.12	60.9 ± 49.5
0.38 m ($n = 54$)	9.30 ± 4.81	13.38 ± 5.67	55.4 ± 45.8
0.51 m ($n = 62$)	8.75 ± 4.76	12.70 ± 4.99	56.0 ± 37.1

population having greater pain levels opposed to asymptomatic subjects. Schapmire, et al. [9] compared the results of 200 consecutive work injury claimants for failed validity criteria with upper extremity hand and pinch grip testing. The results of the 2010 study indicated individuals with surgical intervention had significantly more valid test results compared to those with a non-surgical diagnosis. Their results could be related to the increased biofeedback from nocioreceptors and pressure receptors in the peripheral nervous system [12]. If an individual has increased feedback from nocioreceptors due to pain, the ability to reproduce levels of effort consistently would be enhanced. This mechanism should be present to a lesser degree in the subjects in this study who reported being asymptomatic. When comparing the results from both populations the asymptomatic subjects had greater reproducibility of effort between the crate and lever arm lifts. This contradicts the increased biofeedback mechanism from pain allowing for greater reproducibility but potentially accentuates the secondary gain issues that exists in the patient population [2–5].

Table 8
Valid and non-valid tests using current validity criteria and both 50% and 100% conditions

50% Effort	n	100% Effort	n	Total
True Positive	a = 11	False Positive	b = 0	a + b = 11
False Negative	c = 39	Ture Negative	d = 50	c + d = 89
Total	a + c = 50		b + d = 50	

Table 9
Sensitivity and specificity of current validity criteria identifying 50% effort condition

Sensitivity	22.00%	95% CI: 11.54% to 35.97%
Specificity	100.00%	95% CI: 92.82% to 100.00%
PPV	100.00%	95% CI: 71.33% to 100.00%
NPV	56.18%	95% CI: 42.5% to 66.68%

4.2. *Psychological comparison of patient and control populations*

The differences in psychological states between the asymptomatic non-injured subjects and work injury claimants begin with being in the patient role. The subjects in the under loading study who were asymptomatic and not involved in an injury claim had no secondary gain influences on their performance of the material handling tasks. These subjects had lower percent change measures between baseline crate lifts and lever arm lifts compared to the previous study using injury claimants. It has been well documented that many psychological and socioeconomic factors influence a patient's length of time off work due to injury [4, 5, 13–17]. The secondary gain that exists in the workers' compensation system may influence an injured work's physical performance during functional capacity evaluations and returning to work [5, 16]. The presence of secondary gains to remain in the patient role can influence work injury claimants to perform sub-maximally during physical tests or magnify symptoms that are present [2, 3, 11, 14]. If these attempts to magnify symptoms or perform sub-maximally are purposeful, they could be categorized as attempts of deception. Changes in brain activity have been documented with purposeful deception using an fMRI [18]. Changes in brain activity have also been associated with other psychological / emotional disorders [19]. A correlation with deception, emotional or psychological stress and the ability to reproduce force would indicate a behavioral cause for an inconsistent effort.

4.3. Cross-reference testing results comparison of patient and control populations

The study using 200 consecutive work injury claimants compared percent differences between crate and lever arm lifts between groups based on failed validity of effort criteria with hand and pinch grip testing and the presence of abnormal pain behaviors. The mean percent differences in crate to lever arm lift comparisons for work injury claimants who passed validity criteria for hand strength assessment ranged between 19.6–22.3%. The non-injured subjects of the under loading study had mean percent differences between 6.97–7.87% for the 100% effort trial and 12.1–14.95% for the 50% trail. The protocol used for hand strength assessment was subject to peer review and published in 2002 [20]. The protocol used to identify feigned weakness was public knowledge prior to the data collection on the lever arm study using 200 consecutive work injury claimants. The possibility exists that legal counsel gave instruction to work injury claimants not to display feigned weakness during the hand strength assessment. Therefore, some injury claimant subjects may have given full effort during hand strength assessments and feigned weakness during lifting. This may explain the lower percent differences seen in the asymptomatic subjects in the under loading study.

4.4. Practical implications of results

No other study using the XRTS lever arm for distraction based testing has investigated the accuracy of existing validity criteria for sincerity of effort in a non-patient population. The importance of researching a non-patient population establishes the physical response to testing methods without influence from behavioral components that exist in the patient population. Without the results of this study, speculation on the causation identified invalid efforts being physical or psychological would continue. Using a population with no motivation for secondary gain or kinesiophobia in this study allows for the analysis of primarily a physical response to lever arm testing. The results of this study indicate the probability of being misclassified as invalid or equivocal for effort being very low. The authors of this study feel investigating and verifying a testing method for measuring validity of effort is an ethical responsibility. Cultural bias may be found in any type of assessment instrument, resulting in a systematically overestimation or underestimation of the variable being measured [10]. FCE Practitioners have reported the need to increase evidenced based research to increase credibility and legal defensibility of test results [21, 22]. The results of this study provide additional data for practitioners using distraction base testing during legal testimony in workers compensation cases.

The measured results for the percent difference between crate and lever arm lifts during the 100% effort condition in Table 2 are low when compared to current validity criteria in Table 1. The current validity criteria for an equivocal classification with distraction based lift testing indicates a mean percent change $\geq$20% and $\leq$25% as seen in Table 1. The equivocal classification is also defined as a majority of comparison lifts to be $\leq$25% and no single comparison to be $\geq$30%. Two standard deviations above the mean percent difference between crate and lever arm lifts during the100% effort condition ranges between 18.19–18.72% for the three lift starting heights. The results for mean and standard deviations are in the asymptomatic population 100% effort condition are in Table 2. The equivocal classification requires statistical outlier data for asymptomatic individuals who gave 100% effort during distraction based lifting. The results of the study using asymptomatic subjects indicate percent differences greater than 20% would be greater than two standard deviations above the mean for 100% effort.

Percent differences greater than 20% may indicate non-physical components by work injury claimants influencing comparative lift testing. Non-physical components influencing work injury claimants may exist as secondary gain issues, kinesiophobia or both. The data collected from this study may allow practitioners to begin to identify when non-physical components influencing testing behavior. This will allow for more efficient case management decisions to be made regarding continuation of treatment through physical rehabilitation.

4.5. Limitations and future research

A potential limitation of this study is regarding the instructions to give a perceived 50% effort by the subjects. The intentional act of giving a perceived 50% effort may not represent feigned weakness as it would occur with a work injury claimant. Work injury claimants motivated by secondary gain are more likely to limit their efforts to an absolute workload less than normal job requirements (e.g. 20 kg). The population for this study included 47 of 50 subjects that were students or faculty in a health and human performance major at the university level. The majority population has had experience with resistance training for fitness and performance. These subjects ability to estimate a specific workload (e.g. 20 kg) may be higher than those with less experience handling known workloads during exercise. Frequent handling of known workloads may occur less regularly in subjects with limited to no weight training experience. The instructions to give a perceived 50% effort opposed to reproducing an absolute workload was more suitable for this population. Future research should be performed in the ability to reproduce a specific workload (e.g. 15 kg) using the crate and lever arm lift comparisons. Populations with limited resistance training experience but experience with physical laboring would be ideal for such a study.

Continued investigation of distraction based lifting within a population of individuals with known psychological diagnoses is needed. Further research should take place in investigating if the differences in brain activity could interfere with an individual's ability to reproduce sub-maximal and safe maximal lifting.

5. Conclusion

The methodology of distraction based testing has been used in a psychophysical approach during functional capacity evaluations. The XRTS lever arm has been documented to classify the effort during material handling testing that correlate with other peer reviewed measures of effort [8]. The under loading method using a lever arm was able to identify when individuals gave a 100% safe effort with no false positive for sub-maximal effort. When current validity criteria are failed, indicating a sub-maximal effort, the possibility of these results being inaccurate is statistically insignificant. Behavioral influences in work injury claimants such as fear, anxiety and deception may increase the percent difference between baseline crate lifts and cross-referenced lever arm lifts. A population of asymptomatic individuals who are not in the patient role displayed lower percent variance between the two modes of lifting when giving both 100% and 50% effort. These finding indicate that percent differences between crate and lever arm lifts greater than 20% are potentially due to behavioral or non-physical components.

Acknowledgments

The authors would like to thank Doug Edwards and Bardavon Health Innovations (http://www.bardavon.com) for their kind permission to use both the under loading method protocol and the current XRTS (http://www.xrts.com) validity criteria for categorizing the subjects of this study.

Conflict of interest

The authors have no conflict of interest to report.

References

[1] Reneman MF, Soer R, Gross DP. Developing Research on Performance-Basd Functional Work Assessment: Report on the First Internation Functional capacity Evaluation Research Meeting. J Occup Rehabil 2013;23(4):513-5.

[2] Stokes HM, Landrieu KW, Domangue B, Kunen S. Identification of low-effort patients through dynamometry. J Hand Surg Am 1995;20:1047-56.

[3] Schapmire DW, St. James JD. Functional Capacity Evaluation Part I: Cutting the Gordian Knot of Secondary Gain, Expert Witness Culture, and Valifity of Effort Testing. IAIABC Journal 2011;47:93.

[4] Li-Tsang C, Chan H, Lam CS, Lo-Hui K, Chan CH. Psychological Aspects of Injured Workers' Returning to Work (RTW) in Hong Kong. Journal of Occupational Rehabilitaiton 2007;17:279-88.

[5] Gross D, Battie M. Factors Influencing Results of Functional Capacity Evaluations in Worker's Compensation Claimants with Low Back Pain. Physical Therapy 2005;85:4315-22.

[6] Oliveri M, Jansen T, Oesch P, Kool J. The prognostic value of functional capacity evaluation in patients with chronic low back pain: Part 1: Timely return to work. And part 2: Sustained recovery. Spine (Phila Pa. 1976) 30:1232-3; 2005; author reply 1233-4.

[7] Schapmire DW, St. James JD, Townsend R, Feeler L. Accuracy of visual estimation in classifying effort during a lifting task. Work 2011;40:445-57.

[8] St. James JD, Schapmire DW, Feeler L, Kleinkort J. Simultaneous bilateral hand strength testing in a client population, Part II: Relationship to a distraction-based lifting evaluation. Work 2010;37:395-403.

[9] Townsend R, Schapmire DW, St. James J, and Feeler L. Isometric strength assessment, Part II: Static testing does not accurately classify validity of effort. Work 2010;37:387-94.

[10] Owen TR, Wilkins MJ, Sincerity of Effort Differences in Functional Capacity Evaluations. Journal of Rehabilitation 2014;80(3):53-61.

[11] Waddell G, McCulloch JA, Kummel E, Venner RM. Nonorganic physical signs in low-back pain. Spine (Phila Pa. 1976) 1980;5:117-25.

[12] Guyton Hall. Textbook of Medical Physiology, 9th Edition. W. B. Saunders Company. Philadelphia; 1996.

[13] Hepp U, Moergeli H, Buchi S, Bruchhaus-Steinert H, Sensky T, Schnyder U. The Long-Term Prediction of Return

to Work Following Serious Accidental Injuries: A Follow Up Study. BMC Psychiatry 2011;11:53.

[14] Dionne C, Bourbonnais R, Fremont P, Rossignol M, Stock S, Nouwen A, Larocque I, Demers E. Determinents of "return to work in good health" Among Workers with Back Pain Who Consult in Primary Care Settings: A 2-Year Prospective Study. Eur Spine 2007;16:641-55.

[15] Geisser M, Robinson M, Miller Q, Bade S. Psychosocial Factors and Functional Capacity Evaluation Among Persons with Chronic Pain. Journal of Occupational Rehabilitation 2003;13(4):259-76.

[16] Harris I, Mulford J, Solomon M, et al. Association Between Compensation Status and Outcome After Surgery: A Meta-analysis. JAMA 2005;293(13):1644-52.

[17] Kaplan G, Wurtele S, Gillis D. Maximal Effort During Functional Capacity Evaluations: An Examination of Psychological Factors. Arch Phys Med Rehabil 1996;77:161-4.

[18] Kozel FA, Padgett TM, George MS, A Replication Study of the Neural Correlates of Deception. Behavioral Neuroscience 2004;118(24):852-6.

[19] Vizueta N, Rudie JD, Townsend JD, Torrisi S, Moody TD, Bookheimer SY, Atlshuler LL. Regional fMRI hypoactivation and altered functional connectivity during emotional processing in nonmedicated depressed patients with bipolar disorder. American Journal of Psychiatry 2012;169(8): 831-40.

[20] Schapmire DW, St. James JD, Townsend R, Stewart T, Delheimer S, Focht D. Simultaneous bilateral testing: Validation of a new protocol to detect insincere effort during grip and pinch strength testing. Journal of Hand Therapy 2002;15(3):242-50.

[21] Simons G. Credibility Crisis in FCEs. Physical Therapy Products. October 2006;3.

[22] Mecham J. Under closure scrutiny: Functional capacity evaluations must be evidence based and valid. Advance for Directors in Rehabilitation 2008;47-8.