Journal of Hand Therapy 32 (2019) 368–374

Contents lists available at ScienceDirect

Journal of Hand Therapy

journal homepage: www.jhandtherapy.org

JHT Read for Credit Article #624.
Scientific/Clinical Article

# Development and reliability of the rating of compensatory movements in upper limb prosthesis wearers during work-related tasks

Tallie M.J. van der Laan MD [a,*], Sietke G. Postema MD, PhD [a], Michiel F. Reneman PT, PhD [a], Raoul M. Bongers MSc, PhD [b], Corry K. van der Sluis MD, PhD [a]

[a] Department of Rehabilitation Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[b] University of Groningen, University Medical Center Groningen, Center of Human Movement Sciences, Groningen, The Netherlands

ARTICLE INFO

*Article history:*
Received 7 May 2017
Received in revised form
22 December 2017
Accepted 27 December 2017
Available online 10 February 2018

*Keywords:*
Work capacity evaluation
Prosthesis
Musculoskeletal pain
Upper limb
Compensatory movements

ABSTRACT

*Study Design:* Reliability study.
*Introduction:* Quantifying compensatory movements during work-related tasks may help to prevent musculoskeletal complaints in individuals with upper limb absence.
*Purpose of the Study:* (1) To develop a qualitative scoring system for rating compensatory shoulder and trunk movements in upper limb prosthesis wearers during the performance of functional capacity evaluation tests adjusted for use by 1-handed individuals (functional capacity evaluation-one handed [FCE-OH]); (2) to examine the interrater and intrarater reliability of the scoring system; and (3) to assess its feasibility.
*Methods:* Movement patterns of 12 videotaped upper limb prosthesis wearers and 20 controls were analyzed. Compensatory movements were defined for each FCE-OH test, and a scoring system was developed, pilot tested, and adjusted. During reliability testing, 18 raters (12 FCE experts and 6 physiotherapists/gait analysts) scored videotapes of upper limb prosthesis wearers performing 4 FCE-OH tests 2 times (2 weeks apart). Agreement was expressed in % and kappa value. Feasibility (focus area's "acceptability", "demand," and "implementation") was determined by using a questionnaire.
*Results:* After 2 rounds of pilot testing and adjusting, reliability of a third version was tested. The interrater reliability for the first and second rating sessions were $\kappa = 0.54$ (confidence interval [CI]: 0.52-0.57) and $\kappa = 0.64$ (CI: 0.61-0.66), respectively. The intrarater reliability was $\kappa = 0.77$ (CI: 0.72-0.82). The feasibility was good but could be improved by a training program.
*Discussion:* It seems possible to identify compensatory movements in upper limb prosthesis wearers during the performance of FCE-OH tests reliably by observation using the developed observational scoring system.
*Conclusions:* Interrater reliability was satisfactory in most instances; intrarater reliability was good. Feasibility was established.



## Introduction

People with an upper limb absence (ULA), caused by either congenital deficiency or amputation, need to make compensatory movements with both arms (affected and unaffected) and with their trunk in order to perform daily tasks and work-related tasks with their prosthesis.[1,2] They need to compensate for the loss of wrist, forearm (pronation and supination) and sometimes elbow movements, due to limited degrees of freedom of the prosthesis.[1] In previous studies, prosthesis wearers compensated the most with their trunk and shoulders[1,2]; however, the compensatory movements seem to be task specific.[2,3] Compensatory movements have not yet been studied during execution of work-related tasks in this patient group.

Compensatory movements may cause musculoskeletal complaints (MSCs) due to overuse or increased physical load on the affected or the unaffected arm, neck, and back.[4-7] Postema et al showed that the prevalence of neck, back, and arm (both unaffected and affected) complaints in individuals with ULA is twice as high in comparison with the general population: 57% compared to 29%, respectively.[8] They also identified nonphysical risk factors related to MSC in patients with ULA; these include middle age, being

Conflict of interest: The authors declare that they have no conflict of interest.

* Corresponding author. Department of Rehabilitation Medicine, University of Groningen, University Medical Center Groningen, P.O.Box 30.001, 9700 RB Groningen, The Netherlands. Tel.: +31 50 361 2295; fax: +31 50 361 1708.
*E-mail address:* t.m.j.van.der.laan@umcg.nl (T.M.J. van der Laan).

0894-1130/$ – see front matter © 2018 Hanley & Belfus, an imprint of Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.jht.2017.12.003

divorced or widowed, and lower mental health.[8] Physical risk factors have not yet been identified. In the general population, physical risk factors for MSC are working in awkward positions, static muscle contractions, and forceful and repetitive movements.[9] MSC can affect the quality of life and participation in social roles such as work.[4-6,10]

Quantifying compensatory movements during work-related tasks may help professionals in developing treatment plans to minimize the use of compensatory movements by individuals with ULA, with the aim of preventing MSC. Furthermore, therapists could use this quantification to evaluate treatment effects. To our knowledge, there is currently no system that systematically scores compensatory movements in patients with ULA during execution of work-related tasks. In rehabilitation and occupational medicine, a functional capacity evaluation (FCE), which is a comprehensive test containing work-related tasks,[9] is used to measure functional capacity, assess disabilities, monitor the progress of rehabilitation, and guide return to work recommendations.[9,11,12] Recently, our research group developed the FCE-one handed (FCE-OH), which is a short-form FCE adapted for 1-handed individuals and prosthesis users.[13] The FCE-OH consists of 6 FCE tests; the overhead work test, the repetitive reaching test, the fingertip dexterity test, the overhead lifting test-2 handed, overhead lifting test-1 handed, and the hand grip strength test.

The outcome of each FCE test is a quantitative score, for example, how much weight can be lifted or how much time is needed to complete a task. The FCE clinician also assesses the qualitative aspects of the performance but not always in a standardized way. A standardized qualitative assessment of compensatory movements during FCE tests could provide information to identify potentially dysfunctional approaches to tasks. It would also enable the therapist to intervene which may prevent MSC. Compensatory movements can be assessed in several ways, including observation.[14] Observational scales are already used in FCEs to determine the level of physical effort during FCE tests,[11] but a qualitative scoring system for compensatory movements is lacking. In this study, we therefore aimed (1) to develop a qualitative scoring system (based on observation) for rating compensatory shoulder and trunk movements in upper limb prosthesis wearers during the performance of FCE-OH tests; (2) to determine the interrater and intrarater reliability of this scoring system, and (3) to assess its feasibility.

## Methods

### *Design*

The scale development and testing consisted of 5 phases: planning, construction phase 1, qualitative evaluation phase 1, construction phase 2, and qualitative evaluation phase 2 (Fig. 1). The qualitative evaluation phase 2 not only included reliability testing but also feasibility testing.

### *Ethics*

The local medical ethics committee of the local university medical center approved the study (NL433394.042.13). All participants provided written informed consent before entering the study.

### *Participants*

There were 3 types of participants in this study: (1) raters, (2) patients, and (3) able-bodied controls.

1. Raters:
   - Construction phase 1: 3 groups, each consisting of 7 raters with a medical background, but without FCE experience, participated in 3 consecutive construction-1 tests (Fig. 1). The scale development phases are explained in Scale development section.
   - Qualitative evaluation phase 1: 4 experienced FCE clinicians (FCE experts) from the Netherlands (mean years of FCE experience: 10.8 ± 5.1 years) participated.
   - Construction phase 2: 5 FCE experts from Canada, the United States of America, Switzerland, Hong Kong, and Australia (mean years of FCE experience: 17.6 ± 5.3 years) participated.
   - Qualitative evaluation phase 2: 2 groups of raters were involved: (1) we recruited 12 FCE experts from Canada, the Netherlands, the United States of America, and South Africa (mean years of FCE experience: 16.3 ± 10.5), among participants at an international FCE research meeting[15] and (2) we recruited a convenience sample of 6 physiotherapists/gait analysts (mean years of professional experience: 7.7 ± 8.7 years), without FCE experience, from the rehabilitation department of our local medical center. We based the total sample size of 18 raters in this phase on a previous study testing the reliability of an observational scoring system scoring physical effort during FCE tests.[11]
2. Patients: Inclusion criteria for prosthesis wearers were the following: normal functioning of their sound hand and at least 1 year of experience with prosthesis use. We recruited prosthesis wearers via a questionnaire used for another study,[8,10] in which they all stated that they were interested in follow-up research. All prosthesis wearers (3 females, 9 males, mean age 46.8 ± 11.5 years, mean time of prosthesis use: 30.5 ± 17.2 years, range: 2-50 years) had an amputation or congenital deficiency at the transradial or wrist disarticulation level. Exclusion criteria were invalidating or serious health conditions that might influence the results of the FCE. We analyzed these conditions by using the Physical Activity Readiness Questionnaire.[16]
3. Controls: Inclusion criteria for controls were no joint, muscle, or nerve disorders and a normal hand function in both hands. Exclusion criteria for controls were the same as for patients. We recruited 21 able-bodied controls (3 females, 18 males, mean age 45.8 ± 11.7 years) via advertisements.

We took video recordings of 12 upper limb prosthesis wearers and 21 able-bodied controls performing FCE-OH tests (Text box 1) at the local medical center in 2014.

### *Tests*

We selected 4 out of 6 FCE-OH tests. The overhead lifting test 1-handed and the hand grip strength test were not selected because prosthesis wearers executed these tests only with the nonaffected side; hence, we expected no compensatory movements (Text box 1).

### *Scale development*

We based the development of the qualitative scoring system on the 4 phases described in the guide for instrument development and validation of Benson and Clark,[17] planning, construction, qualitative evaluation, and validation. In this study, the first 3 phases were completed. If the obtained kappa values of the items were <0.6 in qualitative evaluation phases, the construction and qualitative evaluation phases were repeated, see also Figure 1.

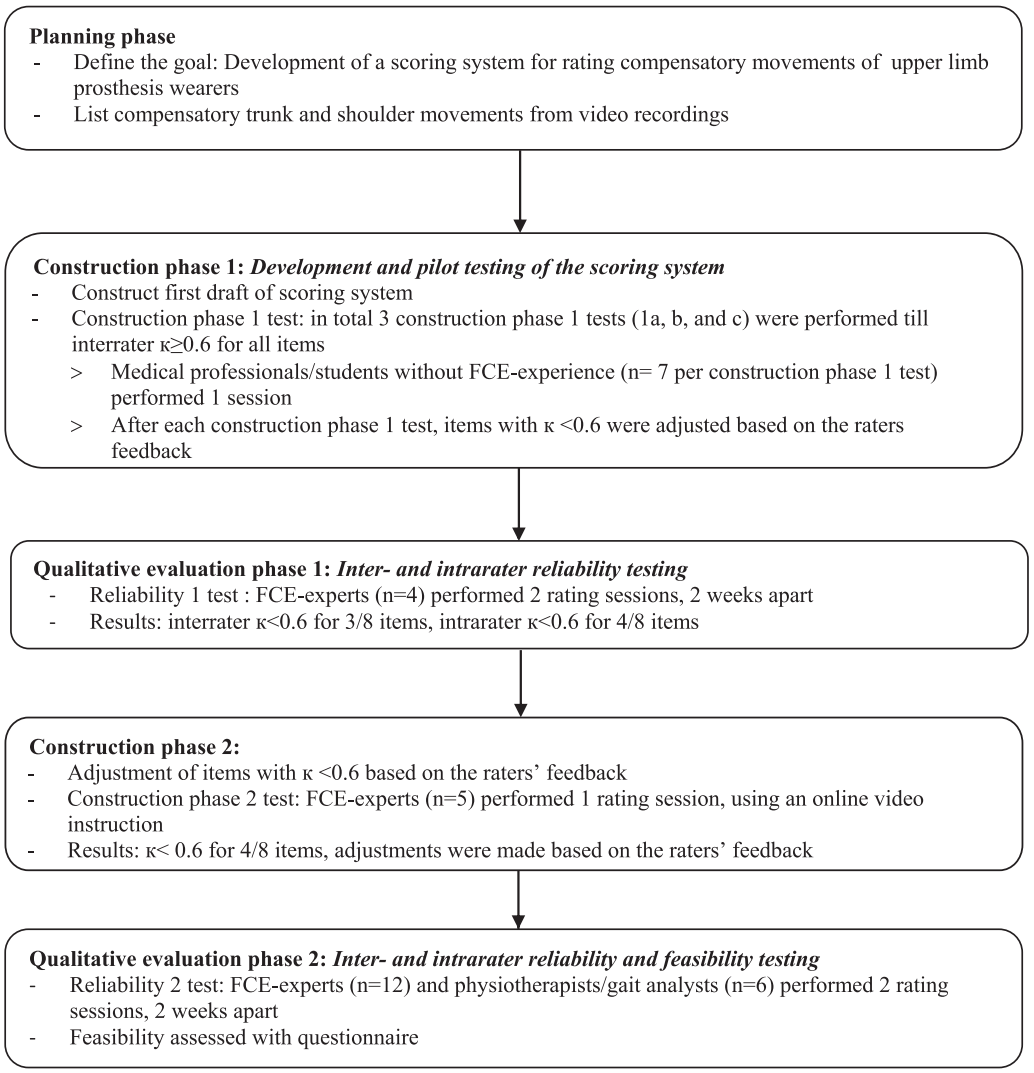


**Fig. 1.** Flowchart of the development, reliability, and feasibility testing of the scoring system for rating compensatory movements in prosthesis wearers during the performance of FCE-OH tests. κ = Cohen's kappa; FCE-OH = functional capacity evaluation-one handed.

Text box 1. FCE tests, FCE objectives, participant instructions, quantitative FCE outcomes, and the used camera position per FCE test.

| FCE test | Objective | Instructions | FCE outcomes | Camera position |
|---|---|---|---|---|
| Overhead lift test | To test the strength of shoulder and arm musculature | Lift the plastic receptacle containing weights from table height to crown height 5 times within 90 s. The start weight is 2 kg for women and 4 kg for men. After every 5 lifts weights of 2 kg for women and 4 kg for man are added until the maximum weight is reached. | The maximum lifted weight | Behind the participant |
| Repetitive reaching test | To test fast repetitive reaching movements of the upper extremity | Sit between 2 clicking systems on wing span and alternate clicking each button for a total of 30 times | Time needed to press each button 15 times | In front of the participant |
| Fingertip dexterity test | To test fingertip dexterity | Sit in front of the pegboard and place the pins in the board as fast as possible. | The number of pins placed in the board in 30 s | In front of the participant |
| Overhead work test | To test the static holding time of shoulder and neck musculature | Stand with the hands on crown height and manipulate nuts and bolts wearing a cuff weight around the nonaffected wrist | Time the position is held | Behind the participant |

FCE = functional capacity evaluation.

*Planning phase*

Four medical professionals listed compensatory movements by comparing the movement patterns of videotaped prosthesis wearers to the movement patterns of videotaped healthy controls. A compensatory movement was defined as a movement different from the control group, if (1) the maximum range of motion in a specific direction deviated more than 2 standard deviations from the mean of the control group or (2) additional movements were observed (that were not observed in the control group) like lateral flexion of the trunk during the fingertip dexterity test. We measured the range of motion manually using frames out of the videotapes in which the range of motion was maximal. In the first draft, we categorized compensatory shoulder movements as subtle or strong compensatory movements. A movement was defined as strong compensatory movement when there was a large difference in degrees of motion compared to the control group; smaller differences were defined as subtle compensatory movements. For example in the overhead work test, shoulder abduction between 97° and 120° was classified as subtle compensation and shoulder abduction over 120° was classified as strong compensation. Four medical professionals determined the cutoff points between strong and subtle compensation per FCE-OH test in a consensus meeting. They based the upper limit on the maximal range of motion per test observed in the patient group. We did not categorize additional movements that were classified as compensatory movements because the range of motion of these movements was too small to distinguish by means of observation.

*Construction phase 1*

We tested and adjusted the first draft of the scoring system in 3 rounds until the interrater kappa values of all items were >0.6. The adjustments were based on feedback of the 21 construction 1 raters (Fig. 1). Examples of adjustments were adding pictures to the scoring system, specifying descriptions of compensatory movements, and combining the categories "no compensation" and "subtle compensation."

*Qualitative evaluation phase 1*

Four FCE experts determined the interrater and intrarater reliability of this first draft. The interrater and intrarater reliability of half of the items needed improvement.

*Construction phase 2*

Construction phase 2 was performed to improve the scoring system using the feedback provided by the FCE experts in qualitative evaluation phase 1. Five international FCE experts tested the improved scoring system online, using the video fragment rating system (a secure online video streaming service). We used a video instruction to explain the use of the scoring system.

*Qualitative evaluation phase 2*

Qualitative evaluation phase 2 consisted of the interrater and intrarater reliability testing and feasibility testing. Two groups of raters performed 2 rating sessions, 2 weeks apart. Before each session, the use of the scoring system was explained by a video instruction lasting 10 minutes, which we used previously in construction phase 2. We included only good-quality videotapes for rating. Subsequently, raters had to score 4 videotapes, each with a duration of 20 seconds per item. The raters reviewed the 2 items per FCE-OH test separately. We instructed the raters to perform the rating without consulting colleagues. The first group of raters (FCE experts) performed both rating sessions online and unsupervised in June and July 2015, using the video fragment rating system. Raters were instructed to remove their ratings from their computer after the first rating session. The second group of raters (physiotherapists/gait analysts) performed both rating sessions with supervision of a researcher at the university hospital in September 2015. After the first rating session, raters had to hand in their ratings.

*Feasibility*

Bowen et al defined feasibility as various focus areas that can be used to determine whether an intervention is worth being tested in daily practice.[18] We developed a 7-item questionnaire in order to evaluate the focus areas "acceptability," "demand," and "implementation." With "acceptability," Bowen et al mean the extent in which raters accept the scoring system as being suitable for rating compensatory movements. The concept "demand" means how likely raters think the scoring system will be used in clinical practice and the concept "implementation" means the extent to which the scoring system can be successfully implemented in clinical practice.[18] After finishing the second rating session, we asked the raters to fill out the questionnaire.

*Statistical analysis*

During testing in construction phase 1, construction phase 2, qualitative evaluation phase 1, Fleiss kappa ($\kappa$) for multiple raters was used to analyze interrater reliability per item. In these phases, we considered reliability sufficient if $\kappa \geq 0.6$, and we adjusted items if $\kappa < 0.6$. Online Kappa Calculator was used to analyze Fleiss kappa.[19,20]

In qualitative evaluation phase 2, we determined the interrater reliability for both sessions using Fleiss kappa. Cohen's kappa was used to determine the intrarater reliability of the final scale, using Statistical Package for the Social Sciences (SPSS) version 22.0 software package (SPSS; IBM, Armonk, NY). In qualitative evaluation phase 2, reliability was considered sufficient if kappa values were $\geq 0.6$, moderate when $0.4 < \kappa < 0.6$, and poor when $\kappa \leq 0.4$.[21] In case of missing data, we omitted the ratings of a rater from analyses.

## Results

*Construction phases 1 and 2, qualitative evaluation phase 1*

We made several adjustments in consensus meetings during the first phases of scale development (Fig. 1). After this developmental process, the observational scoring system consisted of 8 items, 2 items per FCE-OH test (Appendix 1).

*Qualitative evaluation phase 2: interrater and intrarater reliability*

In total, 18 raters participated in both rating sessions. The mean time between first and second rating session was 14.4 ± 1.0 days. The results of reliability testing are presented in Table 1. The ratings of 1 rater for the item trunk movements of the 2-handed overhead lift test were omitted from analysis due to missing data. The overall interrater reliability was moderate in the first rating session ($\kappa = 0.54$, 95% confidence interval [CI]: 0.52-0.57) but sufficient in the second rating session ($\kappa = 0.64$, 95% CI: 0.61-0.66). The overall intrarater reliability was good ($\kappa = 0.71$, 95% CI: 0.72-0.82). The overall kappa values of physiotherapist/gait analysts were similar to the overall kappa scores of FCE experts, although a trend for higher scores in the supervised group was seen.

*Qualitative evaluation phase 2: feasibility*

Eight out of 12 FCE experts and 6 physiotherapists/gait analysts provided feedback on the scoring system. All raters stated that the

**Table 1**
Interrater and intrarater reliability of the developed scoring system

| FCE test | Items | Interrater reliability, first session | | Interrater reliability, second session | | Intrarater reliability | |
|---|---|---|---|---|---|---|---|
| | | κ (%) | 95% CI | Online κ (%) | 95% CI | κ (%) | 95% CI |
| Overhead work test ($n = 18$) | Prosthesis side | 0.53 (79%) | 0.46-0.61 | 0.71 (87%) | 0.63-0.79 | 0.75 (88%) | 0.59-0.91 |
| | Unaffected side | 0.40 (74%) | 0.32-0.48 | 0.56 (81%) | 0.48-0.64 | 0.65 (85%) | 0.45-0.84 |
| Repetitive reaching test ($n = 18$) | Reaching | 0.60 (81%) | 0.52-0.68 | 0.73 (87%) | 0.65-0.80 | 0.79 (90%) | 0.65-0.94 |
| | Press the button | 0.43 (71%) | 0.35-0.51 | 0.44 (72%) | 0.36-0.52 | 0.67 (83%) | 0.50-0.84 |
| Fingertip dexterity test ($n = 18$) | Arm movements | 0.84 (91%) | 0.76-0.92 | 0.82 (95%) | 0.82-0.97 | 0.97 (97%) | 0.92-1.03 |
| | Trunk movements | 0.12 (75%) | 0.04-0.20 | 0.06 (74%) | −0.02 to 0.14 | 0.50 (86%) | 0.47-0.53 |
| Overhead lift test ($n = 18$) | Arm movements | 0.51 (75%) | 0.43-0.59 | 0.74 (87%) | 0.66-0.82 | 0.75 (88%) | 0.60-0.90 |
| | Trunk movements | 0.27 (69%) | 0.19-0.35 | 0.23 (73%)[a] | 0.15-0.31 | 0.72 (94%)[a] | 0.54-0.90 |
| Overall FCE experts ($n = 12$) | | 0.54 (75%) | 0.52-0.57 | 0.66 (81%) | 0.63-0.69 | 0.75 (88%) | 0.68-0.82 |
| Overall physiotherapists ($n = 6$) | | 0.60 (77%) | 0.49-0.70 | 0.72 (86%) | 0.61-0.83 | 0.84 (82%) | 0.75-0.92 |
| Overall FCE experts and physiotherapists ($n = 18$) | | 0.54 (77%) | 0.52-0.57 | 0.64 (82%) | 0.61-0.66 | 0.71 (89%) | 0.72-0.82 |

FCE test = functional capacity evaluation test; κ = Fleiss kappa; % = percentage of absolute agreement; 1st session = first rating session; 2nd session = second rating session; 95% CI = 95% confidence interval; $n$ = number of raters.
[a] The ratings of 1 rater were omitted from analysis due to missing data.

system was easy to use and could be implemented easily in clinical practice. Twelve raters stated that the developed scoring system may provide clinicians with useful additional information for the assessment and treatment of upper limb prosthesis wearers, 2 raters were neutral. Standardization of assessing compensatory movements and easiness to use the scoring system were regarded as important advantages. Ten raters suggested that a training program containing more examples of compensatory movements might improve the scoring system.

## Discussion

We developed a qualitative scoring system for rating compensatory shoulder and trunk movements in upper limb prosthesis wearers during the performance of FCE-OH tests based on the guideline for instrument development.[17] The intrarater reliability of the scoring system was good. The interrater reliability of the first rating session was insufficient but sufficient in the second session. The feasibility of the rating system was established; a training program explaining the use of the scoring system was recommended.

Our study shows that it seems possible to identify compensatory movements reliably by observation using a dichotomous observational scoring system. The current version of the developed scoring system does not rate subtle compensatory movements, since construction 1 tests revealed that differentiation between subtle and strong compensatory movements was unfeasible. As such, the current scale can be used by clinicians to observe and rate large compensatory movements. We expect that when raters would have gained more experience with applying the qualitative scoring system, scoring more subtle compensation strategies should be possible. Then, the system should be extended with a rating scale for advanced raters and tested subsequently.

The interrater reliability increased over the first and second reliability tests, which is similar to the development of an FCE physical effort scale,[11] suggesting a learning effect. Raters mentioned they were accustomed to assess compensatory movements of the entire body and now they had to focus on compensatory movements of a certain body part. Based on the comments of the raters, we recommend a training program explaining the use of the scoring system and providing exercises to rate compensatory movements of certain body parts to increase the reliability.

The interrater kappa values of the trunk movement item of the fingertip dexterity test were unexpectedly low (first session $\kappa = 0.12$, second session $\kappa = 0.06$) compared to the percentage of absolute agreement (first session 75%, second session 74%). This discrepancy between the kappa values and the percentage of agreement is a known phenomenon that can be explained by the lack of variation of column and row totals.[22] In these cases, the kappa value is less meaningful. In our sample, the total of the column “compensation” is higher than the total of the column “no compensation,” causing a high expected percentage of agreement and therefore low kappa values. A more balanced distribution of the scores may be obtained by rating more videotapes per item.

We expected no differences in interrater and intrarater overall kappa values of physiotherapist/gait analysts and FCE experts, since both groups of raters did not have any specific experience in observing compensatory movements, other than might be expected due to their respective professions. We conclude that supervised or unsupervised ratings lead to comparable results, although a trend for better results was seen in the supervised group of physiotherapists/gait analysts. Physiotherapists/gait analysts could get used faster to scoring compensatory movements per body part. FCE experts are familiar with FCE tests and are used to observing movement patterns during FCE tests. Possibly, it is due to this experience that some FCE experts mentioned they also observed compensatory movements of other body parts during rating. Physiotherapists/gait analysts did not mention this. The extra compensatory movements may have distracted the FCE experts during rating and caused that they had more difficulties in rating compensatory movements per body part compared to physiotherapists/gait analysts. As the FCE-OH is rated based on the movements of certain body parts, this may explain why the gait analysts/physiotherapists scored better, even though they had less years of relevant experience, compared to the FCE experts. We expect the results of unsupervised raters to improve if a training program is provided.

Not all compensatory trunk and shoulder movements may have been observed during the planning phase of scale development due to 3 reasons. (1) We defined compensatory movements as movement patterns different from the movement pattern of controls. It may be possible that controls used compensatory movements as well and that compensatory movements of prosthesis wearers were not recognized. (2) Compensatory movements may not have been observed due to the low number of available videotapes of upper limb prosthesis wearers. Possibly there are more compensatory strategies than observed in the studied patient group. The limited number of videotapes suitable for analyses was caused by optical distortion due to insufficient standardized camera positions. We used only videos of good quality. In order to prevent optical distortion in future a standardized camera position is required, we propose that the horizontal center of the image is at shoulder

height and that the spine is in the vertical center of the image. (3) The participants were videotaped in 1 plane; compensatory movements made in other planes may have been missed. The relevance of missing these subtle compensations is unknown at present and should be subject of further (validation) study.

The developed scoring system provides a standardized method for the assessment of compensatory movements in upper limb prosthesis users. It provides insight in the use of compensatory movements during FCE-OH tests which enables therapist to intervene, so that the use of compensatory movements in upper limb prosthesis users could be minimized and MSC could be prevented.

Before the scoring system could be used in practice, the reliability should be established by rating the performance of participants in real time. Furthermore, the validity of the developed scoring system, which is step 4 in instrument development,[17] needs to be tested.

## Conclusion

We developed a scoring system to assess compensatory movements in upper limb prosthesis wearers during the performance of FCE-OH tests by observation. The scoring system was tested in the target population (FCE experts and physiotherapists/gait analysts). The intrarater reliability was good, and interrater reliability was satisfactory in the second rating session. The feasibility of the system was established and could be improved by a training program before the use of the scoring system. The scoring system provides a standardized measurement for compensatory movements and provides information about potentially dysfunctional approaches to tasks and enables therapist to intervene. Furthermore, the scoring system could be used to evaluate treatment effects. Future research is needed to determine the reliability of the scoring system in clinical practice and to determine the validity of the developed scoring system.

## Acknowledgments

The authors thank all the raters who participated in this study.

Funding: Part of this study was funded by Stichting Beatrixoord Noord-Nederland, project number 210.151.

## Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jht.2017.12.003.

## References

1. Metzger AJ, Dromerick AW, Holley RJ, Lum PS. Characterization of compensatory trunk movements during prosthetic upper limb reaching tasks. *Arch Phys Med Rehabil*. 2012;93(11):2029–2034.
2. Carey SL, Highsmith JM, Maitland ME, Dubey RV. Compensatory movements of transradial prosthesis users during common tasks. *Clin Biomech*. 2008;23(9):1128–1135.
3. Adams BD, Grosland NM, Murphy DM, McCullough M. Impact of impaired wrist motion on hand and upper-extremity performance. *J Hand Surg Am*. 2003;28(6):898–903.
4. Datta D, Selvarajah K, Davey N. Functional outcome of patients with proximal upper limb deficiency -acquired and congenital. *Clin Rehabil*. 2004;18(2):172–177.
5. Hanley MA, Ehde DM, Jensen M, Czerniecki J, Smith DG, Robinson LR. Chronic pain associated with upper-limb loss. *Am J Phys Med Rehabil*. 2009;88(9):742–751.
6. Østlie K, Franklin RJ, Skjeldal OH, Skrondal A, Magnus P. Musculoskeletal pain and overuse syndromes in adult acquired major upper-limb amputees. *Arch Phys Med Rehabil*. 2011;92(12):1967–1973.
7. Davidson J. A survey of the satisfaction of upper limb amputees with their prostheses, their lifestyles, and their abilities. *J Hand Ther*. 2002;15(1):62–70.
8. Postema SG, Bongers RM, Brouwers MA, et al. Musculoskeletal complaints in transverse upper limb reduction deficiency and amputation in The Netherlands: prevalence, predictors, and effect on health. *Arch Phys Med Rehabil*. 2016;97(7):1137–1145.
9. Reneman MF, Soer R, Gerrits EHJ. Basis for an FCE methodology for patients with work-related upper limb disorders. *J Occup Rehabil*. 2005;15(3):353–363.
10. Postema SG, Bongers RM, Brouwers MA, et al. Upper limb absence: predictors of work participation and work productivity. *Arch Phys Med Rehabil*. 2016;97(6):892–899.
11. Trippolini MA, Dijkstra PU, Jansen B, Oesch P, Geertzen JHB, Reneman MF. Reliability of clinician rated physical effort determination during functional capacity evaluation in patients with chronic musculoskeletal pain. *J Occup Rehabil*. 2014;24(2):361–369.
12. Soer R, Gerrits EHJ, Reneman MF. Test-retest reliability of a WRULD functional capacity evaluation in healthy adults. *Work*. 2006;26(3):273–280.
13. Postema SG, Bongers RM, Reneman MF, van der Sluis CK. Functional capacity evaluation in upper limb deficiency and amputation: developement and pilot testing. *J Occup Rehabil*. 2017. https://doi.org/10.1007/s10926-017-9703-4.
14. Levin MF, Desrosiers J, Beauchemin D, Bergeron N, Rochette A. Development and validation of a compensations used for reaching in patients with Hemiparesis: the reaching performance scale. *Phys Ther*. 2004;84(1):8–22.
15. Second international FCE research conference in Toronto, Canada, October 2nd 2014.
16. Thomas SS, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Can J Sport Sci*. 1992;17:338–345.
17. Benson J, Clark FA. Guide for instrument development and validation. *Am J Occup Ther*. 1982;36(12):789–800.
18. Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. *Am J Prev Med*. 2009;36(5):452–457.
19. Randolph J. *Online Kappa Calculator*. 2008. Available at: http://justusrandolph.net/kappa. Accessed August 3, 2015.
20. Chang A. *StatsToDo*. 2014. Kappa from rating scores. Available at: https://www.statstodo.com/CohenKappa_Pgm.php#Fleiss%27s. Accessed November 26, 2015.
21. Sim J, Wright CC. Interpretation, and sample size requirements the kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85(3):257–268.
22. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. the problems of two paradoxes. *J Clin Epidemiol*. 1990;43(6):543–549.

# JHT Read for Credit
## Quiz: # 624

**Record your answers on the Return Answer Form found on the tear-out coupon at the back of this issue or to complete online and use a credit card, go to *JHTReadforCredit.com*. There is only one best answer for each question.**

# 1. Raters used _____ to score performance
   a. one-on-one live observation
   b. chart reviews
   c. videography
   d. one-on-one interviews
# 2. Reliability results (k analysis) showed
   a. intrarater scores were better than interrater scores
   b. interrater scores were better than intrarater scores
   c. interrater and intrarater scores were essentially the same
   d. all measures to be unreliable
# 3. Feasibility was determined using
   a. videography
   b. patient interviews
   c. chart reviews
   d. a questionnaire
# 4. Patient subjects' performance was evaluated during
   a. their work day
   b. a strictly regimented group activity
   c. an FCE
   d. weekend recreational activities
# 5. The investigators developed a scoring system which provides standardized measurement for compensatory movements
   a. false
   b. true

When submitting to the HTCC for re-certification, please batch your JHT RFC certificates in groups of 3 or more to get full credit.