Association between social factors and performance during Functional Capacity Evaluations: a systematic review

Jone Ansuategui Echeita, Berry J. van Holland, Douglas P. Gross, Jan Kool, Peter Oesch, Maurizio A. Trippolini & Michiel F. Reneman

To cite this article: Jone Ansuategui Echeita, Berry J. van Holland, Douglas P. Gross, Jan Kool, Peter Oesch, Maurizio A. Trippolini & Michiel F. Reneman (2018): Association between social factors and performance during Functional Capacity Evaluations: a systematic review, Disability and Rehabilitation, DOI: 10.1080/09638288.2018.1448120

To link to this article: https://doi.org/10.1080/09638288.2018.1448120

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 09 Mar 2018.

Article views: 49

View supplementary material

Submit your article to this journal

View related articles

View Crossmark data
Association between social factors and performance during Functional Capacity Evaluations: a systematic review

Jone Ansuategui Echeita, Berry J. van Holland, Douglas P. Gross, Jan Kool, Peter Oesch, Maurizio A. Trippolini, and Michiel F. Reneman

ABSTRACT
Purpose: Determine the association of different social factors with Functional Capacity Evaluation (FCE) performance in adults.

Materials and methods: A systematic literature search was performed in MEDLINE, CINAHL, and PsycINFO electronic databases. Studies were eligible if they studied social factor’s association with the performance of adults undergoing FCE. Studies were assessed on methodological quality and quality of evidence. The review was performed using best-evidence synthesis methods.

Results: Thirteen studies were eligible and 11 social factors were studied. Considerable heterogeneity regarding measurements, populations, and methods existed among the studies. High quality of evidence was found for the association of FCE performance with the country of FCE and examiner’s fear behavior; moderate quality of evidence with previous job salary; and low or very low quality of evidence with compensation status, litigation status, type of instruction, time of day (workday), primary or mother language, and ethnicity. Other social factors were not studied.

Conclusions: Evidence for associations of various social factors with FCE performance was found, but robust conclusions about the strength of the associations cannot be made. Quality of evidence ranged from high to very low. Further research on social factors, also within a biopsychosocial context, is necessary to provide a better understanding of FCE performance.

IMPLICATIONS FOR REHABILITATION
• Research on Functional Capacity Evaluation (FCE) performance and its association with biopsychosocial factors have scarcely addressed the impact of social factors, limiting full understanding of FCE results.
• The social factors, healthcare (examiner’s fear behavior and type of instruction), personal or cultural systems (country of FCE, primary or mother language, and ethnicity), workplace system (previous job salary, time of day (workday)), and legislative and insurance system (compensation and litigation status), have a bearing in FCE performance.
• Better understanding of factors associating with functional capacity provide insights in FCE, allowing clinicians to improve the evaluations and interpretations of the assessment and better design the rehabilitation program.
• Better understanding of factors that influence FCE performance, and of unstudied factors, will allow researchers guidance to further investigate the construct of functional capacity.

Introduction
Several assessment tools have been developed in order to determine patients’ disabilities, to assist in claims’ decisions, and in return-to-work certifications and strategies. In situations when such an assessment is needed, several factors have to be taken into account, such as individual’s proneness to a continued reception of sickness benefit, administration’s developed regulations and financial incentives, and the interaction between healthcare providers, employers, and insurers [1,2]. Therefore, to evaluate an individual’s functional status, disability and readiness for work, the use of a standardized tool has been recommended.

One such instrument is a Functional Capacity Evaluation (FCE), which is a clinical instrument using a battery of standardized performance tests to evaluate an individual’s ability to safely perform work-related activities [3]. For the purpose of using FCE as a decision-making instrument, its reliability and validity should be established. Although some of the tests display robust evidence for reliability and validity, there is still limited evidence for many
protocols and tests [4–6]. A possible reason for this is that functional capacity tests do not only test physical functioning, but they involve personal factors, health status, and environmental factors [7]. The interaction between these factors has been explained through the bio-psycho-social model, where personal or psychological factors, individual health status or biological factors, and environmental or social factors are described [8,9].

Several studies have investigated the potential biopsychosocial factors associated with FCE performance. A Delphi study among scientists, clinicians, and patients examined the most important biopsychosocial factors that were associated with functional capacity results according to the International Classification of Functioning (ICF), Disability and Health framework [10]. Body function, activities, participation, environmental, and personal ICF components were the factors reported to be associated with functional capacity results, but only body function ICF components were part of the main factors. Similarly, two reviews examined evidence on the association of different biopsychosocial factors with performance on functional capacity tests [11,12]. These three studies provided evidence for the association between functional capacity test performance with mainly biological and psychological factors, while social factors were only scarcely described. The social factors were described as, “Factors that make up the physical, social and attitudinal environment in which people live and conduct their lives. These factors are external to individuals and can have a positive or negative influence on the individual’s performance as a member of society, on the individual’s capacity to execute actions or tasks, or on the individual’s body function or structure” [13]. Despite its relevance in patients undergoing FCE testing, it is currently unclear which social factors are associated with FCE performance.

The purpose of the present review was to determine associations between social factors and FCE performance in adults (18–65 years). The findings will provide insights in FCE, which not only will allow clinicians to improve the evaluations and interpretations of the assessment and better design the rehabilitation program, and researchers to further investigate which factors influence FCE performance; but also will add to a better understanding of the construct of functional capacity.

Materials and methods

The present review was performed using best evidence synthesis methods [14,15] and followed the PRISMA reporting guideline [16]. Relevant articles were retrieved from CINAHL, MEDLINE and PsycINFO electronic databases from conception until 31 December 2017. RefWorks (www.refworks.com) was used to remove duplicates and store the retrieved articles.

Systematic search strategy

For FCE terms selection, FCE was split into its composing elements (functional-capacity-evaluation), and synonyms of those elements were searched. Then, the FCE elements, its synonyms and the main FCE capacity tests were combined (i.e., functional capacity, physical performance, and lifting ability). Finally, different types of FCE protocols (i.e., WorkWell, Blankenship, and BTE) were also included in the search terms. The selected terms for social factors were not only related to individuals undergoing assessment; but also to workplace, insurance, legislation, and healthcare environment, and to relationships or interactions developed within those environments. Thus, factors belonging to aspects such as healthcare, sociocultural, socio-demographic, economics and related policies, social and occupational support, media and technologies were included. No biological or psychological factors such as age, pain, psychological traits or signs, or behaviors of the individual were included in the terms. The search was restricted to peer-reviewed articles and adults (18–65 years). A description of the search strategy can be found in Supplementary Table S1. Furthermore, the reference lists of the eligible full-text articles were screened, and experts in FCE were consulted by email for additional relevant published articles.

Study selection

The selection of the articles was examined by two independent reviewers (JAE and BVH). Of the potential articles, the first 100 in author’s alphabetical order were screened on title by both reviewers and the rest by the principal reviewer (JAE). Of the eligible articles following title screening, the first 50 in author’s alphabetical order were screened on abstract by both reviewers and the remaining by the principal reviewer. When in doubt about the inclusion of an article by title or abstract, it was included and further analyzed. The remaining eligible articles were all full-text analyzed by the two reviewers. The reviewers were blinded to the article authors, publishing journal and keywords in the abstract and full-text screening. In case of disagreement or doubt about the inclusion of an article, consensus was reached by discussion between the reviewers.

Inclusion and exclusion criteria

Eligibility of the studies was examined based on the type of article; population, intervention, comparison, outcomes, and study design. A description of the applied criteria can be found in Supplementary Table S2, and its implementation is represented in Figure 1. Articles were included if all six criteria were applied: (1) were peer-reviewed English, Spanish, French, or Dutch written articles (type of article); (2) reported within- or between-group comparison (comparison); (3) were observational or intervention studies (study design); (4) participants were adults (18–65 years), and were undergoing an FCE test to assess their physical capacity or functional performance (population); (5) the reported outcomes incorporated changes in FCE tests performance (outcome) related to social factors (intervention). There was no date restriction in the search. Articles were excluded if: (1) changes in FCE test performance were not measured with the purpose of assessing participants’ activity, i.e., physical capacity or functional performance; but as a means to assess participation as return-to-work, cognitive traits, or pharmacological/treatment effects (outcome); (2) no social factors were involved in FCE test performance changes (intervention); (3) individuals affected by social factors were not the ones undergoing FCE, such as FCE assessors (population). Books, meetings or conferences, interviews, reviews, master dissertations, letters to the editor, guidelines, and editorials (type of article) were excluded.

Assessment of methodological quality

The methodological quality of each of the studies was independently assessed by two reviewers (the principal author and one of the coauthors), composing a total of six reviewers. The division of the articles among the coauthors was based on the social factors studied. Additionally, having papers assessed by an author of the respective paper was avoided. Disagreements were discussed by the reviewers, until consensus was reached. The methodological quality assessment tool employed for this review was the
QualSyst (Table 1), for elaboration on the criteria applied: refer to the original article [17]. The QualSyst tool for quantitative studies is a checklist that assesses both the methodology and reporting quality of the articles. It can be used for both cohort and cross-sectional studies without penalization for study design. This checklist consists of 14 items, each of which was scored depending on the degree to which the criterion was met: yes – 2 points, partial – 1 point, no – 0 points, and not applicable (excluded from the calculation of the summary score). The summary score was obtained by dividing the total score (sum of all the positive items and all the partial items) by the total possible score (28 – (number of N/A x 2)). No further guidance for the classification of the articles was provided; as a result, it was decided that articles were rated high if they scored above the 3rd quartile (75% or over), moderate if they scored between the median and the 3rd quartile (50% and 74%), and low if they scored below the median (49% or less).

**Data extraction**

From the selected articles, the principal reviewer extracted the information. Accuracy was verified by the reviewer who assessed the methodological quality of the corresponding article. The following details were extracted: study characteristics (design, inclusion/exclusion criteria, type of intervention), sample characteristics (sample size, demographic characteristics, and social variables), and outcome characteristics (type of FCE performed, analyses, and test results). All results were presented in agreement with the original author’s judgment on the effect of the social factor on FCE test performance.

**Assessment of quality of evidence**

The quality of evidence of the outcomes was assessed using an adaptation of the GRADE Working Group guideline (Grading of
Recommendations Assessment, Development and Evaluation) for prognosis research [18] by two reviewers independently. The principal reviewer verified data extraction and graded all the outcomes along with one other reviewer who graded the outcomes of the studies they had already assessed for methodological quality. If necessary, additional information was requested from the corresponding author. In case of disagreement on the grading of an article, consensus was reached through discussion. The level of evidence was graded depending on: phase of investigation, limitations, inconsistency, indirectness, imprecision, publication bias, effect size, and exposure–response gradient. The initial quality of evidence was scored according to the phase of investigation of the study: High (++++) for explanatory studies to understand the prognostic pathways (phase 3 explanatory study) and/or to confirm associations between potential prognostic factor, and the outcome (phase 2 explanatory study); and Moderate (+++) for outcome prediction or explanatory studies to identify associations between potential prognostic factors, and the outcome (phase 1 explanatory study). The initial score was downgraded by one point (−1) or upgraded by one (+1), or stayed the same depending on the risk of bias in the different sections. The overall quality of evidence could be High (+++++), Moderate (+++), Low (++), or Very low (+) [18].

**Data synthesis**

The results of the data extraction were synthesized based on the association of the different social factors with FCE test performance per type of population, either patients or healthy participants. We planned to execute a meta-analyses when two or more studies were retrieved per factor and type of test performed, within a certain comparison. However, this criterion was not met; thus, data synthesis was done descriptively.

**Results**

**Study selection**

The results of the literature search and selection are presented in Figure 1. A total of 1634 records were obtained from the search. There were no duplicate records; therefore, 1634 articles were screened on title. This yielded 598 eligible articles. The subsequent abstract screening left 58 articles eligible for full-text assessment. The screening of the reference lists from the eligible full-text articles produced nine articles and consultation of experts in the field. The final selection of 13 articles eligible for quality assessment and data extraction. The articles excluded on abstract and full-text were mainly due to the absence of FCE test and/or social factors. A list with the full-text excluded articles can be found in Supplementary Table S3.

**Assessment of methodological quality**

The results of the quality assessment for the included articles are displayed in Table 2. Overall, nine articles showed high methodological quality [21–26,29–31], three showed moderate methodological quality [20,27,28], and one showed low methodological quality [19]. The main limitations found were: unclear study design to answer the research question, incomplete description of sampling strategy, reduced reproducible objective criteria of the measurements, sample size appropriateness, insufficient results details, and estimates of the main results. No studies were excluded based on result of methodological quality assessment.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Quality*</th>
<th>Study</th>
<th>Design</th>
<th>Quality*</th>
<th>Study</th>
<th>Design</th>
<th>Quality*</th>
<th>Study</th>
<th>Design</th>
<th>Quality*</th>
</tr>
</thead>
</table>

*Quality: low, < 50%; moderate, 50–74%; high, > 75%.
Study description

The retrieved articles were published between 1 January 2000 and 31 December 2017. Most studies were observational with use of primary data; however, two articles used either stored data from a previous study [23], or selected data from the caseload of a rehabilitation counselor and two attorneys [28]. Participants performing FCE were involved in rehabilitation programs, RTW processes, compensation claims or were healthy workers or students. The mean age of the samples ranged from 36 to 56 years, except two studies in which the examined samples were students (mean age 20.5 and 21.6) [20,30], and one study that collected normative data from participants between 3 and 101 years of age [31]. Information extracted from the studies is summarized in Table 3 for patients and in Table 4 for healthy participants.

Eleven social or environmental factors were identified. Six involved patients: compensation status [19,21], litigation status [19], country of FCE [23], previous job salary [25], type of instructions [27], and primary or mother language [28,29]; three involved healthy participants: environmental conditions [20], time of day (workday) [24], and examiner’s fear behavior [30]; and two involved both populations: self-reported workplace social support [22,26] and ethnicity [28,31]. In addition to this, many social characteristics were gathered such as education level, employment situation, and work experience. In most cases, these factors were provided only as sample demographic descriptions; hence, they were not included in the analysis of the studies. FCE measurements were composed of: stoop, climb, crouch, lift, carry, push and pull, forward bent stand, overhead work, grip strength, manual dexterity, ambulation, and sincerity of effort.

Data synthesis

Patients

Compensation status. Two studies analyzed the relation between compensation status and FCE performance [19,21]. One article reported an association between compensated individuals and lower FCE performance (floor-to-waist and waist-to-shoulder lift) [19]; whereas the other article reported that compensation status was associated with a higher rate of failed FCE tests [21]. These associations were maintained for climbing and crouching after adjusting for pain and psychological characteristics. Of the FCE tests the studies measured, only lifting capacity (floor-to-waist lift) was similar in both studies. When the association between compensation status and lifting capacity was analyzed, one of the studies found lower lifting performance, whereas the other article did not find a difference. Overall, FCE performance differed between patients receiving compensation and those who did not. When the relation was further explored it was found that patients receiving compensation had lower performance and more failed FCE tests. However, the only common test in both studies, lifting performance, showed conflicting results between the studies.

Litigation status. Only one study investigated the association between litigation status and FCE performance [19]. Its findings displayed an association between individuals involved in litigation and lower performance of lifting capacity (floor-to-waist and waist-to-shoulder lift) even after accounting for demographic, pain, psychological, and physical characteristics. In general, patients involved in litigation processes had a lower lifting performance.

Self-reported workplace social support. One study addressed this topic and reported no significant association between self-reported workplace social support and FCE performance [22]. The study investigated the relation between the patients’ perceptions of the workplace (measured with Organizational Policies and Practices (OPP) scale) and FCE performance indicators (floor-to-waist lift and number of failed tasks). Patients’ perceptions of the workplace were found to have no significant association with performance on the FCE lift test or number of failed tasks. Hence, there was evidence of no relation between self-reported workplace social support and FCE performance. This was consistent with the results found in the study with healthy participants [26].

Country of FCE. The association between country of FCE and FCE performance was studied in one article [23]. This study compared patients FCE tests’ results from three different countries (The Netherlands, Canada, and Switzerland), and found differences in lifting and carrying capacity between the Dutch sample and the other two samples. The Dutch sample lifted and carried significantly more than the Canadian and Swiss samples; this association remained significant after controlling for physical and psychological characteristics in lifting capacity. Thus, it appeared that differences in FCE performance between different countries exist.

Previous job salary. One study analyzed the relation between FCE performance and the salary of a patient’s previous job [25]. This study reported an association of higher previous job salary with larger walking distance in the 6-min walk test (6MWT), and this association remained significant after correcting for physical and psychosocial variables. The other tests did not show any relationships. In general, a higher salary was related to a larger distance walked, without any other relation to FCE test performances.

Type of instruction. One study investigated the association between different types of instruction for the execution of the 6MWT and its performance [27]. The study reported that the type of instructions given by the examiners had an effect on walked distance. The different instructions were: “walk as far as possible” (standard walk), “walk as fast as you can” (fast walk), “walk at your normal pace” (normal walk), and “walk at a leisurely pace” (leisure walk). Patients walked a mean of 52.7 m more when they were instructed for the fast walk, compared to the standard walk. In general, there was a difference in walked distance when patients were instructed in different ways.

Primary or mother language. Two studies investigated the association of the primary or mother language of the patients on FCE performance [28,29]. Both studies found differences in FCE performance associated with participant’s mother language groups in the German canton of Switzerland (German and non-German) and United States (English and non-English). The study comparing primary German and non-German language speakers evidenced that the German-speaking group performed consistently higher on all FCE tests included in the study, i.e., handgrip strength, waist-to-overhead lift, and overhead work and repetitive reaching task [29]. The study comparing primary English and non-English language speakers evidenced that the English-speaking group showed a moderately higher percentage of valid performance effort (70.2%), with respect to the non-English-speaking group (27.3%) [28]. Therefore, evidence consistently showed FCE performance to be higher in patients whose primary or mother languages was the local language.

Ethnicity. One study analyzed the relation between race or ethnicity and FCE sincerity of effort [28]. The study outcomes showed a moderate difference in the number of valid FCE performances between two ethnic groups, White/Caucasian and non-White. The White/Caucasian group showed higher percentage of valid performance effort (72.0%), compared with the non-White group (38.9%). As a result, the authors concluded that differences in validity of FCE performance were related to patient’s ethnicity.
<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Social Factors</th>
<th>Outcome</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geiser et al. (2000) [19] USA</td>
<td>Chronic disabling back pain assessed at a Spine Program, n = 133</td>
<td>Compensation and litigation status</td>
<td>Progressive Isoinertial Lift Evaluation (PILE)</td>
<td>Compensation and litigation, both showed associations with decreasing floor-to-waist and waist-to-shoulder lifting performances. Litigation was a negative predictor of floor-to-waist lift also in the simultaneous regression model. Workers’ compensation status at admission showed differences in the rates of failure of stopping, climbing, and coughing in climbing and coughing logistic regression models, receiving compensation was associated with higher failure rates. At discharge, workers’ compensation status showed differences in the rates of stopping and coughing. No apparent relationship between claimant’s ratings on OPP scale and PCE performance was found.</td>
</tr>
<tr>
<td>Cutter et al. (2000) [21] USA</td>
<td>CLBP admitted to a multidisciplinary pain treatment center, n = 188</td>
<td>Compensation status</td>
<td></td>
<td>Dutch patients showed higher FCE lifting and carrying performance than Canadian and Swiss patients. In final lifting regression models, this relationship remained significant.</td>
</tr>
<tr>
<td>Gross and Battel (2005) [22] Canada</td>
<td>Low back injured workers for work assessment on claims and RTW decision-making purposes, n = 136</td>
<td>Organizational Policies and Practices (OPP): Claimants’ perception of support at workplace</td>
<td>Dictionary of Occupational Titles Residual Functional Capacity battery (DOT-RFC)</td>
<td>No apparent relationship between claimants’ ratings on OPP scale and PCE performance was found.</td>
</tr>
<tr>
<td>Reneman et al. (2006) [23] Switzerland</td>
<td>CLBP patients: The Netherlands, n = 121; Canada (n = 273); Switzerland (n = 170)</td>
<td>Countries</td>
<td>Isernhagen Work System Functional Capacity Evaluation (IWS-FCE)</td>
<td>No apparent relationship between claimants’ ratings on OPP scale and PCE performance was found.</td>
</tr>
<tr>
<td>Olsich et al. (2013) [24] Switzerland</td>
<td>Patients with chronic NSLBP referred for fitness-for-work evaluation on three rehabilitation centers, n = 126</td>
<td>Previous job salary (CHF)</td>
<td></td>
<td>No other relations were found.</td>
</tr>
<tr>
<td>Weir et al. (2013) [25] USA</td>
<td>Patients recruited from an Advanced Lung Disease Clinic, n = 24</td>
<td>Alternative instruction: Forward walk; Fast walk; Normal walk; Leisure walk</td>
<td>Isernhagen Work System Functional Capacity Evaluation (IWS-FCE)</td>
<td>Greater distance was walked during the fast walk compared with the standard walk.</td>
</tr>
<tr>
<td>Rutherford and Jones Williams (2014) [26] USA</td>
<td>Cases from a rehabilitation counselor and two attorneys, regularly involved in the workers’ compensation system, n = 68</td>
<td>Ethnicity and Primary language: White/Caucasian vs. non-White (n = 30 vs. n = 18); English vs. non-English (n = 57 vs. n = 11)</td>
<td>Sincerity of effort measures for FCE</td>
<td>There were differences in sincerity of effort scores between ethnic and primary language groups. White/Caucasian workers and English speakers showed a higher percentage of valid performance effect. The relation between ethnicity and primary language with sincerity of effort scores showed medium effect sizes.</td>
</tr>
<tr>
<td>Trippolini et al. (2015) [27] Switzerland</td>
<td>WAD grade I or II patients for rehabilitation assessment at the rehabilitation clinic, n = 314</td>
<td>Primary language: German vs. Non-German (n = 152 vs. n = 162)</td>
<td>WorldWell Functional Capacity Evaluation (WWSFCE)</td>
<td>Relationship between FCE performance and patient’s primary language spoken was found. The German-speaking group showed higher performance in all four FCE tests.</td>
</tr>
</tbody>
</table>

LBP: low back pain; CLBP: chronic low back pain; NSLBP: non-specific low back pain; WAD: Whiplash associated disorders. Significance: *p < 0.05; **p < 0.01; ***p < 0.001. 
1Mean (SD). 2Mean (range). 3Cultural background was measured based on the primary language, i.e., the mother language.

\( \phi \): the phi coefficient is a Pearson’s product-moment coefficient of correlation calculated on two nominal-dichotomous variables when the categories of both variables are coded 0 and 1. This measure is used to estimate the magnitude of association in 2 × 2 contingency tables, as in the chi-squared test \( \phi = \sqrt{\frac{\chi^2}{n}} \).
<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Social Factors</th>
<th>Outcome</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reneman et al. (2001) [20]</td>
<td>Healthy students, n = 24</td>
<td>Environmental conditions</td>
<td>Bernhagen Work System Functional Capacity Evaluation (IWS-FCE) – Production rate</td>
<td>The three environmental conditions did not appear to influence FCE performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal</td>
<td>Overhead work (n/min)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.6 (1.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.4 (1.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.6 (1.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Max. production</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.3 (1.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.6 (1.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.4 (1.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kneeling (n/min)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.0 (2.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.4 (1.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.4 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Kyi et al. (2012) [24] Canada</td>
<td>Healthy working age adults, n = 50</td>
<td>Workday</td>
<td>FCE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Morning</td>
<td>Right grip (kg)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Afternoon</td>
<td>47.33 (13.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45.58 (13.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45.92 (13.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Left grip (kg)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45.58 (13.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46.11 (14.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45.92 (5.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Manual dexterity (s)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50.92 (5.6)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46.11 (14.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45.92 (5.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.83 (1.0)</td>
<td></td>
</tr>
<tr>
<td>Lakke et al. (2013) [26] The Netherlands</td>
<td>Healthy workers, n = 403</td>
<td>Perception of workplace support (QAEW)</td>
<td>FCE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Co-worker support</td>
<td>Low lift (kg)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Supervisory support</td>
<td>Overhead lift (kg)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Support</td>
<td>Carry (kg)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Forward bent standing (s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Overhead work (s)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρ = −0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρ = 0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρ = 0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρ = 0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρ = −0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρ = −0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ρ = −0.07</td>
<td></td>
</tr>
<tr>
<td>Lakke et al. (2015) [30] The Netherlands</td>
<td>First- and second-year physical therapy healthy students, n = 256 participants</td>
<td>Examiner’s fear behavior</td>
<td>Lifting capacity (kg)*</td>
<td>Participants with the low-fear examiner scored higher than participants with the high-fear examiner. In the final model, this association remained significant</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low fear</td>
<td>39.6 (16.4)***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High fear</td>
<td>B = 14.41 (3.6)**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>32.1 (13.6)</td>
<td></td>
</tr>
<tr>
<td>McKay et al. (2017) [31] Australia</td>
<td>Healthy population, n = 399</td>
<td>Ethnicity</td>
<td>6-Min Walk Test (m)*</td>
<td>Differences in performance between the ethnic groups were found in one of the FCE tests, British/Europeans walked significantly longer distance. No differences were found for the other FCE tests</td>
</tr>
<tr>
<td></td>
<td></td>
<td>British/European (n = 265)</td>
<td>711.7 (91.4)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>non-British/European (n = 134)</td>
<td>711.7 (91.4)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17.7 (2.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17.5 (2.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20.5 (4.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20.6 (4.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>693.1 (80.6)</td>
<td></td>
</tr>
</tbody>
</table>

QAEW, Dutch Questionnaire on the Appreciation and Evaluation of Work. Significance: *p < 0.05; **p < 0.01; ***p < 0.001. 
*Mean (SD). **Mean (range).
Table 5. Quality of evidence of social factors addressed by eligible articles (GRADE).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Phase of investigation (RCT, Phase 3 or 2/Phase 1)</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Publication bias</th>
<th>Large effect</th>
<th>Exposure-response</th>
<th>Overall quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compensation status</td>
<td>2 (0/2)</td>
<td>– 1</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>– 1</td>
<td>Undetected</td>
<td>No large effect</td>
<td>No exposure-response gradient</td>
<td>+</td>
</tr>
<tr>
<td>Litigation status</td>
<td>1 (0/1)</td>
<td>– 1</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>No large effect</td>
<td>No exposure-response gradient</td>
<td>++</td>
</tr>
<tr>
<td>Country of FCE</td>
<td>1 (0/1)</td>
<td>No serious limitations</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>No large effect</td>
<td>No exposure-response gradient</td>
<td>+++</td>
</tr>
<tr>
<td>Previous salary</td>
<td>1 (0/1)</td>
<td>No serious limitations</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>– 1</td>
<td>Undetected</td>
<td>No large effect</td>
<td>No exposure-response gradient</td>
<td>++</td>
</tr>
<tr>
<td>Alternative instruction</td>
<td>1 (0/1)</td>
<td>No serious limitations</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>– 1</td>
<td>Undetected</td>
<td>+ 1</td>
<td>No exposure-response gradient</td>
<td>++</td>
</tr>
<tr>
<td>Primary or mother language</td>
<td>2 (0/2)</td>
<td>No serious inconsistency</td>
<td>– 1</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>Results pointing in the same direction</td>
<td>No exposure-response gradient</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>2 (0/2)</td>
<td>– 1</td>
<td>– 1</td>
<td>– 1</td>
<td>– 1</td>
<td>Undetected</td>
<td>No large effect</td>
<td>No exposure-response gradient</td>
<td>0</td>
</tr>
<tr>
<td>Workplace</td>
<td>2 (0/2)</td>
<td>– 1</td>
<td>No serious inconsistency</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>+ 1</td>
<td>No exposure-response gradient</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>social support</td>
<td>+ + +</td>
<td>+ + +</td>
<td>+ + +</td>
<td>+ + +</td>
<td>+ + +</td>
<td>+ + +</td>
<td>+ + +</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>Environmental conditions</td>
<td>1 (0/1)</td>
<td>No serious inconsistency</td>
<td>– 1</td>
<td>– 1</td>
<td>Undetected</td>
<td>No large effect</td>
<td>No exposure-response gradient</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Time of day</td>
<td>1 (0/1)</td>
<td>No serious inconsistency</td>
<td>– 1</td>
<td>– 1</td>
<td>Undetected</td>
<td>No large effect</td>
<td>No exposure-response gradient</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Examiner’s behaviors</td>
<td>1 (1/0)</td>
<td>No serious limitations</td>
<td>No serious inconsistency</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>+ 1</td>
<td>No exposure-response gradient</td>
<td>+++</td>
<td></td>
</tr>
</tbody>
</table>

RCT: randomized controlled trial.

+++ (++), high-quality evidence; ++ (+), moderate-quality evidence; +, low-quality evidence; + or 0, very low-quality evidence; – 1, quality of evidence is downgraded by one +; + 1, quality of evidence is upgraded by one +.
Healthy participants

Environmental conditions. One study analyzed the association between the environmental conditions and FCE performance [20]. Environmental conditions were operationalized as: normal (tests performed in consonance with the protocol), maximal production (tests of screws/bolts manipulation performed as fast as possible), and noise (tests performed with annoying volume level). The study results displayed no differences in FCE performance between three environmental conditions. Therefore, the evidence showed there was no relation between these environmental conditions and FCE performance.

Time of day (workday). One study investigated the association of time of day or effect of workday, and FCE performance [24]. FCE measurements were taken in the morning and later in the afternoon after a day's work. Participants were able to perform relevantly faster on the manual dexterity test (5 s), and to walk faster on the 50-feet walk test (0.19 s) in the afternoon. Opposed to this, handgrip performances did not show differences. In general, the afternoon (after work) measurements were significantly associated with an improvement in FCE performance, specifically on manual dexterity and walking velocity; no other FCE test performance showed a significant association with this factor.

Self-reported workplace social support. One study addressed the topic of self-reported workplace social support and the results reported no associations between self-reported workplace social support and FCE performance [26]. This study examined the relation between healthy participants’ self-reported workplace social support: coworker and supervisor support (measured with the Dutch Questionnaire on the Appreciation and Evaluation of Work (QAEW)), and FCE performance. Participant’s coworker and supervisory support was not significantly associated with performance on any of the measured FCE tests. Hence, it was evidenced there was no relation between self-reported workplace social support and FCE performance. This was consistent with the results found in the study with patients [22].

Examiner’s fear behavior. Fear behavior issue was evaluated in one study [30]. The study reported that examiners’ fear behavior during testing significantly influenced lifting performance of the participants. High fear of injury examiners showed a greater guarding behavior by expressing more frequently symptom-focused talks, lifting avoidance, reassurance, ergonomic verbal instructions, procedural talk, and examiner’s decisions. Participants evaluated by a low-fear examiner lifted 7.4 kg more than participants with a high-fear examiner; the association remained after correcting for participant’s physical and personal characteristics. Overall, there was a difference on lifting performance between the groups tested with high- or low-fear examiners.

Ethnicity. One study investigated the association of ethnicity and FCE performance [31]. The study compared a British/European ethnic group with non-British/Europeans in several tests. Of the tests that belong to the FCE battery, differences between the groups were only observed on the 6MWT. The British/European ethnic group (711.7 m) performed significantly better than the non-British/European (693.1 m). In general, the British/European ethnic group walked a longer distance in the 6MWT. No other FCE test performance showed a significant difference between ethnic groups.

Assessment of quality of evidence

The quality of evidence was graded for 11 factors: compensation status, litigation status, self-reported workplace social support, country of FCE, previous job salary, type of instruction, primary or mother language, ethnicity, environmental conditions, time of day (workday), and examiner’s fear behavior. The results of the quality of evidence assessment for the included articles are displayed in Table 5. Overall, the level of evidence for the factors studied in patients was high for country of FCE [23], moderate for previous job salary [25], low for litigation status [19] and type of instruction [27], and very low for compensation status [19,21] and primary or mother language [28,29]. In healthy participants, the level of evidence for the factors studied was high for examiner’s fear behavior [30], and very low for time of day (workday) [24] and environmental conditions [20]. Only two factors, self-reported workplace social support [22,26] and ethnicity [28,31], were studied in both patients and healthy participants, and showed low and very low level of evidence respectively. The initial score was moderate for 10 factors due to being cross-sectional observational studies, and the phase of investigation they belonged to was 1. The subtraction of + from was a consequence of: limitations, inconsistency, indirectness, and imprecision. The addition of + was due a large magnitude of effect.

Discussion

The main purpose of this study was to systematically review and summarize the association between various social factors and FCE performance in adults. Thirteen studies were found to investigate the association between FCE performance and 11 different social factors. There was high quality of evidence that patients performing FCE tests in different countries [23] had significant differences in FCE results, and healthy participants tested in presence of a low-fear examiner [30] lifted significantly greater weight in the floor-to-waist lift test. There was moderate quality of evidence that higher previous job salary was significantly associated with larger distance walked in the 6MWT [25]. There was low or very low quality of evidence that patients not involved in compensation [19,21] and/or litigation [19] processes, and healthy participants tested after a day’s work [24] performed significantly better on FCE tests; also, patients with differences in the instruction [27] and in their primary or mother language [28,29], and both patients and healthy participants with different ethnicity [28,31], had significantly different FCE performances. There was an absence of association of workplace social support reported by patients and healthy participants [22,26] and the different environmental conditions, in which FCE was undertaken [20], with FCE performance.

Although social characteristics such as employment, education, claims status, and disability have served as sample characteristics in FCE research, only two reviews have analyzed the association between psychosocial factors and FCE [11,12]. Both reviews found scarce evidence on the association of compensation, litigation, and secondary gain with FCE performance. To our knowledge, this is the first systematic review that specifically focuses on social factors belonging to the individuals’ legislative and insurance, workplace, healthcare, and personal or cultural systems, thereby providing a broader overview of a large variation of social factors associated with FCE performance. This is consistent with Loisel’s arena of work disability, which emphasizes the importance of these domains and multiple factors [32].

In agreement with the definition of FCE, FCE results should be interpreted considering patients’ physical functioning, personal factors, health status, and environmental factors [7]. The findings of this review contribute to our understanding of associations between social factors and FCE performance. FCE results must be interpreted in a broader context to provide a more comprehensive representation of patients’ performance. As a consequence of
differences between jurisdictions, reference values may not be
generalizable. Country specific reference values may need to be
developed, which may lead to better interpretations of FCE results
for each patient. As a result, optimized recommendations based
on these as well as more effective rehabilitation programs due to
a better targeting of patients may be possible. Additionally, it is
important to outline that some associations between FCE and
social factors may be adaptable. For example, clinicians assessing
FCE, being a social factor themselves, could adapt their attitudes
or behaviors, or instructions which may influence patients’
performance.

There are some limitations in this systematic review that need
to be considered. To begin with, even though a broad search
strategy was applied, we cannot be fully confident that relevant
articles have been missed in the search. However, it is assumed
that the number of absent articles is minimal and that they would
not likely alter the conclusions. Another limitation is the broad-
ness of social and environmental factors and their inherent rela-
tionship with psychosocial factors. As some social factors produce
a psychological effect in the individual, it is difficult to discern the
boundaries between these domains. For example, while fear
avoidance could be considered a psychological factor, this may
(partly) be induced by social factors such as healthcare profes-
Sional behaviors [33]. Therefore, there might be some social–psy-
chosocial factors not included in the review; nevertheless, we trust
that the main and more obvious social factors have been
included. Additionally, the methodological quality assessment tool
used in this review was limited due to the fact that it does not
define the benchmarks between high-, moderate-, and low-quality
studies [17]. For that purpose, we established cutoff values at the
median and 3rd quartile, a choice that has been previously imple-
mented in similar situations [34]. Many different FCE tests and
protoclos were considered in this systematic review, which have
varying evidence of reliability and validity. As a result, variability
across protocols may have influenced the trustworthiness of our
results. A final limitation is the difficulty inherent in measuring
important social factors such as patient’s ethnicity. One study ana-
lyzed the relation of patient’s ethnicity to “valid” FCE perform-
ces instead of to FCE performance itself [28]. However, there was
no definition or description of what constituted a “valid” FCE per-
formance. Secondly, while the authors concluded that differences
in validity of FCE performance were related to patient’s ethnicity,
they also stated: “Examination of actual administrations across eth-
nic and language groups was not undertaken in this study and
may provide information regarding causes of invalid assessments
for non-English speaking workers. In this sample, it is unknown
which assessments were conducted with an English/Spanish trans-
lator” [28, p. 59]. Given these limitations we concluded that there
was no convincing evidence that ethnicity by itself is associated
with FCE performance.

There are also strengths of this study to be accounted for. The
procedure has been transparent, following a systematic search
strategy in various databases, using reliable methodological and
quality appraisal assessment tools, and best evidence synthe-
sis methods.

Quality of evidence is viewed as of paramount importance for
the interpretation of the study results. This was high for associa-
tions with the country where FCE is performed. Neither physical
nor psychological characteristics of the investigated population
explained the performance difference in this study [23]. It can
therefore be hypothesized that certain characteristics of these
countries, such as health care, personal or cultural systems may
explain the performance difference. High quality of evidence was
also found for an association of examiners fear behavior with FCE
performance which can be also viewed as an aspect of the health
care, personal or cultural systems. Moderate quality of evidence
was found for salary in the previous job which is viewed as one
aspect of workplace system factors. All other findings of associa-
tions of legislative and insurance (compensation and litigation)
systems as well as of health care (type of instruction), personal or
cultural (primary or mother language and ethnicity), and work-
place (workplace social support, environmental condition, and
time of day) system factors with FCE performance were of low or
very low quality.

Overall, the findings showed that social factors have a bearing
in FCE performance and these should be carefully considered in
future studies. First, the FCE tests as well as the social factors
examined were not uniformly assessed in all studies: the majority
of the included articles studied a lifting capacity test but only 4
out of the 11 social factors were considered in more than one
study. Second, 5 studies did not measure the typical FCE test tar-
get population, instead healthy workers’, healthy population, or
students’ populations were assessed [20, 24, 26, 30, 31]. Although
the quality of evidence in the factors they studied was downgraded
for this, the generalizations to patients with musculoskeletal dis-
esases should be made with caution. Finally, the studies included
were mainly cross-sectional in its majority and, therefore, any
causal relations should be avoided. Future research should further
investigate on which and how social factors, also within a biopsys-
chosocial context, influence FCE. Because social factors serve as a
description of certain characteristics of the participants under-
going FCE tests, guidance to further investigate FCE construct
could also be obtained. Based on the quality of evidence results
from our review, special attention should be paid to personal or
cultural, health care, and workplace systems (as represented by
country of FCE, clinician’s fear beliefs, and previous job salary).

Conclusions
A variety of social factors were found to be associated with per-
formance during FCEs. The considerable heterogeneity found in
the measures, the populations, and the methods, prevent robust
conclusions about independent associations of social factors on
FCE performance. Further research on social factors is required to
have a more comprehensive understanding of FCE performance.

Acknowledgements
The authors would like to express their gratitude to A. R. den
Otter for his assistance and recommendations to the development
of the review.

Disclosure statement
Five of the review coauthors (D.P. Gross, J. Kool, P. Oesch, M.A.
Trippolini, and M.F. Reneman) were also authors and/or coauthors
in some of the included articles. According to the Cochrane
Review Guidelines and to avoid conflict of interest these coau-
thors were not involved in either the methodological quality or
the quality of evidence appraisals of their studies, nor in the data
analysis of this study [35]. All authors declare they have no con-
lict of interest.

Funding
This systematic review received no funds for its development.
References


