Can a single-item measure assess physical load at work?
An analysis from the GAZEL cohort.
Erika Sabbath, Marcel Goldberg, Qiong Wu, Alexis Descatha

To cite this version:

HAL Id: inserm-00720693
https://www.hal.inserm.fr/inserm-00720693
Submitted on 1 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Can a single-item measure assess physical load at work? An analysis from the GAZEL cohort

Erika Sabbath, MS
1. Department of Society, Human Development, and Health, Harvard School of Public Health, Boston, MA USA
2. Population-based Cohort Platform, Epidemiology and Public Health Research Center, UMRS INSERM 1018, Villejuif, France
3. Université de Versailles St-Quentin, UMRS 1018, France

Marcel Goldberg, MD, PhD
2. Population-based Cohort Platform, Epidemiology and Public Health Research Center, UMRS INSERM 1018, Villejuif, France
3. Université de Versailles St-Quentin, UMRS 1018, France

Qiong Wu, PhD
4. Harvard Center for Population and Development Studies, Cambridge, MA USA

Alexis Descatha, MD, PhD*
2. Population-based Cohort Platform, Epidemiology and Public Health Research Center, UMRS INSERM 1018, Villejuif, France
3. Université de Versailles St-Quentin, UMRS 1018 France
5. Unité de pathologie professionnelle U1018, CHU Poincaré, 104 bd Poincaré 92380 Garches, France
alexis.descatha@inserm.fr
Tel : + 33 01 77 74 74 26
Fax : +33 01 77 74 74 03

* Corresponding author

Running head: Validity of a work-related physical strain measure

Sources of financial support: The GAZEL Cohort Study was funded by EDF-GDF and INSERM and received grants from the ‘Cohortes Santé TGI Program’, Agence nationale de la recherché (ANR) and Agence française de sécurité sanitaire de l’environnement et du travail (AFSSET).

Acknowledgements: The authors express their thanks to EDF-GDF, especially to the Service Général de Médecine de Contrôle, and to the “Caisse centrale d’action sociale du personnel des industries électrique et gazière.” We also wish to acknowledge the Cohortes team of the Unité 1018 Inserm responsible for the GAZEL database management.

Competing interests: None declared
ABSTRACT 135 words

Objective: Assessment of workplace physical load is highly resource-intensive. This study tested whether a single-item measure asking individuals about perceived physical strain (PPS) at work was an acceptable proxy for physical load.

Methods: The study was conducted in a subset of GAZEL cohort (n=2612) undergoing assessment of exposure to 38 occupational biomechanical constraints (representing eight domains) in 1994. Test-retest reliability analyses compared PPS in 1994 and 1995. Validity analyses compared PPS in 1994 to concurrent strains assessed in the more extensive measure.

Results: The measure showed adequate test-retest reliability. Within and across domains of physical load, linear relationships (p<0.0001) existed between n exposures and PPS. Domains considered more strenuous (carrying loads, pulling objects) showed the highest PPS.

Conclusions: PPS approximates physical load in absence of detailed measures. PPS could be used in non-occupational epidemiologic studies.
INTRODUCTION

Assessment of work-related risk factors or exposures is at the core of occupational epidemiology (1). Whether conducted by expert assessment, job-exposure matrices, or self-report (2), precise and accurate evaluation of exposure to occupational hazards is critical for verifying risk factors, establishing exposure limits, and informing policy.

While such detailed assessments are necessary when occupational factors are primary exposures of interest, the time- and resource-intensive nature of these methods means that occupational exposures are difficult to assess when they constitute secondary research questions. Thus, information on physical workplace environments among participants in most general cohorts of working-age populations is limited. However, a general sense of working conditions would be beneficial in studies for which work-related physical exertion could be a confounder or risk factor for outcomes as diverse as absenteeism, disability, depression, and overall health status (3-5). This gap is especially salient for low-income working populations, which tend to have high levels of work-related physical exertion and biomechanical strain and also higher rates of these outcomes (6, 7). Additionally, longitudinal studies of working populations could employ such a question to understand the role of the workplace in the etiology of disease and disability in retired populations (8).

One challenge in understanding workplace biomechanical constraints in a large, diverse cohort or population is the intensive nature of traditional methods for evaluation of such hazards, and thus the prohibitive cost of assessing such hazards on a large scale. Self-reported exposures have been extensively studied as alternatives to in-person assessments of physical or ergonomic hazards. Although direct observation or measurement of workers is considered the gold standard
measure for ergonomic strain, self-report measures—which are less intensive and costly to administer—are often used to assess physical exposures in occupational studies (9-11). Additionally, self-reported measures of physical strain at work are thought to be well-suited to studies for which the goal is to rank workers by relative physical load on a population level or describe patterns in load distribution, rather than evaluate specific exposure-outcome associations that require a more nuanced exposure assessment (12, 13). In a systematic review of the performance of self-report measures for physical work demands, authors concluded that many such instruments demonstrated satisfactory reliability and validity (14).

Given evidence from several occupational population-based studies that self-reported constraints are considered valid when assessing relative levels of physical load (15-17), our study tests whether a single-item measure of perceived physical strain (PPS) at work is an acceptable proxy measure for workplace physical load in large, diverse working population. We hypothesized that higher self-reported PPS would be associated with a higher level of physical workload within and across eight domains of physical constraints in a socioeconomically diverse population of French utility workers.

METHODS

Study population

This analysis was conducted within GAZEL, a French occupational cohort composed of 15,011 men and 5,614 women who were employees at Electricité de France-Gaz de France (EDF-GDF) in 1989. Recruitment strategy and baseline demographics are described elsewhere (18). Subjects complete an annual self-report questionnaire, supplemented by EDF-GDF administrative records. In 1994, a subset of participants, including all blue/white collar workers, half of the managers, and one quarter of the supervisors (n=4,786), was invited to participate in a
comprehensive assessment (Affections Rhumatologiques PEri-articulaires et GEstes professionnels; ARPEGE) of current and past exposure to 38 postural and ergonomic constraints. Of the 2,676 respondents (56% response rate), 2,661 were still working (not yet retired) in the year preceding the survey, and 2,612 also answered the PPS question and so were eligible for inclusion in the present analysis.

**Measures**

The present study tests reliability and validity of the PPS question, “Do you find that your work is physically strenuous?” as a proxy measure of general occupational physical exertion or load. Response options range from 1 (not at all strenuous) to 8 (very strenuous). The entire cohort answers this question annually in self-report questionnaires from baseline (1989) until retirement, death, or loss to follow-up, including 1994, the year in which more extensive biomechanical constraints were measured in the ARPEGE subsample.

Items from the 38-item ARPEGE sub-study were employed to validate the PPS measure. These comparisons were designed to function as a form of construct validity in the present analysis, given that no gold standard or criterion was available and the constructs being tested are related but not identical (actual versus perceived physical strain). In the ARPEGE self-administered questionnaire, individuals were asked about current and past exposure to 38 postural constraints; for each item, individuals were dichotomized as being exposed or not to the constraint during the past year.

From these constraints, study authors empirically identified eight distinct domains of physical strain and the corresponding Cronbach’s $\alpha$ to confirm the unidimensionality of each domain. The identified domains of physical strain were: constraints affecting knees (working in crouching position, working on knees, holding and/or carrying loads with knees bent; $n$
constraints=3, $\alpha=0.86$), holding arms above shoulders (with or without supporting a load; $n$
constraints=2, $\alpha=0.74$), elbow strain (work supported by elbow or heel of hand; $n$
constraints=2, $\alpha=0.62$), carrying heavy loads (carrying loads on shoulder, pushing or pulling heavy loads; $n$
constraints=2, $\alpha=0.72$), use of tools (painting with brush, roller, or spray gun; arc welding; using
pliers or tongs; use of hammering hand tools; use of strong draws, hand rotary cutters, cutting
steel tubes; screwing; hammering; engraving; manually clamping nuts or bolts; opening and
closing valves; $n$ constraints=14, $\alpha=0.94$), pulling (wrapping, pulling cables/hoses/cords,
working with poles; $n$ constraints=3, $\alpha=0.78$), walking (for $>2$ hours/day or on rough terrain; $n$
constraints=2, $\alpha=0.75$), or climbing (climbing or descending stairs, ladder, or stepladder;
standing on ladder for 30 minutes or longer, $n$ constraints=4, $\alpha=0.74$). While the data on all 38
constraints were collected simultaneously, ARPEGE is treated as an index rather than a scale;
indeed, published studies using ARPEGE data have used subsets of items to explore specific
exposure-outcome relationships (19). However, because many tasks inherently involve more
than one domain of constraints (for example, painting while standing on a ladder), the domains
are themselves correlated. Moreover, analysis of the relationship between PPS and each domain
of physical load allows us to understand the types of physical work best captured by a single-
item measure such as PPS.

Other covariates included sex, age in years at testing, and occupational grade at testing
(executive, mid-level professional including foremen, unskilled non-manual, or unskilled
manual).

*Analytical approach*
We first examined the distribution of PPS in the population and by gender and occupational grade. We used one-way ANOVA to test whether mean PPS was significantly different across gender and grade strata.

Next, we determined the measure’s test-retest reliability. Ideally the test-retest period is two weeks to one month because a longer time between testing occasions is generally associated with lower reliability (20); however, PPS was only collected annually. Therefore, the best available measure of test-retest reliability was the Spearman correlation ($r_s$) between PPS in 1994 and 1995. Since a change in job level in the intervening year would likely be associated with a change in actual physical work duties, we partialed out the effect of change in occupational grade and also excluded those who retired between the 1994 and 1995 surveys (n=28).

To assess the construct validity of PPS, we compared the PPS measure in 1994 with each of the eight domains of physical constraints assessed by ARPEGE. Within each domain of ARPEGE, we calculated the mean reported level of PPS by number of reported ARPEGE constraints in that domain and tested for linear trend. We then assessed the number of domains of physical strain to which an individual reported any exposure in the past year, determined the mean reported level of PPS at each level of total reported constraints, and tested for linear trend. Given prior evidence of gender differences in the nature of physical demands at work, the muscular activity involved in a given task, and the perception of workplace strains (21, 22), we conducted the latter analyses stratified by gender.

All analyses were performed using SAS 9.2 (SAS Institute Inc., Cary, NC USA).

RESULTS

The population used for the present validation consisted of 2612 individuals, 57% (n=1487) of whom were men. Thirty percent (n=788) were executives, 41% (n=1059) were
middle managers/foremen, 17% (n=433) were non-manual unskilled (e.g. clerical) workers, and 12% (n=322) were manual unskilled workers. The PPS distribution had a stronger right skew among women, and scores for women were lower overall (Figure 1a). The measure also showed differential distribution by occupational grade, with low-wage manual workers reporting more strain on average than higher-level workers (Figure 1b). One-way ANOVA analyses showed that PPS was highly significantly different (p<0.0001) between occupational groups, men and women, and older versus younger workers.

[Figure 1]

The test-retest reliability between 1994 and 1995 was r=0.61 after partialing out the effect of change in occupational grade in the intervening year (n=487, 18.5% of participants). After stratifying by several factors possibly related both to physically strenuous work and to self-reported strain, the reliability was as follows: for sex, r=0.63 among men and r=0.58 among women; and for occupational grade in 1994, r=0.59 among executives, r=0.60 among midlevel managers, r=0.54 among non-manual unskilled workers, and r=0.65 among manual unskilled workers.

To assess construct validity, PPS was compared against in-depth assessments of eight domains of physical work strains, each of which was composed of two to 14 exposures. Across the study population, for all types of strains there was a highly statistically significant (p<0.0001) monotonic relationship between number of exposures in a given domain and mean PPS score, with each additional strain in a given domain associated with a higher mean PPS. After stratification by gender, the positive linear relationship remained highly statistically significant for men across all domains, but p-values for women were slightly higher; extremely low exposure to high strain in several domains (n=0-10 for several strata) may have contributed to
this pattern. Formal tests of statistical interaction between gender and each domain on PPS were not significant, suggesting that the relationship between constraints and PPS were not different for men and women. We found that, across the population, the greatest average PPS was associated with high exposure (defined as the group with the highest number of exposures in a given domain) to pulling objects (mean PPS 5.15, SD 1.27) and carrying heavy loads (mean 4.76, SD 1.46), and the lowest average PPS was associated with high exposure to walking (mean 4.32, SD 1.65). Among men, the domains associated with highest- and lowest-reported PPS mirrored those of the overall population, while among women the highest-average PPS was associated with carrying heavy loads (mean 5.33, SD 2.08, n participants in strata=3) and use of tools (mean 5.20, SD 0.45, n participants=5); the lowest-average PPS was associated with elbow strain (mean 3.36, SD 1.42, n participants=36).

[Table 1]

In order to characterize the relationship between total physical workload and PPS, we tabulated the number of categories of constraints to which individuals reported any exposure during the past year and calculated mean PPS at each level of total exposure. We found a linear relationship ($p<0.0001$) between total physical load exposure and PPS for the whole sample and stratified by gender, with PPS 54% higher among those with eight versus zero exposures among all participants and 52% higher among men comparing most-exposed to unexposed. As only one woman was in the highest-exposure stratum, we pooled women with four or more exposures (n=25, mean PPS=4.24); comparing this group to the unexposed, the difference in PPS was 39%. In tests for gender differences in the load-PPS relationship, we found that a) the main effect of gender was not significant, suggesting that men and women perceived the same level of physical strain at a given level of load, and b) tests of statistical interaction between total physical load
and gender as predictors of PPS were not statistically significant, suggesting that the effect on PPS of an additional category of exposure was constant by gender.

Because occupational grade and gender are so highly correlated among low-wage workers in GAZEL (97% of unskilled manual workers are men and 80% of unskilled non-manual workers are women), the stratified results for occupational grade largely mirror those of gender and so are not shown here.

Discussion

The present study tested whether a single-item self-report of perceived physical strain (PPS) at work was a reliable and valid proxy measure for physical load at work. We found that PPS had fair test-retest reliability over a one-year period, with reliability varying between gender and occupational groups. Within each of eight domains of physical constraints, there was a linear relationship between number of constraints and PPS score, robust to stratification by gender and occupational grade. There was also a monotonic relationship between number of domains of physical constraints and PPS, also robust to stratification by gender and occupational grade.

The study has several limitations. First, some of the psychometric properties, notably the test-retest reliability, fall below conventional levels of acceptability (r=0.70) (23). The long duration between measurement occasions in the present study (one year versus the recommended two weeks) may have contributed to the discrepancy; actual work tasks could have changed in the intervening year, reducing the underlying correlation between physical strain in 1994 and 1995 and creating a downward bias in the PPS correlation. We partly addressed this concern by partialing out change in occupational grade. However, this likely only captured gross differences in work environment, rather than small variations affecting change in physical tasks. Given that
any bias would likely be downward and that measure was most reliable in the most highly-
exposed subgroups \((r=0.63\) for men and 0.65 for manual workers), we felt confident enough in
the measure’s reliability to test for validity, especially considering the aim of the study, which
was to develop a proxy measure in a diverse population of workers. However, future studies that
are able to assess participants more frequently would be better equipped to analyze the PPS
measure’s true test-retest reliability.

Second, both PPS and physical load at work are self-reported. Although there is a risk of
bias when individuals self-report both exposure and outcome because of potential systematic
over- or under-reporting of both factors, it has been shown that such “rating-behavior bias” is
negligible for studies of non-affective stimuli such as workplace exposures, which was the nature
of the present analysis, particularly the detailed exposure assessment instrument to which we
compared PPS (24). Finally, single-item measures are controversial because they inherently lack
several qualities usually desirable in a measure: internal consistency reliability, ability to capture
a multi-dimensional construct, and ability to determine the relative contribution of component
items to even a uni-dimensional construct (21). However, because single-item measures are easy
and inexpensive to collect and code, they are an attractive alternative to longer scales if their
psychometric properties can be shown to be adequate (25, 26), especially in contexts for which
nuanced exposure assessment is not a primary concern.

Despite these limitations, the analysis has several strengths. First, it was conducted within
an established, socioeconomically diverse cohort with high variation in physical tasks at work.
This allows both for evaluation of differential validity by sociodemographic factors and also for
detection of a linear relationship between level of physical load and level of PPS. Furthermore,
PPS was validated against a comprehensive assessment of biomechanical strains across a wide range of domains.

We found that across each of the eight domains, a higher number of strains were associated with higher PPS, and that mean PPS in each domain at the highest number of strains differed across domains. The domains of work-related strain considered to be the most and least physically strenuous (27) corresponded to the level of PPS reported among the most-exposed in those domains in our analysis (Table 1): for highly strenuous tasks, PPS was 4.76 for carrying heavy loads and 5.15 for pulling objects. Conversely, those tasks considered to be light physical work showed the lowest PPS even at the highest exposure levels: 4.32 for walking and 4.52 related to working on knees. This suggests that PPS can differentiate between more- and less-demanding physical workloads within individuals exposed to some type of biomechanical constraints. Additionally, we found that each additional category of constraints (Table 2) was generally associated with a higher level of PPS. This indicates that PPS is sensitive both to increased levels of exposure within a given domain, and also to overall increases in total physical load. The latter is important because types of tasks to which workers are exposed vary across industries and are not necessarily the same as those evaluated in our validation instrument. While replication of the present study in a different working population would further confirm the monotonic relationship, this study constitutes preliminary evidence that perceived physical strain is linearly associated with the volume of physical work tasks an individual must undertake.

Finally, we found that, although formal tests of interaction between gender and ARPEGE as predictors of PPS were not statistically significant, the relationship between ARPEGE and PPS was strongest in the most highly exposed groups (men and unskilled workers; 97% of the latter were men). Our results found that both within and across domains, men reported higher
levels of PPS given exposure to the same reported level of physical load, a gender pattern opposite that found in prior literature (22, 28). There are several possible explanations for this finding. First, when men and women report exposure to the same task, men might be exposed at a higher level than women (for example, men might lift heavier objects than women); these differences in intensity are not captured by ARPEGE. We also found that men were generally exposed to more domains of physical strain than women in GAZEL (mean n domains 2.09 [SD 2.41] for men, mean 0.62 [SD 1.02] for women), and it is possible that men were also exposed to domains we did not analyze, which would produce the higher reported PPS if PPS is actually an indicator of physical load. Men and women may also have interpreted or answered the PPS question differently with respect to actual physical load. Future research could explore this discrepancy. Additionally, further research is needed in the field of “ergonomic epidemiology” (29) to assess the contexts best suited to self-reported constraints and thus use of this measure (9, 30-32).

A single-item measure could certainly never replace the complex assessment methods used to determine health effects of occupational exposures. However, our analysis has implications for studies in which, due to constraints of time, logistics, and participant attention, a more complete assessment of physical occupational exposures is unfeasible. Although imperfect, and in need of further testing for certain properties such as measure reliability, preliminary results suggest that self-report of PPS is a fairly good indicator of overall physical exertion or load at work within a given population. Inclusion of this question in other epidemiologic studies would provide greater insight into the type and distribution of physical workplace demands within the population, particularly among low-wage manual workers.
References:


27. de Zwart BCH, Broersen JPJ, Frings-Dresen MHW, van Dijk FJH. Musculoskeletal complaints in the Netherlands in relation to age, gender and physically demanding work. *International Archives of Occupational and Environmental Health*. 1997;70:352-360.


Figure 1a: Distribution of PPS by gender

Figure 1b: Distribution of PPS by occupational grade