

# Nurses Work Functioning Questionnaire (NWFQ): refinement and validation of the Italian version

NICOLA MAGNAVITA<sup>1</sup>, CARLO CHIORRI<sup>2</sup>

<sup>1</sup>Istituto di Sanità Pubblica, Università Cattolica del Sacro Cuore, Roma, Italia

<sup>2</sup>Dipartimento di Scienze della Formazione, Università di Genova, Italia

**KEY WORDS:** Work impairment; nurses; job strain; organizational justice

**PAROLE CHIAVE:** Compromissione della capacità lavorativa; infermieri; stress lavoro-correlato; giustizia organizzativa

## SUMMARY

**Background:** *The availability of an assessment measure for work impairment in nurses, or nursing students, is of crucial importance for early detection of workers/students at risk. Recently, a new measure, the Nurses Work Functioning Questionnaire (NWFQ), has become available, but there is no validated Italian version. Objectives:* *The aim of this study was to develop and validate an Italian version of the NWFQ. Methods:* *We evaluated the factor structure, the internal consistency, and the convergent and discriminant construct validity with respect to organizational justice and job strain of the Italian NWFQ using data from 645 nursing students. Results:* *Results suggested that a single-factor, 34-item measurement model could be a more parsimonious alternative (CFI=.915, TLI=.910, RMSEA=.039 e CFI=.907, TLI=.901, RMSEA=.046 in in two random subsamples; median factor loading .50, range .26-.63) to the original seven-factor structure. The score on this version of the NWFQ showed excellent internal consistency and construct validity, as higher scores were significantly associated with lower perceived distributive ( $r=-.30$ ) and interpersonal justice ( $r=-.43$ ), decision latitude ( $r=-.33$ ), and social support ( $r=-.58$ ). Conclusions:* *The Italian refinement of the NWFQ seems to have adequate psychometric properties and it is thus suitable for the assessment of impairment of work functioning in nursing students.*

## RIASSUNTO

«*Nurses Work Functioning Questionnaire (NWFQ): revisione e validazione della versione italiana*». **Introduzione:** *La disponibilità di uno strumento di valutazione dei problemi che riguardano la capacità di lavoro e di giudizio del personale infermieristico, o di quello in formazione, è di cruciale importanza per l'individuazione di lavoratori o studenti a rischio. Recentemente è stata sviluppata una nuova scala, il Nurses Work Functioning Questionnaire (NWFQ), ma attualmente non esiste una sua versione italiana. Obiettivi:* *Lo scopo di questo studio è stato di sviluppare e validare una versione italiana di NWFQ. Metodi:* *La struttura fattoriale, la coerenza interna e la validità di costruito convergente e discriminante rispetto a misure di giustizia organizzativa e stress lavoro-correlato sono state valutate utilizzando dati raccolti su 645 studenti di infermieristica. Risultati:* *I risultati hanno mostrato che un modello di misurazione con un singolo fattore di una versione ridotta a 34 item (CFI=.915, TLI=.910, RMSEA=.039 e CFI=.907, TLI=.901, RMSEA=.046 in due sottocampioni casuali) costituisce un'alternativa più parsimoniosa al modello a sette fattori della versione originale. Il punteggio in questa versione di NWFQ ha mostrato eccellente coerenza interna e validità di costruito, dato che punteggi maggiori erano associati con livelli inferiori*

Pervenuto il 13.5.2017 - Revisione pervenuta il 10.8.2017 - Accettato il 11.10.2017

Corrispondenza: Nicola Magnavita, Istituto di Sanità Pubblica, Università Cattolica del Sacro Cuore, Largo A. Gemelli 8, 00168 Roma, Italia - Tel. +39 3473300367 - Fax +39 0661909399 - E-mail: nicolamagnavita@gmail.com

*di giustizia distributiva* ( $r=-.30$ ) e *interpersonale* ( $r=-.43$ ), *discrezionalità* ( $r=-.33$ ), e *supporto sociale sul lavoro* ( $r=-.58$ ). **Conclusioni:** *L'adattamento italiano di NWFQ sembra possedere adeguate proprietà psicometriche ed è quindi adatto per la valutazione della compromissione della capacità lavorativa negli allievi infermieri.*

## INTRODUCTION

Work functioning has always been a crucial issue in occupational health, as it affects the quantity and quality of production. The individual work functioning capacity can vary over a lifetime: it increases with professional growth, but it can be impaired by many factors, including occupational and non-occupational stress, alcohol and drugs, neurological and psychiatric diseases, and other (26, 29, 32, 34). In some critical sectors, such as health care (HC), impairments in work functioning can have serious consequences not only for the employee and the organization, but also for patients' health (35). Examples of these deleterious effects include medication errors, needle stick injuries, near errors, and decreased patient satisfaction (14). Nursing students can be exposed to the same occupational factors of their professional colleagues (31), and can be impaired during their educational training.

Early detection of impaired workers and their rehabilitation should be a topic for Occupational Health Services (OHS) in HC. Optimal work functioning is also the ideal outcome of nursing education. A valid and reliable tool for assessing work functioning may be useful to promote health in nurses, or to verify the progress of students in nursing schools.

There are a number of measures for assessing impairments in work functioning, such as the Work Ability Index (WAI, (46)), the Work Limitation Questionnaire (WLQ, (23)), the Stanford Presenteeism Scale (SPS, (22)) and the Endicott Work Productivity Scale (EWPS, (7)), but they are designed to be generic. While this allows them to be used in a number of different occupational settings, they may fail to tap into issues specific to nurses' work. In order to address this need and provide a tool for the assessment of nurses' work functioning, Gärtner and coworkers (12, 14) developed the Nurses Work Functioning Questionnaire (NWFQ). Because of its job-specific nature, it successfully connects to the

actual work practice and concrete experiences of the employees. This approach enables specific aspects of work functioning that are impaired to be detected and thus provides valuable information for the development of effective interventions (13).

Given the need to measure HC workers' work functioning in the Italian context, the first aims of this study were to develop an Italian version of NWFQ and to evaluate its psychometric properties. The second aim was to refine the NWFQ, adapting this instrument to the measurement of working capacity in nursing students.

We thus tested the factor structure of the Italian version of the NWFQ on a large sample of nursing students and, using a factor analytic, cross-validation approach, we developed a refined version of the questionnaire, specifically devoted to this population. We then tested (albeit not comprehensively) its construct validity using measures of job strain and organizational justice.

## METHODS

### Participants and procedure

A convenience sample of 645 nursing students (females were 395, 61.1%; mean age 25.08+6.91, range 20-52) of the 2nd and 3rd academic year attending the graduate program in an Italian university, were recruited during the last week of their 40-days practical hospital training and asked to complete the questionnaires described in the next section during an academic lecture. They received no compensation for their participation, which was on a voluntary basis; however, the participation rate was very high (97.4%).

### Measures NWFQ

The original self-report questionnaire consisted of 50 items, referring to seven subscales: (i) Cog-

nitive Aspects of Task Execution and General Incidents; (ii) Impaired Decision Making; (iii) Causing Incidents at Work; (iv) Avoidance Behavior; (v) Conflicts and Annoyances with Colleagues; (vi) Impaired Contact with Patients and Their Families; (vii) Lack of Energy and Motivation. Cronbach's alphas of the subscales in the original version ranged from 0.70 to 0.94. In a study on its psychometric properties, except for Impaired Decision Making, all scales showed adequate reliability and validity (14). Although impaired decision making may be seen as an important aspect of impaired work functioning, the inadequate psychometric properties of its measure discouraged the further use of this scale. For this reason, in the Italian translation we removed the 3 items (items 48, 49 and 50) of the impaired decision making subscale from the questionnaire. The other scales showed adequate test-retest reliability (*rtts* above .70), construct validity (significant positive correlations with the scales of the EWPS and negligible correlations with measures of physical functioning) and criterion validity (significantly higher scores in nurses with mental health complaints).

All items of the NWFQ have a reference period of four weeks. Items 1-32 are scored on a 7-point rating scale, from "no difficulty" to "great difficulty" (items 1-14), or from "totally disagree" to "fully agree" (item 15), or from "(almost) never" to "(almost) always" (items 18-23), or from "not once" to "on average more than 1 per day" (items 24-32). The remaining items are scored on a 5-point rating scale, ranging from "(almost) never" to "(almost) always" (items 33-39), or from "disagree" to "agree" (items 40-47).

The questionnaire has originally been used to screen for work functioning impairments in a randomized controlled trial on the effectiveness of a Workers Health Surveillance program for nurses and allied health professionals in hospitals (11). Gärtner and coworkers had the original Dutch version of the NWFQ professionally translated into English. This English version was translated into Italian by one of the authors (NM), back-translated into English by a native English reader, who was unaware of the questionnaire, and then compared to the original. The critical points were discussed with the authors of the questionnaire (FG) until a consensus was reached.

## Validity measures

Participants were also invited to complete the Italian version (36) of the Demand-Control-Support (DCS) questionnaire and of the Justice Measure (JM, (5, 6); Italian version in (27)).

The DCS is based upon Karasek's model of job strain (19) and assesses the demand, control, and support dimensions described in the Introduction. The job demands-control-support model (DCS; (19)) is currently recognized as an influential theory for understanding how work characteristics relate to employees' well-being, health, and performance. Demand refers to the job's psychological demands, such as work overload, difficulties in tasks, time required to perform tasks, and occurrence of contradictory or controverting orders; Job Control, or Decision Latitude, pertains to the use and development skills and autonomy in decision-making on the job. The Demand/Control weighted ratio is often used as a synthetic indicator of self-perceived job strain. The Social Support at Work taps into the quality of relationships among coworkers and with supervisors. A recent meta-analytic review (25) showed that all these constructs negatively predicted burnout. Hence, we expected that higher levels of work impairment, as measured by the Italian NWFQ, were associated with lower levels of Demand, Job Control and Social Support.

The DCS measure used in this study comprises 17 items to be rated on a 4-point rating scale. The Demand subscale includes 5 items (e.g., "Does your job require too great a work effort?"); the Job Control, or Decision Latitude, subscale includes 6 items (e.g., "Do you have the possibility to decide for yourself how to carry out your work?"); the Social Support at Work subscale includes 6 items (e.g., "There is good collegiality at work"). The Demand/Control weighted ratio is often used as a synthetic indicator of self-perceived job strain. Item scores are summed to yield a total score for each scale, and higher scores indicate higher levels of the relevant trait. The validation study of the Italian version of the DCS showed the questionnaire has adequate factorial validity, as the expected three-factor structure was replicated, and internal consistency (Cronbach's alphas were .76 for Demand, .67 for

Control, and .87 for Support) (25). Furthermore, it has been shown that it has predictive validity on depression (9), absenteeism (30), and metabolic syndrome (10).

The JM is based upon Colquitt's model of organizational justice (5), which includes four main components: distributive justice, which concerns whether individual outcomes are consistent with implicit norms for allocation, such as equity or equality; procedural justice, which concerns the fairness and the transparency of the processes by which decisions are made; interpersonal justice, which measures the treatment people receive as procedures are enacted; informational justice, which assesses whether explanations are perceived to be reasonable, timely, and specific. Meta-analytic findings pointed out the association of organizational justice with work-related outcomes and suggested that organizational justice buffers against work-related stress (see, e.g., (40)). Otto and Mamatoglu (40) reported a negative association of scores of cognitive irritation and emotional exhaustion with scores of interpersonal and informational justice, and showed that the effect of these facets of justice on well-being was mediated by bullying. We thus expected that higher scores on the Italian NWFQ were negatively associated with at least the facets of organizational justice reported by Otto and Mamatoglu (40).

The measure of organizational justice comprises four groups of items. Distributive subscale includes 4 items (e.g., "Does the outcome of your work reflect the effort you have put into your work?"), the Procedural subscale includes 7 items (e.g., "Have you been able to express your views and feelings during the procedures of your work?"), the Interpersonal subscale includes 4 items (e.g., "Has your supervisor treated you with respect?"), and the Informational subscale includes 5 items (e.g., "Has your supervisor communicated details in a timely manner?"). The 20 items are rated on a 5-point rating scale ranging from 1 = 'to a limited extent' to 5 = 'to a large extent'. Item scores are summed to yield a total score for each scale, and higher scores indicate higher levels of the relevant trait. This measure has demonstrated predictive validity for the justice dimensions on important outcomes, including leader evaluation, rule compliance, commitment, and helping behavior (5).

The validation study of the Italian version of the JM showed the questionnaire has adequate factorial validity, as the expected factor structure was replicated, and internal consistency (Cronbach's alphas were .88 for distributive, .79 for procedural, .92 for interpersonal, and .63 for informational) (27).

### Statistical analyses

Statistical analyses were performed with the Statistical Package for Social Sciences IBM/SPSS 20.0 and R. The factor structure of the Italian NWFQ was tested using a cross-validation approach. Since the measurement model of the NWFQ was known in advance thanks to its development study (12, 14), we could have tested its factor structure through confirmatory factor analysis (CFA). However, the psychometric properties of psychological measures are not automatically guaranteed when the measures are adapted into other languages (see e.g., (17)). Moreover, CFA requires each item to load on only one factor, but, as shown by recent studies (e.g., (1)), this assumption might be too restrictive, because items may have secondary loadings significantly different from zero (i.e., item loadings tend to "split" over multiple factors). Items with substantially non-zero secondary loadings have a weak discriminant validity, since an item that is considered an indicator of a specific construct can also be an indicator of another construct. In a CFA, the more the secondary loadings depart from zero, the more the correlations among the factors will be inflated to account for non-zero secondary loadings restricted to zero, thus yielding biased loadings, overestimated factor correlations, distorted structural relations, and lack of fit (1).

We therefore chose a cross-validation approach (8) when testing the factor structure of the Italian NWFQ. We initially carried out exploratory factor analyses (EFA) in a random subsample of participants ( $n=324$ ), and then performed CFAs in the other ( $n=321$ ). This procedure is recommended in order to increase the generalizability of the factor-structure across groups, and to validate both exploratory and confirmatory solution.

The optimal number of factors to be extracted was investigated through the scree-test and Parallel

Analysis (PA, Horn, 1965), which were performed with the *psych* (42) and *hornpa* packages (18) in R (41). With the scree-test, the eigenvalues are plotted against the component number with a simple descending line segment. The optimal number of factors to be extracted corresponds to the component number at which the segment begins to level off. PA is still based on the eigenvalues, but it compares the observed eigenvalues extracted from the correlation matrix to be analyzed with those obtained from the simulation of normal random samples (in this case, 1,000) having the same sample size and number of variables. We retained the factors in which the observed eigenvalues were larger than the 95th percentile of the distribution of the corresponding simulated eigenvalues.

After determining the optimal number of factors to be extracted, we performed EFAs through principal axis factoring with the *fa* function in the R package *psych*. After extraction, factors were Promax rotated.

Once an adequate measurement model for the NWFQ items was found through EFAs, we used CFA to test this model in the other random subsample of participants. We used a robust maximum likelihood estimator (MLR), with standard errors and tests of fit that were robust in relation to the non-normality of observations (median skewness of the item distributions was 0.87, range 0.23-1.25, and median kurtosis was -0.55, range -1.03-0.31). The goodness of fit of the CFA models was evaluated considering the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA), as operationalized in the *lavaan* package (43) in R, in association with the MLR estimator. Following Marsh and coworkers (37), we considered values  $\geq .90$  as acceptable and values of  $\geq .95$  as optimal for TLI and CFI, while for RMSEA, values of  $\leq .08$  were considered as acceptable and values of  $\leq .06$  as optimal.

After testing the fit of the model, we examined modification indices to detect sources of lack of fit. Many authors have cautioned against this practice (e.g., (16)), since the correlations between residual variances may indicate that other factors are needed to account for the intercorrelations of items. However, these factors may be of limited theoretical in-

terest, since they might reflect common (negative) wording and/or content redundancy (the so-called method factors). In these cases, one or more items can be removed from the scale without loss in content validity.

Once a measurement model for the Italian NWFQ items was established, we tested its convergent and discriminant construct validity using a multi-trait mono-method approach. This is a reduced version of the Multitrait-Multimethod Matrix developed by Campbell and Fiske (4). The multi-trait mono-method approach uses measures of different constructs (in this case, work impairment, dimensions of job strain, and organizational justice) assessed by the same method (in this case, self-report questionnaires). Correlations are computed to test the association of scores on the measures. Convergent validity is supported if scores on the measure under investigation have significant correlations with scores on other tests designed to measure theoretically similar concepts. Discriminant validity is supported if scores on the measure under investigation do not have significant correlations with scores on other tests designed to measure theoretically different concepts. Consistent with the hypotheses stated above, we expected that the NWFQ scores would be significantly associated with measures of job strain and with interpersonal and informational justice.

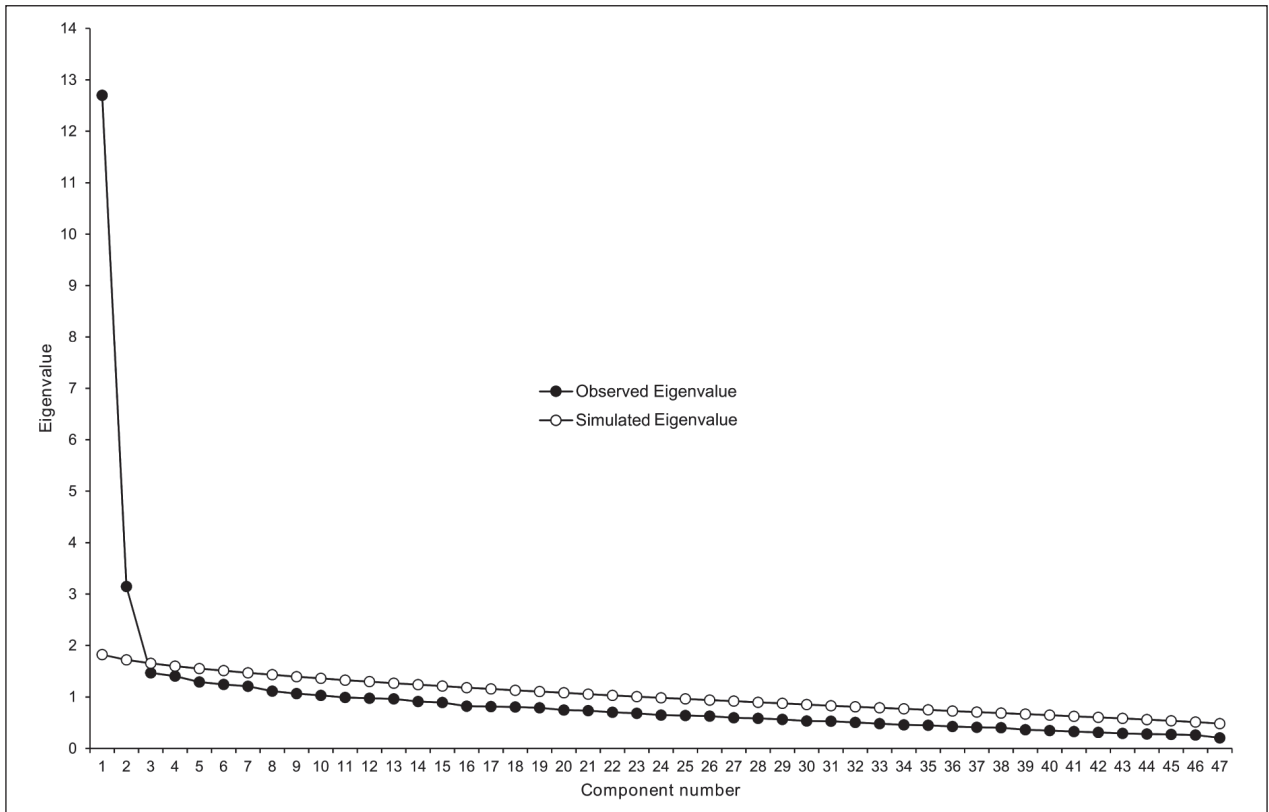
#### *Ethical approval*

The study was approved by the Ethics Committee of the Università Cattolica del Sacro Cuore (Rome, Italy).

## RESULTS

As shown in figure 1, dimensionality analyses on the first random subsample of participants suggested extracting two factors.

The scree-test revealed that the optimal number of factors was two, since the line with of the observed eigenvalues began to flatten out from the third component. Parallel analysis also indicated to extract two factors because only the first two observed eigenvalues were larger than the corresponding simulated



**Figure 1** - Results of the scree-test and of parallel analysis performed in a random subsample of participants ( $n=324$ ). The scree test suggests that the optimal number of factors is two, since the line with of the observed eigenvalues (black dots) begins to flatten out from the third component. Parallel analysis also suggests to extract two factors because only the first two observed eigenvalues are larger than the corresponding simulated ones (white dots)

ones. We thus performed an EFA, setting to two the number of factors to extract. In this two-factor solution, which explained 31% of variance, the 14 items of the cognitive aspects of task execution and general incidents subscale substantively loaded (i.e., loading greater than .30) on one factor and all the remaining items, except item 21 and 40, loaded on the other factor. These two factors had a very high correlation ( $r=.69$ ) suggesting a poor discriminant validity. We then tested a more parsimonious single-factor model that accounted for 25% of variance, and all the items substantively loaded on the single factor. Solutions with 3 and 4 factors did not substantially increase the explained variance (33% and 34%, respectively) and always provided highly intercorrelated factors ( $r$ s were never lower than .50). Hence, we concluded that a one-factor measurement model could be the most adequate.

We then used CFA to test the single-factor model in the other random subsample of participants. The single-factor model did not show an adequate fit ( $\chi^2(1034)=1615.29$ ,  $p<.001$ , Scaling Correction Factor (SCF)=1.097, CFI=.818, TLI=.809, RMSEA=.042). Modification indices suggested that the residual variances of many items should be allowed to correlate, in order to improve model fit. On the basis of modification indices, we identified and removed, one at a time, 13 items that contributed most to the misfit (items 1, 2, 3, 5-10, 13, 14, 29, 31). The fit of the single-factor model with the 34-item pool was acceptable both in the subsample on which we performed the CFA ( $\chi^2(527)=683.09$ ,  $p<.001$ , SCF=1.152, CFI=.915, TLI=.910, RMSEA=.039) and in the subsample on which we performed the EFA ( $\chi^2(527)=745.87$ ,  $p<.001$ , SCF=1.173, CFI=.907, TLI=.901, RMSEA=.046). Factor load-

**Table 1** - Standardized factor loadings for the two random subsamples used in the cross-validation procedure and for the total sample, and corrected item-total correlations ( $r_{it}$ )

NWFQ Item	Subsample 1	Subsample 2	Total sample	$r_{it}$
NWFQ04	.32	.37	.35	.33
NWFQ11	.60	.55	.57	.54
NWFQ12	.42	.30	.36	.35
NWFQ15	.40	.49	.44	.42
NWFQ16	.50	.51	.51	.49
NWFQ17	.42	.44	.43	.41
NWFQ18	.56	.59	.57	.54
NWFQ19	.54	.58	.56	.53
NWFQ20	.45	.51	.48	.46
NWFQ21	.29	.34	.31	.30
NWFQ22	.53	.54	.53	.51
NWFQ23	.61	.63	.62	.59
NWFQ24	.43	.50	.47	.44
NWFQ25	.58	.60	.59	.56
NWFQ26	.57	.59	.58	.54
NWFQ27	.56	.48	.53	.50
NWFQ28	.55	.60	.57	.54
NWFQ30	.57	.50	.54	.52
NWFQ32	.53	.51	.52	.50
NWFQ33	.48	.44	.46	.44
NWFQ34	.40	.39	.39	.38
NWFQ35	.43	.48	.45	.43
NWFQ36	.44	.42	.43	.41
NWFQ37	.52	.50	.51	.49
NWFQ38	.47	.54	.51	.48
NWFQ39	.52	.56	.54	.52
NWFQ40	.36	.26	.32	.30
NWFQ41	.44	.45	.45	.42
NWFQ42	.54	.46	.50	.48
NWFQ43	.52	.49	.50	.48
NWFQ44	.37	.39	.38	.37
NWFQ45	.56	.45	.51	.48
NWFQ46	.37	.28	.34	.32
NWFQ47	.58	.57	.57	.54

ings for both random subsamples are reported in table 1.

The average score of the 34-item NWFQ questionnaire, obtained by adding all the answers, was  $101.86 \pm 29.38$  ( $38.99 \pm 16.89$  when converted to Gärtner and coworkers' (13) suggested 0-100 range). Cronbach's alpha in the total sample was .91 and corrected item-total correlations ranged from .30 to .59 (median=.48, table 1).

Table 2 reports the correlations between the 34-item NWFQ and the measures of job strain and organizational justice, along with descriptive statistics.

Consistent with hypotheses, the NWFQ score showed significant correlations, and in the expected direction, with all the other measures except procedural justice. The effect size of the significant correlations was small ( $.10 < |r| < .30$ ) to strong ( $|r| > .50$ ) with the DCS scales and small to moderate ( $.30 < |r| < .50$ ) with the JM scales.

## DISCUSSION

In this study we have presented the Italian version of the NWFQ, a questionnaire specifically designed to measure the working capacity of nurses, and have developed a shorter version of the NWFQ, which measures the working capacity of nursing students.

The work functioning of nurses is highly critical for health services, given its implications for nurses' and patients' well-being. The availability of a valid and reliable measure of work functioning in this specific population of workers is therefore of paramount importance. The Nurses Work Functioning Questionnaire (NWFQ) addresses this need, as it explores nurses' individual experiences of their own behavior while at work. Unlike other existing instruments for measuring health-related work functioning, the items of the NWFQ do not explicitly refer to (known) health problems, but focus on nurses' common concrete experiences and tasks, that can thus be easily recognized and rated. This provides insight into various aspects of their work that could be improved with appropriate health action.

Studies carried out in the Netherlands have enabled researchers to evaluate the psychometric properties of the NWFQ (11-14). In the original version, the NWFQ proved to have high content validity, largely comprehensive and relevant subscales, acceptable to good internal consistency and temporal stability, as well as construct and criterion validity. Furthermore, its structural validity was good, since the subscale distribution was validated through a confirmatory factor analysis. It also proved useful for measurements of working ability during health promotion programs in the Netherlands (11-13, 20, 21, 39). Unfortunately, so far it has not been applied

**Table 2** - Descriptive statistics and Pearson correlations of the scores on the Nurses Work Functioning Questionnaire (NWFQ) with scores on the Demand-Control-Support (DCS) questionnaire and the Justice Measure (JM)

Scale	NWFQ	DCS - Demand	DCS - Control	DCS - Support	JM - Procedural	JM - Distributive	JM - Interpersonal	JM - Informational
DCS - Demand	-.21**							
DCS - Control	-.33**	.07						
DCS - Support	-.58**	.09*	.27**					
JM - Procedural	.03	-.15**	.04	.04	-.12**			
JM - Distributive	-.30**	.00	.18**	.23**	-.11**	.15**		
JM - Interpersonal	-.43**	.04	.14**	.33**	-.05	.04	.33**	
JM - Informational	-.22**	-.04	.19**	.22**	-.16**	.19**	.22**	.25**
M	101.86	12.74	15.61	16.23	20.88	12.49	13.03	15.36
SD	29.38	2.58	2.80	3.39	3.82	3.27	3.20	3.52

Note M: Mean; SD: Standard deviation; \*\*:  $p < .01$ ; \*:  $p < .05$ .  $n = 645$

in other countries, thus depriving international studies of a useful tool.

In this study, we developed an Italian version of the NWFQ and tested its psychometric properties in a population of nursing students. Using exploratory and confirmatory factor analyses we found support for a single-factor measurement model and, after removing items identified as statistically redundant (i.e., provided a similar information while not substantially improving the content coverage of the construct), we produced a shortened version (34 instead of 47 items), specifically useful for nursing students. Eight of the 13 removed items were included by Gärtner et al. (12) in the Cognitive aspects of task execution and general incidents subscale (items 1, 2, 3, 5, 6, 7, 8, 9); 2 items (items 10, 13) were included in the Impaired contact with patients and their family subscale; the last 3 items (items 14, 29, 31) were included in the Causing incidents at work subscale. These items were found to be redundant (as indexed by modification indices in CFAs) with the other items operationalizing the same facets of impaired work functioning. This result might be due to the idiosyncratic characteristics of the nurse job in Italy, to having used data from nursing students, or, perhaps more possibly, to a combination of these factors.

This version showed high internal consistency (Cronbach's alpha greater than .90) and construct validity, as the total score was significantly correlated with measures of psychological job demands, deci-

sion latitude, workplace social support, and all the facets of organizational justice, except procedural justice. The Job Demand Control model (19) presents combinations of working conditions that may facilitate learning. For instance, high job demands and high job control on performance are assumed to have a positive effect (the so-called active learning hypothesis). However, if job demands exceed job control, then negative effects occur, according to the so-called strain hypothesis (2, 24, 38, 47). The NWFQ assesses nurses' individual experiences of their own behavior while at work and it is not a measure of organizational aspects. The underlying construct of the NWFQ, when applied to nursing students, assumes the presence of impaired work functioning and a learning process due to passive behavior. Consistent with this, the NWFQ scores were inversely related to the Demands and Control subscales, with social support from colleagues (see also, e.g., (44)), and with dimensions of organizational justice (see also, e.g., (3)).

It should be noted that the original NWFQ does not include an overall score of work functioning, since it provided scores for seven subscales. Following the original authors' suggestion, in this study we dropped the three items of the impaired decision making subscale, but we did not find support for a multidimensional measurement model. We also tested a bi-factor model, in which a general work functioning factor and specific factors corresponding to the original subscales were specified, but, as it



is often the case with such models, no convergence was reached. Our results thus suggested the use of a single total score, which provides a synthetic index that may be useful for monitoring population groups. For example, it may very well serve to assess the learning ability of nursing students during their internship.

The job-specificity of the NWFQ items, and the fact that these items do not refer to previous health problems, favor the use of this instrument in the assessment of work functioning of health care workers. Prevention based upon NWFQ results may be designed with many different strategies. Individual workers with functioning problems can be supported, educated or guided. General work organization, or specific tasks, can be changed, and the working environment can be improved, with beneficial effects for both workers and patients.

The main limitation of this study is that it was carried out on nursing students, a population that differs from the one originally used to develop the questionnaire, i.e. nurses with common mental problems. The refined (shorter) version excluded 13 out of the 47 items, and specifically those related to incidents and to the contact with patients and their family, aspects which may be less prominent to training personnel as compared to more experienced nurses. Future studies should therefore investigate the reproducibility of these results on nurses with and without mental problems, together with the test-retest reliability of scores. Knowing that scores are stable in the absence of apparent change is crucial in order to consider the Italian NWFQ suitable as an outcome measure for interventions. The shortage, early retirement, and aging of nurses, and the consequent need for workplace health programs to increase nurses' occupational health and work engagement, call for studies on nurses' work functioning to be conducted in all countries. It is therefore desirable that further validation studies of the NWFQ be carried out.

## CONCLUSIONS

In conclusion, the Italian shorter version of the Nurses Work Functioning Questionnaire (NWFQ) showed a unidimensional factor structure, excel-

lent internal consistency, and construct validity, thus providing Italian practitioners with a valid and reliable tool for the assessment of impairment of work functioning in nursing student and nurses. More specifically, the refined, shorter version proved to be effective in training nurses; the longer, original version, may be useful in experienced nurses. Both questionnaires are freely available upon request.

The NWFQ can be adopted by occupational health physicians as a complementary tool in medical surveillance of nurses. It can be useful in disability management, as well as in longitudinal monitoring of individual capacity. Moreover, it is one of the more promising outcome measures in health promotion intervention studies, as it can be used both as a pathogenic health outcome, i.e., a measure of impairment in relation to injuries or work-related diseases, or as a positive health outcome associated with self-esteem, job satisfaction and job engagement (41,42). In nursing students, it can help to follow the evolution of educational skills (43). Health care accountability can be better guaranteed through the systematic use of NWFQ in hospital workers.

NO POTENTIAL CONFLICT OF INTEREST RELEVANT TO THIS ARTICLE WAS REPORTED BY THE AUTHORS

## REFERENCES

1. Asparouhov T, Muthén B: Exploratory structural equation modeling. *Struct Equ Model* 2009; 16: 397-438
2. Bethge M, Radoschewski FM, Müller-Fahrnow W: Work stress and work ability: cross-sectional findings from the German sociomedical panel of employees. *Disabil Rehabil* 2009; 31: 1692-1699
3. Brunault P, Fouquereau E, Colombat P, et al: Do transactional memory and participative teamwork improve nurses' quality of work life? *West J Nurs Res* 2014; 36: 329-345
4. Campbell DT, Fiske DW: Convergent and discriminant validation by the multitrait-multimethod matrix. *Psych Bull* 1959; 56: 81-105
5. Colquitt JA: On the dimensionality of organizational justice: A construct validation of a measure. *J Appl Psychol* 2001; 86: 386-400
6. Elovainio M, Heponiemi T, Sinervo T, Magnavita N: Organizational justice and health: Review of evidence. *G Ital Med Lav Ergon* 2010; 32: B5-B9
7. Endicott J, Nee J: Endicott Work Productivity Scale (EWPS): A new measure to assess treatment effects. *Psychopharmacol Bull* 1997; 33: 13-16

8. Floyd FJ, Widaman KF: Factor analysis in the development and refinement of clinical assessment instruments. *Psychol Assess* 1995; 7: 286-299
9. Garbarino S, Cuomo G, Chiorri C, Magnavita N: Association of work-related stress with mental health problems in a special police force unit. *BMJ Open* 2013; 3: 1-12
10. Garbarino S, Magnavita N: Work stress and metabolic syndrome in police officers. A prospective study. *PLoS ONE* 2015; 10: e0144318
11. Gärtner FR, Nieuwenhuijsen K, Ketelaar SM, et al: The Mental Vitality @ Work Study: Effectiveness of a mental module for workers' health surveillance for nurses and allied health care professionals on their help-seeking behavior. *J Occup Environ Med* 2013; 55: 1219-1229
12. Gärtner FR, Nieuwenhuijsen K, Van Dijk FJH, Sluiter JK: Impaired work functioning due to common mental disorders in nurses and allied health professionals: The Nurses Work Functioning Questionnaire. *Int Arch Occup Environ Health* 2012; 85: 125-138
13. Gärtner FR, Nieuwenhuijsen K, Van Dijk FJH, Sluiter JK: Interpretability of change in the Nurses Work Functioning Questionnaire: Minimal important change and smallest detectable change. *J Clin Epidemiol* 2012; 65: 1337-1347
14. Gärtner FR, Nieuwenhuijsen K, van Dijk FJH, Sluiter JK: Psychometric properties of the nurses work functioning questionnaire (NWFQ). *PLoS One* 2011; 6: e26565
15. Gärtner FR, Nieuwenhuijsen K, van Dijk FJH, Sluiter JK: The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010; 47: 1047-1061
16. Gerbing DW, Anderson JC: On the meaning of within-factor correlated measurement errors. *J Consum Res* 1984; 11: 572-580
17. Hambleton RK, Merenda PF, Spielberger CD: Adapting educational and psychological tests for cross-cultural assessment. Mahwah, NJ: Lawrence Erlbaum Associates, 2005
18. Huang F: Hornpa: Horn's (1965) Test to Determine the Number of Components/Factors R package version 10 2015
19. Karasek RRA: Job demands, job decision latitude, and mental strain: Implications for job redesign. *Adm Sci Q* 1979; 24: 285-308
20. Ketelaar SM, Gärtner FR, Bolier L, et al: Mental Vitality @ Work - A workers' health surveillance mental module for nurses and allied health care professionals. *J Occup Environ Med* 2013; 55: 563-571
21. Ketelaar SM, Nieuwenhuijsen K, Gärtner FR, et al: The effectiveness of a mental module for workers' health surveillance for nurses and allied health professionals, comparing two approaches in a cluster-randomised controlled trial. *Int Arch Occup Environ Health* 2014; 87: 527-538
22. Koopman C, Pelletier KR, Murray JF, et al: Stanford Presenteeism Scale: Health status and employee productivity. *J Occup Environ Med* 2002; 44: 14-20
23. Lerner D, Amick BC, Rogers WH, et al: The Work Limitations Questionnaire. *Med Care* 2001; 39: 72-85
24. Liu K, You LM, Chen SX, et al: The relationship between hospital work environment and nurse outcomes in Guangdong, China: A nurse questionnaire survey *J Clin Nurs* 2012; 21: 1476-1485
25. Luchman JN, González-Morales MG: Demands, control, and support: A meta-analytic review of work characteristics interrelationships. *J Occup Health Psychol* 2013; 18: 37-52
26. Magnavita N, Bergamaschi A, Chiarotti M, et al: Lavoratori con problemi di alcol e dipendenze. Documento di consenso del gruppo La.R.A. (Lavoratori Rischiosi per gli Altri) [Workers with alcohol and drug addiction problems. Consensus Document of the Study Group on Hazardous Workers] *Med Lav* 2008; 99: 3-58
27. Magnavita N, Bergamaschi A: Lo studio della giustizia sul lavoro. Validazione della versione italiana del JM20 di Colquitt (Justice at the workplace Validation of the Italian version of Colquitt's justice measurement questionnaire). *G Ital Med Lav Ergon* 2008; 30: 449-445
28. Magnavita N, Chiorri C: Academic stress and active learning of nursing students. A cross-sectional study (in press)
29. Magnavita N, Cicerone M, Cirese V, et al: Aspetti critici della gestione dei "lavoratori rischiosi" nei servizi sanitari. Documento di Consenso. [Critical issues of the management of "hazardous" health care workers. Consensus document]. *Med Lav* 2006; 97: 715-725
30. Magnavita N, Garbarino S: Is absence related to work stress? A repeated cross-sectional study on a special police force. *Am J Ind Med* 2013; 56: 765-775
31. Magnavita N, Heponiemi T: Workplace violence against nursing students and nurses. An Italian experience. *J Nurs Scholarsh.* 2011; 43: 203-210
32. Magnavita N: Management of impaired physicians in Europe. *Med Lav* 2006; 97: 762-773
33. Magnavita N: Productive aging, work engagement and participation of older workers. A triadic approach to health and safety in the workplace. *EBPH* 2017; 14: e12436
34. Magnavita N: Sorveglianza sanitaria nei lavoratori della sanità con patologia neurologica, psichiatrica o comportamenti addittivi [Health monitoring in public health workers with neurological and psychiatric disorders and addictive behaviors]. *Med Lav.* 2005; 96: 496-506

35. Magnavita N: The unhealthy physician. *J Med Ethics* 2007; 33: 210-214
36. Magnavita N: Two tools for health surveillance of job stress: The Karasek Job Content Questionnaire and the Siegrist Effort Reward Imbalance Questionnaire. *G Ital Med Lav Ergon* 2007; 29: 667-670
37. Marsh HW, Hau K-T, Wen Z: In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Struct Equ Model* 2004; 11: 320-341
38. Martinez MC, Dias de Oliveira Latorre M do R, Fischer FM, Latorre M: A cohort study of psychosocial work stressors on work ability among Brazilian hospital workers. *Am J Ind Med* 2015; 58: 795-806
39. Noben C, Smit F, Nieuwenhuijsen K, et al: Comparative cost-effectiveness of two interventions to promote work functioning by targeting mental health complaints among nurses: Pragmatic cluster randomised trial. *Int J Nurs Stud* 2014; 51: 1321-1331
40. Otto K, Mamatoglu N: Why does interactional justice promote organizational loyalty, job performance, and prevent mental impairment? The role of social support and social stressors. *J Psychol* 149: 193-218
41. R Core Team: R: A Language and Environment for Statistical Computing R version 320 (2015-04-16) Vienna, Austria, 2015
42. Revelle W: Psych: Procedures for Psychological, Psychometric, and Personality Research 2015
43. Rosseel Y: Lavaan: An R Package for Structural Equation Modeling. *J Stat Softw* 2012; 48: 1-36
44. Sugimura H, Thériault G: Impact of supervisor support on work ability in an IT company. *Occup Med (Chic Ill)* 2010; 60: 451-457
45. Torp S, Vinje HF: Is workplace health promotion research in the Nordic countries really on the right track? *Scand J Public Health* 2014; 42: 74-81
46. Tuomi K, Ilmarinen J, Jahkola A, et al: A Work Ability Index, 2nd ed Helsinki, Finland: Finnish Institute of Occupational Health, 1998
47. Unruh L, Zhang NJ: The role of work environment in keeping newly licensed RNs in nursing: A questionnaire survey. *Int J Nurs Stud* 2013; 50: 1678-1688

*ACKNOWLEDGEMENTS: We thank Fania Gärtner (from the Department of Medical Decision Making, Leiden University Medical Centre, Leiden, the Netherlands) for providing valuable feedback on earlier drafts of this manuscript and Mrs. E. A. Wright for back translating the questionnaire and proofreading the manuscript.*