Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Psychometric properties of an instrument measuring military morale through work engagement and burnout in the Estonian defense forces

Antek Kasemaa [1, 2, *], Ülle Säälik [3]

[1]Department of Leadership and Management Studies, Baltic Defence College, Tartu, Estonia
[2]School of Natural Sciences and Health, Tallinn University, Tallinn, Estonia
[3]Department of Academic Studies, Estonian Military Academy, Tartu, Estonia

ARTICLE INFO

Article history:
Received 12 January 2025
Received in revised form
21 June 2025
Accepted 23 June 2025

Keywords:
Military morale
Burnout
Work engagement
Validation
Estonia

ABSTRACT

Military morale is commonly understood as the level of enthusiasm and commitment among soldiers that supports effective performance in military settings. One clear definition describes morale as a shared sense of dedication to a common goal that unifies a group. This study examines the psychometric properties of the Estonian version of a military morale instrument, conceptualizing morale in relation to burnout and work engagement. Three instruments were employed: a direct question on self-reported morale, a six-item scale assessing motivation and enthusiasm for mission accomplishment, and a multidimensional scale covering dedication, vigour, cynicism, and emotional exhaustion. Psychometric analysis focused on the multidimensional scale, using data from 3,621 members of the Estonian Defence Forces. Confirmatory Factor Analysis tested factorial structure and assessed configural, metric, and scalar invariance across groups based on age, gender, language, survey wave, profession, residence, student status, and education. Results indicated that both a modified four-factor model (dedication, vigour, cynicism, exhaustion) and a two-factor model (morale and burnout) fit the data well, with acceptable invariance across groups. However, the vigour and cynicism dimensions showed insufficient reliability in some subsamples. Therefore, the two-factor model is recommended for research and screening purposes in the Estonian military context, while the four-factor model requires further refinement before practical use.

1. Introduction

Every aspect contributing to success should be a priority for every organization. Morale, both inside and outside the military environment, has traditionally been considered one such concept. In the military, it describes the personal and group motivation or readiness to achieve mission/task-related objectives (Britt and Dickinson, 2006; Fennell, 2014; Manning, 1994; Motowidlo and Borman, 1977; 1978; Peterson et al., 2008) and it has been considered to be essential for improving performance in the military context (Manning, 1991; Britt and Dickinson, 2006). Despite the widespread use of this term in the military environment, surprisingly little empirical research has been published on the construct (Britt and Dickinson, 2006). Neither has this changed much recently. For instance, the PsycINFO database revealed only 16 pre-reviewed texts mentioning military morale in the title, 168 in the abstract, and 363 anywhere in the article (retrieved December 2023). Moreover, between 2006 and 2023, only 18 International Military Testing Association (IMTA) conference presentations were directly related to military morale (www.imta.info). One explanation could be that some researchers use it ambiguously, not defining it well (Fennell, 2014; Hardy, 2010; Manning, 1991) and not distinguishing it from similar constructs like small unit cohesion, commitment, job satisfaction, or self/collective efficacy (Hardy, 2010). The same problem also appears in military texts and understandings. So, we could assume that morale (at least in the military) has not always been precisely understood and measured (Fennell, 2014). This, in turn, indicates the need for validated measurement instruments.

* Corresponding Author.
Email Address: kasemaa@tlu.ee (A. Kasemaa)
https://doi.org/10.21833/ijaas.2025.07.016
Corresponding author's ORCID profile:
https://orcid.org/0000-0003-4073-3383

Military service is potentially very stressful, not only in a combat situation but also in the peacetime context (Fennell, 2014). Different operational stressors could include a high operational tempo, modern technology, the role of civilian actors, and tactics used by friendly and opposing forces (van Boxmeer et al., 2007). Even though not all categories of military personnel (e.g., conscripts, but not only them) may not necessarily experience battlefields during their service, the personnel's well-being, team performance, and non-burnout are still essential for both operational and non-operational contexts (Ivey et al., 2015; Peterson et al., 2008).

Some academic papers in Estonia have attempted to study military morale empirically using samples from the Estonian Defence Forces (EDF), e.g., Kasemaa and Säälik (2021; 2023). One of the most recent (Kasemaa, 2024) adopted a short morale instrument (Britt and Dickinson, 2006) into the EDF context, which could help fill the gap for proper psychometric tools to measure and monitor morale within EDF. However, this instrument tends not to cover all aspects of the concept. It is focused only on energizing soldiers and supporting their efforts to adapt better to stressful conditions (Britt and Dickinson, 2006) and does not cover other possible elements of military morale, such as dedication, vigour, cynicism, or exhaustion. Current research addresses this shortage by providing psychometric properties for the alternative instrument, which conceptualizes military morale through work engagement and burnout. Therefore, our research aims to adopt a new instrument for measuring military morale in the Estonian context. Firstly, this will help the EDF to measure military morale more adequately, offering the possibility of selecting the proper instruments according to needs and the actual situation. Secondly, this will facilitate cross-cultural research as proposed by several authors (Britt and Dickinson, 2006; Hardy, 2010; Ivey et al., 2015), because the same instrument has already been adopted into several languages.

2. Literature overview

2.1. Defining military morale

One of the best-known definitions of military morale is proposed by Manning (1991): The enthusiasm and persistence with which a member of a group engages in the prescribed activities of that group. This definition has formed the basis of several later approaches, taking morale as a psychological state and describing it in terms of energy and enthusiasm (Britt and Dickinson, 2006; Britt et al., 2007; 2013; Dyches et al., 2017; Fennell, 2014; Meyer et al., 2009; Peterson et al., 2008). In summarizing the literature cited above, it could be concluded that enthusiasm, persistence, and in some ways energy and motivation, which are directed toward achieving the group's goals, might be the core components of military morale. The literature identifies several factors contributing to, describing, or resulting from morale (Britt et al., 2007; Dyches et al., 2017; Hardy, 2010; Maguen and Litz, 2006; Peterson et al., 2008). For instance, constructs such as self-efficacy, confidence, self-esteem, a feeling of group cohesion, relatedness, esprit de corps, solidarity, shared beliefs and convictions, all of them are somehow interwoven with the core components of morale (Manning, 1991; van Boxmeer et al., 2007), mentioned above. However, they do not reflect precisely the aforementioned definition and, therefore, contribute to the vagueness of the construct rather than to its clarity.

The question of whether military morale is an individual or group-level construct is still to some extent unanswered. However, this paper considered morale as an individual's psychological state, which becomes aggregated and supported or reduced due to the group's psychological state. Therefore, morale could be viewed as an individual-level construct contextualized by the group.

To summarize, the current paper, therefore, considers morale as a psychological construct. This variable gives the service member energy, directing the person towards more qualitative performance in stressful conditions, emphasizing the enthusiasm and persistence with which a member of a group engages in the prescribed activities of that group.

2.2. Conceptualizing military morale through work engagement and burnout

van Boxmeer et al. (2007) related morale to the constructs of work engagement (Maslach and Leiter, 1997) and burnout (Maslach et al., 2001) to conceptualize it according to Manning's (1991) definition. Both of these are well-known concepts in organizational and motivational psychology. They argued that elements of military morale, such as enthusiasm and persistence, could be very similar to the key elements of work engagement and burnout, even though the latter appear to be opposites. Burnout (BO) is described as exhaustion, cynicism, and reduced professional efficacy, while work engagement (WE) is defined through energy and involvement (Maslach and Leiter, 1997).

WE is a positive, fulfilling, work-related state of mind characterized by vigour, dedication, and absorption. As this definition proposes, the construct has three components: 1) Vigour as a high level of energy and mental resilience, willingness to invest effort in one's work, and persistence even in the face of difficulties; 2) Dedication as strong involvement in one's work and a sense of significance, enthusiasm, inspiration, pride and challenge; 3) Absorption as being entirely concentrated and deeply engrossed in one's work, whereby time passes quickly and one has difficulties with detaching oneself from work (Schaufeli et al., 2002). Vigour and dedication have been found to be core parts of WE, whereas absorption seems to act as a consequence of it (Langelaan et al., 2006).

At the same time, BO is defined as a persistent, negative, work-related state of mind in otherwise

'normal' individuals. It is characterized by exhaustion (the draining of mental energy), cynicism (a negative attitude toward work), and reduced professional efficacy (the belief that one is no longer effective in fulfilling one's job responsibilities) (Maslach et al., 2001). Exhaustion and cynicism are considered to be core parts of BO, whereas professional efficacy seems to develop in parallel to the rest of the components (Bakker et al., 2008).

According to the literature, BO and WE are negatively, but not perfectly, correlated; more precisely, vigour is the opposite of exhaustion, and dedication is the opposite of cynicism (González-Romá et al., 2006; Taris et al., 2017). These correlations allow measuring of morale in both positive and negative directions (van Boxmeer et al., 2007). In summary, WE is related to a high level of energy and strong identification with one's work, and BO is a low level of energy and poor identification with one's work (Schaufeli and Bakker, 2010).

Several studies outside the military have researched relations between WE and BO and various organizational variables. For instance, WE has been found to be positively related to organizational commitment (Huhtala and Feldt, 2016); job satisfaction (Schaufeli et al., 2008; Garg et al., 2017); job involvement; innovativeness (Gorgievski et al., 2010), and person-organization fit (Huhtala and Feldt, 2016). At the same time, Bianchi and Brisson (2019) found an overlap between BO and depression, Roelen et al. (2015) found a positive relation between BO and long-term sickness absence, and Vander Elst et al. (2016) found that BO was positively associated with workload, emotional demands and aggression, while a negative correlation was found between BO and task autonomy, social support and learning opportunities.

Demerouti and Cropanzano (2010) summarized research on WE and BO relations with job performance. They concluded that despite the lack of publications at the time, there is a positive correlation between WE and performance (measured in different ways) and a negative correlation between BO and job performance. Similar results have also been found later (Bal and De Lange, 2015; Sekhar et al., 2017; Corbeanu et al., 2023). The following studies have identified three groups of drivers having an impact on the WE and BO: Organizational, group, and leader level drivers (Edú-Valsania et al., 2022; Vander Elst et al., 2016). In short, these groups of drivers stimulate WE positively and negatively BO, which, in turn, fosters job performance. Such results align with recent studies in the military context (Britt and Dickinson, 2006; Jeppesen and Elrond, 2021), proposing unit, leadership, and personality-related variables as antecedents of morale.

To summarize, the conceptualization of military morale as a combination of WE and BO gives researchers and practitioners the possibility of establishing solid bases for the understanding of this concept. Moreover, it has a solid theoretical base and allows us to make comparisons between the military and the outside world.

2.3. Questionnaires to measure military morale

Traditionally, three types of questionnaires have been used to measure morale in and out of the military environment.

Firstly, several authors have used items asking directly a simple question, 'What is your level of morale?' (Bliese and Britt, 2001; Maguen and Litz, 2006; Anzai et al., 2014; Dyches et al., 2017; Fazal et al., 2024) or 'What level is the morale in your platoon/company?' (Maguen and Litz, 2006; Jeppesen and Elrond, 2021). This approach has great face validity because it asks directly about the phenomenon of interest. However, the possible downside could be that the answer depends heavily on individuals' understanding of what the concept of morale means. Thus, this single-item approach measures rather the individuals' personal concept of the construct. Nevertheless, this approach has also been used for validation purposes of the morale scales (van Boxmeer et al., 2007; Whitesell and Owens, 2012).

Secondly, whatever the field of research, several examples can be found where morale is measured in an indirect way. For instance, especially in traditional military thinking, exploiting consequences for low or high morale, such as a number of accidents or disciplinary acts (Manning, 1991) or combat readiness (Shamir et al., 2000). Sometimes, a score of different constructs (e.g., cohesion, well-being, satisfaction, motivation, etc.) alone or combined has been used (Johnsrud and Rosser, 2002; Chandra et al., 2016; Jeppesen and Elrond, 2021; Fazal et al., 2024; Kras et al., 2024). Although these ways to measure morale have some good points, the question remains of whether they are measuring morale or something else.

Thirdly, some multi-item questionnaires have been developed to encompass morale as a phenomenon more broadly. However, similarly to previous instruments, the question of validity remains, especially the view of whether such a measure reflects the definition of morale or not.

One promising instrument that attempts to address the concern of meeting the definition is proposed by Van Boxmeer et al. (2007). This measure considers military morale as a multidimensional construct based on the definition given by Manning (1991). The authors selected the Utrecht WE and BO measures as the basis for this instrument. Mainly because these have been widely used in various samples in the work context (Vander Elst et al., 2016) and have been well studied from a psychometric point of view (Lesener et al., 2020; De Beer et al., 2024). van Boxmeer et al. (2007) argued that the core dimensions of WE and BO cover well the core elements of military morale (Manning, 1991). Since its introduction, this approach has gained some popularity in subsequent studies and has been adopted by several armed forces (Ivey et

al., 2015). Following a similar approach as van Boxmeer et al. (2007) and Britt and Dickinson (2006) conceptualized morale as motivation and enthusiasm and proposed a short, unidimensional six-item instrument to measure the construct. This instrument has been used in several subsequent studies (Frone and Blais, 2019; Michaud et al., 2025). However, considering the conceptual similarities between this (Britt and Dickinson, 2006) and the instrument introduced above (van Boxmeer et al., 2007), it is clear that both measures reflect respondents' energy, enthusiasm, motivation, and eagerness to achieve mission success or group goals. This becomes especially obvious when comparing the wording of the items. Ivey et al. (2015) analysed this similarity and found a remarkable overlap between them. However, the two instruments were still empirically distinct. The authors claimed that absorption is not conceptually close enough to morale, and therefore, eliminating it might render WE less distinct from military morale.

To summarize the discussion above, we assumed that the instrument measuring military morale through WE and BO is appropriate for adaptation to the Estonian language. Following that, we proposed the following hypotheses: (1) There are negative correlations between WE (dedication and vigour) and BO (cynicism and exhaustion); (2) the model of four correlated latent variables (dedication, vigour, cynicism and exhaustion) demonstrates the best fit to the data compared to the alternative models; (3) WEBO (work engagement and burnout) is invariant across age, gender, language, wave, profession, place of living, student status and educational groups; (4) dedication, vigour, cynicism and exhaustion are all statistically significant predictors of direct morale (DMQ); (5) WE and unidimensional six-item instrument (SMQ) measure the same construct, meaning that SMQ does not add a unique contribution to the WE predicting DMQ.

3. Method

3.1. Sample and procedure

We merged data from seven independent studies, with a total sample of 4152 Estonian service members, with a mean age of 22.46 (SD = 4.58; min 18; max 65 years). Of these, 3591 were male and 168 were female (missing 393), 3434 reported having Estonian as a mother tongue, and 427 other languages, primarily Russian (missing 291). By education, 540 participants had basic, 2826 secondary, and 517 higher education (missing 269). Among all participants, 702 had the status of student. 1232 had lived most of their lives in the countryside, 704 in towns, and 1162 in cities (missing 1054). By military profession, 2845 were conscript soldiers, 443 were conscript NCOs, 104 were conscript reserve officers, 392 were professional soldiers, and 153 were civilians working within the defence system (missing 215). Additionally, the sample was divided between the phases of military training (wave): 729 participants were interviewed in their basic, 1669 specialized, and 1281 during their collective training; however, for 320 professionals and 153 civilians, the wave was not specified.

All participants filled out the questionnaires in classrooms using a paper-and-pencil approach. As a first step before filling out the questionnaires, the authors explained the aim and meaning of the research and asked for informed consent from the participants. Participation in the study was voluntary; participants could stop filling out the questionnaire at any time.

3.2. Measures

Military morale was measured using the 16-item measure of WEBO proposed by van Boxmeer et al. (2007). This instrument was intended to measure both the positive (eight items) and negative (eight items) sides of military morale. Both consisted of two subscales, four items each: Vigour and dedication for the morale, and cynicism and exhaustion for the burnout. A five-point Likert-type scale was used for the answers, ranging from never to always.

Additionally, morale was assessed by the six-item instrument proposed by Britt and Dickinson (2006). Respondents were instructed to think about their work objectives while answering. Items asked them to assess their level of motivation, morale, energy, drive, enthusiasm, and eagerness using a five-point Likert scale (from very low to very high). It was found previously that the measurement properties of this instrument are sufficient for use on Estonian military samples (Kasemaa, 2024).

For validation, morale was also assessed using direct items. To measure individual morale, participants were asked: "My personal morale is …" For collective morale, they were asked: "The morale of my fellow soldiers/platoon/company is …" The correlations between these direct items and the components of military morale measured by other instruments were used to assess the construct validity of the morale questionnaires, as shown in previous studies (Whitesell and Owens, 2012; Dyches et al., 2017; van Boxmeer et al., 2007). CFA for morale items (DMQ_{sum}) demonstrated an acceptable fit of the data: $\chi^2(2) = 16.48$, $p < .001$, RMSEA = .048, CFI = .997, SRMR = .026.

3.3. Strategy of analysis

Firstly, we analysed the construct validity of the military morale instrument (van Boxmeer et al., 2007). For this, we performed a series of CFAs with JASP 0.18.3. A diagonally weighted least squares estimator was used because ordinal variables were not normally distributed.

We compared five alternative models: (1) one-factor model (M1), where all 16 items were specified to load on a 'general' morale factor; (2) two-factor model (M2), where a distinction between morale and

BO was made; (3) four-factor model (M3), where separation between vigour, dedication, cynicism and exhaustion (4 items each) was made; and (4) hierarchical (M4), specifying all four aforementioned components to load into the second level factor named military morale.

The goodness of fit of the CFA models was judged using the comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Additionally, chi-square was reported. For CFI, values ≥ .93 were considered an acceptable fit (Hu and Bentler, 1999), and values ≤ .08 were considered acceptable for RMSEA and SRMR (Marsh et al., 2004).

To assess the measurement invariance (configural, metric, and scalar) of WEBO, we performed a series of multi-group CFAs. The selection of the groups was based on findings from Kasemaa and Säälik (2021), who found meaningful differences in military morale across age, gender, language, place of living, student status, service wave, and position groups. Configural measurement invariance was calculated to assess whether the number of factors (scale factor structure) is the same across groups; metric measurement invariance to demonstrate whether factor loadings are similar across the groups, and scalar measurement invariance to assess whether the residuals are equivalent across the compared groups. To make a comparison between the fit of the models, $\Delta\chi^2$ (Satorra and Bentler, 2001), ΔCFI, ΔRMSEA, ΔSRMR, and ECVI (expected cross-validation index) were used. The change of -.01 in CFI (Cheung and Rensvold, 2002), -.015 in RMSEA, and -.030 in SRMR were used as cut-off criteria (Chen, 2007). For ECVI, a smaller value indicates a better model (Browne and Cudeck, 1993). Additionally, the pattern of factor loadings was evaluated, with loadings ≥ .40 being considered as still meaningful (Stevens, 2002).

The third round of analyses assessed characteristics of WEBO subscales resulting from the CFA. The reliability (McDonald's ω) and arithmetic means, together with the standard deviation, were calculated. Statistical significance of mean differences was examined by Kruskal-Wallis nonparametric dispersion analysis accompanied by effect size (ε^2). Effect sizes between .01-.06 were considered small, and .06-.14 as medium (Field, 2018). The fourth round of analyses examined how well WE, BO, and SMQ might predict individual and collective levels of morale reported directly by respondents. For this, a series of linear regression analyses were conducted, having items directly measuring morale as dependent variables (DV) and WEBO and SMQ as independent variables (IV).

4. Results

4.1. Descriptives and correlations

Descriptive statistics of the WEBO items are presented in Table 1. The average values of the items were between 2.381 (E-3) and 3.645 (V-2), and most items were negatively skewed (highest for item V-2; -.628) and leptokurtic (highest for item E-2; -.775). However, we stuck to Hair et al.'s (2022) suggestion, which proposed sufficient skewness and kurtosis values between -1 and +1.

Table 1: Descriptive statistics of the items of the Estonian version of WEBO

Item[1]	M	SE of M	SD	Median	Skewness	Kurtosis
D-1	2.927	.018	1.076	3	-.149	-.661
V-1	2.788	.017	1.051	3	-.001	-.688
C-1	2.741	.017	1.023	3	.230	-.557
E-1	2.910	.018	1.111	3	.120	-.727
D-2	3.220	.018	1.073	3	-.427	-.533
V-2	3.645	.014	.863	4	-.628	.444
C-2	2.694	.017	1.023	3	.288	-.397
E-2	2.855	.018	1.105	3	.141	-.775
D-3	3.407	.017	1.013	4	-.484	-.120
V-3	3.556	.015	.918	4	-.522	.092
C-3	3.098	.019	1.117	3	-.092	-.723
E-3	2.381	.018	1.061	2	.582	-.209
D-4	3.419	.018	1.080	4	-.541	-.303
V-4	3.537	.016	.962	4	-.566	.049
C-4	3.024	.016	.967	3	-.078	-.213
E-4	3.486	.018	1.072	4	-.372	-.529

n = 3621; D: Dedication; V: Vigor; C: Cynicism; E: Exhaustion; 1: The order of the items in the table follows the sequence of items in the questionnaire

4.2. Estonian version of WEBO: Factorial structure

To assess the internal validity of WEBO, several CFA models were analysed (Table 2). Models 2, 3, and 4 showed a good fit for the data, and M1 failed to meet CFI, RMSEA, and SRMR cut-off criteria; nevertheless, M3 fitted the best. Despite that, the modification indices suggested paths between some items and factors that were not expected to load.

A series of Exploratory Factor Analyses, fixing the number of components extracted at 2, 3, and 4 and using different extraction methods, demonstrated the same result: one item from vigour (V-1) loaded into dedication. The fit of the data of the modified model ($M3_{mod}$) was: $\chi^2(98) = 640.53$; RMSEA .040; CFI .984; and SRMR .040. At the same time, comparison of all models, based on cut-off criteria on $\Delta\chi^2$(df), ΔCFI, ΔRMSEA, ΔSRMR, and ECVI, demonstrated that $M3_{mod}$ differed from M2 and M4.

However, a meaningful difference from M3 was only based on ECVI.

For the WEBO two-factor model, squared factor loadings (R^2) for morale items were between .30 and .57, for burnout items between .11 and .48. For the WEBO four-factor modified model, squared factor loadings (R^2) for dedication items were between .32 and .70 for vigour items between .40 and .53, for cynicism items between .12 and .46, for exhaustion items between .35 and .51. All loadings were statistically significant ($p < .001$). In conclusion, item C-4 (I have grown more cynical about the effects of my work) did not meet the cut-off criteria ($R^2 > .16$). Therefore, due to the low factor loading, we decided to remove this item from the subsequent analyses and conducted an additional CFA for $M2_{mod}$ (Table 2).

Table 3 shows descriptive statistics, correlations (ρ), and reliability of the morale instruments and their subscales. All calculations were made based on the results from CFA analyses. As was expected, negative correlations between morale and burnout subscales were found (between -.31 and -.46). McDonald's ω was between .92 (SMQ) and .60 (cynicism), being the only one below the threshold > .70 (Furr and Bacharach, 2013).

Table 2: CFA for work engagement and burnout (WEBO)

Model	χ^2 (p)	df	RMSEA[90% CI]	CFI	SRMR	ECVI
M1: One F model	3063.59 (< .001)	104	.090[.087-.093]	.908	.087	.898
M2: Two F model	1123.12 (< .001)	103	.053[.050-.056]	.969	.053	.338
$M2_{mod}$: Two F model	967.77 (< .001)	89	.050[.050-.056]	.973	.052	.301
M3: Four F model	766.52 (< .001)	98	.043[.040-.046]	.979	.042	.231
$M3_{mod}$: Four F model	640.53 (< .001)	98	.040[.037-.043]	.984	.040	.213
M4: Hierarchical F model	1449.25 (< .001)	100	.062[.059-.065]	.959	.060	.231

n = 3621; RMSEA: Root mean square error of approximation; CFI: Comparative fit index; GFI: Goodness of fit index; SRMR: Standardized root mean square residual; ECVI: Expected cross validation index; Method: Diagonally weighted least squares

In order to assess the measurement invariance of the Estonian version of WEBO, a series of multi-group CFAs was conducted across age groups (2), mother tongue (2), place of living (2), student status (2), education (3), profession (2), and wave (4).

Firstly, we assessed the two-factor model ($M2_{mod}$) of WEBO. Results are presented in Table 4. Models of configural, metric, and scalar invariance demonstrated a good fit of the data and generally did not statistically differ from their predecessor models. All ΔRSMEA and ΔSRMR values were below the threshold (RMSEA < .015 and SRMR < .030). ΔCFI showed no differences in analysed models, except metric invariance (the factor loadings are constrained to be equal across groups) between the levels of education (-.014 > -.01). The $\Delta\chi^2$ test was significant for most of the models. However, this index is strongly influenced by the number of cases (Tabachnick and Fidell, 2007) and nearly always rejects the model for a large sample size (Jöreskog and Sörbom, 1993). Therefore, this finding was discarded.

In conclusion, two factors of WEBO demonstrated invariance across age, gender, language, profession, wave, place of living, and student status groups, allowing the comparison of military morale for these groups. However, this was partly not true for the educational groups.

Table 3: Descriptive statistics and correlations (Spearman's *rho*) of morale instruments

Variable	m	SD	1	2	3	4	5	6	7	8	9	10
1. Dedication (5 items)	3.15	.84	(.86)									
2. Vigour (3 items)	3.58	.73	.57*	(.71)								
3. Cynicism (3 items)	2.85	.78	-.46*	-.31*	(.60)							
4. Exhaustion (4 items)	2.91	.83	-.37*	-.44*	.56*	(.76)						
5. Morale (8 items)	3.32	.72	.95*	.78*	-.46*	-.43*	(.87)					
6. Burnout (7 items)	2.88	.72	-.46*	-.44*	.83*	.92*	-.50*	(.80)				
7. SMQ (6 items)	3.46	.90	.67*	.53*	-.44*	-.47*	.69*	-.51*	(.92)			
8. DMQ_{sum} (4 items)	3.28	.82	.47*	.37*	-.27*	-.29*	.49*	-.32*	.57*	(.84)		
9. DMQ_{ind} (1 item)	3.36	1.05	.47*	.39*	-.31*	-.34*	.49*	-.37*	.65*	.76*	(n/a)	
10. DMQ_{col} (3 items)	3.25	.88	.39*	.30*	-.20*	-.22*	.40*	-.25*	.46*	.94*	.54*	(.83)

n = 3641; For SMQ n = 981; *: p < .001; DMQ_{sum}: Three collective morale items and one individual level item; DMQ_{ind}: One individual morale item; DMQ_{col}: All three direct collective morale items; McDonald's ω are in brackets

Secondly, we assessed the modified four-factor model ($M3_{mod}$) of WEBO. The results are presented in Table 5, and the mean and SD differences are in Table 6. Models of configural, metric, and scalar invariance demonstrated a good fit of the data and did not statistically differ from their predecessor models. All ΔRSMEA and ΔSRMR values were below the threshold (RMSEA < .015 and SRMR < .030). ΔCFI showed no differences in the analysed models. $\Delta\chi^2$ test was significant in most of the cases. However, the authors disregarded this finding based on the arguments summarized previously. Therefore, we concluded that the modified four-factor model of WEBO is invariant across age, gender, language, wave, profession, place of living, student status, and educational groups and allows the comparison of military morale (dedication and vigour) and burnout (cynicism and exhaustion).

4.3. Estonian version of WEBO: Characteristics

A descriptive analysis (means, SDs, and reliability) was conducted on different socio-demographic groups (Table 6). This analysis highlighted a problem with the reliability of some subscales, which was already identified from Table 3.

A deeper analysis across different socio-demographic groups revealed that for cynicism (3 items), the reliability criteria were not met for almost all subgroups. Additionally, for vigour (3 items), these criteria were not met for some subgroups, such as non-Estonian speakers (.59), basic training participants (.68), training time not specified (.64), professionals (.65), those living in towns (.65), having higher education (.69) and females (.62); or for exhaustion in groups of professionals (.69) and higher education (.68). The reliability figures of the morale (8 items) and burnout (7 points) subscales were above acceptable levels across all socio-demographic groups (McDonald's ω > 0.7).

Table 4: The fit of multi-group models of the modified two-factor model of WEBO

Sample	Model	$\chi^2(df)$*	RMSEA[90% CI]	CFI	SRMR	$\Delta\chi^2$ (df)	ΔCFI	ΔRSMEA	ΔSRMR
Age	Configural	1099.93(178)	.052[.048-.055]	.974	.053				
	Metric	1183.13(191)	.054[.051-.057]	.969	.057	83.20(13)	.005	.002	.004
	Scalar	1235.79(204)	.054[.051-.057]	.968	.058	52.66(13)	.001	.000	.001
Mother tongue	Configural	986.08(178)	.053[.050-.056]	.973	.055				
	Metric	1040.70(191)	.052[.049-.055]	.972	.056	54.62(13)	.001	.001	.001
	Scalar	1198.68(204)	.055[.052-.058]	.967	.058	157.98(13)	.005	.003	.002
Wave	Configural	956.86(267)	.049[.045-.052]	.976	.053				
	Metric	1263.67(293)	.055[.052-.058]	.967	.060	306.81(26)	.009	.006	.007
	Scalar	1427.06(319)	.063[.053-.059]	.962	.056	163.39(26)	.005	.008	.004
Profession	Configural	1035.64(178)	.052[.049-.055]	.971	.054				
	Metric	1214.05(191)	.055[.052-.058]	.965	.057	178.41(13)	.006	.003	.003
	Scalar	1363.46(204)	.057[.054-.060]	.960	.059	151.25(13)	.005	.002	.002
Place of living	Configural	819.97(267)	.050[.046-.053]	.975	.057				
	Metric	883.59(293)	.049[.045-.053]	.973	.059	63.62(26)	.002	.001	.002
	Scalar	901.32(319)	.047[.043-.050]	.974	.059	17.73(26)	.001	.002	.000
Student status	Configural	608.01(178)	.050[.046-.055]	.975	.056				
	Metric	681.59(191)	.052[.048-.056]	.971	.058	73.58(13)	.004	.002	.002
	Scalar	716.41(204)	.051[.047-.056]	.970	.059	34.82(13)	.001	.001	.001
Education	Configural	1023.13(267)	.051[.048-.054]	.976	.054				
	Metric	1433.42(293)	.060[.056-.063]	.963	.062	410.29(26)	.014	.009	.008
	Scalar	1530.06(319)	.059[.056-.062]	.961	.063	96.64(26)	.002	.001	.001
Gender	Configural	932.92(178)	.052[.049-.055]	.973	.054				
	Metric	1018.22(191)	.052[.049-.056]	.970	.055	85.30(13)	.003	.000	.001
	Scalar	1078.20(204)	.052[.049-.055]	.969	.056	59.98(13)	.001	.000	.001

*: p < .001

Additionally, a series of Kruskal-Wallis nonparametric dispersion analyses were conducted, in which previously analysed WEBO subscales (modified) were the dependent variables. The results demonstrated statistically significant differences (p < .05) between the various socio-demographic groups. To summarize: 1) profession and time of training (wave) made a difference across all analysed WEBO subscales; 2) gender and mother tongue showed differences across all WEBO subscales; 3) education differentiated vigour, exhaustion and burnout; 4) student status and living place made a difference in dedication and morale; 5) age differentiated vigour. However, the effect sizes were relatively low, starting with morale across the profession (ε^2 = .051), which was the highest. All other effect sizes were below ε^2 = .051, which are considered as small or very small (Field, 2018).

Table 5: The fit of the multi-group models of the modified 4-factor model of WEBO

Sample	Model	$\chi^2(df)$*	RMSEA[90% CI]	CFI	SRMR	$\Delta\chi^2$ (df)	ΔCFI	ΔRSMEA	ΔSRMR
Age	Configural	701.48(196)	.038[.035-.041]	.985	.044				
	Metric	902.00(208)	.044[.041-.047]	.979	.050	200.52(12)	.006	.006	.006
	Scalar	940.62(220)	.043[.040-.046]	.978	.048	38.62(12)	.001	.001	.002
Mother tongue	Configural	713.33(196)	.040[.037-.043]	.983	.045				
	Metric	747.61(208)	.040[.037-.043]	.982	.046	34.28(12)	.001	.000	.001
	Scalar	849.15(220)	.042[.039-.045]	.980	.045	101.54(12)	.002	.002	.001
Wave	Configural	785.32(392)	.034[.030-.037]	.992	.046				
	Metric	1175.12(428)	.045[.042-.048]	.983	.054	389.80(36)	.009	.011	.008
	Scalar	1385.29(464)	.048[.045-.050]	.978	.054	210.17(36)	.005	.003	.000
Profession	Configural	700.66(196)	.038[.035-.041]	.983	.044				
	Metric	813.19(208)	.041[.038-.044]	.980	.047	112.53(12)	.003	.003	.003
	Scalar	948.17(220)	.043[.041-.046]	.976	.046	134.98(12)	.004	.002	.001
Place of living	Configural	489.55(196)	.034[.031-.038]	.987	.044				
	Metric	513.85(208)	.034[.030-.038]	.987	.045	24.30(12)	.000	.000	.001
	Scalar	520.61(220)	.033[.029-.037]	.987	.043	6.76(12)	.000	.001	.001
Student status	Configural	391.70(196)	.032[.028-.037]	.989	.045				
	Metric	438.74(208)	.034[.030-.039]	.987	.048	47.04(12)	.002	.002	.003
	Scalar	448.15(220)	.033[.029-.037]	.987	.045	9.41(12)	.000	.001	.003
Education	Configural	763.09(294)	.038[.035-.041]	.985	.047				
	Metric	903.14(317)	.041[.038-.044]	.982	.050	140.05(23)	.003	.003	.003
	Scalar	955.63(341)	.041[.037-.044]	.981	.048	52.49(24)	.004	.000	.002
Gender	Configural	622.88(196)	.037[.034-.040]	.985	.044				
	Metric	696.82(208)	.039[.035-.042]	.983	.045	73.94(12)	.002	.002	.001
	Scalar	749.17(220)	.039[.036-.042]	.982	.043	52.35(12)	.001	.000	.002

*: p < .001

4.4. Incremental validity of SMQ over WEBO

The next set of analyses answered the question of how well WE, BO subscales, and SMQ (IVs: Dedication, vigour, cynicism, exhaustion, and SMQ) would predict directly-asked morale (DVs: Summarized, collective, and individual morale). Table 3 shows the correlations between the independent and dependent variables. BO subscales were negatively (ρ between -.20 and -.34) and WE positively (ρ between .30 and .47) correlated with direct morale indicators. Nevertheless, dedication had the highest correlations with direct morale (ρ between .39 and .47), and all morale subscales had the highest correlations with individual direct morale. At the same time, SMQ demonstrated the strongest correlations with direct morale indicators (ρ between .46 and .65) compared with WEBO subscales. Table 7 presents the results of the regression analyses. Firstly, the data was checked for assumptions and outliers. Outliers were not found, and assumptions of linearity, normality, homoscedasticity, and collinearity were found to be met (Field, 2018): VIF < 2.249; Tolerance > .441; Durbin-Watson statistic between 1.663 and 2.266 (p > .001).

Only dedication (β between .34 and .41, p < .01) and exhaustion (β between -.20 and -.26, p < .01) were significant predictors of all direct morale indicators when BO and WE subscales were included in the model. Among direct morale indicators, individual morale was best predicted by BO and WE subscales.

The model explained 38% of the variance of the IV (R^2 = .38, F(4,503) = 79.63, p < .01). Adding SMQ into the model, all BO and WE subscales lost their power to predict direct morale indicators. However, dedication and exhaustion both remained statistically significant (p < .05). Similarly to the previous results, individual direct morale was the best (R^2 = .48, F(5,502) = 96.23, p < .01) and collective morale the worst (R^2 = .22, F(5,503) = 29.71, p < .01) predicted by the WEBO and SMQ.

Table 6: Differences in military morale across various socio-demographic groups

Group	Dedication (5)	Vigour (3)	Cynicism (3)	Exhaustion (4)	Morale (8)	Burnout (7)
	m(*SD*)	*m*(*SD*)	*m*(*SD*)	*m*(*SD*)	*m*(*SD*)	*m*(*SD*)
Age	$p > .05$*	$p < .05$; $\varepsilon^2 = .002$	$p > .05$	$p > .05$	$p > .05$	$p > .05$
Younger** (n = 1884)***	3.16(.83)	3.54(.74)	2.83(.71)	2.92(.85)	3.31(.72)	2.88(.74)
Older (n = 1710)	3.14(.86)	3.61(.72)	2.86(.71)	2.90(.80)	3.32(.71)	2.89(.70)
McDonalds ω****	.84; .87	.70; .72	.59; .59	.78; .73	.86; .86	.79; .76
Mother tongue	$p < .05$; $\varepsilon^2 = .002$	$p < .05$; $\varepsilon^2 = .020$	$p < .05$; $\varepsilon^2 = .001$	$p < .05$; $\varepsilon^2 = .007$	$p < .05$; $\varepsilon^2 = .007$	$p < .05$; $\varepsilon^2 = .005$
Estonian (n = 2971)	3.14(.86)	3.60(.72)	2.85(.77)	2.89(.83)	3.32(.73)	2.88(.71)
Other (n = 367)	3.06(.69)	3.30(.75)	2.94(.79)	3.12(.81)	3.15(.64)	3.03(.71)
McDonalds ω	.87; .71	.73; .59	.57; .62	.76; .74	.87; .75	.79; .80
Wave	$p < .05$; $\varepsilon^2 = .031$	$p < .05$; $\varepsilon^2 = .021$	$p < .05$; $\varepsilon^2 = .025$	$p < .05$; $\varepsilon^2 = .012$	$p < .05$; $\varepsilon^2 = .031$	$p < .05$; $\varepsilon^2 = .020$
Basic training (n = 714)	3.25(.72)	3.57(.72)	2.81(.76)	2.93(.85)	3.38(.65)	2.88(.72)
Spec. training (n = 1561)	3.01(.87)	3.48(.74)	2.91(.76)	2.90(.82)	3.19(.74)	2.95(.70)
Collect. training (n = 1072)	3.19(.85)	3.67(.70)	2.87(.78)	2.85(.81)	3.37(.71)	2.86(.71)
Other (n = 247)	3.56(.79)	3.83(.68)	2.38(.84)	2.64(.81)	3.67(.66)	2.52(.73)
McDonalds ω	.77; .87;.86; .86	.68;72; .72;.64	.57; .56;.59; .64	.80; .74; .75; .71	.80; .88; .86; .84	.80; .79; .80; .79
Profession	$p < .05$; $\varepsilon^2 = .049$	$p < .05$; $\varepsilon^2 = .030$	$p < .05$; $\varepsilon^2 = .044$	$p < .05$; $\varepsilon^2 = .026$	$p < .05$; $\varepsilon^2 = .051$	$p < .05$; $\varepsilon^2 = .042$
Conscripts (n = 3184)	3.09(.84)	3.54(.73)	2.90(.76)	2.96(.82)	3.26(.72)	2.93(.70)
Professionals (n = 410)	3.65(.73)	3.91(.60)	2.38(.78)	2.55(.76)	3.75(.60)	2.47(.68)
McDonalds ω	.85; 85	.71; .65	.56; .64	.76; .69	.86; .83	.79; .78
Place of living	$p < .05$; $\varepsilon^2 = .004$	$p > .05$	$p > .05$	$p > .05$	$p < .05$; $\varepsilon^2 = .003$	$p > .05$
Countryside (n = 1041)	3.19(.82)	3.61(.70)	2.90(.75)	2.91(.79)	3.35(.70)	2.91(.67)
Towns (n = 568)	3.13(.83)	3.60(.66)	2.90(.75)	2.93(.79)	3.31(.69)	2.92(.68)
Cities (n = 983)	3.06(.86)	3.57(.72)	2.88(.77)	2.88(.83)	3.25(.73)	2.88(.71)
McDonalds ω	.87; .86; .86	.72; .65; .72	.55; .61; .58	.73; .74; .75	.87; .87; .86	.78; .79; .79
Student status	$p < .05$; $\varepsilon^2 = .021$	$p > .05$	$p > .05$	$p > .05$	$p < .05$; $\varepsilon^2 = .014$	$p > .05$
Student (n = 649)	2.89(.90)	3.53(.73)	2.92(.77)	2.91(.81)	3.13(.74)	2.91(.69)
Non-student (n = 1333)	3.17(.81)	3.57(.70)	2.91(.69)	2.95(.81)	3.32(.70)	2.93(.70)
McDonalds ω	.87; .86	.70; .71	.57; .58	.70; .76	.86; .87	.76; .80
Education	$p > .05$	$p < .05$; $\varepsilon^2 = .006$	$p > .05$	$p < .05$; $\varepsilon^2 = .002$	$p > .05$	$p > .05$
Basic (n=442)	3.23(.82)	3.47(.82)	2.86(.77)	2.97(.85)	3.33(.75)	2.92(.73)
Secondary (n=2455)	3.14(.83)	3.58(.72)	2.84(.77)	2.91(.83)	3.31(.71)	2.88(.71)
Higher (n=465)	3.16(.95)	3.69(.71)	2.80(.86)	2.82(.80)	3.36(.75)	2.81(.72)
McDonalds ω	.84; .86; .90	.76; .72; .69	.53; .59; .64	.80; .76; .68	.87; .87; .88	.81; .80; .77
Gender	$p < .05$; $\varepsilon^2 = .017$	$p < .05$; $\varepsilon^2 = .009$	$p < .05$; $\varepsilon^2 = .009$	$p < .05$; $\varepsilon^2 = .003$	$p < .05$; $\varepsilon^2 = .017$	$p < .05$; $\varepsilon^2 = .007$
Male (n = 3070)	3.13(.85)	3.58(.72)	2.87(.76)	2.91(.81)	3.30(.72)	2.87(.70)
Female (n = 157)	3.65(.78)	3.88(.65)	2.48(.92)	2.68(.85)	3.73(.63)	2.59(.78)
McDonalds ω	.86; .86	.72; .62	.58; .72	.75; .72	.87; .83	.79; .81

*: Non-significance (p > .05) or significance (p <. 05) of the dispersion analysis (Kruskal-Wallis), accompanied by effect size (ε^2) if p < .05; **: Cut-off point age of 21.43; ***: Number in parentheses shows sample size (n); ****: All McDonalds ω figures are presented according to the order presented in the groups column

5. Discussion

This study aimed to adopt an instrument measuring military morale into the Estonian context. The aim was justified by the need for alternative measurement tools, which would be based on a broader approach to military morale compared to the existing ones. This study analysed an instrument that assesses morale through work engagement and burnout (WEBO). Both are considered to be closely related to the elements of military morale: Work engagement (dedication and vigour) as a positive, work-related state of mind that enhances professional efficacy, and burnout (cynicism and

exhaustion) as a persistent, negative, work-related state of mind that reduces professional efficacy (van Boxmeer et al., 2007).

The results supported the proposed first hypothesis, meaning that the anticipated pattern of correlations was found as predicted (van Boxmeer et al., 2007). There were negative correlations between WE and BO components, which were in the range of previous studies summarized by Schaufeli and Bakker (2010), confirming the expected relations between those two.

Table 7: Summary of hierarchical regression analyses for military morale variables predicting direct morale

	DV: Direct morale all			DV: Direct morale collective			DV: Direct morale individual		
Variable	B	SE B	β	B	SE B	β	B	SE B	β
Intercept	2.50	.24	-	2.56	.29	-	2.63	.30	-
Dedication	.40	.05	.41**	.36	.06	.34**	.49	.06	.38**
Vigour	.04	.05	.04	.01	.06	.01	.12	.06	.09
Cynicism	.00	.05	.001	.04	.06	.04	-.10	.06	-.07
Exhaustion	-.21	.05	-.22**	-.21	.06	-.20**	-.33	.06	-.26**
R^2 *(adjusted)*		.289**			.175**			.383**	
F(df)		52.54**(4;504)			27.85**(4;504)			79.63**(4;503)	
Intercept	1.71	.25	-	1.90	.30	-	1.44	.30	-
Dedication	.20	.05	.21**	.20	.06	.19**	.19	.06	.15**
Vigour	-.03	.05	-.03	-.05	.06	.04	.02	.06	.02
Cynicism	-.03	.05	.02	.07	.06	.06	-.07	.06	-.05
Exhaustion	.-.10	.05	-.11*	-.12	.06	-.11*	-.17	.06	-.13*
SMQ	.38	.05	.43**	.32	.06	.33**	.57	.06	.48**
R^2 *(adjusted)*		.368**			.220**			.484**	
ΔR^2		.08**			.05**			.10**	
F(df)		60.21**(5;503)			29.71**(5;503)			96.23**(5;502)	

The enter method was used to include all WEBO components and SMQ in the model; n = 508; *: p < .05; **: p < .01

The second hypothesis proposed the four-factor model as the most suitable for the data. The CFA confirmed this proposition, in which the modified four-factor model (dedication, vigour, cynicism, and exhaustion) best fitted the data. Nevertheless, one item from vigour tended to load into dedication. This item (at my work, I feel bursting with energy) was perceived by the respondents as a part of dedication, together with the items measuring enthusiasm, purpose, pride, and challenge. The remaining items from vigour touched resilience and perseverance, into which energy did not fit. Additionally, item C-4 (I have grown more cynical about the effects of my work) demonstrated low loadings into the cynicism factor, perhaps a reason why the reliability figure for this factor was below the threshold ($\omega > .70$). A more in-depth analysis across different socio-demographic groups showed that the reliability of cynicism is low in most of the subgroups analysed. At the same time, for vigour, this was the case in some groups. Nevertheless, CFA results were in line with previous studies (van Boxmeer et al., 2007; Ivey et al., 2015), providing evidence that both the four-factor and two-factor modified models fit the data well (indicating the construct validity).

We could not identify one single reason for such a discrepancy. However, in general, cynicism has been found to have lower reliability figures than exhaustion (De Beer et al., 2024). One possible explanation would involve the wording of the items, some of which contain complex words (e.g., cynicism). Therefore, we suggest rewording those items from vigour and cynicism for future studies. This could enhance the reliability and factor loading figures and allow more adequate group comparisons. We do not recommend dismissing these items, because they cover important parts of the concept they measure. Thus, eliminating those items will lead to the reduction of the concept, which has to be avoided. The third hypothesis anticipated WEBO being invariant across various socio-demographic groups to allow its use as a diagnostic and research tool in the Estonian language. We analysed modified versions of the two- and four-factor models (an item from vigour was added to dedication, and an item from cynicism was removed). The modifications were based on the previous analyses. The selection of the groups (age, gender, language, place of living, student status, service wave, and position groups) was based on findings from Kasemaa and Säälik (2021), who found meaningful differences in military morale across these groups. It is worth emphasizing here that the current study is one of the first to test the measurement invariance of the WEBO instrument across different socio/demographic groups (at least to the best of the authors' knowledge).

The modified two-factor model of WEBO demonstrated invariance in most cases, except for the educational groups (metric invariance), which did not meet ΔCFI cut-off criteria. However, we ignored this result (it was the only one) because other indicators did not support this. Therefore, we concluded that the modified two-factor WEBO is suitable for measuring military morale across age, gender, language, profession, place of living, wave, and student status groups. However, our results did not fully confirm the suitability of the modified two-factor WEBO to measure military morale across the educational groups. Considering the explanation of metric invariance (Putnick and Bornstein, 2016), this means that at least one morale or burnout item is not contributing to the designated latent factor to a similar degree across the educational groups. Although we eliminated an item (C-4) from the analysis, a problem remained. A closer look at the details revealed that the problematic items are those measuring cynicism. We have already discussed some 'weak' items in the previous paragraph, so rewording items could help to solve the problem.

Secondly, we analysed the modified four-factor WEBO model, which demonstrated full invariance across the analysed groups. Therefore, we conclude that the modified four-factor WEBO is invariant across the age, gender, language, profession, place of living, wave, education, and student status groups.

The fourth proposition stated that WEBO predicts well directly-measured morale. Results from regression analyses demonstrated that only dedication and exhaustion (from WEBO) were statistically significant predictors of all direct morale variables. Therefore, we could confirm that the fourth proposition is accepted while considering that some morale and burnout components did not contribute statistically significantly to the prediction model.

These results indicate that dedication alone might form the core part of military morale. At the same time, exhaustion seems to form the core part of burnout, predicting directly-measured military morale. However, van Boxmeer et al. (2007) found that exhaustion did not make a remarkable contribution to the concept of morale, while cynicism and vigour did. Our result might be biased by the low reliability of the cynicism and vigour subscales compared with van Boxmeer et al.'s (2007) study. An alternative explanation, however, could be that our study sample is somehow unique because it involves a remarkable proportion of conscripts and perhaps is not fully comparable with the samples from professional armies. This explanation is in some way supported by contrasting conscripts' and professionals' reliability figures (Table 6). Additionally, comparing the regression models predicting the collective and individual levels of direct morale demonstrated that WEBO worked better to describe individual morale. This might indicate that this measure is unsuitable for assessing collective energy and enthusiasm (for instance, aggregating individual-level data to squad or platoon level), which are needed to achieve mission objectives.

The sixth proposition stated that the short measure of morale does not add predictive power to the WE and BO-based components. The results did not support this hypothesis. The squared correlation between SMQ and dedication explains approximately 50% of the variance of both constructs, meaning that the other half remains unexplained. Additionally, adding SMQ into the regression models reduces about 50% of the prediction power of dedication and exhaustion (both remained significant at $p < .05$). However, the model R^2 increased statistically significantly ($p < .05$) for all three models. At the same time, looking at the pattern of statistically significant differences between various socio-demographic groups, the configuration was different for WE and SMQ. Reviewing the literature, Ivey et al. (2015) found that WE and SMQ are closely related but empirically still diverse constructs. They explained this using the WE component of absorption, proposing that eliminating it would make WE and SMQ empirically closer to each other. However, our study did not support this claim because absorption was not involved, although the distinction between WE and SMQ remained. Therefore, we concluded that WE and SMQ measure different constructs despite the involvement of absorption.

6. Conclusion and further directions

Military morale is considered to be one of the key factors predicting performance in a military setting. Taking it as a combination of dedication, vigour, cynicism, and exhaustion, the concept pays attention to the personnel's well-being. At the same time, it avoids burnout (supporting the positive side of service, such as work engagement, and at the same time reducing negative effects, such as burnout). Thus, the question of how to measure and screen military morale and its development over time becomes increasingly important. The more adequately we can measure the concept, the more successful we might be in identifying the factors that could influence morale in both positive (related to work engagement) and negative (related to burnout) ways.

This article analyses the suitability of the instrument of military morale (WEBO) for the EDF as a screening and research tool. We are convinced that the approach to measuring morale through dedication and vigour (work engagement) as a positive side and cynicism and exhaustion (burnout) as a negative side of morale is working well. Of the approaches we used, this is the most informative approach, giving descriptions about both the positive and negative aspects of morale. Despite that, we do not recommend the four-factor model to measure military morale in the EDF, due to the low reliability figures for some subscales. To use it requires rewording some items, especially from the cynicism subscale. However, we recommend the modified two-factor WEBO instrument for measuring military morale across age, language, time of training (wave), profession, place of living, student status, gender, and education groups.

At the same time, our research demonstrated that the previously adopted short six-item measure of military morale (Britt and Dickinson, 2006) is closely related to work engagement, although it measures a slightly different construct. The decision to use one or the other instrument should be based on the needs of researchers–for instance, if the focus is on the positive and motivational construct, the short morale instrument should be chosen; however, if there is a desire to measure morale more broadly, including the negative side, WEBO should be chosen. Nevertheless, both are suitable to be included in the annual survey, either as an outcome or as a mediating variable.

At the same time, we recognized some limitations, which should be taken into account when reading the findings of this article. Firstly, since the instrument under investigation was tested in the Estonian military environment, its use in other

possible contexts (such as police, rescue service, etc.) would require additional investigation. Secondly, the results indicated rather moderate reliability of some subscales of work engagement and burnout, which may be caused by linguistic reasons and/or cultural aspects of perceiving item meanings differently. Thirdly, due to the limits of the scope of this paper, we did not deeply analyse why measurement invariance did not work for some socio-demographic groups, not allowing the two-factor instrument to measure military morale in educational groups.

From a psychometric point of view, the next subject for analysis could be how well all military morale instruments used in this study will work to measure collective military morale. We propose to use a similar approach as in the current study; however, as a first step, aggregating individual-level data to the squad and platoon level. In addition, since this article analysed between-group differences only from the point of view of measurement invariance, it would be necessary in the future to find precise cut-off values for high and low morale separately for each social-demographic group. This would allow an easier comparison of morale between the groups and, therefore, more precise monitoring of it.

Finally, we encourage researchers to undertake longitudinal research designs because they provide a better understanding of the dynamics of military morale over the course of service; however, taking into account the possible differences between professional, voluntary, and conscripted soldiers, and also between different organizations. A similar approach was also suggested previously, for example, by Ivey et al. (2015).

List of abbreviations

BO	Burnout
WE	Work engagement
EDF	Estonian Defence Forces
CFA	Confirmatory factor analysis
RMSEA	Root mean square error of approximation
CFI	Comparative fit index
SRMR	Standardized root mean square residual
ECVI	Expected cross-validation index
DV	Dependent variable
IV	Independent variable
WEBO	Work engagement and burnout
SMQ	Short morale questionnaire
DMQ	Direct morale question
DMQ_{sum}	Summarized direct morale question (three collective morale items and one individual level item)
DMQ_{ind}	Individual direct morale question (one individual morale item)
DMQ_{col}	Collective direct morale question (all three direct collective morale items)
IMTA	International Military Testing Association
VIF	Variance inflation factor
D	Dedication
V	Vigor
C	Cynicism
E	Exhaustion
M1, M2, M3, M4	Model 1, Model 2, Model 3, Model 4 (CFA models)
$M3_{mod}$	Modified four-factor model
$M2_{mod}$	Modified two-factor model
$\Delta\chi^2$	Change in chi-square
ΔCFI	Change in comparative fit index
ΔRMSEA	Change in root mean square error of approximation
ΔSRMR	Change in standardized root mean square residual
ε^2	Epsilon squared (effect size measure)
r^2	Squared correlation coefficient
β	Standardized regression coefficient
B	Unstandardized regression coefficient
SE B	Standard error of the regression coefficient
p	Probability value
n	Sample size
SD	Standard deviation
M	Mean

Compliance with ethical standards

Ethical considerations

Participation in this study was voluntary, and informed consent was obtained from all participants. Anonymity and confidentiality of participant data were strictly maintained throughout the research process.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Anzai E, Douglas C, and Bonner A (2014). Nursing practice environment, quality of care, and morale of hospital nurses in Japan. Nursing and Health Sciences, 16(2): 171-178. https://doi.org/10.1111/nhs.12081 **PMid:23855754**

Bakker AB, Schaufeli WB, Leiter MP, and Taris TW (2008). Work engagement: An emerging concept in occupational health psychology. Work and Stress, 22(3): 187-200. https://doi.org/10.1080/02678370802393649

Bal PM and De Lange AH (2015). From flexibility human resource management to employee engagement and perceived job performance across the lifespan: A multisample study. Journal of Occupational and Organizational Psychology, 88(1): 126-154. https://doi.org/10.1111/joop.12082

Bianchi R and Brisson R (2019). Burnout and depression: Causal attributions and construct overlap. Journal of Health Psychology, 24(11): 1574-1580. https://doi.org/10.1177/1359105317740415 **PMid:29139312**

Bliese PD and Britt TW (2001). Social support, group consensus and stressor–strain relationships: Social context matters. Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 22(4): 425-436. https://doi.org/10.1002/job.95

Britt TW and Dickinson JM (2006). Morale during military operations: A positive psychology approach. In: Britt TW, Castro CA, and Adler AB (Eds.), Military life: The psychology of serving in peace and combat: Military performance: 157–184. Praeger Security International, Westport, USA.

Britt TW, Adler AB, Bliese PD, and Moore D (2013). Morale as a moderator of the combat exposure-PTSD symptom

relationship. Journal of Traumatic Stress, 26(1): 94-101. https://doi.org/10.1002/jts.21775 **PMid:23371305**

Britt TW, Dickinson JM, Moore D, Castro CA, and Adler AB (2007). Correlates and consequences of morale versus depression under stressful conditions. Journal of Occupational Health Psychology, 12(1): 34-47. https://doi.org/10.1037/1076-8998.12.1.34 **PMid:17257065**

Browne MW and Cudeck R (1993). Alternative ways of assessing model fit. In: Bollen KA and Long JS (Eds.), Testing structural equation models: 136-162. SAGE, Newbury Park, USA.

Chandra S, Wright SM, Ghazarian S, Kargul GM, and Howell EE (2016). Introducing the hospitalist morale index: A new tool that may be relevant for improving provider retention. Journal of Hospital Medicine, 11(6): 425-431. https://doi.org/10.1002/jhm.2543 **PMid:26969890**

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling: A Multidisciplinary Journal, 14(3): 464-504. https://doi.org/10.1080/10705510701301834

Cheung GW and Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2): 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Corbeanu A, Iliescu D, Ion A, and Spînu R (2023). The link between burnout and job performance: A meta-analysis. European Journal of Work and Organizational Psychology, 32(4): 599-616. https://doi.org/10.1080/1359432X.2023.2209320

De Beer LT, van der Vaart L, Escaffi-Schwarz M, De Witte H, and Schaufeli WB (2024). Maslach burnout inventory—general survey: A systematic review and meta-analysis of measurement properties. European Journal of Psychological Assessment, 40(5): 360–375. https://doi.org/10.1027/1015-5759/a000797

Demerouti E and Cropanzano R (2010). From thought to action: Employee work engagement and job performance. In: Bakker AB and Leiter MP (Eds.), Work engagement: A handbook of essential theory and research: 147-163. Psychology Press, London, UK.

Dyches KD, Saboe KN, Anderson JA, Wilk JE, Hinman SJ, Sipos ML, and Quartana PJ (2017). Modeling the indirect association of combat exposure with anger and aggression during combat deployment: The moderating role of perceived unit morale. Military Psychology, 29(4): 260-270. https://doi.org/10.1037/mil0000165

Edú-Valsania S, Laguía A, and Moriano JA (2022). Burnout: A review of theory and measurement. International Journal of Environmental Research and Public Health, 19(3): 1780. https://doi.org/10.3390/ijerph19031780 **PMid:35162802 PMCid:PMC8834764**

Fazal TM, Sumner JL, Korona-Bailey J, and Koehlmoos TP (2024). Military medicine and morale: Perceptions of inequities in triage impact national security. Journal of Health and Social Sciences, 9(3): 432-456.

Fennell J (2014). In search of the 'X' factor: Morale and the study of strategy. Journal of Strategic Studies, 37(6-7): 799-828. https://doi.org/10.1080/01402390.2013.846856

Field AP (2018). Discovering statistics using IBM SPSS statistics. 5th Edition, SAGE, Newbury Park, USA.

Frone MR and Blais AR (2019). Work fatigue in a non-deployed military setting: Assessment, prevalence, predictors, and outcomes. International Journal of Environmental Research and Public Health, 16: 2892. https://doi.org/10.3390/ijerph16162892 **PMid:31412599 PMCid:PMC6721391**

Furr RM and Bacharach VR (2013). Psychometrics: An introduction. 2nd Edition, SAGE, Thousand Oaks, USA.

Garg K, Dar IA, and Mishra M (2018). Job satisfaction and work engagement: A study using private sector bank managers. Advances in Developing Human Resources, 20(1): 58-71. https://doi.org/10.1177/1523422317742987

González-Romá V, Schaufeli WB, Bakker AB, and Lloret S (2006). Burnout and work engagement: Independent factors or opposite poles? Journal of Vocational Behavior, 68(1): 165-174. https://doi.org/10.1016/j.jvb.2005.01.003

Gorgievski MJ, Bakker AB, and Schaufeli WB (2010). Work engagement and workaholism: Comparing the self-employed and salaried employees. The Journal of Positive Psychology, 5(1): 83-96. https://doi.org/10.1080/17439760903509606

Hair JF Jr, Hult GTM, Ringle CM, and Sarstedt M (2022). A primer on partial least squares structural equation modeling (PLS-SEM). 3rd Edition, SAGE, Newcastle upon Tyne, UK. https://doi.org/10.1007/978-3-030-80519-7

Hardy B (2010). Morale: Definitions, dimensions and measurement. Ph.D. Dissertation, Apollo-University of Cambridge Repository, Cambridge, UK.

Hu LT and Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1): 1-55. https://doi.org/10.1080/10705519909540118

Huhtala M and Feldt T (2016). The path from ethical organisational culture to employee commitment: Mediating roles of value congruence and work engagement. Scandinavian Journal of Work and Organizational Psychology, 1(1): 3. https://doi.org/10.16993/sjwop.6

Ivey GW, Blanc JRS, and Mantler J (2015). An assessment of the overlap between morale and work engagement in a nonoperational military sample. Journal of Occupational Health Psychology, 20(3): 338-347. https://doi.org/10.1037/a0038559 **PMid:25528688**

Jeppesen LS and Elrond AF (2021). Exploring morale in an active warzone: A study of the predictors of morale during deployment to Afghanistan. Scandinavian Journal of Military Studies, 4(1): 267-279. https://doi.org/10.31374/sjms.100

Johnsrud LK and Rosser VJ (2002). Faculty members' morale and their intention to leave: A multilevel explanation. The Journal of Higher Education, 73(4): 518-542. https://doi.org/10.1353/jhe.2002.0039

Jöreskog KG and Sörbom D (1993). LISREL 8: Structural equation modeling with the SIMPLIS command language. Scientific Software International, Chapel Hill, USA.

Kasemaa A (2024). Psychometric properties of the Estonian version of the six-item military morale instrument. Journal on Baltic Security, 10(1): 1-25.

Kasemaa A and Säälik Ü (2021). Personality traits and select socio-demographic variables as predictors of military morale: Longitudinal research in the Estonian defence forces. Journal of Military Studies, 10(1): 22-37. https://doi.org/10.2478/jms-2021-0003

Kasemaa A and Säälik Ü (2023). Personality predicting military morale and the role of positive and negative affectivity. Current Issues in Personality Psychology, 11(4): 269-282. https://doi.org/10.5114/cipp/156046 **PMid:38075467 PMCid:PMC10699293**

Kras KR, Meyer KS, Dmello JR, and Rudes DS (2024). Individual and organizational correlates of morale among workers in a youth and family court agency. Crime and Delinquency, 70(2): 491-518. https://doi.org/10.1177/00111287221115646

Langelaan S, Bakker AB, Van Doornen LJ, and Schaufeli WB (2006). Burnout and work engagement: Do individual differences make a difference? Personality and Individual Differences, 40(3): 521-532. https://doi.org/10.1016/j.paid.2005.07.009

Lesener T, Gusy B, Jochmann A, and Wolter C (2020). The drivers of work engagement: A meta-analytic review of longitudinal evidence. Work and Stress, 34(3): 259-278. https://doi.org/10.1080/02678373.2019.1686440

Maguen S and Litz BT (2006). Predictors of morale in US peacekeepers. Journal of Applied Social Psychology, 36(4): 820-836. https://doi.org/10.1111/j.0021-9029.2006.00045.x

Manning FJ (1991). Morale, cohesion, and esprit de corps. In: Gal R and Mangelsdorff AD (Eds.), Handbook of military psychology: 453–470. John Wiley and Sons, Hoboken, USA.

Manning FJ (1994). Morale and cohesion in military psychiatry. In: Jones FD, Sparacino LR, Wilcox VL, and Rothberg JM (Eds.), Military psychiatry: Preparing in peace for war: 1-18. TMM Publications, Washington, USA.

Marsh HW, Hau KT, and Wen Z (2004). 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. Structural Equation Modeling, 11(3): 320-341. https://doi.org/10.1207/s15328007sem1103_2

Maslach C and Leiter MP (1997). The truth about burnout: How organizations cause personal stress and what to do about it. Jossey-Bass, San Francisco, USA.

Maslach C, Schaufeli WB, and Leiter MP (2001). Job burnout. Annual Review of Psychology, 52: 397-422. https://doi.org/10.1146/annurev.psych.52.1.397 **PMid:11148311**

Meyer MJ, Macmillan RB, and Northfield S (2009). Principal succession and its impact on teacher morale. International Journal of Leadership in Education, 12(2): 171-185. https://doi.org/10.1080/13603120802449660

Michaud K, Barbeau-Julien K, and Slinger M (2025). Transformational leadership, well-being, morale, and readiness: The mediating role of empowerment. Military Psychology, 37(2): 117-126. https://doi.org/10.1080/08995605.2024.2319522 **PMid:38386695 PMCid:PMC11849952**

Motowidlo SJ and Borman WC (1977). Behaviorally anchored scales for measuring morale in military units. Journal of Applied Psychology, 62(2): 177-183. https://doi.org/10.1037//0021-9010.62.2.177

Motowidlo SJ and Borman WC (1978). Relationships between military morale, motivation, satisfaction, and unit effectiveness. Journal of Applied Psychology, 63(1): 47-52. https://doi.org/10.1037//0021-9010.63.1.47

Peterson C, Park N, and Sweeney PJ (2008). Group well-being: Morale from a positive psychology perspective. Applied Psychology, 57: 19-36. https://doi.org/10.1111/j.1464-0597.2008.00352.x

Putnick DL and Bornstein MH (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review, 41: 71-90. https://doi.org/10.1016/j.dr.2016.06.004 **PMid:27942093 PMCid:PMC5145197**

Roelen CA, van Hoffen MF, Groothoff JW, De Bruin J, Schaufeli WB, and van Rhenen W (2015). Can the Maslach burnout inventory and Utrecht work engagement scale be used to screen for risk of long-term sickness absence? International Archives of Occupational and Environmental Health, 88: 467-475. https://doi.org/10.1007/s00420-014-0981-2 **PMid:25212752**

Satorra A and Bentler PM (2001). A scaled difference chi-square test statistic for moment structure analysis. Psychometrika, 66(4): 507-514. https://doi.org/10.1007/BF02296192

Schaufeli WB and Bakker AB (2010). Defining and measuring work engagement: Bringing clarity to the concept. In: Bakker AB & Leiter MP (Eds.), Work engagement: A handbook of essential theory and research: 10–24. Psychology Press, New York, USA.

Schaufeli WB, Salanova M, González-Romá V, and Bakker AB (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. Journal of Happiness Studies, 3: 71-92. https://doi.org/10.1023/A:1015630930326

Schaufeli WB, Taris TW, and Van Rhenen W (2008). Workaholism, burnout, and work engagement: Three of a kind or three different kinds of employee well-being? Applied Psychology, 57(2): 173-203. https://doi.org/10.1111/j.1464-0597.2007.00285.x

Sekhar C, Patwardhan M, and Vyas V (2017). Linking work engagement to job performance through flexible human resource management. Advances in Developing Human Resources, 20(1): 72-87. https://doi.org/10.1177/1523422317743250

Shamir B, Brainin E, Zakay E, and Popper M (2000). Perceived combat readiness as collective efficacy: Individual-and group-level analysis. Military Psychology, 12(2): 105-119. https://doi.org/10.1207/S15327876MP1202_2

Stevens J (2002). Applied multivariate statistics for the social sciences. 4th Edition, Lawrence Erlbaum, Mahwah, USA.

Tabachnick BG and Fidell LS (2007). Using multivariate statistics. 5th Edition, Pearson Allyn and Bacon, Boston, USA.

Taris TW, Ybema JF, and van Beek I (2017). Burnout and engagement: Identical twins or just close relatives? Burnout Research, 5: 3-11. https://doi.org/10.1016/j.burn.2017.05.002

van Boxmeer F, Verwijs C, de Bruin R, Duel J, and Euwema M (2007). A direct measure of morale in the Netherlands' armed forces morale survey: Theoretical puzzle, empirical testing and validation. In the Proceedings of the 49th Annual Conference of the International Military Testing Association, International Military Testing Association, Gold Coast, Australia.

Vander Elst T, Cavents C, Daneels K, Johannik K, Baillien E, Van den Broeck A, and Godderis L (2016). Job demands–resources predicting burnout and work engagement among Belgian home health care nurses: A cross-sectional study. Nursing Outlook, 64(6): 542-556. https://doi.org/10.1016/j.outlook.2016.06.004 **PMid:27427405**

Whitesell AA and Owens GP (2012). The impact of patriotism, morale, and unit cohesion on mental health in Veterans of Iraq and Afghanistan. Traumatology, 18(1): 1-7. https://doi.org/10.1177/1534765610395625