



OPEN

## Psychometric properties of the Bern illegitimate tasks scale using classical test and item response theories

Beata Aleksandra Basinska<sup>1✉</sup> & Anna Maria Dåderman<sup>2</sup>

Combining a classical test theory and an item response theory (IRT), this study aimed to investigate the psychometric properties of the Bern Illegitimate Tasks Scale (BITS) by measuring two conceptually separate dimensions capturing unnecessary tasks (perceived by employees as pointless) and unreasonable tasks (perceived as unfairly or inappropriately assigned). Data collected among Polish employees in two samples ( $N=965$  and  $N=803$ ) were analysed. Results from the classical test theory (parallel analysis, exploratory and confirmatory factor analyses) indicated two correlated factors with four items each, confirming the theory of illegitimate tasks. This study is the first to report item and scale functioning using IRT analysis of each of the two dimensions of BITS. All items on each dimension had acceptable discrimination and difficulty parameters. Moreover, items had measurement invariance between men and women. All levels of unnecessary and unreasonable tasks were reliably captured by BITS items. Convergent and discriminant validities of both dimensions of BITS were confirmed in relation to work overload, work performance and occupational wellbeing. We conclude that BITS, in the case of the Polish version, is psychometrically suitable to use with the working population.

Contemporary organisations operate in a competitive, uncertain and volatile environment<sup>1</sup>. Employees working in such organisations are often asked to perform duties (illegitimate tasks) that are not in agreement with their professional identity<sup>2,3</sup>. They usually report that tasks are illegitimate due to either employees feeling that a task should not have to be carried out by anybody, or that a task could have been avoided with better organizational systems<sup>4</sup> (unnecessary tasks); or tasks being unfairly assigned to them, that is, “tasks that may be outside the range of an employee’s occupation or status, or professional role”<sup>4</sup> (p. 73) (unreasonable tasks). In the subjective view of employees (that reflect normative prescriptions of what is expected, or seen as legitimate, in a given role or position), unnecessary tasks are perceived as pointless, whereas unreasonable ones are seen as inappropriate, given employees’ competencies or professional resources. Illegitimate tasks lead to greater levels of administrative work and reduce the time spent on core responsibilities. For instance, rather than focusing on providing medical care to patients, teaching students, producing more research or facilitating citizens’ daily lives, employees may be asked to prepare double documentation (i.e., both on paper and in electronic form), numerous reports, and summaries<sup>5,6</sup>. Illegitimate tasks have been recognised as an important social stressor that threatens the self and violates one’s professional identity, according to Stress-as-Offence-to-Self (SOS) theory<sup>7,8</sup>. SOS theory assumes that maintaining a positive self-view is a primary human concern, and any threat to self-esteem elicits strain<sup>7</sup>. Illegitimate tasks cause stress because employees feel they are not valued and respected when they receive these tasks<sup>3,8</sup>. Thus, recipients feel that either no one should be assigned such tasks because they are pointless or they are left feeling “I should not have to do this! It is not my job!”. Illegitimate tasks can also be understood as job demands due to their energy-related features because they require effort and drain employees’ physical and psychological wellbeing<sup>9</sup>. Therefore, as illegitimate tasks mount over time, employees may lose their professional identity and engagement, which can cause further negative consequences for individuals and organisations<sup>2,10–12</sup>. What is needed, then, is a valid and effective tool that will respond to challenges which appear at rapid speed. The Bern Illegitimate Tasks Scale (BITS) is the most frequently used instrument to evaluate unnecessary and unreasonable tasks<sup>4,13</sup>. To date, nevertheless, few psychometric property studies have been conducted<sup>13–15</sup>. The

<sup>1</sup>Faculty of Management and Economics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland. <sup>2</sup>Department of Social and Behavioral Studies, University West, Trollhättan, Sweden. ✉email: Beata.Basinska@pg.edu.pl

contribution of the current study is its simultaneous application of a classical test theory and an item response theory (IRT) to assess the validity of BITS. IRT is a family of associated mathematical models that relate latent traits (e.g., perceived illegitimate tasks) to the probability of responses to items in an assessment. Thus, IRT is used to evaluate the item performance and precision of the scale measuring a latent trait<sup>16</sup>. It is the first study to report item functioning of BITS using IRT consisting of two conceptually separate dimensions capturing unnecessary and unreasonable tasks.

**Illegitimate tasks in the workplace.** In line with SOS theory<sup>7,8</sup>, illegitimate tasks decrease employees' self-esteem<sup>2,11,17,18</sup>, as they may imply a lack of appreciation for their professional role<sup>5,19</sup>. Therefore, illegitimate tasks threaten one's personal and social self, and professional identity<sup>7,8</sup>. In particular, intrinsically motivated employees suffered more from illegitimate tasks when they were not recognised for their job efforts<sup>14</sup>. Illegitimate tasks contribute to aversive and counterproductive behaviours<sup>4,20</sup> or poorer job performance<sup>10,21</sup>. Illegitimate tasks can be accompanied by other job demands (e.g., work overload) which threaten both the wellbeing and health of employees<sup>22</sup>. They are also related to a higher level of burnout<sup>23,24</sup>, lower job satisfaction<sup>12,25,26</sup>, work engagement and meaning of work<sup>27</sup>, and higher turnover intentions<sup>28,29</sup>.

Significantly, illegitimate tasks have been identified as a severe problem across various workplaces, including the healthcare sector<sup>6,15,30,31</sup>, higher education<sup>32</sup>, IT professionals<sup>33</sup>, teachers<sup>5</sup>, engineers<sup>17</sup>, administrative staff<sup>11</sup>, blue-collar workers<sup>34</sup>, and also Red Cross volunteers<sup>29</sup>.

In terms of other effects, illegitimate tasks may also be transferred to the domain outside work. Due to this, illegitimate tasks have been shown to be associated with work-family conflict<sup>13,24</sup>, which may be exacerbated by psychological detachment<sup>35</sup>. Moreover, illegitimate tasks are related to biological stress indicators, such as worse sleep quality<sup>36</sup> and a higher level of cortisol, especially among male employees who evaluated their health as poor<sup>37</sup>.

Illegitimate tasks constitute two distinct facets: unnecessary and unreasonable tasks<sup>2,4</sup>. For example, unnecessary tasks are frequently observed<sup>30</sup> in the healthcare sector. However, unreasonable tasks, such as tasks outside one's occupational role, are also widely reported<sup>6</sup>. Unnecessary tasks are related to job dissatisfaction<sup>38</sup>, while unreasonable tasks consume time and energy<sup>30,39</sup>. Furthermore, unnecessary tasks result in a fluctuation of negative affect within individuals, while unreasonable tasks lead to negative affect at the between-person level<sup>18</sup>. Unnecessary tasks predict a higher level of distress, while unreasonable tasks cause a broader spectrum of mental health issues, such as distress, anxiety, and depression<sup>40</sup>.

Concluding, illegitimate tasks are commonplace at work. They may negatively affect the effectiveness of an organisation. They may also harm employee wellbeing and organisational behaviour and even encroach on other areas of employees' lives.

**Bern Illegitimate tasks scale.** Illegitimate tasks are usually measured with BITS, which consists of two conceptually separate dimensions capturing unnecessary and unreasonable tasks. Initially, BITS comprised nine items that were classified into two dimensions<sup>13</sup>. Later, however, their number was reduced to eight<sup>4</sup>. The two dimensions are moderately correlated<sup>2,13,29</sup>.

The popularity of BITS is underlined by the fact that it has been translated into several languages, including English<sup>11</sup>, Swedish<sup>10,41</sup>, Norwegian<sup>39</sup>, Finnish<sup>6,27</sup>, Turkish<sup>42</sup>, Indian<sup>43</sup>, Chinese<sup>21,35</sup> and Spanish<sup>15</sup>. Existing studies have mostly focused on the scale's convergent and discriminant validity. Moreover, most studies have shown that BITS has good internal consistency with regard to both subscales (unnecessary and unreasonable tasks)<sup>15</sup>.

However, few studies have demonstrated the psychometric properties of BITS in detail based on classical test theory. Using confirmatory factor analysis (CFA), Jacobshagen<sup>13</sup>, Muntz and Dormann<sup>14</sup>, and Portilla et al.<sup>15</sup> observed two moderately correlated subscales. Even though some studies on the construct validity of BITS have conceived illegitimate tasks as having only a single construct<sup>26,27</sup>, they concordantly emphasise that this scale captures two substantially different types of illegitimate tasks.

As mentioned above, the original version of BITS consisted of nine items<sup>13</sup>. Some data indicated that Item 4: 'Do you have work tasks to take care of which keep you wondering if they would not exist (or could be done with less effort), if some other people made less mistakes?' loaded on both factors<sup>13</sup> (see Table 9, p. 75). Thus, Jacobshagen suggested that Item 4 should be omitted, and Semmer et al.<sup>4</sup> started applying the eight-item version of BITS. Unfortunately, it was not emphasised enough that Item 4 had been removed, and a sole description of four items for both unnecessary and unreasonable tasks was revealed. Consequently, this change was confusing to researchers, who have been applying different versions since. The nine-item BITS is still used<sup>25,32</sup>. Moreover, a few reports of factor analysis revealed that occasionally Item 5 was removed<sup>25,44</sup>. In addition, BITS was reduced to a few items<sup>45</sup> or a single-item scale for unnecessary and unreasonable tasks<sup>23</sup>. Furthermore, we have found different response formats in the existing studies. In addition to the original five-point response format scale, we have also observed a four- and seven-point scale ranging from 'strongly disagree' to 'strongly agree'<sup>17,20,23,42</sup>.

**The current study.** In the 16-year history of BITS, there have been too few solid reports on the psychometric properties of the scale. It seems necessary to test if the eight items of BITS sufficiently measure the unidimensionality of the construct or two separate dimensions. Moreover, no published studies have investigated the item and scale functioning of BITS using IRT. Therefore, to date, our study is the first to report the item and scale functioning of this instrument using IRT.

An overarching purpose of the current study was to evaluate the factor structure of BITS, to determine whether the items and dimensions of BITS function well, and to validate the Polish version of BITS as an important job-related stressor.

We have formulated the following research questions:

RQ1 What is the structure of BITS, i.e., how many factors does it represent and is it homogeneous (unidimensional)?

RQ2 How effectively does each item discriminate between employees with different levels of illegitimate tasks (i.e., unnecessary tasks and unreasonable tasks), as well as do the items adequately measure illegitimate tasks across low to high levels?

RQ3 What is the construct validity, namely, convergent and discriminant validity, of the Polish version of BITS?

**Methods.** To assess the psychometric properties of BITS we used two heterogeneous samples of employees. In sample 1, first we examined the factor structure of BITS and then assessed item discrimination and local dependence. In sample 2, we replicated the construct validity and examined the convergent and discriminant validity, as well as the reliability of BITS.

**Participants and Procedure.** *Sample 1.* To conduct a two-parameter logistic model (2PLM) IRT test with a graded response model (GRM)<sup>46</sup>, the recommended minimum sample size is 500 to 750 participants for 10–20 items<sup>47</sup>. In 2019, 966 fully completed protocols were received. One outlier by having a high value of the Mahalanobis distance ( $p < 0.001$ ) was detected. Thus, the sample consisted of 965 employees at Polish organisations in the education, public administration and IT sectors. Respondents (65% women) were aged between 21 and 68 years with a mean age of 42.4 years ( $SD = 10.2$  years), and they had job tenures of 1 to 40 years. Most had high levels of education (bachelor's or master's degree).

*Sample 2.* In 2020, 821 fully completed protocols were obtained (response rate 68%). 18 multivariate outliers were excluded. Thus, the final sample consisted of 803 employees at different Polish organisations in the healthcare, education, public administration and IT sectors. Participants (59% women) were aged between 23 and 66 years with a mean age of 43 years ( $SD = 10.2$  years). They had an average job tenure of 19.7 years ( $SD = 10.8$  years, range 2–45) and an average job tenure in their current organisation of 12.9 years ( $SD = 10.1$ , range 0.5–41).

**Ethics statement.** The study was performed in line with the principles of the Declaration of Helsinki. All data were collected using anonymous online surveys. All participants were informed of the nature of the current study and gave their informed consent to participate. According to national recommendations, issued by the Psychological Committee of the Polish Science Academy (21st June 2013) and the National Science Centre, ethics approval to conduct this study was not required. More details about the study procedure are presented in the Supplementary material.

**Measures.** We took the original Bern Illegitimate Tasks Scale (BITS)<sup>13</sup> instrument including nine items (all of which are positively worded) in order to replicate previous findings from past research, and we clarified how each item functioned. Unnecessary tasks are assessed on BITS with five items. Each item is introduced with the lead-in question, 'Do you have work tasks to take care of which keep you wondering if...,' followed by phrases such as '...they have to be done at all?' There are four items assessing unreasonable tasks, and each of these is introduced with the lead-in question, 'Do you have work tasks to take care of which you believe...,' followed by phrases such as '...should be done by someone else?' Responses are marked on a five-point Likert scale ranging from 1 (never) to 5 (frequently). Higher scores indicate that employees are more often required to do unnecessary and unreasonable tasks.

A six-stage translation and adaptation process was used to adapt BITS from English to Polish<sup>48</sup>. This process included forward translation into Polish by two different translation agencies, synthesis, back translation from Polish to English, harmonisation, cognitive interviews, revisions, and pilot data sampling. The Polish version can be found in the Supplementary material.

To assess the convergent and discriminant validity of BITS (sample 2), we applied related constructs such as job demand (work overload) and outcomes (work performance and occupational wellbeing, i.e., job burnout and job satisfaction). Instruments to evaluate these constructs are reported in Table 1.

**Statistical analyses.** *Preliminary analyses.* Scores on the total BITS and its two dimensions were calculated according to the manual<sup>13</sup>. Using sample 1, the exploratory factor analysis (EFA; the principal axis factoring extraction method with oblimin rotation) and parallel analysis (PA) identified two factors. Each item belongs to

| Constructs and citation               | Item | Sample items   | Response format                                     |
|---------------------------------------|------|--|---|
| Work overload <sup>49</sup>           | 4    | Do you have too much work to do?                               | 5-point (1 = never, 5 = always)                     |
| Burnout Assessment tool <sup>50</sup> | 23   | I want to be active at work, but somehow I am unable to manage | 5-point (1 = never, 5 = always)                     |
| Job satisfaction <sup>51</sup>        | 1    | To what extent are you satisfied with your job?"               | 5-point (1 = very dissatisfied, 5 = very satisfied) |
| Work performance <sup>49</sup>        | 1    | How would you rate your current job performance?               | 11 point (from 0 to 10, the higher, the better)     |

**Table 1.** Study instruments: item number, sample item, response format and citation.

one of the two factors. The factor loadings for the items were above 0.68, with the exception of Item 4 (0.44) (see Supplementary Table S1). Indeed, the content of this item refers to both unnecessary and unreasonable tasks. IRT for the original nine-item BITS was also carried out and, following Baker's<sup>16</sup> cut-offs, almost all items on the unnecessary tasks dimension were designated very high, whereas Item 4 was moderate in terms of its precision to discriminate between respondents with different levels of illegitimate tasks. Thus, Item 4 was relatively different. Considering jointly the preliminary analyses, Item 4 was removed from further analyses. Consequently, two subscales containing four items each were assigned to the analyses (unnecessary tasks: Items 1–3, 5; unreasonable tasks: Items 6–9).

**Classical test theory analyses.** To ensure that illegitimate tasks measured by items of BITS met the IRT assumption of unidimensionality, EFA and PA<sup>52</sup> were conducted. Using EFA, the most commonly used indicators that support the unidimensionality assumption include having at least 20% of the variance on the first factor or having a ratio between the eigenvalues of the first factor compared to the next factor that is greater than 4<sup>53</sup>.

Factor validity was also assessed with CFA based on a maximum likelihood estimation method in sample 1, and we replicated these findings in sample 2. The assumptions of data being continuous with multivariate normal distribution were met. The goodness-of-fit of the CFA model was assessed by the chi-squared test ( $\chi^2$ ), comparative fit index (CFI), root mean square error of approximation (RMSEA), standardised root mean square residual (SRMR) and Bayesian information criterion (BIC). Values of CFI > 0.90<sup>54</sup>, RMSEA and SRMR < 0.08, and a lower BIC<sup>55</sup> indicate an acceptable goodness-of-fit. Measurement reliability was established using two coefficients, which were Cronbach's alpha and composite reliability (CR) with the value > 0.70<sup>56</sup>.

Using sample 2, convergent validity was measured in terms of average variance extracted (AVE) with the value > 0.50 for each latent variable<sup>57</sup>. Discriminant validity was assessed by examining whether AVE was higher than the squared coefficients of the bivariate correlation between work overload, job burnout, job satisfaction and work performance<sup>57</sup>. Construct validity was assessed by running multiple linear regressions (enter method). SPSS and AMOS softwares were used to conduct the statistical analyses.

**Item response theory analyses.** Before applying IRT models with GRM, we evaluated the basic assumption of the model's unidimensionality, monotonicity, and item local independence. The 2PLM IRT analyses were conducted using IRTPRO. Given BITS' five response categories, a GRM was selected<sup>46</sup>. A key assumption of the GRM is that scores for each item are ordered consistently across items, such that the lowest response category (labelled 'never') is indicative of the lowest level of  $\theta$  (where  $\theta$  indicates different levels of the illegitimate tasks, measured by BITS). The GRM estimates a common item discrimination parameter for each item ( $a$ ) and also estimates the location or difficulty/severity threshold parameters ( $b_i$ ) for each response category within the item<sup>58</sup>. Item discrimination describes how well an item can differentiate between examinees at different trait levels. Each item of BITS has an average discrimination parameter  $a$  across five response categories. A higher  $a$  value indicates that an item discriminates more precisely between respondents with different levels of  $\theta$ . Each item was evaluated in terms of its function relative to the other items, which was assessed by both its scale score and its contribution to the overall information gathered from each scale score<sup>59</sup>. Item difficulty is known as "location" on the difficulty range (the location on the latent trait) and describes where the item functions best along the trait scale, namely, lower values will be expected to be endorsed at lower trait levels. Difficulty threshold parameters  $b_i$  are expressed as the standard deviation away from the mean of the value 0, and have a common range of  $-3$  to  $+3$ <sup>59</sup>. Both discrimination and difficulty threshold parameters have been represented graphically as item characteristic curves (ICCs). In addition, measurement invariance of the scale between men and women was estimated by differential item functioning (DIF). Gender differences of the discrimination and difficulty parameters for items were examined using chi-squared tests.

These two approaches, i.e., classical test theory modelling with CFA and modelling using IRT, reinforce each other. CFA is confirmative and able to test the conceptual model of BITS, while IRT produces sample-independent estimates of parameters.

## Results

**Factor structure of BITS.** The total BITS (8-item) failed to sufficiently meet the assumptions of unidimensionality based on the ratio between the eigenvalues of the first factor and the next factor being greater than 4<sup>53</sup>. The results of the PA suggested two factors (see Supplementary Table S2).

The extraction of two factors using EFA showed that none of the items cross-loaded on the second factor with a value greater than 0.30 (see Supplementary Table S3). Therefore, two factors were retained as two conceptually-derived dimensions (four items each), which each showed good reliability (Table 2) and supported unidimensionality. Table 2 also presents descriptive statistics of the Polish version of BITS.

We also assessed factor validity for BITS using CFA. We analysed two models: a single-factor model and a model of two correlated latent factors (Table 3). The findings revealed that the modified model with the two correlated latent factors, including two pairs of error correlation (Item 1 and Item 2, as well as Item 8 and Item 9), yielded an excellent fit. Those factors were moderately correlated (0.66) (see Supplementary Fig. S1A). This result confirmed the result from the PA, indicating that the eight-item BITS, applied on the current sample, comprises two separate dimensions, and therefore separate analyses were performed on both IRT models, and validation analyses.

**Item functioning.** Table 4 shows that discrimination parameter  $a$  ranged from 1.96 to 4.93 for items relating to unnecessary tasks and from 2.42 to 3.27 for those relating to unreasonable tasks. Following Baker's<sup>16</sup>

| Composite scale score | No items | Item     | <i>M</i> / <i>SD</i> | Sk    | K     | $\alpha/M_{ic}$ | Correlation with Illegitimate tasks | % variance first factor; second factor | First eigenvalue /Second eigenvalue = Ratio | UD  | Factor correlation (EFAs) |
|-----------------------|----------|----------|----------------------|-------|-------|-----------------|-------------------------------------|--|---|-----|---------------------------|
| Illegitimate tasks    | 8        | 1–3, 5–9 | 2.93/0.75            | 0.01  | –0.12 | 0.89/0.50       | –                                   | 57;16                                  | 4.53/1.26 = 3.60                            |     | 0.63                      |
| Unnecessary tasks     | 4        | 1–3, 5   | 3.10/0.84            | –0.04 | –0.14 | 0.87/0.64       | 0.88                                | 73;13                                  | 2.91/0.50 = 5.82                            | Yes | 1-factor                  |
| Unreasonable tasks    | 4        | 6–9      | 2.75/0.86            | 0.18  | –0.24 | 0.87/0.62       | 0.89                                | 72;12                                  | 2.87/0.50 = 5.74                            | Yes | 1-factor                  |

**Table 2.** Descriptive statistics of the Polish version of the BITS. Sample 1 ( $N=965$ ). Sk = Skewness. K = Kurtosis.  $M_{ic}$  = Mean inter-item correlation. UD = Unidimensionality according to Reeve et al.<sup>54</sup>.

| Model                                    | $\chi^2$ (df) | CFI   | SRMR  | RMSEA (90%CI)        | BIC      | $\Delta\chi^2$ | $\Delta df$ |
|--|---------------|-------|-------|----------------------|----------|----------------|-------------|
| Sample 1 $N=965$                         |               |       |       |                      |          |                |             |
| M1: single factor model                  | 1117.553 (20) | 0.750 | 0.106 | 0.239 (0.227, 0.251) | 1227.507 | –              | –           |
| M2: correlated two factor model          | 179.644 (19)  | 0.963 | 0.044 | 0.094 (0.810, 0.106) | 179.644  | 937.909***     | 1           |
| M3: modified correlated two factor model | 43.404 (17)   | 0.994 | 0.021 | 0.040 (0.025, 0.055) | 173.974  | 136.240***     | 2           |
| Sample 2 $N=803$                         |               |       |       |                      |          |                |             |
| M1: single factor model                  | 1104.742 (20) | 0.725 | 0.112 | 0.260 (0.247, 0.273) | 1211.756 | –              | –           |
| M2: correlated two factor model          | 248.706 (19)  | 0.942 | 0.066 | 0.123 (0.109, 0.137) | 362.408  | 849.348***     | 1           |
| M3: modified correlated two factor model | 34.225 (17)   | 0.996 | 0.019 | 0.036 (0.018, 0.053) | 161.304  | 201.104***     | 2           |

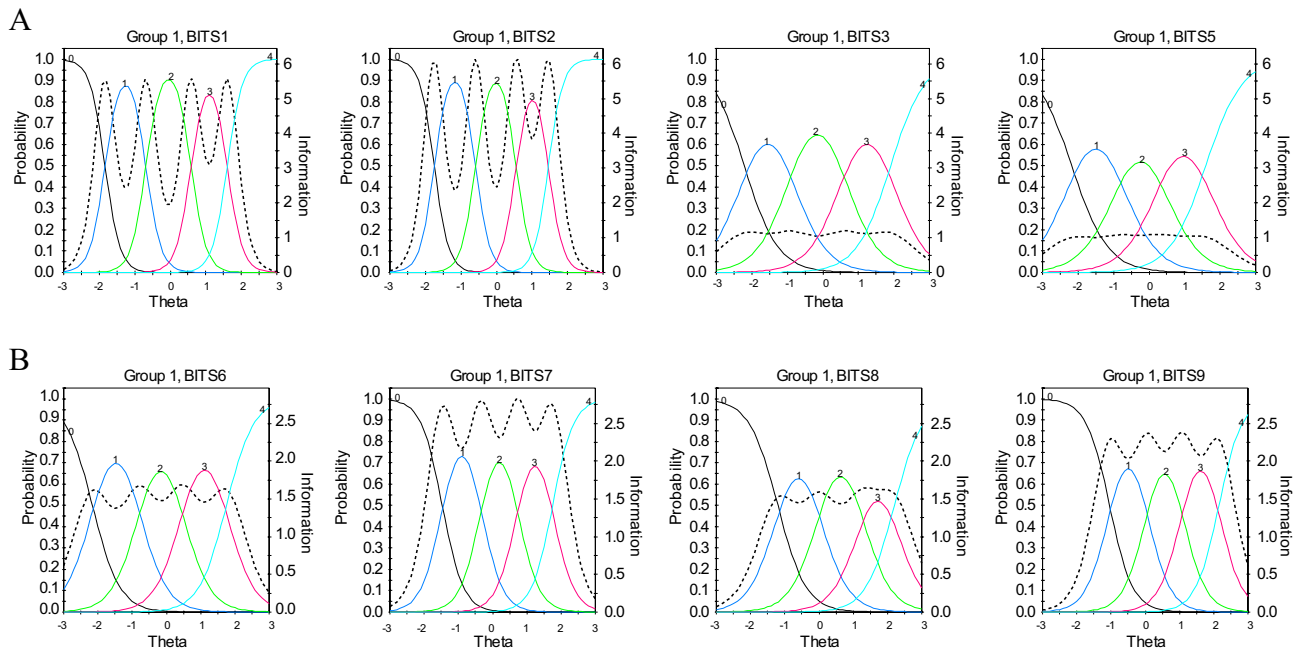
**Table 3.** Confirmatory factor analysis of the single—and two- factor model of BITS. CFI = confirmatory fit index, SRMR = standardized root mean square residual, RMSEA = root mean square error of approximation, 90%CI = 90% confidence interval, BIC = Bayesian information criterion. \*\*\*  $p < 0.001$ .

| Item  | <i>a</i> / <i>s.e</i> | Quality of <i>a</i> | $b_1$ | $b_2$ | $b_3$ | $b_4$ | <i>p</i> for DIF* |       |
|---|-----------------------|---------------------|-------|-------|-------|-------|-------------------|-------|
|   |                       |                     |       |       |       |       | $a_i$             | $b_i$ |
| Unnecessary tasks   |                       |                     |       |       |       |       |                   |       |
| Do you have work tasks to take care of which keep you wondering if ...                          |                       |                     |       |       |       |       |                   |       |
| 2...they make sense at all?   | 4.93/0.41             | Very high           | –1.75 | –0.59 | 0.56  | 1.46  | 0.34              | 0.44  |
| 1...they have to be done at all?  | 4.70/0.38             | Very high           | –1.84 | –0.69 | 0.59  | 1.61  | 0.19              | 0.29  |
| 3...they would not exist (or could be done with less effort), if it were organized differently? | 2.09/0.12             | Very high           | –2.21 | –0.89 | 0.58  | 1.91  | 0.04              | 0.28  |
| 5...they just exist because some people simply demand it this way?                              | 1.96/0.11             | Very high           | –2.18 | –0.83 | 0.34  | 1.59  | 0.34              | 0.12  |
| Unreasonable tasks  |                       |                     |       |       |       |       |                   |       |
| Do you have work tasks to take care of, which you believe ...                                   |                       |                     |       |       |       |       |                   |       |
| 7...are going too far, which should not be expected from you?                                   | 3.27/0.23             | Very high           | –1.47 | –0.34 | 0.73  | 1.75  | 0.04              | 0.11  |
| 9...are unfair that you have to deal with them?   | 2.98/0.20             | Very high           | –1.04 | 0.05  | 1.08  | 2.15  | 0.39              | 0.21  |
| 6...should be done by someone else?   | 2.51/0.15             | Very high           | –2.16 | –0.79 | 0.47  | 1.75  | 0.44              | 0.38  |
| 8...put you into an awkward position?   | 2.42/0.15             | Very high           | –1.21 | 0.01  | 1.25  | 2.20  | 0.67              | 0.79  |

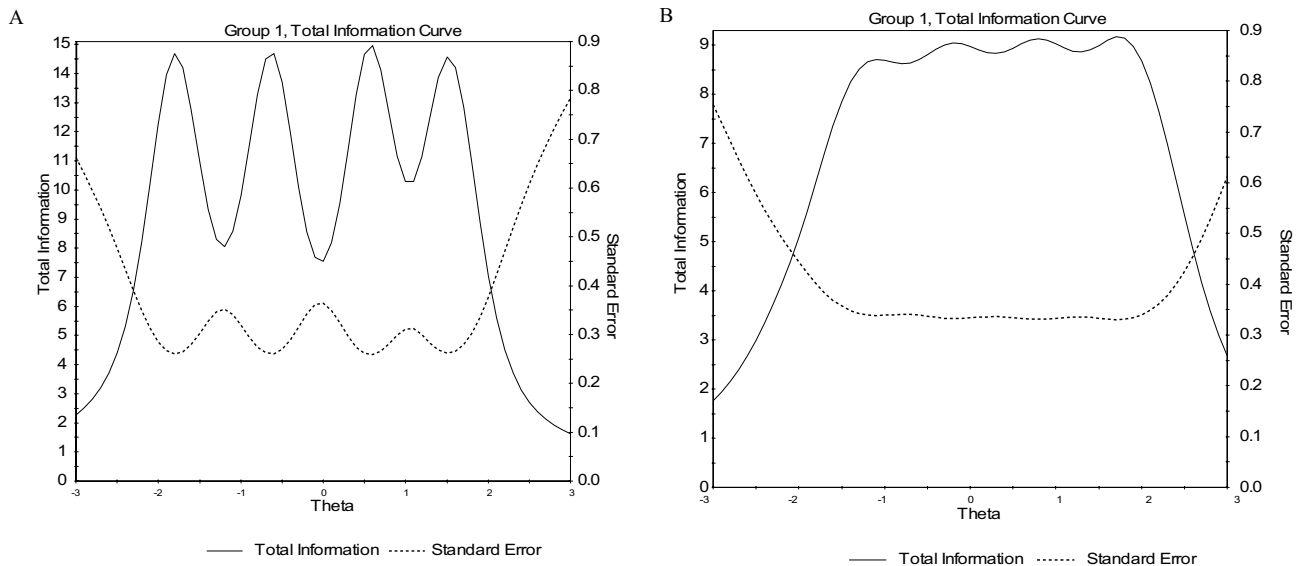
**Table 4.** Item Response Analysis of the Polish version of the BITS. Sample 1 ( $N=965$ ). *a* = discrimination parameter. Items were ranked by the *a* value. Very high  $\geq 1.69$ .  $b_1$ – $b_4$  = item difficulty/threshold parameters. Shaded *b<sub>i</sub>* values > 95 percentiles. DIF = differential item functioning. \* *p* values of chi-square tests to examine DIF with respect to gender-specific discrimination and difficulty parameters (Bonferroni adjusted for multiple comparisons  $0.05/4 = 0.0125$  indicating that no significant DIF were found).

cut-offs, all items were designated ‘very high’ ( $\geq 1.69$ ). Therefore, the items were deemed able to adequately distinguish between employees with different levels of unnecessary and unreasonable tasks.

An inspection of the item difficulty parameters *b* indicates that both dimensions of illegitimate tasks in the Polish version of BITS adequately measured these traits across low to high levels. DIF showed that all items had measurement invariance between males and females. ICCs combined with IIFs for the four unnecessary task items and four unreasonable task items in the Polish version of BITS are presented in Fig. 1.



**Figure 1.** Item characteristics curves (ICCs, colored lines), combined with item information functions (IIFs, dashed lines) for the unnecessary tasks’ subscale (A) and for the unreasonable tasks’ subscale (B) of the Polish version of the BITS (Sample 1 N = 965).



**Figure 2.** Test information function (TIF) of the dimension of unnecessary tasks (A) and the unreasonable tasks (B) of the Polish version of the BITS under the graded response model (GRM) (Sample 1 N = 965). Note. Latent trait  $\theta$  is shown on the horizontal axis, and the amount of information and the standard error yielded by the test at any trait level are shown on the vertical axis. Ranging from about  $-2.30$  SDs below the mean to about  $2.00$  SDs above the mean, the amount of test information was at least  $6.5$  (which yields a standard error of estimate about  $0.38$ ). We can interpret the information magnitude by computing the associated reliability ( $r = 1 - 1/\text{information}$ ). Thus, reliability was equal to or greater than  $0.85$ , and within the range described. Note. Latent trait  $\theta$  is shown on the horizontal axis, and the amount of information and the standard error yielded by the test at any trait level are shown on the vertical axis. Ranging from about  $-2.00$  SDs below the mean to about  $2.50$  SDs above the mean, the amount of test information was at least  $4.8$  (which yields a standard error of estimate about  $0.47$ ). We can interpret the information magnitude by computing the associated reliability ( $r = 1 - 1/\text{information}$ ). Thus, reliability was equal to or greater than  $0.73$  within the range described.

Concerning specific reliability, the test information function indicated that both factors were sufficiently informative (i.e., reliable) for a broad range of the trait illegitimate tasks (see Fig. 2).

**Further validity of BITS: replication in sample 2.** We replicated construct validity for BITS using CFA in sample 2 (Table 3). The single-factor model showed a poor fit to the data and was worse than the two-factor models. A modified model with the two correlated latent factors obtained an excellent fit. Those factors were moderately correlated (0.67) (see Supplementary Fig. S1B).

Convergent validity was assessed with AVE, and the value for both factors in the model exceeded the criterion of 0.50 (0.62 for unnecessary and unreasonable tasks). In sample 2 unnecessary tasks were more frequently observed than unreasonable tasks ( $t(802) = 11.03, p < 0.001$ ).

Discriminant validity was assessed with the Fornell–Larcker<sup>57</sup> criterion in regard to work overload, occupational wellbeing (i.e., job burnout and job satisfaction) and work performance. Descriptive statistics, reliabilities, and correlations for study variables are presented in Table 5.

AVE for unnecessary tasks (0.62) was higher than the value of its squared correlations with work overload (0.12), job burnout (0.19), job satisfaction (0.11) and work performance (0.04). In addition, AVE for unreasonable tasks (0.62) was higher than the value of its squared correlations with work overload (0.21), job burnout (0.11), job satisfaction (0.10) and work performance (0.01). Therefore, both types of illegitimate tasks can be regarded as distinct from measures assessing the level of work overload, occupational wellbeing and work performance. Further to this, we evaluated the reliability of BITS. The Cronbach's alpha value was 0.89 for unnecessary tasks and 0.87 for unreasonable tasks, and similarly, the CR was 0.87 for both subscales. Thus, it confirmed the internal consistency and reliability of BITS.

Construct validity was also assessed by running multiple linear regressions, including work overload and both types of illegitimate tasks as predictors, and two indicators of occupational wellbeing (namely, job burnout and job satisfaction), as well as work performance (Table 6).

Unnecessary tasks, unreasonable tasks, and work overload explained 29% of the variance in job burnout. Additionally, both unnecessary and unreasonable tasks, but not work overload, explained 14% of the variance in job satisfaction. Therefore, as expected, both types of illegitimate tasks explain more variance in job burnout, which is a negative indicator of wellbeing, compared with job satisfaction, which is a positive indicator of wellbeing. Moreover, unnecessary tasks, but not unreasonable tasks or work overload, explained 4% of the variance in work performance.

## Discussion

Combining the classical test and item response theories, our results showed that BITS is *not* a unidimensional scale. It measures two conceptually different separate dimensions, comprising four items each. IRT analyses showed that all items on each factor on different levels were very highly informative (reliable). All levels of the unnecessary and unreasonable tasks were reliably captured by the BITS items, suggesting that the questionnaire has very good psychometric properties. Moreover, BITS had measurement invariance with respect to

| Variables             | <i>M</i> | <i>SD</i> | <i>Sk</i> | <i>K</i> | 1        | 2        | 3        | 4        | 5       | 6 |
|-----------------------|----------|-----------|-----------|----------|----------|----------|----------|----------|---------|---|
| 1. Unnecessary tasks  | 2.97     | 0.78      | -0.04     | 0.02     | (0.89)   |          |          |          |         |   |
| 2. Unreasonable tasks | 2.68     | 0.82      | 0.23      | -0.07    | 0.57***  | (0.87)   |          |          |         |   |
| 3. Work overload      | 3.59     | 0.62      | -0.37     | 0.02     | 0.34***  | 0.46***  | (0.75)   |          |         |   |
| 4. Job burnout        | 2.31     | 0.59      | 0.32      | -0.01    | 0.44***  | 0.44***  | 0.40***  | (0.95)   |         |   |
| 5. Job satisfaction   | 3.70     | 0.75      | -0.90     | 1.76     | -0.33*** | -0.32*** | -0.21*** | -0.47    | -       |   |
| 6. Work performance   | 7.75     | 1.50      | -1.12     | 2.40     | -0.21*** | -0.11**  | -0.06    | -0.39*** | 0.32*** | - |

**Table 5.** Descriptive statistics, reliabilities, and correlations for study variables. Sample 2  $N = 803$ . *Sk* = Skewness. *K* = Kurtosis. Scale reliabilities listed along diagonal. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

| Job demands        | Job burnout        |           |         |          | Job satisfaction  |           |         |          | Work performance  |           |         |          |
|--------------------|--------------------|-----------|---------|----------|-------------------|-----------|---------|----------|-------------------|-----------|---------|----------|
|                    | <i>B</i>           | <i>SE</i> | $\beta$ | <i>t</i> | <i>B</i>          | <i>SE</i> | $\beta$ | <i>t</i> | <i>B</i>          | <i>SE</i> | $\beta$ | <i>t</i> |
| Constant           | 0.59               | 0.11      |         | 5.35***  | 5.00              | 0.15      |         | 32.69*** | 8.85              | 0.32      |         | 27.47*** |
| Unnecessary tasks  | 0.20               | 0.03      | 0.26    | 7.15***  | -0.21             | 0.04      | -0.22   | -5.36*** | -0.42             | 0.08      | -0.22   | -5.14*** |
| Unreasonable tasks | 0.13               | 0.03      | 0.18    | 4.66***  | -0.15             | 0.04      | -0.17   | -3.90*** | 0.03              | 0.08      | 0.02    | 0.35     |
| Work overload      | 0.22               | 0.03      | 0.23    | 6.70***  | -0.08             | 0.05      | -0.06   | -1.70    | 0.02              | 0.10      | 0.01    | 0.18     |
| $R^2$              | 0.29               |           |         |          | 0.14              |           |         |          | 0.04              |           |         |          |
| <i>F(df)</i>       | 106.43*** (3, 799) |           |         |          | 42.87 (3, 799)*** |           |         |          | 11.94*** (3, 799) |           |         |          |

**Table 6.** Job demands, occupational well-being and work performance—results of multiple linear regressions. Sample 2  $N = 803$ . *B* = Unstandardized Regression Coefficients.  $\beta$  = standardized Regression Coefficients. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

gender-specific discrimination and difficulty parameters. Additionally, the results from validity tests revealed that both dimensions of unnecessary and unreasonable tasks were distinct from other work overload and occupational wellbeing indicators.

The present study investigated for the first time the psychometric properties of an eight-item BITS, using IRT, to provide evidence that this questionnaire reliably measures illegitimate tasks among employees without producing biased data. In addition, a new language version was verified. Thus far, a German and a Spanish version of BITS have been verified using CFA<sup>13–15</sup>. Therefore, our study links advantages of the classical test and item response theories.

We used IRT to analyse the extracted two factors rather than the single scale of illegitimate tasks, as the latter did not meet the unidimensional criterion. Moreover, the results of parallel analysis and CFA, as well as the findings of past research on BITS<sup>2,4,13–15</sup>, showed that the construct of illegitimate tasks consists of two dimensions, namely, unnecessary and unreasonable tasks.

This study also provides evidence that Item 4 may be excluded from the original version of BITS. Indeed, using classical test theory, this item obtained weak validity, psychometrically and semantically. Conceptually, it refers to both subscales (i.e., unnecessary and unreasonable tasks). IRT analysis has shown Item 4 had moderate discrimination power<sup>16</sup>, and due to this, this item produced relatively lower and flatter information curves. Additionally, scarce previous research on BITS has resulted in researchers applying the eight-item BITS, although Item 4 was not always excluded<sup>25,44</sup>. Therefore, different content versions are implemented. Finally, some studies have been based on different response formats<sup>17,20,23,42</sup>. Due to this confusion in the literature, the existing findings cannot be accurately compared. Overall, we present new evidence for improving BITS, which replicates in great depth previous findings<sup>4,13</sup>.

There are no previous IRT findings on BITS to compare with the current study's results. However, significantly, all eight items with very high discrimination power indicate that the items of the Polish version of BITS can adequately distinguish between people with different levels of illegitimate tasks. Moreover, in our study, Item 2 and Item 1 (for unnecessary tasks) and Item 7 (for unreasonable tasks), demonstrated the highest discrimination power (see Table 4 and also observe the dashed lines in Fig. 1). Thus, we may propose that in future studies requiring single items, e.g., in diary studies, researchers may use the three mentioned items (1, 2 and 7). However, we have found a diary study<sup>23</sup> that used Item 3 and Item 6 as representative of the unnecessary and unreasonable tasks, respectively. No rationale was presented for using these items. We also suggest that a reduced scale, with only three items, can be useful in screening tests or surveys comprising a multitude of scales. However, fully established evidence requires further IRT analysis among other language versions.

Convergent and discriminant analyses revealed that the Polish version of BITS is valid for both its factors. Furthermore, BITS is a reliable instrument, as earlier studies have found<sup>14,15,24</sup>. Our study adds to prior research by presenting reliability on the whole scale range, using IRT (see Figs. 1 and 2). The results of our study, like previous studies, confirmed that unnecessary and unreasonable tasks are separate constructs in terms of stressful job demands. Similar to existing studies<sup>12,24,26</sup>, illegitimate tasks were moderately correlated with work overload and jointly predicted occupational wellbeing. However, both dimensions of illegitimate tasks in our sample contributed more to occupational wellbeing beyond and above work overload. What is more important, illegitimate tasks related to job burnout to a greater extent than job satisfaction. It is consistent with the assumptions of Job Demands-Resources theory<sup>60</sup>, that excessive social demands negatively affect people's health more than decrease their motivation.

Additionally, we have observed that unnecessary tasks occurred more often than unreasonable tasks among Polish employees. The same was noted for Swedish employees<sup>30</sup>. Moreover, our findings demonstrated that only the unnecessary tasks were negatively related to employees' self-rated work performance. However, to a lesser extent than they affected their wellbeing. Unnecessary tasks go hand-in-hand with 'red tape', are more often assigned at random<sup>17</sup> and may create unproductive organisational behaviour and diminish the meaningfulness of work among teachers<sup>5</sup>, administrative staff<sup>11</sup> and IT specialists<sup>33</sup>.

Our study has some shortcomings. The samples were not representative of the employed Polish population, and they were limited to four economic sectors, i.e., IT, administration, education and healthcare. Nonetheless, samples included respondents from each of Poland's 16 voivodeships. We should emphasise that less-educated employees were underrepresented. In addition, self-reporting methods were used, therefore in further research it is advisable to obtain data from a variety of sources, as was done by Muntz and Dorman<sup>14</sup>. Above all, the cross-sectional design is an important limitation for estimating the predictor-outcomes relationship and, also, for controlling the common method variance<sup>59</sup>. However, Harman's single-factor method showed that CMV was not a problem in this study (the first factor explained 35.1% of the total variance in sample 2, which constituted less than half of the covariance among measures).

Our findings may have practical implications. BITS has proven to be a valid and reliable scale that can be used in managerial practice. Illegitimate tasks are widespread, therefore this is still a real problem in work management<sup>5,6,17</sup>. Thus, a valuable diagnostic instrument to identify illegitimate tasks is expected. Our findings provided new evidence that BITS is an exceptional scale. In addition, based on the IRT results, we may suggest that a few items of BITS could be useful during screening examination and diary surveys to control the rationale of work tasks. And above all, in day-to-day interviews (e.g., during staff interviews), to help managers focus on a qualitative understanding of illegitimate tasks.

We conclude that the Polish version of BITS, completed by employees from different Polish organisations, is best measured by its two dimensions, namely, unnecessary and unreasonable tasks, which include four items each. The validity and reliability results of BITS demonstrate that the scale is outstanding and worth applying in scientific research and managerial practice.



## Data availability

The data underlying the results presented in the study are available on Mendeley (<https://doi.org/10.17632/7wfgz62xgs>).

Received: 27 February 2022; Accepted: 22 April 2023

Published online: 03 May 2023

## References

- Baran, B. E. & Woznyj, H. M. Managing VUCA: The human dynamics of agility. *Organ. Dyn.* **100787**, 1. <https://doi.org/10.1016/j.orgdyn.2020.100787> (2020).
- Semmer, N. K. *et al.* Illegitimate tasks as a source of work stress. *Work Stress* **29**, 32–56. <https://doi.org/10.1080/02678373.2014.1003996> (2015).
- Ding, H. & Kuvaas, B. Illegitimate tasks: A systematic literature review and agenda for future research. *Work Stress* <https://doi.org/10.1080/02678373.2022.2148308> (2022).
- Semmer, N. K., Tschan, F., Meier, L. L., Facchin, S. & Jacobshagen, N. Illegitimate tasks and counterproductive work behavior. *Appl. Psychol. Int. Rev.* **59**, 70–96. <https://doi.org/10.1111/j.1464-0597.2009.00416.x> (2010).
- Faupel, S., Otto, K., Krug, H. & Kottwitz, M. U. Stress at school? A qualitative study on illegitimate tasks during teacher training. *Front. Psychol.* **7**, 1410. <https://doi.org/10.3389/fpsyg.2016.01410> (2016).
- Kilponen, K., Huhtala, M., Kinnunen, U., Mauno, S. & Feldt, T. Illegitimate tasks in health care: Illegitimate task types and associations with occupational well-being. *J. Clin. Nurs.* **30**, 2093–2106. <https://doi.org/10.1111/jocn.15767> (2021).
- Semmer, N. K., Jacobshagen, N., Meier, L. L. & Elfering, A. Occupational stress research: The “stress-as-offense-to-self” perspective. In *Occupational health psychology: European perspectives on research, education and practice* 43–60 (eds Houdmont, J. & McIntyre, S.) (ISMAI Publishing, London, 2007).
- Semmer, N. K. *et al.* Stress as offense to self: A promising approach comes of age. *Occup. Health Sci.* **3**, 205–238. <https://doi.org/10.1007/s41542-019-00041-5> (2019).
- Fila, M. J. & Eatough, E. Extending the boundaries of illegitimate tasks: The role of resources. *Psychol. Rep.* **123**, 1635–1662. <https://doi.org/10.1177/0033294119874292> (2020).
- Björk, L., Bejerot, E., Jacobshagen, N. & Härenstam, A. I shouldn't have to do this: illegitimate tasks as a stressor in relation to organizational control and resource deficits. *Work Stress* **27**, 262–277. <https://doi.org/10.1080/02678373.2013.818291> (2013).
- Eatough, E. M. *et al.* You want me to do what? Two daily diary studies of illegitimate tasks and employee well-being. *J. Organ. Behav.* **37**, 108–127. <https://doi.org/10.1002/job.2032> (2016).
- Omansky, R., Eatough, E. M. & Fila, M. J. Illegitimate tasks as an impediment to job satisfaction and intrinsic motivation: Moderated mediation effects of gender and effort-reward imbalance. *Front. Psychol.* **7**, 1818. <https://doi.org/10.3389/fpsyg.2016.01818> (2016).
- Jacobshagen, N. *Illegitimate tasks, illegitimate stressors: Testing a new stressor-strain concept* (Bern University, 2006).
- Muntz, J. & Dormann, C. Moderating effects of appreciation on relationships between illegitimate tasks and intrinsic motivation: A two-wave shortitudinal study. *Eur. J. Work Organ. Psy.* **29**, 391–404. <https://doi.org/10.1080/1359432X.2019.1706489> (2020).
- Portilla, D. L. V., Rosero, A. G., Alvarado-Villa, G. & Moncayo-Rizzo, J. Psychometric properties of the Bern illegitimate tasks scale-Spanish version. *Front. Psychol.* **12**, 593870. <https://doi.org/10.3389/fpsyg.2021.593870> (2021).
- Baker, F. B. *The basics of item response theory* 2nd edn. (ERIC Clearinghouse on Assessment and Evaluation, London, 2001).
- Pindek, S., Demircioğlu, E., Howard, D. J., Eatough, E. M. & Spector, P. E. Illegitimate tasks are not created equal: Examining the effects of attributions on unreasonable and unnecessary tasks. *Work Stress* **33**, 231–246. <https://doi.org/10.1080/02678373.2018.1496160> (2019).
- Sonnentag, S. & Lischetzke, T. Illegitimate tasks reach into afterwork hours: A multilevel study. *J. Occup. Health Psych.* **23**, 248–261 (2018).
- Pfister, I. B. *et al.* Appreciation and illegitimate tasks as predictors of affective well-being: Disentangling within- and between-person effects. *J. Work Organ. Psychol.* **36**, 63–75. <https://doi.org/10.5093/jwop2020a6> (2020).
- Zhao, L., Lam, L. W., Zhu, J. N. & Zhao, S. Doing it purposefully? Mediation of moral disengagement in the relationship between illegitimate tasks and counterproductive work behavior. *J. Bus. Ethics* <https://doi.org/10.1007/s10551-021-04848-7> (2021).
- Ma, J. & Peng, Y. The performance costs of illegitimate tasks: The role of job identity and flexible role orientation. *J. Vocat. Behav.* **110**, 144–154 (2019).
- Hirschle, A. L. T. & Gondim, S. M. G. Stress and well-being at work: A literature review. *Cien Saude Coletiva* **25**, 2721–2736. <https://doi.org/10.1590/1413-81232020257.27902017> (2020).
- Bianchi, R., Manzano-García, G. & Rolland, J. P. Is burnout primarily linked to work-situated factors? A relative weight analytic study. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2020.623912> (2021).
- Meier, L. L. & Semmer, N. K. Illegitimate tasks as assessed by incumbents and supervisors: Converging only modestly but predicting strain as assessed by incumbents, supervisors, and partners. *Eur. J. Work Organ. Psy.* **27**, 764–776. <https://doi.org/10.1080/1359432X.2018.1526785> (2018).
- Ilyas, A., Hassan, R., Khan, A. & Khan, W. Illegitimate tasks and job satisfaction among employees of micro informal enterprises. *Manag. Sci. Lett.* **11**, 291–296. <https://doi.org/10.5267/j.msl.2020.8.004> (2021).
- Kottwitz, M. U. *et al.* SOS-Appreciation overboard! Illegitimacy and psychologists' job satisfaction. *Ind. Health* **57**, 637–652. <https://doi.org/10.2486/indhealth.2018-0008> (2019).
- Mäkikangas, A., Minkinen, J., Muotka, J. & Mauno, S. Illegitimate tasks, job crafting and their longitudinal relationships with meaning of work. *Int. J. Hum. Resou. Man.* <https://doi.org/10.1080/09585192.2021.1987956> (2021).
- Ping, Z. L., Fu, H. Y., Ye, Z. X. & Zhao, S. Illegitimate tasks and employees' turnover intention: A serial mediation model. *Front. Psychol.* **12**, 739593. <https://doi.org/10.3389/fpsyg.2021.739593> (2021).
- van Schie, S., Güntert, S. T. & Wehner, T. (2014) How dare to demand this from volunteers! The impact of illegitimate tasks. *Voluntas: International Journal of Voluntary and Nonprofit Organizations* **25**, 851–868. <https://doi.org/10.1007/s11266-013-9375-4>
- Anskär, E., Lindberg, M., Falk, M. & Andersson, A. Legitimacy of work tasks, psychosocial work environment, and time utilization among primary care staff in Sweden. *Scand. J. Prim. Health.* **37**, 476–483. <https://doi.org/10.1080/02813432.2019.1684014> (2019).
- Stein, M., Vincent-Höper, S., Schümann, M. & Gregersen, S. Beyond mistreatment at the relationship level: Abusive supervision and illegitimate tasks. *Int. J. Environ. Res. Public Health* **17**, 2722. <https://doi.org/10.3390/ijerph17082722> (2020).
- Bramlage, J. K., Julmi, C., Pereira, J. M. & Jackenkroll, B. When enough is enough: modelling the path from unreasonable tasks to the intention to leave academia. *Eur. J. High. Educ.* **11**, 386–407. <https://doi.org/10.1080/21568235.2021.1873160> (2021).
- Apostel, E., Syrek, C. & Antoni, C. H. Turnover intention as a response to illegitimate tasks: The moderating role of appreciative leadership. *Int. J. Stress Manage.* **25**, 234–249. <https://doi.org/10.1037/str0000061> (2018).
- Mauno, S., Minkinen, J. & Shimazu, A. Do unnecessary tasks impair performance because they harm living a calling? Testing a mediation in a three-wave study. *J. Career Assess.* **30**, 94–104. <https://doi.org/10.1177/10690727211018977> (2022).

35. Zhou, Z. E., Eatough, E. M. & Che, X. X. Effect of illegitimate tasks on work-to-family conflict through psychological detachment: Passive leadership as a moderator. *J. Vocat. Behav.* **121**, 103463. <https://doi.org/10.1016/j.jvb.2020.103463> (2020).
36. Pereira, D., Semmer, N. K. & Elfering, A. Illegitimate tasks and sleep quality: an ambulatory study. *Stress Health* **30**, 209–221. <https://doi.org/10.1002/smi.2599> (2014).
37. Kottwitz, M. U. *et al.* Illegitimate tasks associated with higher cortisol levels among male employees when subjective health is relatively low: an intra-individual analysis. *Scand. J. Work Environ. Health* **39**, 310–318. <https://doi.org/10.5271/sjweh.3334> (2013).
38. Muntz, J., Dormann, C. & Kronenwett, M. Supervisors' relational transparency moderates effects among employees' illegitimate tasks and job dissatisfaction: A four-wave panel study. *Eur. J. Work Organ. Psy.* **28**, 485–497. <https://doi.org/10.1080/1359432X.2019.1625888> (2019).
39. Thun, S., Halsteinli, V. & Løvseth, L. A study of unreasonable illegitimate tasks, administrative tasks, and sickness presenteeism amongst Norwegian physicians: an everyday struggle?. *BMC Health Serv. Res.* <https://doi.org/10.1186/s12913-018-3229-0> (2018).
40. Graf-Vlachy, L., Sun, S. & Zhang, S. X. Predictors of managers' mental health during the COVID-19 pandemic. *Eur. J. Psychotraumatol.* <https://doi.org/10.1080/20008198.2020.1834195> (2020).
41. Aronsson, G., Bejerot, E. & Härenstam, A. Onödiga och oskåliga arbetsuppgifter bland läkare: Samband mellan illegitima arbetsuppgifter och stress kartlagt i enkätstudie. *Lakartidningen* **109**, 2216–2219 (2012) ((in Swedish)).
42. Akyurek, S. S., Can, O. Illegitimate tasks and occupational outcomes: The impact of vertical collectivism. In *Evidence-based HRM: A global forum for empirical scholarship* <https://doi.org/10.1108/EBHRM-02-2021-0025> (2021).
43. Ahmed, S. F., Eatough, E. M. & Ford, M. T. Relationships between illegitimate tasks and change in work-family outcomes via interactional justice and negative emotions. *J. Vocat. Behav.* **104**, 14–30. <https://doi.org/10.1016/j.jvb.2017.10.002> (2018).
44. Aronsson, G. & Mellner, C. Illegitima arbetsuppgifter och identitet-En introduktion. *Arbetsmarknad & Arbetsliv* **22**, 28–46 (2016) ((in Swedish)).
45. von Thiele Schwarz, U. *et al.* The work of having a chronic condition: development and psychometric evaluation of the distribution of co-care activities (DoCCA) scale. *BMC Health Serv. Res.* <https://doi.org/10.1186/s12913-021-06455-8> (2021).
46. Samejima, F. Estimation of latent ability using a response pattern of graded scores. *Psychom. Monogr. Suppl.* **17**, 386–415 (1969).
47. Şahin, A. & Anil, D. The effects of test length and sample size on item parameters in item response theory. *Edu. Sci-Theor Pract.* **17**, 321–335. <https://doi.org/10.12738/estp.2017.1.0270> (2017).
48. Brislin, R. W. Back-translation for cross-cultural research. *J. Cross Cult. Psychol.* **1**, 185–216. <https://doi.org/10.1177/13591045700100301> (1970).
49. Schaufeli, W. B. Applying the job demands-resources model. *Organ. Dyn.* **2**, 120–132. <https://doi.org/10.1016/j.orgdyn.2017.04.008> (2017).
50. Schaufeli, W. B., Desart, S. & De Witte, H. Burnout assessment tool (BAT)—development, validity, and reliability. *Int. J. Environ. Res. Public Health* **17**, 9495. <https://doi.org/10.3390/ijerph17249495> (2020).
51. Wanous, J. P. & Hudy, M. J. Single-item reliability: A replication and extension. *Organ. Res. Methods* **4**, 361–375 (2001).
52. O'Connor, B. P. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behav. Res. Meth. Ins. C.* **32**, 396–402. <https://doi.org/10.3758/BF03200807> (2020).
53. Reeve, B. B. *et al.* Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the patient-reported outcomes measurement information system (PROMIS). *Med. Care* **45**, 22–31 (2007).
54. Hu, L. T. & Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struc. Equ. Model.* **6**, 1–55. <https://doi.org/10.1080/10705519909540118> (1999).
55. Kline, R. B. *Principles and practice of structural equation modeling* (Guilford, 2015).
56. Hair, J. F., Hult, G. T. M., Ringle, C. & Sarstedt, M. *A primer on partial least squares structural equation modeling (PLS-SEM)* (SAGE, 2014).
57. Fornell, C. & Larcker, D. F. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* **18**, 39–50. <https://doi.org/10.1177/002224378101800104> (1981).
58. Penfield, R. D. An NCME instructional module on polytomous item response theory models. *Educ. Meas. Issues Pra.* **33**, 36–48. <https://doi.org/10.1111/emip.12023> (2014).
59. Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. & Podsakoff, N. P. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol.* **88**, 879–903 (2003).
60. Bakker, A. B. & Demerouti, E. Job demands-resources theory: taking stock and looking forward. *J. Occup. Health Psych.* **22**, 273–285. <https://doi.org/10.1037/ocp0000056> (2017).

## Author contributions

Conceptualization: [B.A.B.], [A.M.D.]; Methodology: [B.A.B.], [A.M.D.]; Formal analysis and investigation: [B.A.B.], [A.M.D.]; Writing - original draft preparation: [B.A.B.], [A.M.D.]; Writing - review and editing: [B.A.B.], [A.M.D.]; Funding acquisition: [B.A.B.]; Resources: [A.M.D.].

## Funding

This study was funded by the National Science Centre, Poland (grant number 2017/26/M/HS6/00451).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-34006-0>.

**Correspondence** and requests for materials should be addressed to B.A.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023