ASSESSMENT OF RELIABILITY, VALIDITY AND RESPONSIVENESS OF THE DETACHMENT AND RELAXATION AT WORK (DRAW) SCALE

Coffeng JK, Hendriksen IJM, Terwee CB, Twisk JWR, Van Mechelen W & Boot CRL.

Submitted.
ABSTRACT

Background: Work breaks help to counteract the negative effects of work-related stress and positively influence feelings of energy and restoration. As there is no scale to measure detachment and relaxation breaks within a workday, we adapted the subscales detachment and relaxation from the Recovery Experience Questionnaire into a within workday setting.

Objective: To evaluate the measurement properties of the Detachment and Relaxation At Work (DRAW) scale.

Methods: The DRAW scale consists of four subscales: within workday detachment (1) and relaxation (2), after workday detachment (3) and relaxation (4). Data from a randomized controlled trial were used (n=412). The following measurement properties were evaluated: internal consistency, floor- and ceiling effects, test-retest reliability, construct validity and responsiveness (i.e., formulation of hypotheses with similar/dissimilar instruments).

Results: A 4-factor model fitted the data (i.e., Confirmatory Factor Analysis). The internal consistency ranged from 0.87-0.94. The ICCs for test-retest reliability ranged from 0.57-0.84. Construct validity was moderate. Overall responsiveness was moderate to poor.

Conclusions: Based on the results, no firm conclusions can be drawn and we recommend to be cautious when using the DRAW in interventions studies. Further research is necessary (i.e., preferably in other populations/intervention studies and elaborating items with work engagement strategies).
BACKGROUND

According to a recent report in the Netherlands, almost one in eight employees (13%) suffers from work-related stress. The cost of work days lost to presenteeism (i.e., employees tend not to call in sick, but are less productive at work due to health problems) and sickness absence totalled to 2.7 billion Euros (1,2). One of the most important factors influencing physical and mental health of an employee is whether an individual is able to recover from fatigue and stress at work (3). Insufficient recovery has been found to be associated with an increased risk for developing work-related diseases such as burnout, cardiovascular disease (CVD) and musculoskeletal disorders (3-7). Therefore, researchers have made significant strides toward understanding factors and processes needed to recover from work (8,9). Not surprisingly, research on recovery revealed that people should take enough breaks, which can be vacations, weekends, evenings, but also breaks at work (10).

Research has shown that breaks consisting of detachment and relaxation are related to increasing one’s feeling of recovery (11-15). Moreover, it was shown that lower levels of detachment and relaxation are associated with weaker health, emotional exhaustion and sleeping problems (12). The term detachment was introduced by Etzion, Eden and Lapidot (1998) and defined as “an individual’s sense of being away from the work situation” (p. 579) (16). Relaxation poses no further demands on the functional system (e.g., neuro-endocrine and cardiovascular systems) and internal resources (e.g., self-regulation). Relaxation activities are associated with a decreased resting heart rate, lowered resting breath rate and alleviated resting muscle tension (17).

So far, research has highlighted the importance of recovery activities undertaken after work. While research examining recovery processes within a workday is sparse, it is an emerging and fruitful field of research for two main reasons. First, the rationale for within workday breaks is that breaks are helpful to reduce fatigue, injury, and boredom (e.g., employees with repetitive tasks) (18). Moreover, breaks help to counteract the negative effects of work-related stress and physical inactivity, and can have positive effects on feelings of energy and restoration (19). Second, because people spend one third to one half of their day at work, making breaks part
of the workday, could have a positive spill-over effect to recovery after a workday (i.e., a reduction of the need for recovery). This is supported by a study investigating a group of service employees, which showed that enjoyable and restful breaks within a workday improved the recovery (20).

A valid and reliable questionnaire is required to measure the effects of work break interventions on feelings of recovery (e.g., detachment and relaxation) during work. Thus far, there is no questionnaire that focuses on feelings of recovery within a workday. The Recovery Experience Questionnaire developed by Sonnentag & Fritz (12), measures detachment and relaxation after a workday. However, this questionnaire is not suitable for assessing within workday detachment and relaxation. To facilitate research on interventions for recovery at work, we adapted the subscales detachment and relaxation from the Recovery Experience Questionnaire (12) into a within workday context by means of changing the instruction and response categories. This has led to the Detachment and Relaxation At Work (DRAW) scale, which consists of four subscales: within workday detachment, within workday relaxation, after workday detachment and after workday relaxation. The main purpose of this study was to evaluate the measurement properties of the DRAW scale by applying guidelines from the COSMIN (Consensus-Based Standards for the selection of Health Measurements Instruments) Group (21). In this study, we investigated internal consistency, floor- and ceiling effects, test-retest reliability, construct validity (i.e., hypotheses testing) and responsiveness of the DRAW scale.

METHODS

Participants
For the present study, data from a randomized controlled trial alongside the Be Active & Relax project were used (22). The main aim of the Be Active & Relax project was to reduce the need for recovery in employees by stimulating physical activity and relaxation. It was also hypothesized that the intervention groups would improve in detachment and relaxation compared to the control group. Full details of the design of the Be Active & Relax project have been reported elsewhere (22).
In September 2011, 1182 office workers of a financial service provider were invited to participate. Those on sick leave for more than four weeks were not eligible to participate. A total of 412 employees from 19 departments completed the baseline questionnaire, signed the informed consent form, and were included in the Be Active & Relax project.

To assess the measurement properties internal consistency, floor- and ceiling effects and construct validity, baseline data were used. For responsiveness, baseline and 6 months follow-up data were used. Because of missing values at 6 months, 274 (66.7% of the 412 participants at baseline) cases were available for assessing responsiveness. To assess test-retest reliability, respondents who completed the scale at 12 months follow-up (N=255; 61.8% of baseline) were invited five days later by e-mail with a link to the web-based DRAw scale to complete it again within 0-2 days. Respondents were approached until at least 100 participants had completed the DRAw scale for the second time (N=121). The study was approved by the Medical Ethics Committee of the VU University Medical Center Amsterdam.

The Detachment and Relaxation At Work (DRAw) scale

The DRAw scale consists of four subscales (i.e., the two subscales detachment and relaxation consist within two domains namely within a workday and after a workday), which will be explained below. All items of these four subscales are listed in table 1.

**Detachment and Relaxation within a workday.** The original items detachment and relaxation from the Recovery Experience Questionnaire (12) were adapted for the purpose of the current study into a during work context, starting each item with “During a break at work...”, instead of “During time after work...”. Furthermore, to optimize responsiveness, the response categories were changed from a 5-point scale to a 7-point scale, and the response answers were changed from “I don’t agree at all” - “I fully agree” to “never” - “always”.

**Detachment and Relaxation after a workday.** Detachment and Relaxation after a workday were assessed with two 4-item subscales, i.e., detachment and relaxation from the Recovery Experience Questionnaire (12). Content validity of the Recovery Experience Questionnaire was assessed by 16 advanced psychology students who were asked to categorize the items which were derived from the literature and...
brainstorming. The validation study of the Recovery Experience Questionnaire (12) resulted in four items measuring detachment (e.g., “I don’t think about work at all”) and four items measuring relaxation (e.g., “I use the time to relax”). The Dutch version of this questionnaire has shown a high internal consistency for the items on detachment (α=0.88) and for relaxation (α=0.81) (12;23). For the present study, the response categories were changed from a 5-point scale to a 7-point scale, and the response answers were changed from “I don’t agree at all” - “I fully agree” to “never” - “always” to optimize responsiveness.

Table 1. Translated items of the four subscales of the DRAW scale

**Subscale 1: Within workday detachment.**
*Instruction: During a break at work...*
1. I forget about work
2. I don’t think about work at all
3. I distance myself from my work
4. I get a break from the demands of work

**Subscale 2: Within workday relaxation.**
*Instruction: During a break at work...*
5. I kick back and relax
6. I do relaxing things
7. I use the time to relax
8. I take time for leisure

**Subscale 3: After workday detachment.**
*Instruction: During time after work...*
9. I forget about work
10. I don’t think about work at all
11. I distance myself from my work
12. I get a break from the demands of work

**Subscale 4: After workday relaxation.**
*Instruction: During time after work...*
13. I kick back and relax
14. I do relaxing things
15. I use the time to relax
16. I take time for leisure

*Items 1-16: 7-point scale “never” - “always”*
Comparator measurements

Need for recovery. Need for recovery refers to the need to recuperate from work-induced effort after a working day and was assessed using the Need for Recovery after Work scale (24). This scale consists of eleven dichotomous items (yes/no), representing short-term effects of a day at work, with questions like “I find it hard to relax at the end of a working day” and “When I get home, people should leave me alone for some time”. The need for recovery score is a percentage score (0 to 100) of positive answers of those providing data for at least 8 of the 11 items. The need for recovery has shown good internal consistency ($\alpha=0.87$) and validity (24). Validity was studied by analyzing the association of need for recovery with psychosocial risk factors (e.g., emotional load and physical exertion) (24).

Exhaustion. Exhaustion is defined as intense physical, affective and cognitive strain that is a long-term consequence of prolonged exposure to certain job demands. It was measured with the OLdenburg BurnOut Inventory (OLBI). The OLBI consists of eight items on a 4-point scale ranging from “totally agree” to “don’t agree” (25). The OLBI has shown satisfactory internal consistency ($\alpha=0.85$) and reasonable validity in different occupational groups (i.e., health care workers experienced higher levels of exhaustion than white collar workers) (26).

Work engagement. Work engagement refers to a state of energy, being connected to work positively and feeling able to cope with job demands. It was measured using the validated Utrecht Work Engagement Scale (UWES), which consists of three scales; vigour (6 items), dedication (5 items) and absorption (6 items) measured on a 7-point scale from “never” to “always” (27). The internal consistency of this questionnaire has been tested and results indicated an acceptable internal consistency of vigour ($\alpha=0.83$), dedication ($\alpha=0.92$) and absorption ($\alpha=0.80$). Validity was investigated by analyzing the association with exhaustion, cynicism, and professional inefficacy (i.e., burnout) (28).

Assessment of measurement properties

The following measurement properties were investigated for the four subscales of the DRAW; internal consistency, floor- and ceiling effects, test-retest reliability, construct validity (i.e., hypotheses testing) and responsiveness.
Chapter 5

**Internal consistency.** Internal consistency concerns the interrelatedness among the items in a subscale (29). For Cronbach’s alpha, the widely accepted cut-off of higher than 0.7 was used (30). Dimensionality was assessed by a confirmatory factor analysis.

**Floor- and ceiling effects.** The variability of all subscales of the DRAW was investigated to assess floor- and ceiling effects. In case more than 15% of the participants chose the highest or lowest scores, floor and/or ceiling effects were considered present (31;32).

**Test-retest reliability.** Test-retest reliability measures to which extent scores of participants who have not changed are the same at repeated measurement (29). Test-retest reliability was determined by means of an Intraclass Correlation Coefficient (ICC) (range: 0-1); which indicates the ability to distinguish between participants (33). ICCs below 0.40 were considered to be poor, between 0.40-0.59 as fair, between 0.60-0.74 as good and between 0.75-1.0 as excellent (34;35). The measurement error was assessed by the Standard Error of Measurement (SEM) and the Smallest Detectable Change (SDC) (33). The measurement error represents the systematic and random error of a score that cannot be attributed to true changes in the DRAW scale (29). The SEM concerns the standard deviation of repeated measures in one participant. The SDC represents the minimal change a participant has to show to ensure that the observed change is relevant and not just measurement error (27).

**Construct validity: Hypotheses testing.** In the domain construct validity as defined by the COSMIN manual (29), three aspects of construct validity are specified. That is structural validity (i.e., the degree to which scores of a measurement are an adequate reflection of the dimensionality of the construct to be measured), cross-cultural validity (i.e., the degree to which the performance of the items on a translated culturally adapted measurement are an adequate reflection of the performance of the items of the original version of the measurement) and hypotheses testing. In the present study, we focus on hypotheses testing, based on the assumption that the DRAW scale validly measures the construct to be measured. The following comparator measurements were used to assess construct validity with hypotheses testing; need for recovery, OLBI (exhaustion) and UWES (work engagement; vigour, dedication and absorption). A priori hypotheses were formulated for each of the four
subscles based on three general assumptions. First, correlations between the DRA\V scale and related questionnaires should be moderate (r: 0.30-0.50). Second, the correlations between the DRA\V scale and distantly related questionnaires should be weak (r<0.30). Third, as the DRA\V scale is expected to be most closely related to the need for recovery, correlations of the DRA\V scale with need for recovery should be higher than correlations with exhaustion, vigour, absorption, and dedication. If more than 75% of the hypotheses were confirmed, the construct validity of a subscale was considered to be good (21;36;37), when 50-75% of the hypotheses were confirmed it was moderate and if less than 50% of hypotheses were confirmed it was poor (38). The following hypotheses were formulated for construct validity:

1. A moderate negative correlation (0.3-0.5) was expected between need for recovery and the four DRA\V subscales.
2. A moderate negative correlation (0.3-0.5) was expected between exhaustion and the four DRA\V subscales.
   These moderate correlations for need for recovery and exhaustion with the DRA\V scale were chosen based on previous research, in which correlations between 0.27 and 0.56 were found for after workday detachment and relaxation (12).
3. Vigour was expected to correlate negatively weak (<0.3) with the two subscales within workday detachment and relaxation, and positively weak (<0.3) with the two subscales after workday detachment and relaxation.
4. Dedication was expected to correlate negatively weak (<0.3) with the two subscales within workday detachment and relaxation, and positively weak (<0.3) with the two subscales after workday detachment and relaxation.
5. A moderate negative correlation (0.3-0.5) was expected between absorption and the four DRA\V subscales.

Previous research on work engagement (i.e., UWES; scale of vigour, dedication and absorption) and after workday detachment showed a weak positive correlation of 0.27 (39). For relaxation after a workday, a weak positive correlation of 0.10 was found (40). Only for absorption, a moderate negative correlation was expected with both within and after workday detachment and relaxation, because these items were likely to be related mostly to the DRA\V (e.g., items were: ‘It is difficult to detach myself from work’ and ‘I am immersed in my work’) (27;28).
6. The correlations between need for recovery and all four subscales of the DRAW were expected to be higher than the correlations with other variables.

**Responsiveness.** Responsiveness refers to “the ability of an instrument to detect change over time in the construct to be measured” (29). We applied the construct approach of responsiveness, which refers to the assessment of whether changes in scores of a scale over time (scores at six months follow-up compared to baseline) are consistent with predefined hypotheses (29). The same hypotheses and three general assumptions formulated for construct validity were applied to measure responsiveness. The hypotheses for responsiveness were extended by focusing on correlations between changes in mean differences of the scale under study (DRAW), and changes in scores on need for recovery, exhaustion, vigour, dedication and absorption on the other hand. Further, change scores of all four subscales of the DRAW with the change score of the need for recovery were expected to correlate more than the correlations of change with exhaustion, vigour, dedication and absorption.

**Statistical analyses**

Internal consistency was first analysed by assessing dimensionality, using confirmatory factor analysis for categorical items, and was performed in MPlus (Meuthen and Meuthen, Los Angeles, CA) by use of the method of weighted least squares with mean and variance adjustment. Factor loadings and model fit were examined. Factor loadings are generally considered to be meaningful when they exceed 0.30 - 0.40 (41). For model fit, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI) and the Root Mean Square Error of approximation (RMSEA) were used as measures. A CFI and TLI of greater than 0.95 and a RMSEA of less than 0.05 were considered as adequate fit (38). Subsequently, internal consistency was assessed by calculating Cronbach’s alphas for each unidimensional subscale. Test-retest reliability was examined by calculating Intraclass Correlation Coefficients (ICCs) with a 95% confidence interval. A 2-way random-effects ANOVA model for absolute agreement was used. The measurement error was assessed by the standard error of measurement (SEM) and the smallest detectable change (SDC) (35). The SEM was calculated as SD x \sqrt{1-ICC} and the SDC as 1.96 x \sqrt{2}*SEM. Since the data were not
normally distributed, construct validity (i.e., hypotheses testing) and responsiveness were assessed by calculating Spearman correlation coefficients. Statistical analyses were performed with SPSS 20.0 (SPSS Inc. Chicago, Illinois, USA) and MPlus. The level of significance was set at p<0.05.

RESULTS

Sample characteristics
Participants had a mean age of 41.3 years (SD=10.3), worked 36 hours per week (SD=5.1), most were male (60%), and most were highly educated (57%). On average, participants rated their general health as good (M=3.8, SD=0.8, on a 5-point scale). The overall mean score of the need for recovery (range 0-100) in the total study population was 32.2 (SD=29.3).

Measurement properties
The internal consistency is shown in table 2. A 4-factor model fitted the data. The CFI was 0.960, TLI was 0.952, and RMSEA was <0.001. All items had high factor loadings (>0.8). The Cronbach’s alphas ranged from 0.87 to 0.94.

Table 2. Factor loadings and Cronbach’s alphas of the four subscales of the DRAW scale

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Within workday Estimate</th>
<th>SE</th>
<th>Alpha</th>
<th>Within workday Estimate</th>
<th>SE</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detachment</td>
<td>(Alpha=0.93)</td>
<td></td>
<td></td>
<td>(Alpha=0.94)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 I forget about work</td>
<td>0.88</td>
<td>0.01</td>
<td></td>
<td>0.91</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>2 I don’t think about work at all</td>
<td>0.89</td>
<td>0.01</td>
<td></td>
<td>0.91</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>3 I distance myself from my work</td>
<td>0.92</td>
<td>0.01</td>
<td></td>
<td>0.94</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>4 I get a break from the demands of work</td>
<td>0.92</td>
<td>0.01</td>
<td></td>
<td>0.95</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Relaxation</td>
<td>(Alpha=0.87)</td>
<td></td>
<td></td>
<td>(Alpha=0.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 I kick back and relax</td>
<td>0.86</td>
<td>0.02</td>
<td></td>
<td>0.89</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>2 I do relaxing things</td>
<td>0.85</td>
<td>0.01</td>
<td></td>
<td>0.90</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>3 I use the time to relax</td>
<td>0.83</td>
<td>0.02</td>
<td></td>
<td>0.90</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>4 I take time for leisure</td>
<td>0.80</td>
<td>0.02</td>
<td></td>
<td>0.92</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>
None of the four subscales of the DRAW showed floor (<=3%) - and ceiling effects (<=8%). The results are presented in table 3.

**Table 3.** Floor- and ceiling effects of the DRAW scale

<table>
<thead>
<tr>
<th>Subscales</th>
<th>Range</th>
<th>Floor (%)</th>
<th>Ceiling (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within workday detachment</td>
<td>4-28</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Within workday relaxation</td>
<td>4-28</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>After workday detachment</td>
<td>4-28</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>After workday relaxation</td>
<td>4-28</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 4 presents the results of the data for the baseline test and the retest after 12 months. The ICC of after workday detachment was highest (0.84), followed by within workday detachment (0.70).

**Table 4.** Test-retest reliability and descriptive statistics of the DRAW scale (n=121)

<table>
<thead>
<tr>
<th></th>
<th>Test M (SD)</th>
<th>Retest M (SD)</th>
<th>Change M (SD)</th>
<th>SEM</th>
<th>SDC</th>
<th>ICC (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within workday</td>
<td>2.51 (1.45)</td>
<td>2.59 (1.37)</td>
<td>0.08 (1.10)</td>
<td>0.71</td>
<td>1.97</td>
<td>0.70 (0.59;0.78)</td>
</tr>
<tr>
<td>detachment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within workday</td>
<td>2.68 (1.39)</td>
<td>2.58 (1.41)</td>
<td>-0.10 (1.29)</td>
<td>0.81</td>
<td>2.25</td>
<td>0.57 (0.44;0.68)</td>
</tr>
<tr>
<td>relaxation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After workday</td>
<td>3.83 (1.38)</td>
<td>3.81 (1.26)</td>
<td>-0.00 (0.74)</td>
<td>0.51</td>
<td>1.41</td>
<td>0.84 (0.78;0.89)</td>
</tr>
<tr>
<td>detachment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After workday</td>
<td>4.29 (1.13)</td>
<td>4.09 (1.16)</td>
<td>-0.20 (0.92)</td>
<td>0.60</td>
<td>1.66</td>
<td>0.67 (0.55;0.76)</td>
</tr>
<tr>
<td>relaxation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

M, mean; SD, standard deviation; SEM, standard error of measurement; SDC, smallest detectable change; ICC, intra class correlation; CI, confidence interval.

For construct validity, all subscales showed that 67% of the hypotheses were confirmed (table 5).

Regarding responsiveness, 42% of hypotheses were confirmed (table 6). The subscales after workday detachment and relaxation showed the highest responsiveness, 67% and 50% respectively.
Table 5. Construct validity (i.e., hypotheses testing) of the DRAW scale expressed by Spearman correlations

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Within workday detachment r</th>
<th>Within workday relaxation r</th>
<th>After workday detachment r</th>
<th>After workday relaxation r</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Need for recovery</td>
<td>0.3 - 0.5</td>
<td>-0.33</td>
<td>-0.32</td>
<td>-0.33</td>
</tr>
<tr>
<td>2. Exhaustion</td>
<td>0.3 - 0.5</td>
<td>-0.22</td>
<td>-0.24</td>
<td>-0.25</td>
</tr>
<tr>
<td>3. Vigour</td>
<td>&lt;0.3</td>
<td>-0.02</td>
<td>-0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>4. Dedication</td>
<td>&lt;0.3</td>
<td>-0.03</td>
<td>-0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>5. Absorption</td>
<td>0.3 - 0.5</td>
<td>-0.15</td>
<td>-0.08</td>
<td>-0.14</td>
</tr>
<tr>
<td>6. &gt; Need for recovery</td>
<td>2.5 &gt; 1</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Confirmed hypotheses per subscale (%)

r: Spearman correlation coefficient; confirmed hypotheses are in bold; + hypothesis on construct validity confirmed; - hypothesis on construct validity rejected.

Table 6. Responsiveness of the DRAW scale expressed by Spearman correlations

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Within workday detachment r</th>
<th>Within workday relaxation r</th>
<th>After workday detachment r</th>
<th>After workday relaxation r</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Need for recovery</td>
<td>0.3 - 0.5</td>
<td>-0.16</td>
<td>-0.16</td>
<td>-0.33</td>
</tr>
<tr>
<td>2. Exhaustion</td>
<td>0.3 - 0.5</td>
<td>-0.06</td>
<td>-0.05</td>
<td>-0.26</td>
</tr>
<tr>
<td>3. Vigour</td>
<td>&lt;0.3</td>
<td>-0.02</td>
<td>0.13</td>
<td>0.10</td>
</tr>
<tr>
<td>4. Dedication</td>
<td>&lt;0.3</td>
<td>-0.01</td>
<td>0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>5. Absorption</td>
<td>0.3 - 0.5</td>
<td>-0.12</td>
<td>0.08</td>
<td>-0.05</td>
</tr>
<tr>
<td>6. &gt; Need for recovery</td>
<td>2-5 &gt; 1</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Confirmed hypotheses per subscale (%)

r: Spearman correlation coefficient; confirmed hypotheses are in bold; + hypothesis on responsiveness confirmed; - hypothesis on responsiveness rejected.

DISCUSSION

The purpose of this study was to evaluate the measurement properties of the Detachment and Relaxation At Work (DRAW) scale; i.e., internal consistency, floor-and ceiling effects, test-retest reliability, construct validity (i.e., hypotheses testing) and responsiveness. The four subscales of the DRAW were internally consistent, reliable and construct validity was moderate (i.e., 67% confirmation of hypotheses). Results for responsiveness for within workday detachment and relaxation were
disappointing (25%), but for after workday detachment and relaxation moderate responsiveness (59%) was found.

**Explanations of results**
All but one subscale (i.e., within workday relaxation) showed good to excellent test-retest reliability. Additionally for reliability, the measurement error was determined in terms of SEM and SDC. The SDC (smallest detectable change) score helps to interpret the change in scores. A change must be larger than the SDC before being considered real (i.e., not due to measurement error) (29; 42). To illustrate, we found a SDC of 1.97 for within workday detachment. This SDC means that a difference of at least 1.97 on within workday detachment is needed to ensure a real change and not one that is due to a measurement error. However, this is the first study on the DRAW scale and therefore we cannot compare the scores with other populations. As a consequence, it is difficult to state how many points of change represent a clinical relevant change. We need to identify a minimal important change (MIC) and to obtain a clinically relevant change, the SDC must be smaller than the MIC (43).

Our data offer support for moderate construct validity. We expected relatively low correlations, because there is no recognized gold standard present to compare the DRAW scale with. Regarding the size of the correlations, the highest correlations for the DRAW scale were found with need for recovery. Because the size of the correlations is not high (i.e., <0.35) there is sufficient ground to state that the DRAW scale is distinct from need for recovery. Also, the small correlations of the DRAW scale (i.e., from -0.02 to -0.35) with exhaustion, vigour, dedication and absorption show that the DRAW measures a different construct.

The present study showed a poor overall responsiveness of the DRAW scale. However, when examining the subscales after workday detachment and relaxation, moderate responsiveness was found (67% and 50%, respectively). An explanation for poor responsiveness could be lack of content validity (i.e., judgement about the relevance and comprehensiveness of the items) (29). Content validity of detachment and relaxation after a workday was assessed in a study (12) with a population of 16 advanced psychology students. However, within workday detachment and relaxation were not assessed. In addition, the population of students in the study might differ from the population of employees in our study.
Strengths and limitations

One of the main strengths of the present study is that all relevant measurement properties of the DRAW scale were investigated at once. For appropriate assessment of a scale, it is important to investigate reliability, validity and responsiveness (37). Another advantage is the large sample size (n=412) that is used, which is in line with the COSMIN manual that recommends a sample size larger than seven times the number of items and at least 100 for executing factor analyses (21). Further, the sample was representative for all employees in the financial service provider, based on the gender and age distribution. Moreover, a strength is that the measurement error was calculated. De Vet and colleagues (2006) advocate that when evaluation is central, researchers should not rely on ICCs solely, but also on the measurement error (SEM/SDC) to avoid misinterpretation of results (33). In addition, test-retest reliability is often poorly assessed, because studies use large time intervals (44) (e.g., 3-4 weeks). For the present study, two measurement points were chosen with only 5-7 days in between (at 12 months and 5-7 days after completion of the 12 months scale). Because within these 5-7 days it is also possible that participants change their detachment and relaxation behavior, two questions on detachment and relaxation were included to control for changes. However, no differences were found, which contributes to the robustness of our results. Ideally, the test-retest reliability measurement was conducted 5-7 days after completing the baseline questionnaire or at the first follow-up. Due to practical issues, it had to be postponed to the follow-up measurement at 12 months. Another contribution is that the responsiveness of the DRAW scale was investigated, which is important for future use of the DRAW scale when evaluating intervention studies. Remarkably, there is a limited amount of studies that paid attention to responsiveness (44;45), while this is very important to assess treatment effects.

There are also several limitations to be mentioned. A weakness is that only self-reported measurements were used, which might give rise to the common method variance. Consequently, the variance in observed scores may be partially explained to a methods effect rather than to a true correlation between scores (46). Another common flaw is that for construct validity, comparator measurements were used that are not measuring the same construct. More evidence for validity is demonstrated
when hypotheses are formulated with comparator measurements that are similar (47).

**Recommendations for future practice and research**

We want to point out several recommendations for future practice and research. The first point concerns the responsiveness. It should be investigated whether the responsiveness is truly low or because of the use of a relatively healthy population (i.e., general health were rated by the participants as good: M=3.8, SD=0.8, on a 5-point scale). Thus, it is recommended for future studies to investigate the responsiveness among a sample of employees with (work-related) physical or mental health problems.

Further, regarding within workday relaxation, this subscale performed the weakest in our evaluation of the measurement properties. The subscale had the lowest internal consistency, test-retest reliability and responsiveness. Therefore, for future studies it is recommended to set up focus groups with employees to judge the content validity of the within workday relaxation items. For now, no inferences based on changes in scores of individuals on the DRAW could be made as no information is available about the MIC. An anchor-based approach is preferred in which an external criterion is formulated to operationalize a MIC (48), e.g., an anchor can be extracted from subjective judgment of employees and/or clinicians or an anchor can be extracted through investigating dose-response relationships (33;43;49).

As this was the first attempt to measure detachment and relaxation within a workday, it is important to expand the research on this topic. It seems that the real key to recovery lies in the types of activities (i.e., low effort activities such as relaxation/sitting quietly or physically tiring activities such as exercising) people engage in during their work breaks. For each of the different types of work breaks studied, it was concluded that activities that removed job demands resulted in more positive outcomes (50). For more comprehensive understanding, questions about various types of work breaks should be added to the DRAW scale.

Another recommendation is to add questions concerning work engagement strategies to the DRAW scale. According to Parkinson and Totterdell’s (51), affect-regulation classification exists of two strategies: diversion (i.e., moving away from the stressor),
and engagement (i.e., ongoing involvement with the stressor). The DRAW scale is based on the diversion strategy (i.e., relaxation and pleasure oriented). Ongoing involvement is likely to disrupt recovery and therefore we did not incorporate items about this strategy. However, research showed that for shorter breaks work-related strategies (e.g., engagement strategies such as answering work-email) are more helpful whereas for longer work breaks non-work related strategies (e.g., diversion strategies such as walking) appear more helpful (10). Based on this, we can argue that work engagement strategies (i.e., problem solving, acceptance, or cognitive re-appraisal) might facilitate recovery in shorter breaks and therefore items that also measure work engagement strategies within work hours should be developed.

Further, the conclusions drawn from the presented data are only valid for the Dutch version of the DRAW and for employees working in a financial service provider setting. It is therefore recommended to investigate the DRAW also in other (cultural) work settings.

**Concluding remarks**

The four subscales of the DRAW were internally consistent, reliable and had moderate construct validity. The responsiveness of the DRAW was moderate. Further studies on responsiveness are necessary. We recommend to be cautious when using the DRAW in intervention studies. Further research is needed to improve the items to enhance responsiveness (i.e., preferably in other populations/intervention studies and elaborating items with work engagement strategies).
REFERENCES


(4) Sluiter JK, De Croon EM, Meijman TF, Frings-Dresen MH. Need for recovery from work-related fatigue and its role in the development and prediction of subjective health complaints. Occup Environ Med 2003;60:i62-i70.


(15) van Hooff MLM. The daily commute from work to home: examining employees’ experiences in relation to their recovery status. Stress Health, 2013 (epub ahead of print).


(22) Coffeng JK, Hendriksen JIM, Dujits SF, Proper Kl, van Mechelen W, Boot CRL. The development of the Be Active & Relax ‘Vitality in Practice’ (VIP) project and design of an RCT to reduce the need for recovery in office employees. BMC Public Health 2012;12:592-617.


