Efficacy of an internet-based problem-solving training for teachers: results of a randomized controlled trial

by Ebert DD, Lehr D, Boß L, Riper H, Cuijpers P, Andersson G, Thiart H, Heber E, Berking M

Internet-based interventions may be a promising approach to address the adverse effects of occupational stress. However, evidence from RCT for internet-based occupational mental-health interventions is scarce. This study may be the first that evaluates an internet-based guided self-help intervention aimed at reducing depressive symptoms among teachers (as well as employees in general).

Affiliation: Leuphana University Lüneburg, Innovation Incubator, Division Health Trainings Online, Rotenbleicher Weg 67, 21335 Lüneburg, Germany. ebert@inkubator.leuphana.de


Key terms: depression; depressive symptom; internet-based; mental health; occupational health; randomized controlled trial; RCT; teacher; training; web-based
Efficacy of an internet-based problem-solving training for teachers: results of a randomized controlled trial

By David Daniel Ebert, PhD,1,2,3 Dirk Lehr, PhD,1,2 Leif Boß, MSc,2 Heleen Riper, PhD,2,4,5 Pim Cuijpers, PhD,5 Gerhard Andersson, PhD,6,7 Hanne Thiart, MSc,2 Elena Heber, MSc,2 Matthias Berking, PhD2,3,8


Objective The primary purpose of this randomized controlled trial (RCT) was to evaluate the efficacy of internet-based problem-solving training (iPST) for employees in the educational sector (teachers) with depressive symptoms. The results of training were compared to those of a waitlist control group (WLC).

Methods One-hundred and fifty teachers with elevated depressive symptoms (Center for Epidemiologic Studies Depression Scale, CES-D ≥16) were assigned to either the iPST or WLC group. The iPST consisted of five lessons, including problem-solving and rumination techniques. Symptoms were assessed before the intervention began and in follow-up assessments after seven weeks, three months, and six months. The primary outcome was depressive symptom severity (CES-D). Secondary outcomes included general and work-specific self-efficacy, perceived stress, pathological worries, burnout symptoms, general physical and mental health, and absenteeism.

Results iPST participants displayed a significantly greater reduction in depressive symptoms after the intervention (d=0.59, 95% CI 0.26–0.92), after three months (d=0.37, 95% CI 0.05–0.70) and after six months (d=0.38, 95% CI 0.05–0.70) compared to the control group. The iPST participants also displayed significantly higher improvements in secondary outcomes. However, workplace absenteeism was not significantly affected.

Conclusion iPST is effective in reducing symptoms of depression among teachers. Disseminated on a large scale, iPST could contribute to reducing the burden of stress-related mental health problems among teachers. Future studies should evaluate iPST approaches for use in other working populations.

Key terms depression; depressive symptom; mental health; occupational health; RCT; web-based

Occupational stress has been implicated as a risk factor for mental health problems, such as depression (1), impaired sleep (2) and early retirement (3). In particular, depression accounts for a considerable burden of disease for individuals, society, and employers (4, 5). Depressive symptoms are highly prevalent in working populations, with a 12-month prevalence rate of 7.1% for men and 6.2% for women (6, 7). Depression is associated with considerable costs due to productivity loss and absenteeism (8, 9), and it negatively affects workplace safety (10).

Recently, internet-based interventions for working populations were introduced as a promising approach to address the adverse effects of occupational stress (11, 12). Internet-based interventions are advantageous because they: (i) are easily accessible at any time and place, (ii) assure anonymity when participants want to avoid stigmatization or self-disclosure in group settings, (iii) allow participants to work at their own pace and review materials as often as they want, and (iv) may reach affected employees earlier than traditional mental

1 Joint first authorship.
2 Leuphana University, Innovation Incubator, Division Health Trainings Online, Lüneburg, Germany
3 Philipps University Marburg, Department of Psychology; Clinical Psychology and Psychotherapy, Marburg, Germany
4 GGZ inGeest, Regional Mental Health Service Centre, VU University Medical Centre, Amsterdam, the Netherlands
5 Department of Clinical Psychology and EMGO Institute for Health and Care Research, VU University, Amsterdam, The Netherlands
6 Department of Behavioral Sciences and Learning, Linköping University, Sweden
7 Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden
8 Institute for Psychology, Division for Clinical Psychology and Psychotherapy, University of Erlangen-Nürnberg, Nürnberg, Germany.

Correspondence to: David Daniel Ebert, Leuphana University Lüneburg, Innovation Incubator, Division Health Trainings Online, Rotenbleicher Weg 67, 21335 Lüneburg, Germany. [email: ebert@inkubator.leuphana.de]
health services, hence preventing the onset of more severe mental health problems (13, 14).

Internet-based interventions have been shown to be effective in clinical applications, including the treatment of depression (15–17), anxiety (18), and sleep disorders (19). In particular, internet-based interventions for depression that include support from a healthcare professional (guided self-help) lead to greater effects and stronger adherence than self-help treatments on their own (15, 17). Internet-based interventions are also effective in changing negative health behaviors (eg, reducing alcohol consumption) (20), although mixed results were obtained for smoking cessation (21). Internet-based psychotherapy has been investigated in a large number of randomized controlled trials (RCT) (16–18). However, only a few interventions have been developed and evaluated to address the specific needs of working populations.

Evidence from RCT for internet-based occupational mental-health interventions is scarce. In what may be the first RCT of an internet-based stress management training, Zetterqvist and colleagues (11) found that an internet self-help intervention could be effective in reducing symptoms of stress, anxiety, and depression. In another study, computerized cognitive behavioral therapy for employees with recent stress-related absenteeism was found to be effective in reducing depression post-treatment compared to conventional care, but it was not found to be effective at the three-month follow-up (22). Ruwaard and colleagues (12) demonstrated that an e-mail-based cognitive behavioral treatment for work-related stress was more effective in reducing stress, depression, and anxiety than a waiting control condition. However, e-mail-based interventions might require a high degree of (health) literacy, and they do not make full use of the advantages of website-driven interventions, such as interactivity and multimedia interfaces, which can improve adherence, efficiency and cost-effectiveness. Another study evaluating a self-help online resilience training for sales managers observed no effects in terms of depression or productivity (23). Lastly, the efficacy of a mindfulness-based intervention at the workplace was investigated by comparing an online virtual to an in-person classroom. Interestingly, that study found no differences regarding the reduction of depression between the two environments. However, the online group had a significantly lower attrition rate compared to the face-to-face intervention (24). To the best of our knowledge, there has been no published study to evaluate internet-based guided self-help interventions in an occupational context with depressive symptoms as the primary outcome.

In contrast to internet-based interventions, traditional in-person stress management interventions have much stronger evidence, as demonstrated in several reviews and meta-analyses (25–27). When designing internet interventions for employees, some challenges and findings from this field should be considered. For example, Richardson and Rothstein (26) analyzed 19 RCT that considered the general working population, in which participants did not have a psychiatric diagnosis or stress-related somatic disorder. One of the main findings was that cognitive-behavioral interventions were more effective than other approaches (eg, relaxation, multimodal, organizational). Interventions with one or two treatment components were more effective than broader interventions with three or more components. Furthermore, shorter interventions were more effective than longer ones. Van der Klink and colleagues (27) observed the strongest effects for anxiety symptoms and the weakest for depressive symptoms. Similarly, Martin and colleagues (25) found that different types of interventions had a positive but small effect on depressive symptoms, which leaves room for improvement in terms of the design of effective treatments to reduce depression in a workplace setting.

Interventions based on problem-solving therapy (PST) have shown to be effective in reducing depression and several other mental health problems (28). According to D’Zurilla and Nezu (29), PST is based on the assumption that ineffective coping behavior causes psychopathology. Adverse health effects, especially depression and the creation of further problems, are expected if a person is not able to resolve stressful problems. PST aims to increase problem-solving skills and facilitate successful problem solving. As a consequence, a reduction in depression is expected. Besides this, worrying about problems should decrease and self-efficacy should increase as participants experience their ability to cope with problems and stressful situations more successfully.

The purpose of this study was to evaluate the efficacy of internet-based problem-solving training (iPST) for teachers with a heightened level of depressive symptoms. We chose this occupational group because school teachers face a high risk of work-related mental health issues and mood disorders (3, 30, 31).

The primary objective was to test whether participating in iPST would lead to a greater reduction in depressive symptom severity compared to a waitlist control group (WLC). We further hypothesized that compared to the WLC group, participants in the iPST group would be more likely to report a reliable change in depressive symptom severity and be classified as being free of symptoms after having completed the iPST. Additionally, we explored whether the iPST group fared better than the WLC group in terms of secondary outcomes, such as perceived stress, worry, problem-solving skills, work-related and general self-efficacy, health-related quality of life, and work absenteeism.
Methods

Study design

We evaluated the efficacy of iPST in a RCT including 150 participants with elevated symptoms of depression. Participants were randomly assigned to the intervention or WLC group. Self-reported outcome assessments were assessed at baseline (T1) and seven weeks (T2), three months (T3), and six months (T4) after randomization. The sample size was calculated to be able to detect a moderate effect at the post-treatment time point based on a power (1-β) of 0.80 in a two-tailed test, α=0.05. Participants who met the study criteria and provided informed consent were randomly allocated to either the iPST or WLC group. The groups were randomized by an independent researcher using a computer-based random integer generator (randlist). Participants of the iPST group received access to the training immediately after randomization, participants of the control group received access to iPST after the six-month follow-up.

The Ethical Committee of the University of Marburg approved all procedures involved in the study which were consistent with generally accepted standards of ethical practice (reference number 2012-06K). The trial was registered at ISRCTN15635876.

Participants and recruitment

Participants were recruited from April 2012 to January 2013 with different recruitment strategies: (i) all over Germany, regional associations of school psychologists and educational institutes for teacher training were contacted and asked to relay information material about the study to potential participants; (ii) an invitation for study participation was posted on several online teacher forums; and (iii) announcements were made at several conferences concerning health promotion for teachers. To be eligible for the study, participants had to meet the following criteria: (i) a score of ≥16 on the Center for Epidemiologic Studies Depression Scale (CES-D) (32, 33), (ii) be a working teacher, (iii) have sufficient German language (reading and writing) skills, (iv) have no notable suicidal risk as indicated by a score of <2 on item 9 of the Beck Depression Inventory (BDI) (34); (2=“I’d like to kill myself”, 3=“I’d kill myself if I had a chance”). To maximize the external validity of the findings, there were no further exclusion criteria. The cut-off score for clinically relevant symptoms of depression in the German version of the CES-D is 23 whereas a cut-off score of 16 has been recommended for indicated prevention settings (32). As we aimed to reach employees suffering from elevated symptoms of depression and not restrict the intervention to employees suffering from mood disorders, we chose a cut-off score of 16. Individuals interested in participating in the study received notification via e-mail and were asked to complete an online screening questionnaire.

Intervention

Internet-based problem-solving training (iPST). The iPST is based on an empirically evaluated Dutch online-based intervention “Alles onder controle” (everything under control) (35, 36). The intervention was translated from Dutch into German and adapted for use by teachers. The iPST is composed of five lessons, in which participants acquire different problem-solving techniques. Additionally, the training includes components for behavioral activation with respect to important values in life. The following components were added to the original program: techniques for coping with rumination, video introductions for each lesson produced by a nationally known expert in mental health training for teachers and example teacher characters who depict the targeted problems and demonstrate implementation of problem-solving techniques in a variety of situations.

The iPST is structured as follows. First, participants describe what really matters to them (eg, values, life-goals). Second, the participants write down their current worries and problems, which are then divided into three categories: unimportant, important but solvable, and unsolvable problems. Third, for each of the three types of problems, a different strategy is developed to either solve or cope with the problem if it is unimportant or unsolvable. The solvable problems are approached by a six-step procedure: describing the problem, brainstorming possible solutions, choosing the best solution, making a plan for carrying out the solution, actually carrying out the solution, and evaluating the success. Unsolvable and unimportant problems are handled by different coping techniques for rumination. Participants were advised to conduct one lesson per week and practice problem-solving techniques between lessons. Within 48 hours, participants received personalized written feedback from an eCoach on the exercises they had completed. The eCoaches were psychologists and trained master’s-level psychology students who followed feedback guidelines according a standardized manual. Guidance was conceptualized according to a theoretical model for providing guidance in eHealth interventions (37). Thus, guidance was aimed at improving adherence to the web-based intervention, and the eCoaches were supervised to ensure that they did not teach techniques that go beyond those of the internet-based intervention.

Waitlist control (WLC). Teachers on the waiting list received no online intervention but had full access to treatment as offered by the workplace occupational health management programs and routine mental health...
services. Additionally, these teachers were granted access to iPST six months after the randomization.

Primary outcome measure

**Depressive symptoms.** The primary outcome was the level of depressive symptoms, as measured by the widely used CES-D (32, 33). This scale consists of 20 items. Subjects rate the frequency of symptoms during the past week on a 4-point Likert-scale (0=rarely, less than one day, 1=sometimes, 2=more often, 3=most of the time, 5–7 days) (eg, “During the past week I felt sad”). The item values can be summarized to a total score that ranges between 0–60. The scale has been shown to be valid, sensitive to change (38), and highly consistent (α=0.87–0.92 in various German samples (32). Cronbach’s α for this study was 0.88.

**Secondary outcome measures**

**General and work-specific self-efficacy.** The 10-item General Self-Efficacy Scale (GSE) developed by Schwarzer and Jerusalem (39) assesses a general sense of perceived self-efficacy, with the goal of predicting the ability to cope with daily problems and adapt after experiencing stressful life events. The participant is asked to evaluate statements on a 4-point Likert-type scale (1=not at all true, 2=hardly true, 3=moderately true, 4=completely true) (eg, “I can typically handle whatever comes my way”). A higher score indicates higher self-efficacy. The item values can be summarized to a total score that ranges between 10–40. Additionally, work-related self-efficacy (WSE) was assessed using the Teacher Self-Efficacy Scale (40). The scale assesses 10 items of perceived self-efficacy with regard to occupational performance. Participants respond to every statement using a four-point Likert-type scale (1=not at all true, 2=hardly true, 3=moderately true, 4=completely true) (eg, “If my lessons are disturbed, I am confident and stay calm”). A higher score indicates higher perceived work-specific self-efficacy. The item values can be summarized to a total score that ranges between 10–40. Both scales have been shown to have good reliability and validity in a number of studies (41). Cronbach’s α for the GSE and WSE, respectively, were adequate (44). In this study, the Cronbach’s α for the PSWQ was 0.90.

**Worrying.** The Penn State Worry Questionnaire (PSWQ) (46, 47) is a commonly used, self-rating measure of (pathological) worrying. The 16 items are directed at the excessiveness, duration, and uncontrollability of worry on a 7-point Likert-type scale (1=almost never, 2=sometimes, 3=often, 4=usually) during the past four weeks. Higher scores on the subdomains of worries, tension, joy, and demands denote a higher level of perceived stress, whereas a higher score for joy denotes a lower level of perceived stress. An average score of all item scores can be calculated. After a linear transformation the overall score ranges between 0–1. The reliability and validity of the scale are adequate (44). In this study, the Cronbach’s α for the PSWQ was 0.95.

**Health-related quality of life.** Health-related quality of life was measured with the SF-12 Health Survey (48). The SF-12 consists of 20 items measuring perceived stressful situations and stress reactions in four subdomains: worries, tension, joy, and demands. The participant is asked to rate how often he or she experienced certain situations or reactions on a 4-point Likert-type scale (1=almost never, 2=sometimes, 3=often, 4=very often, 6=almost always) for events that occurred during the past week, where a higher score indicates more worrying. The item values can be summarized to a total score, varying between 0–96. The scale displayed good internal consistency [α=0.86 (47)]. In this study, Cronbach’s α for the SF-12 was 0.85.
a mental health component score (MCS). Each score ranges from 0–100, with a higher score indicating a higher degree of functioning. Cronbach’s α was 0.80 and 0.68 for the physical and mental health component scores, respectively, in the present study.

**Absenteeism:** Absence from work due to sickness was measured as the self-reported sick leave of participants during the past four weeks (yes/no) and the self-rated amount of total days on sick leave during the past four weeks. Table 1 provides an overview of the measurement tools for all assessments.

**Data analysis**

All analyses were performed according to the recommendations of the CONSORT-Statement for RCT. Analyses were based on intention-to-treat (ITT) procedures.

**Missing data.** Baseline data on primary and secondary outcomes were provided by all participants. Absences from follow-up assessments were imputed using a Markov Chain Monte Carlo multivariate imputation algorithm (missing data module in SPSS 20) with 10 estimations per missing value. To assess potential systematic effects of non-ignorable missing data, pattern mixture analyses for multi-level longitudinal approaches (49) were conducted to test whether the intervention effect was systematically related to missing data. For this purpose, the missing-data pattern of each participant was first coded in 1 of 16 possible missing patterns (2^4: [Missing Yes/No number of assessments]. The created between-subject variable was then included in a three-way interaction (missing pattern×condition×change in symptom severity over time) in the main outcome analyses.

The assumed superiority of the iPST compared to the WLC group was tested with regard to (i) changes in depressive symptom severity and secondary outcomes from baseline (T1) to post-intervention (T2) and follow-up assessments (T3, T4), (ii) number of participants with a reliable change in depressive symptom severity, and (iii) number of participants who reached symptom-free status.

**Change in depressive symptom severity and secondary outcomes.** To assess the differences between conditions in terms of the degree of change in outcomes over time, mixed-effects models (MEM) of change were calculated for the primary and secondary outcome variables. In the models, time was included at level 1 (dummy-coded indicator variables for changes in symptoms between T1–T2, T1–T3, and T1–T4), treatment conditions (0=WLC; 1=iPST) were included at level 2 (ie, participants), and all cross-level interaction effects were included (condition×T1–T2; condition×T1–T3; condition×T1–T4). We imposed no restrictions whatsoever on the covariance matrix; therefore, slopes indicating changes between measurement occasions were allowed to vary freely between participants. All effects were estimated using the full information maximum likelihood (FIML) procedure. T-tests were used to assess differences between groups in terms of absenteeism at T2, T3, and T4. For all outcomes, Cohen’s d (50) was calculated based on the imputed dataset by standardizing the pre-post differences between groups based on the pooled standard deviation of change scores. Moreover, 95% confidence intervals (95% CI) for effect sizes were calculated according to Rosnow and Rosenthal (51). Additional per protocol analyses were conducted based on the sample of participants who adequately adhered to the intervention protocol (by completing at least four out of five intervention sessions).

**Reliable change:** To assess improvements of the primary outcome (depressive symptom severity) on an individual level, we examined the number of participants who displayed a reliable change according to the widely used reliable change index of Jacobsen and Truax (52). Participants were defined as reliably improved if their CES-D-score declined from baseline to post-assessment with a reliable change index >1.96 (8.65 points in the CES-D).

**Absence of subclinical depression:** Because subclinical symptoms of depression are the key risk factor for developing severe depression (53, 54), the number of participants who were free of depressive symptoms was
counted in both conditions. The absence of subclinical depression was predefined as a drop below the cut-off value of 16 on the CES-D.

We also calculated the number needed to treat (NNT) and 95% CI that indicate the number of teachers that needed to participate to generate one additional positive outcome.

All analyses were performed with SPSS version 20 (IBM Corp, Armonk, NY, USA). All reported P-values are two-sided with a significance level of 0.05.

**Results**

**Participants**

Figure 1 summarizes the enrollment and flow of participants throughout the study. Out of 245 individuals assessed for eligibility, 63 (25.7%) were excluded because of a screening CES-D score of >16. Another 30 (12.2%) individuals did not provide informed consent. Two individuals (1%) withdrew from the study before
the baseline assessment and randomization procedure. After these exclusions, 150 (61.2%) teachers were randomized to either the iPST or WLC group.

Missing data
There were 133 (88.7%) participants who completed the online-based questionnaire seven weeks after randomization. There were 121 (80.7%) participants who completed the three-month follow-up, and 127 (84.7%) provided data at the six-month follow-up. The respective numbers of responses for the iPST group were 64 (85.3%), 53 (70.6%), and 61 (81.3%) compared to 69 (92%), 68 (90.6%), and 66 (88%), respectively, for the WLC group. Groups did not differ with regard to missing data (all P>0.10), except for the three-month follow-up (Chi²=9.62, P<0.05). Participants who did not provide data at one of the follow-up assessments did not differ from participants without missing data on baseline depression severity scores or any other baseline characteristics, except for pathological worries (P=0.02). All other differences were not significant (all P>0.10). Little’s overall test of randomness indicated that the missing data were completely random, and thus, multiple imputations of the missing data could be conducted (55). Pattern mixture analyses did not indicate a significant interaction between the pattern of missing data and the primary outcome. Thus, missing data did not appear to bias the results.

Baseline characteristics
As indicated in table 2, the sample consisted of 150 teachers with an average age of 47.1 years [standard deviation (SD)=8.2]. The majority of participants (N=125; 83.3%) were female, and 93 participants (62.0%) were married or in a steady cohabiting relationship. The participants had an average of 19 years of teaching experience (SD=9.6). Only 30 (24.2%) had participated in any traditional mental health training programs before the study. Whereas 55 (44.4%) of the participants had never taken part in a mental health training program or psychotherapy, 61 (48.8%) participants had received psychotherapy in the past. Relatively few participants reported being currently on sick leave (N=7; 4.7%), and the number of total sick days within the last four weeks was also comparatively low (mean 2.2, SD 5.6). Table 3 provides descriptive data for all outcome variables. No meaningful differences were found between the iPST and WLC group for any of the baseline variables, which indicates that the randomization procedure was successful.

Intervention usage
Out of 75 participants in the iPST group, 70 (93%) completed at least one lesson, 62 (83%) completed two lessons, 56 (75%) completed three lessons, and 52 (70%) completed four lessons in the training program. Only 45 (60%) completed all five lessons of the training. Attrition was not significantly associated with any specific session (Chi²=1.67; df=4; P=0.79). Qualitative analyses of problems indicated that both work and non-work problems were reported and addressed during the training. On average, participants reported 4.15 (SD=2.09) different problems; 1.46 (SD=1.40) of the problems were reported as being work-related (eg, lack of esteem from supervisor), 1.97 (SD=1.98) were related to private life (eg, arguing with own kids), and 0.72 (SD=1.34) could be assigned to both domains (eg, difficulties to relax).

### Table 2. Demographic characteristics: means/counts and standard deviations/percentages before the treatment. [iPST=internet-based guided problem-solving training; WLC=waitlist control group; SD=standard deviation]

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>iPST</th>
<th>WLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>47.1 (8.2)</td>
<td>46.4 (9.2)</td>
<td>47.8 (7.3)</td>
</tr>
<tr>
<td>Years of occupational experience (years)</td>
<td>19.0 (9.6)</td>
<td>18.7 (10.4)</td>
<td>19.2 (9.0)</td>
</tr>
<tr>
<td>Weekly working hours</td>
<td>42.0 (14.6)</td>
<td>42.4 (16.9)</td>
<td>41.7 (12.1)</td>
</tr>
<tr>
<td>Days on sick leave during the past four weeks</td>
<td>2.2 (5.6)</td>
<td>2.8 (6.5)</td>
<td>1.6 (4.5)</td>
</tr>
<tr>
<td>Females</td>
<td>125 (83.3)</td>
<td>62 (83.3)</td>
<td>63 (84.0)</td>
</tr>
<tr>
<td>Married/partnership</td>
<td>93 (74.4)</td>
<td>44 (75.9)</td>
<td>49 (73.1)</td>
</tr>
<tr>
<td>Experiences with occupational mental health training</td>
<td>30 (24.2)</td>
<td>17 (29.3)</td>
<td>13 (19.7)</td>
</tr>
<tr>
<td>Experiences with psychotherapy</td>
<td>61 (48.8)</td>
<td>26 (44.8)</td>
<td>35 (52.1)</td>
</tr>
</tbody>
</table>

* Due to missing data, some means/counts refer to a subsample with N=125.

* Due to missing data, some means/counts refer to a subsample with N=120.

* Due to missing data, some means/counts refer to a subsample with N=124.
Table 3. Means and standard deviations of outcome variables at baseline, post-treatment, and at three-month and six-month follow-ups (intention-to-treat sample). [iTST=internet-based guided problem-solving training; WLC=waitlist control group; SD=standard deviation; CES-D=Center For Epidemiological Studies Depression Scale; GSE=General Self-Efficacy Scale; WSE=Work-specific Self-Efficacy Scale; MBI=Maslach Burnout Inventory; EE=emotional exhaustion; DP=depersonalization; PA=personal accomplishment; PSWQ=Penn State Worry Questionnaire; PSQ=Perceived Stress Questionnaire; SF-12-PCS=Short Form Health Survey - physical component summary; MCS=mental component summary]

| Outcome               | iPST T1 Mean | SD | WLC T1 Mean | SD | iPST T2 Mean | SD | WLC T2 Mean | SD | iPST T3 Mean | SD | WLC T3 Mean | SD | iPST T4 Mean | SD | WLC T4 Mean | SD |
|-----------------------|--------------|----|-------------|----|--------------|----|-------------|----|--------------|----|-------------|----|--------------|----|-------------|----|-------------|----|
| WSE                   | 26.95        | 4.32 | 27.43       | 3.98 | 28.57        | 4.78 | 26.79        | 4.12 | 29.41        | 5.11 | 27.79       | 4.45 | 28.63        | 5.37  | 27.45       | 4.62  |
| MBI-EE                | 38.39        | 7.39 | 38.85       | 6.53 | 36.10        | 7.82 | 38.14        | 7.31 | 33.75        | 10.65| 38.63       | 8.22 | 34.51        | 9.84  | 38.01       | 7.52  |
| MBI-PA                | 30.99        | 3.89 | 30.91       | 3.38 | 31.97        | 4.02 | 30.97        | 3.41 | 31.96        | 4.70 | 31.00       | 3.50 | 32.52        | 5.06  | 31.24       | 4.16  |
| PSWQ                  | 58.49        | 16.30| 55.90       | 15.70| 42.86        | 23.01| 53.74        | 19.40| 33.66        | 27.54| 47.57       | 24.46| 39.45        | 26.88 | 50.13       | 21.37 |
| PSQ                   | 0.66         | 0.15 | 0.67        | 0.14 | 0.55         | 0.19 | 0.62         | 0.17 | 0.48         | 0.27 | 0.56        | 0.25 | 0.53         | 0.19  | 0.60        | 0.16  |
| SF12-PCS              | 44.25        | 10.81| 46.76       | 8.76 | 45.48        | 13.97| 45.78        | 10.50| 46.55        | 18.47| 46.08       | 11.25| 47.29        | 11.37 | 45.36       | 9.59  |
| SF12-MCS              | 34.01        | 8.62 | 34.25       | 9.51 | 40.48        | 13.45| 35.72        | 10.20| 44.03        | 21.36| 37.80       | 13.90| 42.59        | 11.83 | 39.55       | 10.35 |
| Absenteeism*          | 0.00         | 0.00 | 0.00        | 0.00 | 0.00         | 0.00 | 0.00         | 0.00 | 0.00         | 0.00 | 0.00        | 0.00 | 0.00         | 0.00  | 0.00        | 0.00  |

Primary outcome analyses – changes in depressive symptom severity

Figure 2 displays the estimated changes in the primary outcome (CES-D) over time based on the ITT MEM analyses. The first row in table 4a lists the results of the MEM analyses in detail. The intercept represents the estimated initial level of depressive symptoms before treatment (T1) in the WLC group. The regression coefficient T1–T2 represents an estimator for change in depressive symptoms between the baseline and seven-week post-treatment assessment in the WLC group. The regression coefficients T1–T3 and T1–T4 represent estimates for changes between the baseline and follow-up assessments after three and six months, respectively, in the WLC group. The regression coefficient for “group” represents the estimated difference between the WLC and iPST groups before treatment. The regression coefficients T1–T2×group, T1–T3×group, and T1–T4×group are estimates of cross-level interaction effects. These coefficients represent estimates of the group differences in “change from baseline to post-treatment”, for the three- and six-month follow-ups, respectively.

There was no significant effect on T1–T2 and T1–T3, indicating no significant change from baseline to post-intervention and three-month follow-up in the WLC group. The significant effect on T1–T4 indicates a small decrease in the severity of depressive symptoms over time. The significant negative interaction effect for T1–T2×group indicates a greater decrease in depressive symptom severity in the iPST compared to the WLC group. Moreover, the larger decrease in the iPST group was maintained at both follow-ups (as indicated by significant negative interaction effects for T1–T3×group and T1–T4×group), although the interaction effects declined slightly over time. Effect sizes were medium to large for the differences between iPST and WLC in “change from baseline to post-treatment”, and the effects were small to moderate for “change from baseline to three- and six-month follow-up” (see table 5).

A separate analysis examining the intervention effects for the per protocol subsample (participants who adhered to at least four training lessons; N=52) was conducted. This per protocol analysis revealed larger effects compared to the ITT sample for changes from baseline to post treatment (d=0.90, 95% CI 0.53–1.27),
to the three-month \((d=0.51, 95\% \text{ CI } 0.16–0.87)\) and six-month \((d=0.58, 95\% \text{ CI } 0.22–0.94)\) follow-up.

Reliable and significant change in depressive symptom severity

The iPST group was superior to the WLC group in terms of the number of participants with reliable improvements in depressive symptom severity from before the treatment to (i) seven weeks after the treatment (iPST: \(N=37 \ [49.3\%]\), WLC: \(N=16 \ [21.3\%]\), \(P<0.01\)), (ii) three months after the treatment (iPST: \(N=35 \ [46.7\%]\), WLC: \(N=25 \ [33.3\%]\), \(P=0.10\)), and (iii) six months after the treatment (iPST: \(N=34 \ [45.3\%]\), WLC: \(N=17 \ [22.7\%]\), \(P<0.01\)). The NNT to achieve one reliable improvement from baseline to post treatment was 3.5 (95% CI 2.3–6.4). This result indicates three teachers had to participate in the intervention iPST to result in one more teacher having reliable improvement in depressive symptom severity compared to the WLC group. At the three- and six-month follow-ups, the NNT values were 7.5 (95% CI 3.4–10.6) and 4.4 (95% CI 2.6–12.5), respectively.

Further analyses indicated that significantly more participants in the iPST group (\(N=44, 58.7\%\)) reached symptom-free status (CES-D<16) seven weeks after randomization compared to the WLC group (\(N=16, 21.3\%\)) (NNT=2.6; 95% CI 1.9–4.3, \(P<0.01\)). After three months, 45 (60%) teachers in the iPST group and 31 (41.3%) teachers in the WLC group reached symptom-free status (NNT=5.3; 95% CI 2.9–33.9, \(P=0.02\)). After six months, 38 (50.7%) teachers in the iPST group and 27 (36%) in the WLC group (NNT=6.8; 95% CI 3.2–10.6, \(P=0.07\)) had passed the cut-off point on the CES-D and thus reached symptom-free status.

Secondary outcomes

Table 4 presents the results of the ITT MEM analyses for changes in secondary outcomes. Significant cross-level interaction effects indicated that there were between-group differences in change from baseline to post-treatment (T1–T2×group) that favored the iPST group for all outcomes, except for the MBI scales (EE: \(P=0.13\); PA: \(P=0.13\); DP: \(P=0.15\)) and the physical health scale of the SF-12 (\(P=0.21\)). Effect sizes for secondary outcomes with significant interactions ranged from \(d=0.36\) (95% CI 0.04–0.69) for perceived stress to \(d=0.63\) (95% CI 0.30–0.96) for worrying (table 5). Moreover, the significant cross-level interaction effects T1–T3×group and T1–T4×group indicate that between-group differences from baseline to the three- and six-month follow-ups favor the iPST group for of the majority of outcomes. Interactions were not significant for changes from baseline to the three-month follow-up in the MBI subdomains of DP (\(P=0.43\)) and PA (\(P=0.22\)), and a trend toward significance was observed for perceived stress (\(P=0.09\)). Changes from baseline to six-month follow-up were only non-significant for the MBI scale of PA (\(P=0.22\)). Effect sizes of significant interactions were moderate, ranging from \(d=0.38\) (95% CI 0.06–0.70) for GSE to \(d=0.62\) (95% CI 0.29–0.95) for worrying at the three-month follow-up. At the six-month follow-up, the effect sizes of significant interactions were small to moderate, ranging from \(d=0.33\) (95% CI 0.01–0.65) for DP to \(d=0.54\) (95% CI 0.21–0.68) for worrying.

Table 4a. Differences in changes of primary and secondary outcomes over time in the mixed-effects models (intention-to-treat sample).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>CES-D</th>
<th>GSE</th>
<th>WSE</th>
<th>MBI-EE</th>
<th>MBI-DP</th>
<th>MBI-PA</th>
<th>PSQ</th>
<th>SF12-PCS</th>
<th>SF12-MCS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>B</td>
<td>SE</td>
<td>B</td>
<td>SE</td>
<td>B</td>
<td>SE</td>
<td>B</td>
</tr>
<tr>
<td>Intercept T1</td>
<td>22.81</td>
<td>1.05</td>
<td>-1.61</td>
<td>1.11</td>
<td>-2.94</td>
<td>a</td>
<td>1.56</td>
<td>-2.90</td>
<td>a</td>
</tr>
<tr>
<td>T1–T2</td>
<td>25.09</td>
<td>0.52</td>
<td>0.90 c</td>
<td>0.45</td>
<td>1.05 c</td>
<td>0.52</td>
<td>1.31 c</td>
<td>0.54</td>
<td>0.05</td>
</tr>
<tr>
<td>T1–T3</td>
<td>27.43</td>
<td>0.47</td>
<td>-0.65</td>
<td>0.44</td>
<td>0.35</td>
<td>0.50</td>
<td>0.02</td>
<td>0.46</td>
<td>-0.50</td>
</tr>
<tr>
<td>T1–T4</td>
<td>38.89</td>
<td>0.80</td>
<td>-0.75</td>
<td>0.69</td>
<td>-0.26</td>
<td>0.82</td>
<td>-0.83</td>
<td>0.89</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>12.03</td>
<td>0.51</td>
<td>0.00</td>
<td>0.47</td>
<td>-0.32</td>
<td>0.52</td>
<td>-0.07</td>
<td>0.46</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>30.89</td>
<td>0.42</td>
<td>0.07</td>
<td>0.40</td>
<td>0.11</td>
<td>0.46</td>
<td>0.35</td>
<td>0.44</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>56.19</td>
<td>1.84</td>
<td>-2.45</td>
<td>2.23</td>
<td>-8.62 b</td>
<td>2.79</td>
<td>-6.06 c</td>
<td>2.60</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>0.67</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.02</td>
<td>-0.11 a</td>
<td>0.03</td>
<td>-0.07 b</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>46.76</td>
<td>1.13</td>
<td>-0.98</td>
<td>1.12</td>
<td>-0.68</td>
<td>1.35</td>
<td>-1.40</td>
<td>1.18</td>
<td>-2.51</td>
</tr>
<tr>
<td></td>
<td>34.35</td>
<td>1.04</td>
<td>1.37</td>
<td>1.33</td>
<td>3.45 c</td>
<td>1.60</td>
<td>5.20 a</td>
<td>1.39</td>
<td>-0.34</td>
</tr>
</tbody>
</table>

\(^a P<0.001\)

\(^b P<0.01\)

\(^c P=0.05\)

\(^d P<0.10\)
Table 5. Between-group effect sizes of primary and secondary outcomes for changes from baseline to post and follow-up assessments (intention-to-treat sample). [iPST=internet-based guided problem-solving training; WLC=waitlist control group; 95% CI=95% confidence interval; CES-D=Center For Epidemiological Studies Depression Scale; GSE=General Self-Efficacy Scale; WSE=Work-specific Self-Efficacy Scale; MBI=Maslach Burnout Inventory; EE=emotional exhaustion; DP=depersonalization; PA=personal accomplishment; PSWQ=Penn State Worry Questionnaire; PSQ=Perceived Stress Questionnaire; SF12-PCS=Short Form Health Survey; PCS=physical component summary; MCS=mental component summary]
a short intervention can produce substantial and enduring effects. Qualitative analyses of participant problems indicated that problems from both the work and non-work domains were reported and addressed during the intervention. This result supports findings emphasizing the importance of non-work determinants for the prediction of a worker’s mental health (56). The changes in depressive symptoms correspond to results from earlier studies of worker-directed interventions, which had effect sizes of $d=0.13$ (24), $d=0.20$ (23), $d=0.40$ (12), and $d=0.82$ (22) when targeting employees with heightened levels of occupational stress. Only a few earlier studies have investigated the intermediate-term effects of occupational interventions; for instance, Grime (22) reported $d=0.60$ after three months and $d=0.30$ after six months. In the present study, the positive effects on depressive symptoms and stress-related outcomes (eg, worrying, work-related and general self-efficacy) were moderate to high in the medium term (after three or six months). The effects of iPST on depressive symptoms exceeded the relatively small effects of traditional occupational health interventions reported in meta-analyses [$d=0.28$ (25) or $d=0.33$ (27)]. In this study, an effect size of $d=0.59$ with regard to the reduction in depressive symptoms from baseline to seven weeks after treatment was observed, which is slightly higher than the reduction observed in other occupational health interventions. It could be that participants with a high level of education display greater responses, and the sample of teachers may have been biased towards individuals with a high level of education. Some studies have found that a high education level is associated with better treatment outcomes in internet-based interventions (57, 58). Other studies (59–61) did not find such an association. Thus, future research should attempt to clarify how education influences treatment outcomes in internet-based self-help interventions.

Our results are also in agreement with studies that evaluated non-worker-directed internet-based self-help interventions for depressive symptoms. For example, in a recent meta-analysis of RCT of computer and internet-based interventions for depressive symptoms, Richard and Richardson (17) found a mean effect size of $d=0.56$. The results of this study also agree with those of studies evaluating the original Dutch version of iPST in the general population, which had a post-intervention effect size on depressive symptoms of $d=0.50$ (35) and $d=0.47$ (36).

The present study found significant effects on secondary outcomes that are explicitly related to the occupational context. Compared to the WLC group, the iPST group displayed greater improvements in work-specific self-efficacy and the burnout dimension of DP, with small to moderate effects, as well as moderate effects on the burnout dimension of EE. No improvements were observed for absenteeism. One possible reason for this lack of improvement for absenteeism may be that the number of participants on sick leave and the reported number of sick days were small for all time points [eg, for the six-month follow-up: iPST=1.18 (4.27); WLC=2.76 (7.75)].

As in many other internet-based interventions focusing on depression and health behavior, women were over-represented in this trial (83.3%), even when considering the actual proportion of female teachers in Germany (70.9%) (62). Future studies should replicate the results in samples with a higher proportion of male participants.

**Study limitations**

The results should be viewed in the context of the study population which was comprised of a sample of highly educated school teachers. The findings may only be valid for populations with comparable demographic and job-related characteristics (high job security, relatively high income, and a high level of job autonomy); the findings may not be applicable to other types of occupations. In addition, the assessments relied exclusively on self-reported impairment at particular time points. Future studies should also include independent outcome evaluation, such as observer-based depression measures or biological outcome indicators. We did not assess treatment-as-usual utilization (eg, psychological or pharmacological co-treatment) of participants during the study period. Thus, we were not able to control for co-treatment in the analysis and cannot rule out the possibility that results are biased. The potential bias could result in an over- or underestimation of the true treatment effect.

It is often assumed that internet-based self-help interventions offer a more cost-effective treatment alternative than traditional face-to-face interventions because they typically require fewer resources (eg, travel time and costs are eliminated for both participants and trainers through the use of self-help materials). However, this study did not measure all necessary costs to conduct proper cost-effectiveness analyses (eg, health-care utilization, presenteeism). Thus, future studies are needed to clarify whether providing low-threshold guided self-help interventions to employees with depressive symptoms indeed offers a good return on investment. Future trials should replicate our findings in other working populations, particularly in populations that include a higher proportion of men or individuals of low socioeconomic status. Future studies should also explore different ways to improve the current intervention in future trials. For example, the current trial only included limited interactive elements, such as videos, diaries and persuasive design elements. A recent systematic review demonstrated that such persuasive elements are associated with higher adherence and may increase the effectiveness.
of an intervention (63). Another possible improvement would be the implementation of an emotion-regulation skill module. At work, people often must cope with difficult situations and solve problems while also addressing the challenging emotions that arise in those situations. Because deficits in the ability to adaptively cope with difficult emotions are related to various mental health problems (64), the use of effective emotion-regulation techniques (65) in low-threshold internet-based stress-management training appears promising (66, 67).

Not all employees may benefit from this particular intervention delivery to the same extent (eg, participants with low internet literacy or with only mild depressive symptoms may not benefit greatly). A recent study found that mild depressive symptoms were associated with lower effect sizes in low-intensity interventions when compared to effects for individuals with at least moderate symptoms (68). Moderator analyses should thus clarify which subgroups of participants benefit the most and the extent to which they benefit from iPST. Future studies should also investigate whether a more targeted approach, based on either the work situation of the employee (eg, cognitive and emotional task demands) or individual risk factors (eg, high effort-reward imbalance), would further improve intervention efficacy. Moreover, future studies should explore potential negative effects of internet-based occupational mental health interventions (69, 70).

Internet-based problem solving was effective in reducing depressive symptoms and other relevant aspects of occupational mental health among teachers. Thus, this study demonstrates that strategies and techniques typically used in traditional occupational health and stress management programs can be successfully adapted to internet-based occupational health programs. Internet-based interventions appear to have several benefits for working populations, particularly for individuals whose occupations would otherwise make interventions costly in terms of time and personal resources. Such interventions are easily scalable, and thus, only a small increase in resources is required to reach a greater proportion of the eligible population. If disseminated on a large scale, such interventions could help to reduce the burden of stress-related health problems for many employees.

Acknowledgments

The European Union funded this study (EU EFRE: ZW6-80119999, CCI 2007DE161PR001). We would like to Claudia Buntrock, Stephanie Nobis, Jo Annika Reins, and Torsten Tarnowski for providing guidance to the participants. We would further like to thank all participating associations of school psychologists and educational institutes for teacher training, as well as all employees who participated in this study.

The authors declare no conflicts of interest.

Trial Registration ISRCTN15635876.

References

11. Zetterqvist K, Maanmies J, Strom L, Andersson G. Randomized controlled trial of internet-based stress...


Internet-based occupational health intervention


Received for publication 25 December 2013