Introduction

As awareness of the importance of workplace mental health issues increases, the importance of stress management programs and their integration in the workplace have become topics of international concern\(^1\). In Japan, where an increasing number of workers have recently shown symptoms of depression\(^2\), mental health-care is needed in the workplace. According to Murphy, such programs, commonly termed ‘stress management’, are individual-oriented and seek to educate workers about the nature of stress, teaching workers specific techniques to reduce the physiological and psychosomatic consequences of stress\(^3\).

Corporate expectations regarding the benefits of stress management interventions include enhanced productivity, lower medical and disability costs, and reduced absenteeism and turnover\(^4\).

With computers and Internet access becoming ubiquitous, public interest in stress management programs using the World Wide Web has been increasing. Such programs are reported to be as effective as conventional programs provided directly by healthcare specialists or counselors\(^5\). Web-based programs are convenient for participants because they can be accessed 24 h a day, anytime and anywhere, and participants receive interactive feedback instantaneously\(^6\). Another reason why Web-based programs are attracting public interest is that from the viewpoint of companies, they can be provided to employees at a low cost and can be useful to make up for the recent
shortage of health-care specialists\(^2\).

Goldenhar \textit{et al.} indicated that for a stress management program to be effective, it is important that the outcome of the program should be evaluated, and that the details of the program implementation process and its elements are examined\(^7\). Evaluating the process will help to resolve the \textit{black box} of how the program operates and is expected to produce feedback that will enable its implementation with greater performance\(^8\).

In the evaluation of the process, the variables mediating the outcome are important. These variables, which include the proximal outcome and the change in direct response to exposure to the program and represent the factors essential to clarifying the mechanism of the program’s effects\(^9,10\). Known mediating variables, including psychological changes as positive \textit{intention} obtained by participation in the program\(^11\), \textit{expectation} from the effects of the program\(^12\), \textit{self-efficacy belief}\(^13,14\), and the acquisition of \textit{knowledge}\(^15\). The formative evaluation implemented using a small sample group before the present study indicated that the participants who assessed the program as \textit{informative} tended to enhance their \textit{self-efficacy belief} and \textit{intention to act}, and that this enhancement influenced the program’s outcome\(^16\).

According to recent findings concerning the evaluation of a program’s effectiveness, it is important to show the program impact theory, which is causal theory, to systematically evaluate it\(^8\). In the program impact theory, the outcome is expressed as a part of the logic ranging from the proximal outcome to the distal outcome; thus it is effective to examine the mechanism that produces either an outcome or no outcome\(^17\).

The purpose of this study was to evaluate its fidelity, in which the intervention was delivered as planned, of the Web-based stress management program created in the development phase by Kawai \textit{et al.}, and to clarify the mechanism of the program’s effect based on the program impact theory (Fig. 1).

\section*{Subjects and Methods}

\subsection*{Design}

The design of this study was a single group pre-, post-test design. After completing the baseline questionnaire (pre-test), participants were given IDs and passwords and asked to access a website at their workplace or home, and to then carry out the programs of sessions 1 through 4 within two weeks (Fig. 2). The period of two weeks was determined based on the results of the development phase. An evaluation questionnaire was given at the end of each session. A post-test was conducted approximately 2.5 wk after the pre-test. All measures and questionnaires were conducted on the web. The Research Ethics Committee of the University of Tokyo reviewed and approved the study protocol.

\subsection*{Subjects}

We recruited participants in cooperation with a business magazine, \textit{P}, in March, April, and May 2006. We placed announcements on the magazine’s Web site, in its semimonthly issues, and in its e-mail messages, which are sent out twice monthly. Two hundred and fifty-one participants had access to the host server of the program and agreed to participate in the survey (Fig. 2). Of this number, 239 completed the pre-test, 168 started on the program, 28 dropped out after session 1 (16.6\% of those who started), 9 after session 2 (5.4\%), and 5 after session 3 (3.0\%). The number completing the program totaled 126 (75\%). A total of 121 people completed the post-test, or 72.0\% of the number that started the program. Those who did not complete the program within the specified two weeks were regarded as dropouts, and were counted among the dropouts after the individual sessions they each had completed.

\subsection*{Program}

The program used in this study aimed to improve psychological well-being as a positive psychological function of humans\(^18,19\) (Table 1). The program consisting of four sessions, provided with modifications, to incorporate the knowledge obtained in the development phase\(^16\). All sessions started with session’s objectives, and to gain the attention of the participant graphics and music were used. The web pages also contained some quizzes to assess participant’s stress level. A part of this program used cognitive behavior therapy, and the results of the formative evaluation implemented in the development phase before this study confirmed the potential effectiveness of the program\(^16\).

Although well-being has conventionally expressed the feelings of happiness and satisfaction, it is known from previous studies that Ryff’s psychological well-being is a psychological function formed by the environmental factors and experiences of the person and is useful in predicting the mental and physical conditions of that person\(^18\). Measuring these parameters can be effective in understanding the resilience of a person encountering an adverse situation\(^19\) and a few researchers have reported the effectiveness of psychological well-being therapy that targeted people with mood and anxiety disorders\(^20,21\).

\subsection*{Measurement items}

\textbf{Basic attributes}

Gender, age, educational background, occupation, size of employer, the frequency of using a personal computer, and whether stress has been experienced within the
Evaluation of the sessions

Because any Web-based program needs to be carried out alone, it is important to encourage participants to continue the program. The present one was created after the ARCS model, which Keller called an indicator of how encouraging learning materials can be for those who use them\textsuperscript{22, 23}). According to ARCS model, learning material should be found interesting by the learner (attention), evaluated as useful for the learner’s life and as worth doing (relevance), increase the learner’s self-efficacy belief (confidence—"this can be done"), and provide the learner with the feeling that it is good that he or she has completed it (satisfaction). The ARCS evaluation factors are similar to the psychological changes identified as the mediating variables of a program outcome in previous studies, such as positive feelings obtained by participation in the program (intention)\textsuperscript{24}), hope regarding the

past week.

Fig. 1. The Program Impact Theory (above) and Hypothetic model (below).

The changes in psychological well-being were calculated by subtracting T1 points from T2 points (post-test). Greater changes indicate a greater improvement. The changes in CES-D scores were obtained by subtracting T2 points from T1 points. The higher scores indicate a greater improvement in depression.

Fig. 2. Flow of the intervention.
effects of the program (expectation), and feelings of self-efficacy (self-efficacy belief).

Based on the above discussion, this study used the following four items, as mediating variables, as well as for evaluation of the sessions. Responses were solicited on a five-point scale, ranging from “strongly agree” (5 points) to “strongly disagree” (1 point). The higher the score, the more positively the program was evaluated.

- Informative—Do you believe this session was useful to increase your stress-coping ability?
- Enjoyment—Did you enjoy working on this session?
- Intention to act—Did you feel like you wanted to increase your stress-coping ability by working on this session?
- Self-efficacy belief—Did you believe you could increase your stress-coping ability by working on this session?

Psychological well-being

Ryff’s PWB scale was used to measure psychological well-being. This is a self-rated 43 item inventory that covers six dimensions of well-being: Autonomy (8 items), Environment Mastery (6 items), Personal Growth (8 items), Positive Relations with Others (6 items), Purpose in Life (8 items), and Self-acceptance (7 items). Responses were solicited on a five-point scale, ranging from ‘strongly disagree (1)’ to ‘strongly agree (6)’ (total score range =43–258).

Depression

We used the 13-item version of the Center for Epidemiologic Studies Depression Rating Scale (CES-D) (range = 0–39). This scale consists of the 12-item abridged version of the 20-item scale by Radloff and one positive item. We decided to use the 13-item version because Kinoshita indicated that it was more desirable after examining the reliability and validity of the 12-item version.

Analysis method

For analysis, we used SPSS™ 12.0 for Windows in combination with Amos™ 5.0.

To statistically assess the average values of independent statistical populations, the $\chi^2$ test was used when the variable was binary, and the Mann-Whitney U test was used when it was continuous. Wilcoxon matched-pairs signed-rank test was used to reflect the evaluation of the sessions by the participants. To compare the average values of the pre-intervention and post-intervention program, paired $t$-test was used. To consider the possible mechanism of the intervention, which would reflect the causal relationship of program effects, we built a hypothetic model based on the program impact theory and conducted path analysis. This method was based upon a study conducted by Nowacek et al. In this model, the sessions evaluation shows the total of the five-point scores of four questions of all sessions. The higher scores indicate that the participants were more affirmative. The changes in psychological well-being were calculated by subtracting T1 points (pre-test) from T2 points (post-test). Greater changes indicate a greater improvement. The changes in CES-D scores were obtained by subtracting T2 points (post-test) from T1 points (pre-test). The higher scores indicate a greater improvement in depression.

To test the goodness of fit of the model, we used the $\chi^2$ test, RMSEA (root mean square error of approximation, 0.05 or smaller), CFI (comparative fit index, 0.95 or greater), and NFI (normed fit index, 0.95 or greater).
Results

Basic attributes of program participants

Table 2 presents the basic attributes of the 168 people who started on the program. The average age was 39.3 ± 8.7 yr, most of whom were male. Of these, 93.5% responded that they used personal computers every day, which indicated they had the equipment and skills to work on a Web-based program. A CES-D score was 10.3 ± 8.3, which was significantly lower than the 71 who answered the pre-test but did not participate in the program ($p<0.05$). We also found that 16.7% of the 168 who started answered ‘yes’ to the experience of stress, was significantly less than the 36.6% who did not participate in the program ($p<0.05$).

Participant evaluations of the sessions

The observed evaluation score for each session are shown in Table 3. The median was 4 to 5 points and the interquartile range was 0.0 to 1.0 for all items (range=1–5). For the informative and self-efficacy belief dimension, the first quartile was 4 points in session 2, session 3 and session 4, and the observed evaluation score was statistically significant increased between session 1 and other sessions. For the item enjoyment, the first quartile was 3 points in session 2 and session 4, and the session 1 was evaluated most affirmatively by participants. For the item intension to act, the first quartile was 4 points in all sessions and no significant differences were found between session 1 and other sessions. Additionally, moderate to large correlations according to Cohen’s criteria were found among all evaluation measures. No significant difference were found in the total scores of four questions between session 1 and others.

Figure 3 presents the average values of psychological well-being and CES-D scores of the pre-, and post-test. The psychological well-being scores showed positive and statistically significant increase ($p=0.000$). In terms of the CES-D scores, no significant differences were found between before and after intervention; however, the scores after intervention were lower than those before intervention.

The characteristics of the dropouts

Those who dropped out after session 1 had an average CES-D score of 13.1 ± 8.9, which was significantly higher than that of those who continued (Table 4). No differences were found between the dropout group and the group that continued the program in terms of participant attributes or their evaluations of the session after experiencing session 2 and later sessions.

The mechanisms of program effect

The model depicted in Fig. 1 was tested and represented an acceptable, but not good, model-data fit ($\chi^2 (df)=0.758 (3), p=0.110$, RMSEA=0.07, CFI=0.986, NFI=0.965). The paths from before-intervention depression (CES-D) to participants’ evaluations about the session and changes in psychological well-being was not statistically significant, we fixed these paths zero and the model provided a better fit. Figure 4 represents the final results of the hypothetic model. The figure shows only the statistically sig-
significant paths. Whether the participants had experienced stress produced no significant path and is not shown in this figure. The gender is not shown for the same reason. Based on the following figures, the goodness of fit of the model was satisfactory: \( \chi^2 (df)=0.592 \) (4), \( p=0.441 \), RMSEA=0.000, CFI=1.00, NFI=0.998). Participant evaluations about the sessions shows the total of the five-point scores of all sessions. Higher scores indicate that the participants were more affirmative (range=16–80). The changes in psychological well-being were calculated by subtracting the before-intervention points from the after-intervention points. Greater changes mean that the psychological well-being had greater improvement. The CES-D scores were obtained by subtracting the after-intervention points from the before-intervention points. The higher scores indicate a greater improvement in depression.

Participants with higher psychological well-being before intervention, evaluated the sessions more positively (\( \beta=0.41, p<0.001 \)). No direct influence of depression prior to intervention (CES-D T1) was found in the participant evaluations of the sessions and changes in psychological well-being. Participant evaluations of the sessions directly predicted changes in psychological well-being (\( \beta=0.29, p<0.01 \)), and changes in the psychological well-being predicted changes in the CES-D (\( \beta=0.50, p<0.001 \)).

Discussion

As a result of the evaluation of the process of a Web-based stress management program that is being implemented to improve psychological well-being, the fidelity and the mechanism of the program effect were confirmed. Moreover, several important results were obtained to allow greater performance in implementation of the program. First, the majority of participants evaluated the program affirmatively after all sessions, and no significant differences were found in the evaluations between the dropout group and the group of those who continued the

---

**Table 3. The sessions evaluation**

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=168</td>
<td>n=140</td>
<td>n=131</td>
<td>n=126</td>
</tr>
<tr>
<td>Informative</td>
<td>median</td>
<td>median</td>
<td>median</td>
<td>median</td>
</tr>
<tr>
<td></td>
<td>4 1.0</td>
<td>4 0.0</td>
<td>4 0.0</td>
<td>4 0.0</td>
</tr>
<tr>
<td></td>
<td>(3, 4)</td>
<td>(4, 4)</td>
<td>(4, 4)</td>
<td>(4, 4)</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>median</td>
<td>median</td>
<td>median</td>
<td>median</td>
</tr>
<tr>
<td></td>
<td>4 1.0</td>
<td>4 1.0</td>
<td>4 0.0</td>
<td>4 0.0</td>
</tr>
<tr>
<td></td>
<td>(4, 5)</td>
<td>(3, 4)</td>
<td>(4, 4)</td>
<td>(4, 4)</td>
</tr>
<tr>
<td>Intensity to act</td>
<td>median</td>
<td>median</td>
<td>median</td>
<td>median</td>
</tr>
<tr>
<td></td>
<td>5 1.0</td>
<td>4 1.0</td>
<td>4 1.0</td>
<td>4 1.0</td>
</tr>
<tr>
<td></td>
<td>(4, 5)</td>
<td>(4, 5)</td>
<td>(4, 5)</td>
<td>(4, 5)</td>
</tr>
<tr>
<td>Self-efficacy belief</td>
<td>median</td>
<td>median</td>
<td>median</td>
<td>median</td>
</tr>
<tr>
<td></td>
<td>4 1.0</td>
<td>4 0.0</td>
<td>4 0.0</td>
<td>4 0.0</td>
</tr>
<tr>
<td></td>
<td>(3, 4)</td>
<td>(4, 4)</td>
<td>(4, 4)</td>
<td>(4, 4)</td>
</tr>
<tr>
<td>Total score</td>
<td>median</td>
<td>median</td>
<td>median</td>
<td>median</td>
</tr>
<tr>
<td></td>
<td>16 3.0</td>
<td>16 2.0</td>
<td>16 3.0</td>
<td>16 2.0</td>
</tr>
<tr>
<td></td>
<td>(15, 18)</td>
<td>(15, 17)</td>
<td>(15, 18)</td>
<td>(15, 17)</td>
</tr>
</tbody>
</table>

\( IQR(\text{Inter Quartile Range})=Q_{25}-Q_{75} \)

Wilcoxon matched-pairs signed-rank test.

\( p^*=\text{session 1 vs. session 2} \)

\( p^1=\text{session 1 vs. session 3} \)

\( p^2=\text{session 1 vs. session 4} \)

Each analysis was done by the number of people with correspondence.

\(* p<0.05 \), \(** p<0.01 \), \(*** p<0.001 \), \( ns = \text{non significant} \).
program. These findings suggest that this program satisfies the requirements as a self-study tool following the method of the ARCS model. The number of participants who evaluated the self-efficacy belief positively increased after session 1. This result might indicate that the negative cognition of participants may have switched to positive cognition as they proceeded with the program. High intension to act scores were reported by the majority of participants, suggesting that the program effectively motivated participants to take measures to improve their psychological well-being. As for the evaluation of the item enjoyment, session 1 was the most positively evaluated session. Previous studies indicated that enjoying the program can be a motivation factor to increase self-efficacy belief\(^\text{14}\). Therefore it is possible that the evaluation of enjoyment may trigger a rise in the evaluation of self-efficacy belief.

Second, participants who dropped out after session 1 had significantly higher depression score. This tendency was also observed among participants who completed the pre-test but did not undertake the program. Previous studies that implemented Web-based programs for those in depressive conditions indicated that a Web-based program had the benefit that it was easier for them to participate in than a face-to-face program\(^\text{31–33}\). However, the results of the present study indicate that participants in a depressive condition are easily discouraged from continuing the program, or they tend to feel confused about the program, despite the fact that they might be interested in it. Furthermore, 25.0% of the participants dropped out before completing the program. In a Web-based antismoking program for college students, participants were recruited through magazines and street posters, and the program process was evaluated similarly. Of those participants, 57.1% dropped out before completing the program\(^\text{34}\). This suggests that although the dropout rate of the present program needs to be examined in the future, the rate of dropouts was not extremely high.

### Table 4. The characteristics of dropouts

<table>
<thead>
<tr>
<th></th>
<th>Session 1 (N=168)</th>
<th>Session 2 (N=140)</th>
<th>Session 3 (N=131)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dropouts</td>
<td>Continued Group</td>
<td>Dropouts</td>
</tr>
<tr>
<td></td>
<td>n=28</td>
<td>n=140</td>
<td>n=9</td>
</tr>
<tr>
<td>Psychological well-being</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Psychological well-being</td>
<td>181.0 ± 28.8</td>
<td>184.9 ± 29.2</td>
<td>192.2 ± 25.7</td>
</tr>
<tr>
<td>CES-D</td>
<td>13.1 ± 8.9</td>
<td>9.7 ± 8.9*</td>
<td>6.0 ± 6.2</td>
</tr>
<tr>
<td>Age</td>
<td>37.8 ± 8.0</td>
<td>39.7 ± 8.8</td>
<td>40.4 ± 8.8</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Male</td>
<td>18 (64.3)</td>
<td>99 (70.7)</td>
<td>7 (77.8)</td>
</tr>
<tr>
<td>Female</td>
<td>10 (35.7)</td>
<td>40 (28.6)</td>
<td>2 (22.2)</td>
</tr>
<tr>
<td>Un-answered</td>
<td>-</td>
<td>1 (0.7)</td>
<td>-</td>
</tr>
<tr>
<td>Experienced stress</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>6 (21.4)</td>
<td>22 (15.7)</td>
<td>1 (11.1)</td>
</tr>
<tr>
<td>No</td>
<td>22 (78.6)</td>
<td>115 (82.1)</td>
<td>8 (88.9)</td>
</tr>
<tr>
<td>Un-answered</td>
<td>-</td>
<td>3 (2.2)</td>
<td>-</td>
</tr>
<tr>
<td>Program evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informative AG</td>
<td>17 (60.7)</td>
<td>87 (62.1)</td>
<td>8 (88.9)</td>
</tr>
<tr>
<td>NG</td>
<td>11 (39.2)</td>
<td>53 (37.9)</td>
<td>1 (11.1)</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>AG</td>
<td>22 (78.6)</td>
<td>124 (88.6)</td>
</tr>
<tr>
<td>NG</td>
<td>6 (21.4)</td>
<td>16 (11.4)</td>
<td>2 (22.2)</td>
</tr>
<tr>
<td>Intension to act</td>
<td>AG</td>
<td>27 (96.4)</td>
<td>135 (96.4)</td>
</tr>
<tr>
<td>NG</td>
<td>1 (3.6)</td>
<td>5 (3.6)</td>
<td>0</td>
</tr>
<tr>
<td>Self-efficacy belief</td>
<td>AG</td>
<td>15 (53.6)</td>
<td>101 (72.1)</td>
</tr>
<tr>
<td>NG</td>
<td>13 (46.4)</td>
<td>39 (27.9)</td>
<td>1 (11.1)</td>
</tr>
</tbody>
</table>

\(^*p<0.05\)  
AG=affirmative group, NG=negative group.
Affirmative group includes participants who answered ‘strongly agree’ and ‘agree’.
Negative group includes participants who answered ‘strongly disagree’, ‘disagree’ and ‘neither’. 

\(^\chi^2\) tests for binary data. Mann-Whitney U test for continuous data.
As a result of path analysis, based on the impact theory, the path from the participants’ evaluation of the sessions to changes in psychological well-being was significant, confirming that the evaluation of the program was a mediating variable of the outcome. This finding agrees with the knowledge obtained in previous studies that psychological changes and enjoyment are the mediating variables of program effects. The path from depression before intervention was insignificant with regard to changes in participant evaluations of the sessions or psychological well-being. Previous studies indicated that the stress level before intervention could easily influence the program effects. However, such knowledge must be used for outcome evaluation and examined in comparison with the control group. The path from changes in psychological well-being to changes in the CES-D was significant. This result reflects that psychological well-being is an internal resistance resource to cope with stress, following the concept of salutogenesis developed by Antonovsky.

It should be noted that this study has several limitations that must be discussed along with future directions. First, the study implemented only quantitative examinations and did not interview participants. Therefore, we couldn’t know the real reasons why participant dropped out. When participants placed inquiries by e-mail, we handled them appropriately (for example, screens or sounds might not have started, depending on the system configuration of the personal computer). It is still possible that something unpredictable would occur before program implementation and that these factors were not sufficiently addressed, or that participants dropped out because of an unanticipated factor. It is desirable to send a simple questionnaire to dropouts where possible to obtain knowledge for the future, thus improving the program as much as possible. Second, we recruited participants through a business magazine. Almost all of the participants were white-collar workers, relatively healthy and had high educational backgrounds, meaning that careful consideration is necessary before generalizing the results of this study. Third, of the 251 people who agreed to participate and registered in the program, only 121 (48.2%), fewer than half, completed the program up to the post-test. Those participants who were analyzed may represent a group with high interest in stress management or were more highly motivated among all those who registered. Forth, this study was based on a single group pre-, post-test design, and there was no control group. Therefore, no firm conclusions of the effectiveness of the program can be drawn from the results, and randomized controlled studies will be required to outcome evaluation to be conducted in the next stage.

**Conclusion**

This study has evaluated the process of implementing a Web-based program aiming to improve psychological well-being, originally developed by Kawai et al. As a
result, it was found that the participants whose evaluation of the program sessions was more positive achieved greater psychological well-being improvement and relieved depression. More specifically, the program can be effective if the participants feel enjoyment, find the program informative to increase their ability to cope with stress, and increase their self-efficacy belief that they want to increase this ability and that the program will enable it.

Acknowledgements

This research was partially supported by Grant in Aid for Scientific Research for 2009: Foundation Research (Topic number 21243033) the Ministry of Education, Culture, Sports, Science and Technology.

References

25) Bunce D (1997) What factors are associated with the outcome of individual-focused worksite stress manage-


