UNIVERSITEIT VAN AMSTERDAM
Faculty of Psychology

Evaluatie studie

Columbus en Mandela training
Corinor BV

Periode van uitvoering: herfst 2011- herfst 2012
Rapportage: maart 2013

Contact:
Dr. Kai J. Jonas
Sociale Psychologie
Universiteit van Amsterdam
Weesperplein 4
1018XA Amsterdam
k.j.jonas@uva.nl
020 525-6888

UNIVERSITEIT VAN AMSTERDAM
Faculty of Psychology

Executive summary

The Columbus and Mandela training supplied by Corinor BV were evaluated during a one year time course in both a waiting control group design and a longitudinal assessment on 5 central variables (positive and negative mood, self-efficacy, work locus of control and personal abilities), based on the training goals, and 8 exploratory variables (assertiveness, introspection and clinical symptom questionnaires). In total 177 participants of the training took part in the evaluation.

The general picture reveals a set of trainings that provides positive effects for the participants during and closely around the training time-span. The data shows that most of the effects occur with the Columbus training and that the Mandela training serves as a stabilizing function. For the 5 central variables significant long-term effects are found (with the sole exception of positive mood, for which only a short-term effect could be obtained). The set of exploratory variables revealed a mixed pattern of results, but given the training no effects were predicted here. In fact, the trainings showed a potential for enhanced effects also on these exploratory variables. The results of the control group study fit with this picture.

The evaluation also revealed that two personality factors, namely conscientiousness (negatively) and neuroticism (positively) overall contribute to training success. The training is more successful for females on burnout indices. Also motivation and introspection contribute to training success and should be included in the future to be able to provide more tailor-made trainings.

In sum, the Columbus and Mandela trainings are effective trainings in the context of the improvement of individual well-being and self-determined abilities to work. The fact that such an evaluation trajectory was embarked on is a very noteworthy positive sign and in and of itself a statement of quality management of the training products.

UNIVERSITEIT VAN AMSTERDAM
Faculty of Psychology

Executive summary

De Columbus en Mandela training van Corinor BV zijn over een tijdsperiode van een jaar geëvalueerd in een longitudinale meting gecombineerd met een controlegroep die de training afwachtte. Er zijn 5 centrale variabelen gemeten (positieve en negatieve stemming, zelf-effectiviteit, werk-gerelateerde beheersingsoriëntatie en vaardigheden op persoonlijk gebied) die direct aan de doelen van het training gerelateerd zijn. Bovendien zijn er 8 exploratieve variabelen (assertiviteit, introspectie en klinische symptoom vragenlijsten) afgenomen. In totaal hebben er 177 deelnemers van de training mee gedaan.
Over het algemeen laat de evaluatie zien dat de trainingen positieve effecten hebben voor de deelnemers, zowel tijdens als na de training. Het onderzoek laat ook zien het grootste deel van de effecten door de Columbus training worden gedreven, en dat de Mandela training de effecten stabiliseert. Voor de 5 centrale variabelen zijn er significante lange-termijn effecten gevonden (met uitzondering van positieve stemming, waarvoor alleen een korte-termijn effect is gevonden). Voor de exploratieve variabelen zijn er gemengde resultaten gevonden, hoewel er ook geen verwachtingen waren geformuleerd voor deze exploratieve variabelen. Toch suggereert de uitkomst van dit onderzoek dat de trainingen mogelijk ook effect hebben op de exploratieve variabelen. De uitkomsten van de metingen bij de controle groep ondersteunen de conclusie dat de trainingen effectief zijn.
De evaluatie laat ook zien dat twee persoonlijkheidseigenschappen, namelijk nauwgezetheid (op negatieve manier) en neuroticisme (op positieve manier) bijdragen aan het succes van de trainingen. De training is bovendien succesvoller voor vrouwen met burnout indicatoren. Daarnaast dragen motivatie en introspectie bij aan het succes van de training, zodanig dat deze factoren zouden moeten worden meegenomen in toekomstige, meer op het individu afgestemde, trainingen.
Samenvattend zijn de Columbus en Mandela trainingen effectieve trainingen voor het vergroten van het individuele welzijn en zelf-effectieve vaardigheden op het werk. Het is ook belangrijk om op te merken dat het feit dat Corinor BV de aanzet heeft gegeven voor dit evaluatie-traject aangeeft dat dit bedrijf ernaar streeft hoogwaardige trainingen te ontwikkelen en aan te bieden.

UNIVERSITEIT VAN AMSTERDAM
Faculty of Psychology

Results in Detail

Design and Sample

To evaluate the effects of both the Columbus and the Mandela training a waiting control group design was employed. In the experimental group four points of measurement were realized: before the Columbus training (T1, app 2 weeks before the training), after the Columbus training (T2, three weeks after), after the Mandela training (T3, three weeks after) and a final measurement T4 (3-4 months after the Mandela training). The T2 measurement also served as a pre-measure for the Mandela assessment. A control group was assessed out of participants who where planned to attend a Columbus training with two points of measurement, T1 and T2 matching in time distance with the experimental group.

In total the measurement period for the experimental group covers at minimum 6 months, at maximum 9 months, depending on when participants took part in the Mandela training (participants could choose to not follow both training sequentially but opt for a subsequent training). Participant who did not take part in the Mandela training at all where called up for the T4 measurement as if the had followed the Mandela training, i.e. 4 months after the completion of the Columbus training.

24 participants took part in the control group, 15 males and 9 females. Mean age was 42, range 32-58. Mean work experience was 22 years, range 7-35. Education level was 16,7% VMBO, 20,8% MBO, 25% HBO and 25% WO, with the rest adding up to 100% defined as “other”.

In the experimental group at T1 there are 124 valid cases, 59 males, 65 females, mean age was 40, range 23-62. At T2 there are 153 valid cases, 67 males, 86 females, mean age was 39, range 23-62. At T3 there are 112 valid cases, mean age 40, range 23-60 and at T4 there are 75 valid cases, mean age 40, range 25-61. Based on the valid cases there is an attrition of 40 – 51% depending on the reference n (124 or 153). In both cases the level of attrition is moderate, compared to other longitudinal training evaluations.
Mean work experience was 18 years, range 2-44. Education level was 9,7% VMBO, 12,4% MBO, 29,6% HBO and 11,8% WO, with the rest adding up to 100% defined as “other”.
There were no differences in terms of education level, work experience and age in terms of gender.

The control group does not fully match the experimental group in terms of work experience, and education. Yet at the same time, it is a randomly chosen group of "waiting" participants, with a similar motivational background.

At T2 the n is higher since participants from previous Columbus trainings could take part in this training, but were not measured previously. In addition, some participants who did not provide answers at T1 chose to do so at T2. Actual cell sizes in the repeated measures analyses may vary since not all participants took part in all measurement points. In most cases an overall cell size of 54 cases over all measurement points (T1-T4) could be reached and this constitutes a suitable sample size for the desired analyses.

Measures

In the following all dependent measures will be reported with their respective internal reliability measures for all measurement points.

General self-efficacy questionnaire (Chen, Gully & Eden, 2001) was measured on a 5-point Likert type scale with 1 for "zeer oneens" and 5 for "zeer eens". Internal reliability was good overall, at T1 (α = .869), T2 (α = .893), T3 (α = .907), and T4 (α = .897).

Assertiveness was measured with a subset of the scale by Lorr and More (1980) on a 5-point Likert type scale with 1 for "helemaal niet" and 5 for "helemaal wel". Since the two subscales, independence and defending ones rights and interest, did not form reliable scales, the two scales were used as one scale. Still internal reliability was moderate at all points of measurement, T1 (α = .696), T2 (α = .591), T3 (α = .729), and T4 (α = .644).

Burnout symptoms were assessed with the *Utrechtse Burnout Schaal* (UBOS; Schaufeli & van Dierendonck, 2000, from : Schaufeli et al., 2008) and measured on a 7-point Likert type scale, with 0 for "never" and 6 for "always". There are three subscales in the instrument, emotional depletion, T1 (α = .875), T2 (α = .892), T3 (α = .915), and T4 (α = .906), depersonalization, T1 (α = .771), T2 (α = .796), T3 (α = .844), and T4 (α = .790), personal abilities, T1 (α = .759), T2 (α = .848), T3 (α = .832), and T4 (α = .887).

In addition, to assess further symptoms, we used the *4 Dimensional Symptoms Questionnaire* (4DKL, Terluin et al., 2004). The scale consists of four subscales and was measured on a 5-point scale from 1 for "not at all" and 5 for "very often or continuous". The distress scale has a very internal reliability, T1 (α = .93), T2 (α = .914), T3 (α = .911), and T4 (α = .939). The same holds for the somatization subscale, T1 (α = .861), T2 (α = .818), T3 (α = .829), and T4 (α = .886). The internal reliability of the depression subscale is good for T1 (α = .901), T2 (α = .898), and T4 (α = .946), but not for T3 (α = .274). The reason for this drop

in reliability at T3 remains unclear. There were no problems for the internal reliability of the fear subscale, T1 (α = .821), T2 (α = .789), T3 (α = .840), and T4 (α = .954).

To assess positive and negative affect, the Positive Affect Negative Affect Scale (PANAS; Watson, Clark & Tellegen, 1988) was used, employing a 5-point Likert type scale, with 1 for "geheel niet" and 5 for "absoluut wel". The positive affect subscale shows good internal reliability, T1 (α = .857), T2 (α = .857), T3 (α = .9), and T4 (α = .883), as does the negative affect subscale, T1 (α = .874), T2 (α = .88), T3 (α = .903), and T4 (α = .884).

Work Locus of Control was measured with a translation Work Locus of Control Scale from Spector (1988; see Siu & Cooper, 1998; Strauser et al., 2002). The 5-point Likert type scale ranged from 1 for "do not agree at all" to 5 for "fully agree". Internal reliability was good: T1 (α = .741), T2 (α = .807), T3 (α = .857), and T4 (α = .852).

We also assessed level of introspection with a scale of 23 items (sample "Ik sta vaak bij de wijze stil hoe ik dingen doe"). It was assessed on a 5-point Likert type scale, with 1 for "very untypical for me" to 5 for "very typical for me". Internal reliability was good at all measurement points, T1 (α = .827), T2 (α = .807), T3 (α = .812), and T4 (α = .782).

A Big Five (BFI-10) personality measure (Rammstedt & John, 2007) was assessed, too. The five constructs (Extraversie, Meegaandheid, Zorgvuldigheid, Emotionele Stabiliteit en Openheid voor Ervaringen) were measured with two items each, with one items always being reverse coded. Measurement took place on a 5-point Likert-type scale (1= "Helemaal wel", 5 = "Helemaal niet"). Correlation of the items was in most cases satisfactory (Extraversion r = .523, Agreeableness r = .26, Conscientiousness r = .145, Neuroticism r = .556, Openness, r = -.019)

Regulatory Focus (RFT), was measured with the Regulatory Focus Questionaire by Semin et al. (2005, based on Higgins et al., 2001). The scale has two subscales Prevention Focus ("gericht op het voorkomen van verlies"), and Promotion Focus ("gericht op het behalen van winst"). The internal reliability of the scales was satisfactory (Promotion α = .697 and Prevention α = .739).

Procedure

Participants were informed about the evaluation procedure by Corinor and received an invitation email. Previous to all measurement points participants received an invitation email again. The time window for measurement was 7 days and was sometimes, to account for public holidays or vacation periods, extended to 14 days.

Results

The results section is split up in three parts. Part 1 contains the analysis of the waiting control group design. Part 2 reports the results of the longitudinal analysis and Part 3 contains the analysis of intraindividual determinants on the training effect.

Level of significance is reported in p-values (with $p < .05$ as test level) and partial eta squared (η_p^2) as indicator of effect size (0 for no effect, 1 for maximum effect size). In regressions the standardized β is reported as indicator (again ranging from 0 to +/- 1). Correlations are reported with the r parameter, again ranging from 0 to +/- 1.

Waiting control group

In this analysis we compared T1 and T2 for both groups, waiting control group and training group for all dependent measures. ANCOVAs were run with experimental and control group as between subjects factor and T1 point of measurement as a covariate for all repeated dependent measures.

For *self-efficacy*, the analysis revealed a significant main effect for condition, $F(1,123) = 4.15$, $p = .044$, $\eta_p^2 = .03$, and a significant covariate, $F(1,123) = 103.12$, $p = .0001$, $\eta_p^2 = .45$., with higher means at T2 for the training group, $M_{2\_training} = 3.97$ ($SE = .04$), versus the control group, $M_{2\_control} = 3.78$ ($SE = .08$).

For *assertiveness*, the analysis revealed a non-significant main effect for condition, $F < 1$, and a significant covariate, $F(1,123) = 36.0$, $p = .0001$, $\eta_p^2 = .22$, with the following mean at T2 for the training group, $M_{2\_training} = 3.97$ ($SE = .04$),and for the control group, $M_{2\_control} = 3.78$ ($SE = .08$) was found.

For UBOS *emotional depletion*, the analysis revealed a non-significant main effect for condition, $F(1,123) = 1.51$, $p = .22$, $\eta_p^2 = .012$, and a significant covariate, $F(1,123) = 96.43$, $p = .0001$, $\eta_p^2 = .43$, with the following mean at T2 for the training group, $M_{2\_training} = 2.25$ ($SE = .065$),and for the control group, $M_{2\_control} = 2.43$ ($SE = .13$) was found. Here the control group mean has to be higher, since it indicates more symptoms.

For UBOS *depersonalization*, the analysis revealed a non-significant main effect for condition, $F < 1$, and a significant covariate, $F(1,123) = 76.92$, $p = .0001$, $\eta_p^2 = .38$, with the following mean at T2 for the training group, $M_{2\_training} = 2.4$ ($SE = .06$),and for the control group, $M_{2\_control} = 2.53$ ($SE = .13$) was found.

For UBOS *personal abilities*, the analysis revealed a significant main effect for condition, $F(1,123) = 5.9$, $p = .017$, $\eta_p^2 = .04$, and a significant covariate, $F(1,123) = 93.46$, $p = .0001$, $\eta_p^2 = .43$, with higher means at T2 for the training group, $M_{2\_training} = 5.90$ ($SE = .06$), versus the control group, $M_{2\_control} = 4.82$ ($SE = .14$).

For 4DKL *distress*, the analysis revealed a non-significant main effect for condition, $F(1,123) = 1.37$, $p = .24$, $\eta_p^2 = .011$, and a significant covariate, $F(1,123) = 61.95$, $p = .0001$, $\eta_p^2 = .33$, with the following mean at T2 for the training group, $M_{2\_training} = 1.4$ ($SE = .041$),and for the control group, $M_{2\_control} = 1.51$ ($SE = .08$) was found.

For 4DKL *depression*, the analysis revealed a non-significant main effect for condition, $F(1,123) = 1.17$, $p = .28$, $\eta_p^2 = .009$, and a significant covariate, $F(1,123) = 62.35$, $p = .0001$, $\eta_p^2 = .33$, with the following mean at T2 for the training group, $M_{2\_training} = 1.11$ ($SE = .035$),and for the control group, $M_{2\_control} = 1.20$ ($SE = .074$) was found.

For 4DKL *angst*, the analysis revealed a non-significant main effect for condition, $F(1,123) = 1.75$, $p = .18$, $\eta_p^2 = .014$, and a significant covariate, $F(1,123) = 85.71$, $p = .0001$, $\eta_p^2 = .41$, with the following mean at T2 for the training group, $M_{2\_training} = 1.08$ ($SE = .018$),and for the control group, $M_{2\_control} = 1.14$ ($SE = .037$) was found.

For 4DKL *somatization*, the analysis revealed a non-significant main effect for condition, $F < 1$, and a significant covariate, $F(1,123) = 67.33$, $p = .0001$, $\eta_p^2 = .35$, with the following mean at T2 for the training group, $M_{2\_training} = 1.27$ ($SE = .02$),and for the control group, $M_{2\_control} = 1.39$ ($SE = .05$) was found.

For *positive affect*, the analysis revealed a significant main effect for condition, $F(1,123) = 7.98$, $p = .006$, $\eta_p^2 = .061$, and a significant covariate, $F(1,123) = 19.45$, $p = .0001$, $\eta_p^2 = .137$, with higher means at T2 for the training group, $M_{2\_training} = 3.84$ ($SE = .04$), versus the control group, $M_{2\_control} = 3.51$ ($SE = .1$).

For *negative affect*, the analysis revealed a significant main effect for condition, $F(1,123) = 6.99$, $p = .009$, $\eta_p^2 = .054$, and a significant covariate, $F(1,123) = 58.25$, $p = .0001$, $\eta_p^2 = .321$, with lower means at T2 for the training group, $M_{2\_training} = 1.82$ ($SE = .05$), versus the control group, $M_{2\_control} = 2.17$ ($SE = .12$).

For *Work Locus of Control*, the analysis revealed a significant main effect for condition, $F(1,123) = 4.71$, $p = .032$, $\eta_p^2 = .037$, and a significant covariate, $F(1,123) = 95.61$, $p = .0001$, $\eta_p^2 = .437$, with higher means at T2 for the training group, $M_{2\_training} = 3.84$ ($SE = .02$), versus the control group, $M_{2\_control} = 3.71$ ($SE = .05$).

For *introspection*, the analysis revealed a non-significant main effect for condition, $F < 1$, and a significant covariate, $F(1,123) = 40.33$, $p = .0001$, $\eta_p^2 = .24$, with the following mean at T2 for the training group, $M_{2\_training} = 2.69$ ($SE = .03$),and for the control group, $M_{2\_control} = 2.61$ ($SE = .08$) was found.

Taken together, no significant effects were found on eight variables and significant effects were found on 5 variables. It is noteworthy that with the exception of UBOS personal abilities no significant effects were found on the central symptom related variables. The significant differences relative to the control condition play out on “softer” variables such as self-efficacy, positive and negative affect and work locus of control.

UNIVERSITEIT VAN AMSTERDAM
Faculty of Psychology

Longitudinal Analysis

In this analysis we compared all four points of measurement for the experimental group only in a repeated measures ANOVA with T1- T4 as within factor.

The analysis revealed a significant effect for *self-efficacy* over time, $F(3,72) = 7.92$, $p = .0001$, $\eta_p^2 = .24$. Self-efficacy is increasing over time: The mean pattern shows the following significant differences: $M_1 = 3.86$ ($SE = .09$) is significantly different from $M_2 = 4.16$ ($SE = .09$), $M_3 = 4.19$ ($SE = .1$), and $M_4 = 4.19$ ($SE = .08$), all ps < .003, while there are no significant differences between T2 to T4, all ps > .68.

A similar effect was found for *assertiveness*, $F(1.99,47.85) = 10.72$, $p = .0001$, $\eta_p^2 = .30$ (with Greenhouse-Geisser correction). Assertiveness is increasing over time: The mean pattern shows the following significant differences: $M_1 = 3.55$ ($SE = .13$) is significantly different from $M_2 = 3.86$ ($SE = .09$), $M_3 = 3.98$ ($SE = .11$), and $M_4 = 3.92$ ($SE = .1$), all ps < .004, while there are no significant differences between T2 to T4, all ps > .077.

For the UBOS subscales the effects are as following. There is a significant decrease of emotional depletion, $F(3,54) = 9.55$, $p = .0001$, $\eta_p^2 = .34$. *Emotional depletion* decreased over time: The mean pattern shows the following significant differences: $M_1 = 3.38$ ($SE = .25$) is significantly different from $M_2 = 2.56$ ($SE = .25$), $M_3 = 2.43$ ($SE = .31$), and $M_4 = 2.49$ ($SE = .24$), all ps < .0001, while there are no significant differences between T2 to T4, all ps > .509. For the subscale of *depersonalization* the effect is not reaching conventional levels of significance, $F(3,54) = 2.51$, $p = .068$, $\eta_p^2 = .12$, but is trending. Again there is a decrease of depersonalization from $M_1 = 2.91$ ($SE = .24$) to $M_2 = 2.55$ ($SE = .22$), $p = .072$ (again trending), to $M_3 = 2.36$ ($SE = .23$), $p = .013$ and $M_2 = 2.61$ ($SE = .20$), all other ps > .176. It is noteworthy that T1 and T4 were not significantly different from each other, thus the improvement dynamic was not stable. The pattern of results for *personal abilities* shows an increase over time that remained stable, $F(3,54) = 9.22$, $p = .0001$, $\eta_p^2 = .33$, with the following mean pattern: $M_1 = 4.82$ ($SE = .18$) was significantly different from $M_2 = 5.48$ ($SE = .18$), $M_3 = 5.58$ ($SE = .18$), and $M_4 = 5.48$ ($SE = .17$), all ps < .002, while there were no significant differences between T2 to T4, all ps > .472.

In case of the *4DKL* scale, there was a stable decrease for perceived *distress*, $F(3,54) = 7.4$, $p = .0001$, $\eta_p^2 = .29$, with the following mean pattern: $M_1 = 1.8$ ($SE = .15$) which was significantly different from $M_2 = 1.45$

(*SE* = .12), M_3 = 1.37 (*SE* = .14), and M_4 = 1.49 (*SE* = .14), all *p*s < .017, while there were no significant differences between T2 to T4, all *p*s > .075. For the *somatization* subscale there was an overall significant effect, $F(3,54) = 3.29$, $p = .027$, $\eta_p^2 = .15$, but the significant decrease started only after T2. M_1 = 1.43 (*SE* = .09) was not significantly different from M_2 = 1.31 (*SE* = .10), $p = .067$, but was significant from M_3 = 1.37 (*SE* = .14), and M_4 = 1.49 (*SE* = .14), all *p*s < .049, while there were no significant differences between T2 to T4, all *p*s > .287. Since the reliability for the *depression* subscale at T3 was so low, two analyses were run here, one with and one without T3. If T3 is included the overall effect was just reaching conventional levels of significance, $F(3,54) = 2.74$, $p = .052$, $\eta_p^2 = .13$, but the only significant difference is between T1 and T3, $p = .007$. An analysis without T3, thus only T1, T2, and T4 revealed a non-significant effect, $F(2,42) = 1.86$, $p = .167$, $\eta_p^2 = .08$. Taken together, a conservative interpretation should rather be that there is no effect (M_1 = 1.20 (*SE* = .07), M_2 = 1.13 (*SE* = .08), M_3 = 1.01 (*SE* = .12), and M_4 = 1.13 (*SE* = .08)). A similar effect was found for *fear*, but with more dynamic, $F(2.15,38.74) = 4.46$, $p = .006$, $\eta_p^2 = .20$ (with Greenhouse-Geisser correction). Fear was first found to decrease, but then reached again the level of T1: The mean pattern shows the following significant differences: M_1 = 1.28 (*SE* = .10) is significantly different from M_2 = 1.07 (*SE* = .04), $p = .01$, and also from M_3 = 1.11 (*SE* = .08), $p = .025$, but not different from M_4 = 1.16 (*SE* = .08), $p = .081$, while there are no significant differences between T2 to T4, all *p*s > .126.

We also tested changes in *positive* and *negative* affect. For the positive affect subscale the same pattern was found as for fear, first an increase (i.e. more positive affect) but then no difference with T1, $F(1.74,31.36) = 4.17$, $p = .01$, $\eta_p^2 = .18$ (with Greenhouse-Geisser correction). Positive affect was first found to increase, but then reached again the level of T1: The mean pattern shows the following significant differences: M_1 = 3.52 (*SE* = .14) is significantly different from M_2 = 3.94 (*SE* = .12), $p = .015$, and also from M_3 = 3.92 (*SE* = .15), $p = .045$, but not different from M_4 = 3.73 (*SE* = .14), $p = .229$. T2 was actually different from T4, $p = .045$, while there were no further significant differences, all *p*s > .065. The effect turned out to be more robust for negative affect, $F(3,54) = 8.78$, $p = .0001$, $\eta_p^2 = .328$, with the following mean pattern: M_1 = 2.37 (*SE* = .19) which was significantly different from M_2 = 1.73 (*SE* = .18), M_3 = 1.55 (*SE* = .17), and M_4 = 1.78 (*SE* = .16), all *p*s < .013, while there were no significant differences between T2 to T4, all *p*s > .113.

For the *Work Locus of Control* scale there was a significant effect, $F(2.10,130.761) = 4.91$, $p = .003$, $\eta_p^2 = .73$ (with Greenhouse-Geisser correction), with the following mean pattern: M_1 = 3.76 (*SE* = .05) which

was significantly different from M_2 = 3.88 (*SE* = .05), M_3 = 3.95 (*SE* = .05), and M_4 = 3.93 (*SE* = .07), all *p*s < .013, while there were no significant differences between T2 to T4, all *p*s > .066.

There was also an increase in *introspection* over time, $F(3,54) = 4.86$, $p = .005$, $\eta_p^2 = .21$, with the following mean pattern: M_1 = 2.41 (*SE* = .10) which was significantly different from M_2 = 2.68 (*SE* = .10), M_3 = 2.63 (*SE* = .08), and M_4 = 2.71 (*SE* = .11), all *p*s < .044, while there were no significant differences between T2 to T4, all *p*s > .410.

Taken together, the pattern of results show an effect of the Columbus training (which has to be seen in critical conjunction with the control group effects), while the Mandela training leads to a stabilization of effects (with the exclusion of somatization where the effect depends on the Mandela training). Still, on certain, clinically relevant variables, a remission to pre-training levels has to be pointed to.

Further determinants of treatment effects.

We further tested if certain parameters had an effect on treatment success. For certain measures there was an indication that females had more treatment success than males. This was the case for 4DKL/somatisation where there is no effect comparing T1-T3 for males, $p = .457$, but for females $p = .0001$. It is also noteworthy that females and males differed on this measure at T1, $p = 018$ (based on a significant overall interaction ($F(1,59) = 4.45$, $p = .039$, $\eta_p^2 = .07$). A similar effect was found for 4DKL/distress, where both genders revealed significant changes, but females showed a much bigger effect size, $\eta_p^2 = .23$, compared to $\eta_p^2 = .04$ for males.

Yet none such differences were found for the UBOS, assertiveness, self-efficacy, positive and negative effect, work locus of control and introspection.

In addition we tested the influence of personality and motivation strategy on treatment effects. First of all there were few gender differences in terms of personality and motivation strategy. Females are scored higher on Agreeableness ($p = .001$), which is a normal effect. There was a significant correlation of work experience on conscientiousness, $r = 233$, $p = .01$, which lends specific relevance to the influence of the conscientiousness predictor, see below.

Males scored higher on Prevention focus ($p = .002$) which is unusual and may be attributed to a specific male sample. No further effects were found.

For the influence of those moderators we computed differences scores on the effect level for T1 and T2. Those scores were computed in that way, that higher values always indicate a higher effect.

RFT did only have marginal effect on the UBOS subscale *depersonalization* where individuals with a higher promotion focus show a smaller effect, β =-.246, p = 013. A similar effect was found for 4DKL subscale *distress*, β =-.22, p = 026, and for *depression*, β =-.224, p = 024, *somatization*, β =-.207, p = 037. For positive affect, β =.258, p = 009, the reverse holds. Here individuals with a high promotion focus show a greater increase in positive affect.

For *personality factors*, there is an overall negative prediction for Conscientiousness and a positive prediction for Neuroticism. The conscientious effects were found for self-efficacy, β =-.238, p = 016, UBOS depersonalisation, β =-.254, p = 013, 4DKL distress, β =-.19, p = 051, 4DKL depression, β =-.307, p = 001, 4DKL somatization, β =-.206, p = 04. The neuroticism effect was found for assertiveness, β = .272, p = 005, UBOS emotion depletion, β = .195, p = 055, UBOS personal abilities, β = .196, p = 059, 4DKL distress, β = .31, p = 002, 4DKL depression, β = .296, p = 002, 4DKL somatization, β = .22, p = 029.

We furthermore tested whether introspection (measured at T1) has an effect on the training over time. In including introspection in the analysis turned almost all effects non-significant, except for assertiveness, where the training effect remained (and also a significant interaction with introspection. Yet this covariate analysis over all four points of measurement may lead to overconservative interpretations, given the invariability of certain effects from T2-T4. A correlation analysis on the effect difference scores gives a better view on the effect. Higher introspection correlated positively with the following outcome variables: UBOS emotion depletion, r = .194, p = .05, 4DKL distress, r = .367, p = .0001, 4DKL depression, r = .260, p = .008, 4DKL angst, r = .264, p = .007, 4DKL somatization, r = .226, p = .022, negative affect, r = .344, p = .0001. Taken together, higher levels of introspection are related to an improved training effect.

Extreme groups analysis

In addition to the previous analyses we ran an extreme groups analysis on participant groups who scored higher and lower on the 4DKL variable at T1 for effect on T2-T4. The only significant, but highly relevant effect was found for the 4DKL subscale depression. The split in high (> 1.5) , medium (>1 and <= 1.5), and low (= 1) depression at T1 revealed a significant interaction, $F(4,32) = 7.52$, p = .0001, η_p^2 = .48. The

simple main comparisons reveal a significant drop from T2, M_2 = 2.13 (*SE* = .05), to T3, M_3 = 1.16 (*SE* = .0), *p* = .0001, but a remission to the same level as T2 for T4, M_4 = 1.83 (*SE* = .18), *p* = .164, also indicated by a significant difference between T3 and T4, p = .003. Given that the internal reliability was low for T3, the latter effect needs to be interpreted with caution, but the lack of significant difference between T2 and T4 is not touched by this.

In sum, there is evidence for future tailor made training potential, mostly along the lines of gender, personality traits, and also based on initial level of symptoms.

References

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, *4*, 62-83.

Higgins, E. T., Freidman, R. S., Harlow, R. E., Chen Idson, L., Ayduk, O. N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology*, *31*, 3-23.

Lorr, M., & More, W. W. (1980). Four dimensions of assertiveness. *Multivariate Behavioral Research*, *2*, 127-138.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five inventory in English and German. *Journal of Research in Personality*, *41*, 203-212.

Schaufeli, W. B., Taris, T. W., & van Rhenen, W. (2008). Workaholism, burnout, and work engagement: Three of a kind or three different kinds of employee well-being? *Applied Psychology: An International Review*, *57*, 173-203.

Semin, G. R., Higgins, T., de Montes, L. G., Estourget, Y., & Valencia, J. F. (2005). Linquistic signatures of regulatory focus: How abstraction fits promotion more than prevention. *Journal of Personality and Social Psychology*, *89*, 36-45.

Siu, O., & Cooper, C. L. (1998). A study of occupational stress, job satisfaction and quitting intention in Hong Kong firms: The role of locus of control and organizational commitment. *Stress Medicine*, *14*, 55-66.

Strauser, D. R., Ketz, K., & Keim, J. (2002). The relationship between self-efficacy, locus of control and work personality. *Journal of Rehabilitation*, *68*, 20-26.

Terluin, B., van Rhenen, W., Schaufeli, W. B., & de Haan, M. (2004). The four dimensional symptom questionnaire (4DSQ): Measuring distress in a working population. *Work and Stress*, *18*, 187–207.

Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063-1070.

UNIVERSITEIT VAN AMSTERDAM
Faculty of Psychology