Kieser Training works. Study results.

Kieser Training AG, Zurich

Contents

Kieser Training works. Study results.

Kieser Training AG, Zurich

The results of «Kieser Training works» are presented below. The study, which was conducted between April and October 2009, was a prospective, randomised, controlled and multi-centre study involving 118 Kieser Training facilities in Germany.

Kieser Training AG in Zurich was responsible for its planning, overall coordination and analysis. An independent group of experts[1] evaluated the design of the study, the selection of participants and the results.

1 Paul Eigenmann (Qualitop), Prof. Dr. Walter Krämer (Economic and Social Statistics, Technical University of Dortmund), Rolf Dieter Müller (Health industry consultant, lecturer at the Technical University of Berlin), Prof. Dr. Dietmar Schmidtbleicher (Professor in Training and Exercise Sciences at the University of Frankfurt), Prof. Dr. Ulrich Smolenski (Institute for Physiotherapy at the University Clinic of Jena and Chairman of DGMM (German Association for Manual Medicine)).

1 Design of study

The study reflected the twin-track approach offered by Kieser Training, i.e. training and therapy, and so participants were divided into two intervention groups: The training group (PST group) completed a 6-month programme of Preventive Strength Training. The therapy group (MST group) also completed a programme of the same duration; however, this group started with Medical Strengthening Therapy followed by Preventive Strength Training. In addition, a control group (C) for the PST Group was formed. The study was designed primarily to determine the net training effects, i.e. comparison PST/C and for ethical reasons, we did not form a control group for the MST group[2]. If we had, this group would by definition have consisted of individuals for whom therapy was indicated but who would have not received any other intervention during the course of the study, i.e. they would have been a waiting-list group. Encouraging this type of behaviour from persons with pain symptoms would only be ethical if it were likely to increase substantially our level of knowledge. In addition, the results from controlled intervention studies for Medical Strengthening Therapy are already available and its contribution to alleviating back and neck pain is already known[3]. As a result, the study did not include a controlled evaluation of therapy.

We used an optimum sample size, which was based on the study criteria and the envisaged effect sizes[4]. This reflected a drop-out rate of 30% split between the three groups as follows: approx. 96 persons in the MST group, approx. 322 in the PST group and approx. 82 in the C group.

The allocation of participants to individual groups was based on the answers to questions contained in the application. They identified whether it would be appropriate – in health terms – to do strength training. Based on the answers, each applicant was given a «training» or «therapy» status. From the group given therapy status, we then drew lots to obtain the MST group. From the group with training status we also drew lots but the resultant group was then randomised to form a PST group and a control group. Having selected participants in this way, the Kieser Training doctors in each participating facility reviewed the status of all participants.

All participants were examined by a doctor before and after the six-month training period. In addition, we used computer-aided machines to test the back strength of participants and asked participants to complete a written questionnaire at the start of training (0 months), at midterm (3 months) and on completion (6 months). We used a range of parameters to check the effectiveness of strength training: pain, strength, general health and wellbeing together with performance and quality of life.

We used SPSS Statistics 17.0 software to evaluate the data. As some variables had no normal distribution (verified using the Shapiro-Wilk-Test), we used the Wilcoxon non-parametric signed-rank test to verify differences between paired means. For non-paired means, we used the U test by Mann Whitney. To evaluate the practical significance of differences, we calculated the effect sizes d[5] and the corrected effect sizes d_{corr}[6]; the latter reflect group differences in the pre-test and sample size. As the measure of

2 See «Kontrollproblem» (Control Problem) in BORTZ J., DÖRING N. (2002): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. (Research methods and evaluation for human and social scientists) Heidelberg: Springer, Page 116.

3 E.g. RISCH S., NORVELL N., POLLOCK M., RISCH E., LANGER H., FULTON M., GRAVES J., LEGGETT S. (1993): Lumbar strengthening in chronic low back pain patients. Spine 18, 232-238. NELSON B., O'REILLY E., MILLER M., HOGAN M., WEGNER J., KELLY C. (1995): The clinical effects of intensive, specific exercise on chronic low back pain: a controlled study of 895 consecutive patients with 1-year follow up. Orthopedics 18, 971-981.

4 Estimates for sample mean and sample deviation already exist for most of the parameters. Similarly, we were aware of other studies of the expected effects of training intervention. We therefore used these values to calculate the required sample size (BORTZ J. & DÖRING N. (2002): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler (Research methods and evaluation for human and social scientists), Heidelberg: Springer, Section 9.2.2).

5 See Section 10 for explanation of effect sizes.

6 JACOBS B. (1999): Einführung in die Versuchsplanung – Praktische Bedeutsamkeit empirischer Ergebnisse (Introduction to the design of experiments – practical significance of empirical results.) Online query of http://www.phil.uni-sb.de/~jakobs/seminar/vpl/bedeutung/eskorr.htm. performed on May 25th, 2010.

location for distributions, we selected the arithmetic mean for representing data approaching the norm distribution and the median for skew distributions. To verify statistical significance, we set the significance level in accordance with current practice at α = 5% (p<0.05, two-sided)[7]. The verifiable effects with the current sample sizes are shown in Diagram 12 (Chapter 11) (power of test (1-β) = 0.8).

2 Intervention

Both intervention groups completed progressive strength training on training machines with a variable resistance. The control group underwent no training during this period but were given the opportunity to train free of charge for 6 months after the study had ended (waiting-list control group).

2.1 Preventive strength training

At their first three sessions, participants were given instruction by qualified staff. At their 10th session and then at every subsequent 20th training session, participants were given a check session and if necessary modifications were made to their programme. Participants started with a basic programme (primarily single-joint exercises) and after 20 – 40 sessions switched to a build-up programme (an increasing number of multi-joint exercises). Table 1 shows the standard loading conditions used in the strength-training programmes.

2.2 Medical Strengthening Therapy

The MST group first completed a programme of strengthening therapy under the control of a doctor using a special test and therapy machine for isolated lumbar extensions[8]. They also did a maximum of six other exercises drawn from the strength training programme described in Section 2.1. On completion of strengthening therapy – after 12 or 18 therapy sessions (as decided by the doctor and the documented therapy results) – participants transferred to the strength training described in Section 2.1.

3 Sample

From the initial 627 participants, 531 remained until the end of the study. The drop-our rate (percentage who failed to complete the study) for each group was as follows: PST: 15.6%, MST: 21.7%, C: 6%.

The average drop-out rate in German fitness chains is 35.3%[9]. The most frequent reason cited for drop-out from our study was illness or injury in daily life (N = 24). 3 participants experienced pain after training and so stopped training. 6 participants had to stop for professional or private reasons. 11 participants said they no longer wanted to train and 22 trained irregularly or with long gaps. For 30 participants, we were unable to find out the reason for stopping.

3.1 Description of sample groups

The following table shows the size and composition of the sample. There may be some variation in the numbers shown because some individuals may not have completed a specific test or answered an individual question.

7 BORTZ J., DÖRING N. (2002): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. (Research methods and evaluation for human and social scientists), Heidelberg, Springer, page 30f.

8 CARPENTER D., NELSON B. (1999): Low back strengthening, for the prevention and treatment of low back pain. Medicine and Science in Sports and Exercise, 31, 18-24.

9 DELOITTE (2009): Der deutsche Fitness- und Wellnessmarkt (German fitness and wellness market).

Table 1: Standard loading conditions for the strength training

	Training (PST)	Strengthening Therapy (MST)
Load magnitude	approx. 60% of dynamic maximum strength	approx. 50% to 60% of isometric maximum strength
Number of repetitions	6-9	approx. 12
Number of sets	1	1
Training frequency	1.6 times per week	1.4 times per week
Duration of training period	24.5 weeks	12.5 weeks
Duration of each type of contraction per repetition	4 sec. concentric 2 sec. isometric 4 sec. eccentric	4 sec. concentric 2 sec. isometric 4 sec. eccentric
Time under load	training session 1-20: 60-120 sec. from training session 21: 60-90 sec.	training session 1-6: 120-135 sec. from training session 7: 75-120 sec.
Muscle fatigue	training session 1-20: sub-maximal number of repetitions to repetition maximum from training session 21: repetition maximum to volitional muscular failure.	training session 1-6: sub-maximal number of repetitions to repetition maximum from training session 7: repetition maximum to volitional muscular failure.
Range of motion	individual's maximum (pain-free) possible movement of joint	individual's maximum (pain-free) possible movement of joint
Recovery time between training sessions	minimum 48 hrs	minimum 48 hrs
Exercises	whole-body programme of 10 exercises consisting of lumbar extension plus exercises for the hip, abdomen, thigh, back, shoulder and pectoral muscles	isolated lumbar extension + supplementary programme: 6 exercises with the standard load norms used for training (PST)
Order of exercises	large muscle groups before small muscle groups; «problem exercises» given precedence	isolated lumbar extension before supplementary exercises

Table 2: Description of sample (n = 531)

		PST	MST	C
Number		317	109	105
Age (years)		43.1 ±13.7	44.7 ±11.2	41.8 ±12.3
BMI (kg/m²)		24.9 ±4.2	25.0 ±4.5	24.3 ±4.2
Gender	**Female**	49.2 %	51.4 %	64.8 %
	Male	50.8 %	48.6 %	35.2 %
Mainly sedentary lifestyle		62.1 %	65.1 %	60.2 %
No physical exercise		30.4 %	30.4 %	32.4 %
Experience of strength training		35.6 %	50.5 %	39.4 %

The control group contained a higher percentage of women. All three groups had a similar BMI and it was in the normal range. We also checked whether the BMI distribution in the sample was representative of Germany as a whole[10]. We found that the sample contained an above average percentage of persons who were underweight or normal weight (BMI<25) and the percentage of those who were obese (BMI≥30) was significantly below the average. This indicates an element of self selection in terms of who applied to participate in the study. It is assumed that the fitness nature of the study was less attractive to those who were obese. An above average percentage of persons of normal weight or less than normal weight applied and so formed a larger percentage of the sample group.

3.2 Training aims

The most frequent aims cited by the PST group were «general strengthening» (36%) and «back strengthening» (35%). The MST group mainly cited – apart from «back strengthening» (31%) – «pain reduction» (22%) and «elimination of pain» (24%).

3.3 Pain situation

Participants were asked whether they had experienced pain in the last 4 weeks: 419 out of 531 participants (78%) said that they had. Moderate to very severe pain was much more common in the MST group than that in the other two groups (Table 3). Their pain was mainly in the lumbar spine, cervical spine, knees and shoulders.

Prevalence of back pain

74% of participants indicated that they had back problems (N = 396). Lumbar pain was the most common.

All those in the MST group had back pain, whereas in the PST and control group, it was 69% and 64% respectively.

While one in two of the MST group were seeing a doctor for their back pain, this ratio was only one in five for the PST and control groups.

10 Comparison with data from German Federal Health Monitoring (1998): www.gbe-bund.de.

Table 3: Pain ratings

	PST		MST		C		Total	
	N	%	N	%	N	%	N	%
Very severe	--	--	4	3.7	--	--	4	0.8
Severe	16	5.1	21	19.4	5	4.9	42	8.0
Moderate	85	27.0	45	41.7	29	28.2	159	30.2
Slight	78	24.8	24	22.2	21	20.4	123	23.4
Very slight	57	18.1	10	9.3	24	23.3	91	17.3
No pain	79	25.2	4	3.7	24	23.3	107	20.3

Table 4: Prevalence of back pain in each group

	PST		MST		C		Total	
	N	%	N	%	N	%	N	%
Total	317	100	109	100	105	100	531	100
Receiving treatment for back pain (in last 3 months)	70	22	62	56	20	19	152	28
Back pain – total	221	69	107	98	68	64	396	74
■ Back pain – lumbar area[11]	160	50	97	88	51	48	308	58
■ Back pain – dorsal spine	50	15	22	20	17	16	89	16
■ Back pain – cervical spine	92	29	51	46	28	26	171	32

Character of back pain in the MST Group
37% (N = 41) of the MST group reported continuous pain for at least three months and 42% (N = 46) had had chronic recurring pain for at least two years. This pain primarily affected the lumbar area (79%, N = 87) with 25% (N = 28) reporting pseudo-radicular radiation and 9% (N = 10) radicular radiation.

The most frequent cause of symptoms were functional problems, e.g. facet joint dysfunctions (55%, N = 60), followed by degenerative changes (52%, N = 57). From these persons 24 had prolapsed discs, 17 disc protrusions and 30 osteochondroses.

Spinal malformations such as scoliosis were only occasionally cited as the cause of the pain (17%, N = 19).

We used the «Mainz Pain Staging System for pain chronification»[12] to evaluate the degree of pain chronification. This showed that 54.1% (N = 59) of the group were in Stage I (slight chronification), 22.9% (N = 25) in Stage II and 8.3% (N = 9) in Stage III (advanced chronification).

11 Some participants cited pain in several areas of the back, i.e. cervical, lumbar or dorsal spine.
12 See: www.dgss.org/fileadmin/pdf/mpss_testanweisungen.pdf.

4 Results – pain and its effects

The following diagram shows the percentage of participants who had no pain at the start and end of the intervention.

340 of those who trained as part of the study (PST and MST combined) indicated at the start of the study that they had experienced pain in the last month (80%). By the end of the study, 102 of the 340 (30%) had no pain; 135 (39%) still had very slight to slight pain and only 77 (22.6%) had moderate to very severe pain. 26 (7%) did not answer this question.

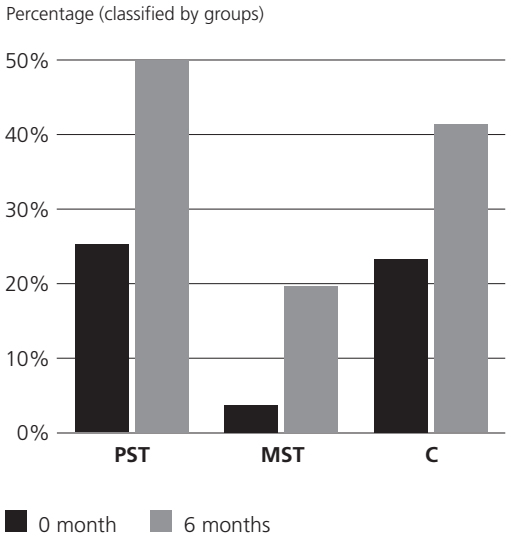
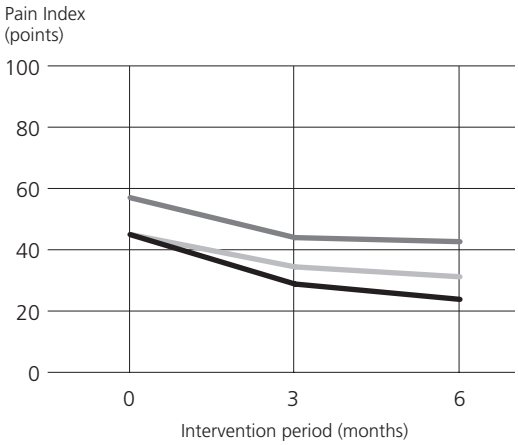


Diagram 1:
Percentage of pain-free participants in each group

The following evaluation is based on data for participants who reported some pain before the study started (419 of the 531 participants, 78%).

4.1 Pain index

In order to express the change in pain distribution numerically within groups, we transformed the six-stage rating scale used for an approximate pain assessment into a numeric scale from 0–100[13]. The following diagram shows the change in the pain index in the three groups.



	0	3	6
PST	44.95 [42.59;47.32]	28.81 [25.49;32.13]	23.76 [20.50;7.02]
C	45.00 [40.39;49.61]	34.41 [28.63;40.20]	31.18 [25.07;37.29]
MST	57.14 [53.14;61.15]	43.96 [39.18;48.73]	42.64 [36.84;48.44]

Diagram 2:
Pain Index on the 3 test dates (arithmetic mean with confidence interval (95%), range 0-100 all changes in relation to 0 months significant), data base: complete data sets for participants reporting pain at 0 months (N_{PST} = 218, N_C = 68, N_{MST} =91).

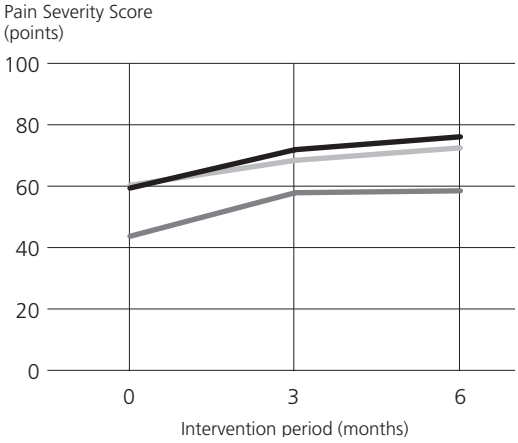
13 To transform to a higher level of measurement, the rating scale can be interpreted as a «measurement by fiat» and treated as interval scaled. «It is standard practice in research to forego empirical verifications of relevant scale axioms. Most measurements are measurements «by fiat» (measurement by presumption based on survey instruments (questionnaires, tests, rating scales, etc.)), whereby it is assumed that the relevant attribute will be measured on an interval scale. This means that the entire statistical «inventory» for interval scales can be used, so allowing much greater differentiation compared with the process for ordinal and nominal data (…). The reasoning behind this «liberal» interpretation is a conviction that it would be difficult to confirm a research hypothesis if an incorrect scale level were assumed.» (BORTZ J., DÖRING N. (2002): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler (Research methods and evaluation for human and social scientists), Heidelberg, Springer, page 74.

Table 5 shows a significant drop in the pain index for all three groups. The benefit from strength training (PST group) compared with no intervention (C group) was of practical significance with a net effect of d_{corr} = -0.30. This was apparent even after 3 months of training (d_{corr} = -0.23).

4.2 Severity of pain

In order to determine the severity of pain experienced by participants, we used the «Pain Severity» (PS) sub-scale from the Medical Outcomes Study[14]. This sub-scale consists of a range of questions about the severity of pain within the last 4 weeks, including frequency of pain, its duration, average severity and maximum severity of pain. Scores range from 0 to 100: the higher the score, the less severe the pain.

The following diagram shows the changes to the PS scale as a result of the intervention. The MST group had experienced the most severe pain. In the PST and control groups, levels of pain were similar. All three groups showed a significant increase in the mean PS, i.e. a reduction in pain. The benefit from strength training (PST) compared with no intervention (C group) was of practical significance with a net effect of d_{corr} = -0.22. This was apparent even after 3 months of training (Table 6).



	0	3	6
PST	59.38 [57.48;61.27]	71.89 [68.85;74.94]	76.08 [72.96;79.21]
C	60.27 [56.57;63.96]	68.41 [63.63;73.19]	72.48 [67.22;77.74]
MST	43.67 [40.16;47.19]	57.86 [53.33;62.38]	58.49 [53.14;63.84]

Diagram 3:
Pain Severity on the three test dates (arithmetic mean with a confidence interval (95%), range 0 to 100, all changes compared with 0 months significant), data base: complete datasets for individuals with pain at 0 months (N_{PST} = 219, N_C = 71, N_{MST} = 91).

Table 5: Effects and changes to the Pain Index within the groups

		Effects		Change	
Group	**N**	**d (d_{corr}) 3 months**	**d (d_{corr}) 6 months**	**absolute 6 months**	**percentage 6 months**
PST	218	-0.76* (-0.23)	-1.01* (-0.30)	-21.19 ±25.48*	-58*
C	68	-0.49*	-0.62*	-13.82 ±25.92*	-29*
MST	91	-0.63*	-0.62*	-14.51 ±29.52*	-26*

d_{corr}: net effect of PST compared with C, *significant change compared with start (0 months)

14 HAYS RD., SHERBOURNE CD., MAZEL RM. (1995): User's Manual for Medical Outcomes Study (MOS) Core Measures of Health-Related Quality of Life. RAND Corporation, MR-162-RC. Manual, detailed questionnaire and scoring instructions at http://www.rand.org/health/surveys_tools/mos/.

Table 6: Effects and changes to the Pain Severity Scale in each group

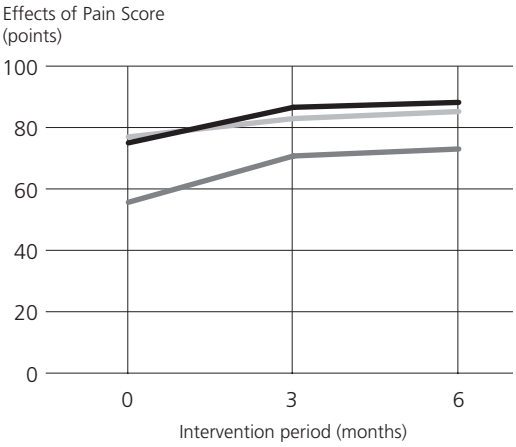
		Effects		Change	
Group	**N**	**d (d_{corr}) 3 months**	**d (d_{corr}) 6 months**	**absolute 6 months**	**percentage 6 months**
PST	219	0.67* (0.22)	0.89* (0.22)	16.70 ± 21.91*	28*
C	71	0.45*	0.65*	12.21 ± 22.94*	16*
MST	91	0.73*	0.70*	14.82 ± 23.84*	38*

d_{corr}: Net effect of PST compared with C, *significant change compared with start (0 months)

4.3 Effects of pain

The «Effects of Pain» scale (EP, Medical Outcomes Study) expresses the degree to which pain impacts on certain aspects of life (mood, ability to walk or move about, sleep, recreational activities and enjoyment of life). Test scores range from 0 to 100 whereby the higher the score, the less the effect of the pain.

The following diagram shows changes to the EP scale as a result of the intervention. The effects experienced by the MST group were greater than those experienced by the PST and C groups. All three groups showed a significant increase in the mean EP during the study; this represents a reduction in the effects of pain. The benefit from strength training (PST) compared with no intervention (C group) was of practical significance, whereby the net effect was d_{corr} = 0.33 after 3 training months and d_{corr} = 0.27 after 6 training months (Table 7).



	0	3	6
PST	75.00 [72.37;77.64]	86.59 [84.50;88.67]	88.19 [85.95;90.44]
C	76.95 [72.42;81.48]	82.92 [78.96;86.89]	85.21 [80.54;89.88]
MST	55.65 [50.70;60.60]	70.74 [65.55;75.93]	73.01 [67.73;78.29]

Diagram 4:
Effects of Pain on the three test dates (arithmetic mean with a confidence interval (95%), range 0-100, all changes compared with 0 months significant), base data: complete data sets for individuals with pain at 0 months (N_{PST} = 215, N_C = 71, N_{MST} = 90).

Table 7: Effects and changes to the Effects of Pain scale in each group

		Effects		Change	
Group	**N**	**d (d_{corr}) 3 months**	**d (d_{corr}) 6 months**	**absolute 6 months**	**percentage 6 months**
PST	215	0.66* (0.33)	0.73* (0.27)	13.19 ± 22.88*	12*
C	71	0.33*	0.43*	8.26 ± 17.61*	9*
MST	90	0.62*	0.71*	17.36 ± 25.32*	22*

d_{corr}: net effect of PST compared with C, *significant change compared with start (0 months)

4.4 Treatment for pain

The PST and C groups recorded similar reductions in the percentage of participants receiving medical treatment for the pain (12% and 11%): in contrast, 26% of the MST group were able to dispense with treatment.

5 Results – strength and mobility

The study examined the range of motion of the lumbar spine and the strength of the back extensor muscles. For analysis, we used a Delphex LE machine[15]. This test and therapy machine meets the technical criteria required for precise measurements[16].

5.1 Lumbar extension strength

We tested the maximum isometric strength of back extensor muscles of all participants at a maximum of seven different angles before and after completion of the study. For reasons of clarity, we have not displayed the individual results for each angle but have calculated the mean. The following diagram shows the changes to lumbar extension strength in the three groups.

Table 8: Percentage in each group receiving treatment for back pain in the three months before the start and the end of the study

	PST		MST		C	
	N	**%**	**N**	**%**	**N**	**%**
Group as a whole	317	100	109	100	105	100
Receiving treatment (start)	70	22	62	56	20	19
Receiving treatment (end)	30	9	33	30	8	7
Difference	40	12	29	26	12	11

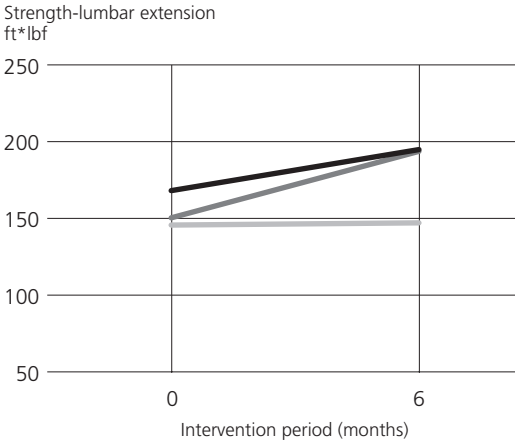
15 Design identical to the MedX Lumbar Extension machine described in: CARPENTER D., NELSON B. (1999): Low back strengthening for the prevention and treatment of low back pain. Medicine and Science in Sports and Exercise, 31, 18 – 24.

16 For additional information see http://www.kieser-training.com/de/medizinische-kraeftigungstherapie/therapiemaschinen.

The strength increase recorded by the MST Group, who had done intensive strength training on the Delphex LE machine, was more (+36%, d = 0.66 moderate effect) than the corresponding increase in the PST group (+20%, d = 0.4 small effect). Existing scientific literature indicates that strength increases following use of the LE machine are likely to be in the range 24% – 42%[17]. The percentage increase achieved by the MST group was within this range. Strength did not change in the control group.

We compared the strength level recorded for participants at each angle with the equivalent level for norm groups of a similar age and gender. The norm values calculated by the University of Florida currently apply to the interpretation of test results[18].

As a result of the training, the percentage of participants with a significant strength deficit declined and the percentage with above average strength levels increased. Overall, distribution in both training groups approached the norm distribution. The improvement was particularly evident in the MST group. Before therapy, 35% had significantly weaker back muscles than the age and gender-specific norm groups. On completion of the study, the comparable figure was only 11%. (Diagram 6).



	0	6
PST	168.07 [160.21;175.93]	194.83 [186.65;203.00]
C	145.65 [134.33;156.97]	147.09 [135.82;158.36]
MST	150.35 [136.09;164.61]	193.67 [179.14;208.20]

Diagram 5:
Strength of back extensors at start and end of intervention (arithmetic mean with confidence interval (95%), change in control group not significant, base data: N_{PST} = 265, N_C = 84, N_{MST} = 82).

Table 9: Effects and changes to back strength (lumbar extension) in each group

		Effects	Change	
Group	**N**	**d (d_{corr}) 6 months**	**absolute 6 months**	**percentage 6 months**
PST	265	0.40* (0.38)	26.75 ± 33.11*	20.27 ± 29.09*
C	84	0.03	1.45 ± 24.27	2.82 ± 18.81
MST	82	0.66*	43.32 ± 34.94*	36.16 ± 33.83*

d_{corr}: net effect of PST compared with C, *significant change compared with start (0 months)

17 See Bibliography, footnote 3, and: HOLMES B., LEGGETT S., MOONEY V., NICHOLS J., NEGRI S., HOEYBERGHS A. (1996): Comparison of female geriatric lumbar-extension strength: asymptotic versus chronic low back pain patients and their response to active rehabilitation. Journal of Spinal Disorders, 9, 17-22.

18 CARPENTER D., POLLOCK M., GRAVES J., FOSTER D., LEGGETT S., GARZARELLA L., FEURTADO D. (1993): Gender and age-specific isometric strength norms of the isolated lumbar extensor muscles. Conference abstract.

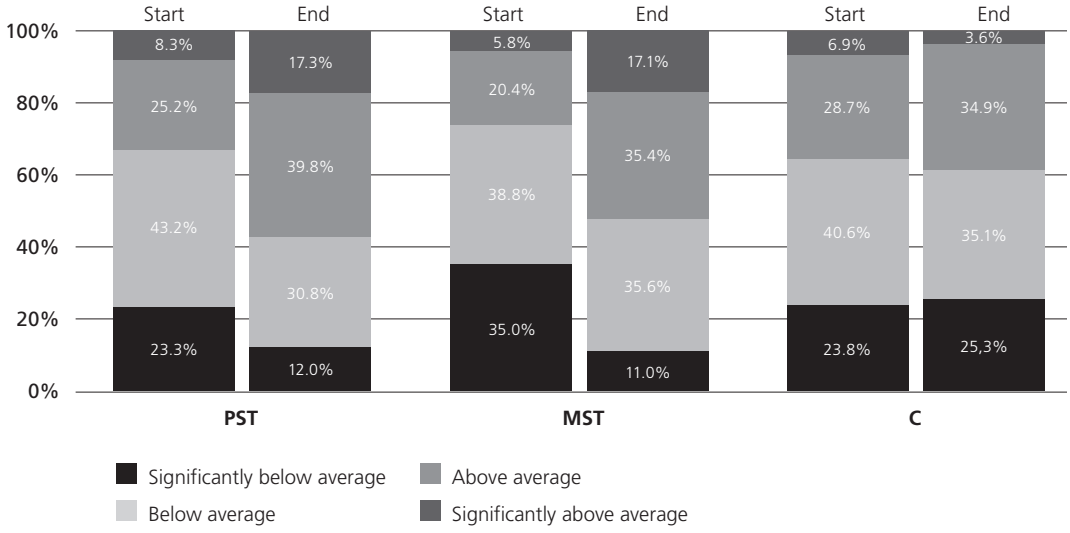


Diagram 6:
Comparison between the norm sample and individual maximum strength levels for participants. If strength level is denoted as significantly below average (above average), this means that strength is more than one standard deviation below (above) the mean norm. A strength level below average (above average) denotes strength within the standard deviation immediately below (above) the mean norm.

5.2 Lumbar mobility

The maximum range of motion of the lumbar spine on the test and therapy machine is 72°. For those with back pain, range is frequently reduced. As symptoms improve, mobility can be gradually increased. All three groups had restricted lumbar mobility at the beginning of the study.

The improvement in the PST and MST groups was statistically significant. However, the improvement only had a significant effect in the MST group (improved by 2.8 ± 8.9 degrees or 7.8 ± 23.4%, significant effect: d = 0.26, N = 82).

In the PST group, the percentage with normal mobility in extension (69-72°) increased by 4% and by 9% in the MST group. In the control group, the percentage declined by 3%.

6 Results – physical health and wellbeing

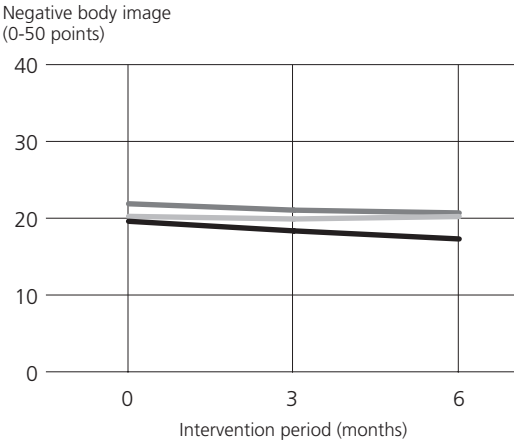
6.1 Body image

We used a body image questionnaire (Body Image Questionnaire 20)[19] to determine how participants perceived their body image. This questionnaire contains 20 five stage items and distinguishes between two dimensions (scales):

- Negative body image (NBI) evaluates external body appearance and whether participants feel in tune with their body and at ease with themselves; here, a lower score means an improvement.
- Vital body dynamics (VBD) evaluates aspects of body image relating to energy and movement and whether participants feel strong, fit and healthy; here, a higher score means an improvement.

19 CLEMENT U., LÖWE B. (1996): FKB-20 Fragebogen zum Körperbild [Body Image Questionnaire) Göttingen: Hogrefe.

In the PST group, NBI declined significantly by -12.6% (d = -0.44, small effect). In the MST group, the score declined significantly by -6.8% (d = -0.20, small effect). There were no changes in the control group. The benefit from strength training (PST) compared with no intervention (C group) was of practical significance, whereby the net effect was d_{corr} = 0.43 and was evident after training for 6 months. (Diagram 7, Table 10).



	0	3	6
PST	19.60 [18.85;20.36]	18.35 [17.64;19.06]	17.30 [16.62;17.98]
C	20.24 [18.70;21.78]	19.90 [18.38;21.41]	20.22 [18.65;21.80]
MST	21.88 [20.54;23.22]	21.05 [19.50;22.61]	20.69 [19.34;22.04]

Diagram 7:
Negative body image (NBI) on the three measurement dates (arithmetic mean with confidence interval (95%), range 0-50), data base: complete data sets (N_{PST} = 208, N_C = 67, N_{MST} = 75).

Table 10: Effects and changes to the scale «Negative Body Image» in each group

		Effects		Change	
Group	**N**	**d (d_{corr}) 3 months**	**d (d_{corr}) 6 months**	**absolute 6 months**	**percentage 6 months**
PST	208	-0.23* (-0.17)	-0.44* (-0.43)	-2.3 ± 3.5*	-12.6*
C	67	-0.05	0.00	0 ± 4.5	-4
MST	75	-0.13	-0.20*	-1.2 ± 3.5*	-6.8*

d_{corr}: net effect of PST compared with C, *significant change compared with start (0 months)

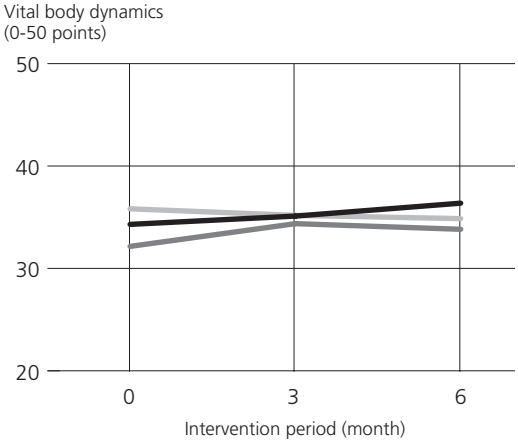
In the PST and MST groups, VBD increased significantly by 6% (d = 0.38 and 0.37, small effect). There were no changes in the control group. The benefit from strength training (PST) compared with no intervention (C group) was evident after just 3 months of training at which time there was a minor net effect (d_{corr} = 0.27). After 6 months, the effect had increased to moderate (d_{corr} = 0.56). (Diagram 8).

In terms of body image, both training groups reported a reduction in negative body image and an increase in vital body dynamics at the end of the 6 months. In contrast, no change was recorded by the control group. (Table 11).

6.2 Performance and quality of life

At the end of the study, participants were asked to rate on a scale of 1 – 4 how various aspects of performance (Diagram 9) and Quality of Life (Diagram 10) had changed during the course of the study (1 = totally inapplicable, 2 = more or less inapplicable, 3 = more or less applicable, 4 = totally applicable).

A very high percentage of those in the PST and MST groups felt that performance and aspects of quality of life had improved considerably following Kieser Training for 6 months. The corresponding percentage increase amongst the control group was less than half.



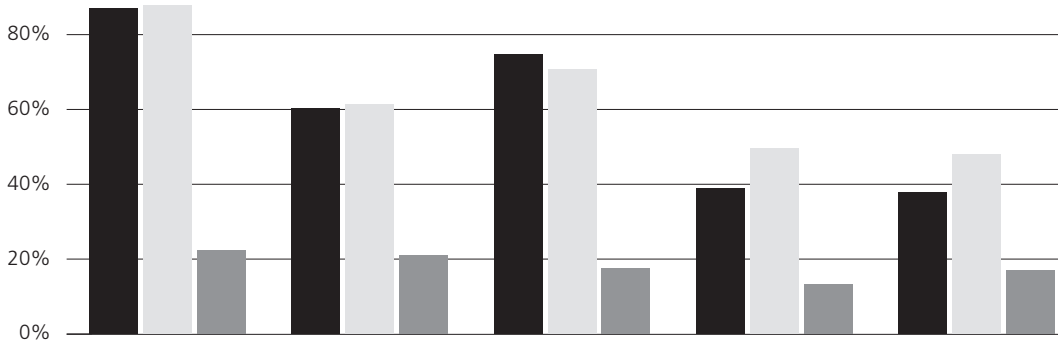
	0	3	6
PST	34.30 [33.48;35.12]	35.11 [34.29;35.92]	36.37 [35.56;37.17]
C	35.82 [34.51;37.13]	35.18 [33.81;36.55]	34.86 [33.39;36.33]
MST	32.15 [30.95;33.35]	34.37 [33.30;35.43]	33.87 [32.67;35.07]

Diagram 8:
Vital body dynamics (VBD) on the three measurement dates (arithmetic mean with confidence interval (95%), range 0-50), data base: complete datasets (N_{PST} = 179, N_C = 50, N_{MST} = 60).

Table 11: Effects and changes to the scale «Vital Body Dynamics» in each group

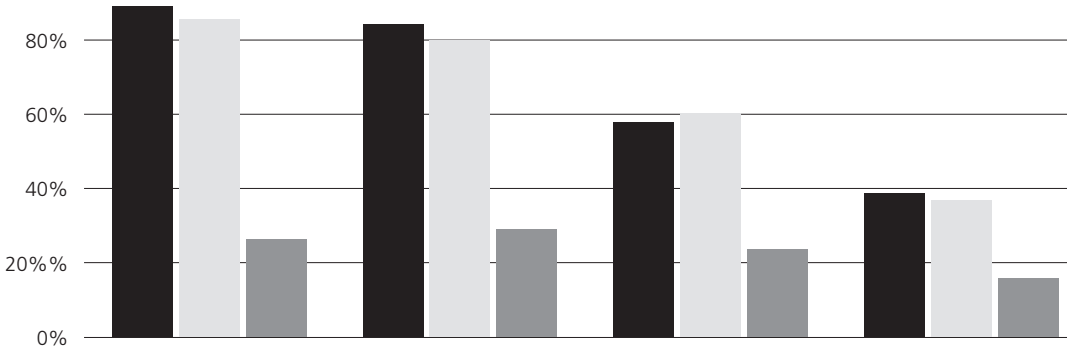
		Effects		Change	
Group	**N**	**d (d_{corr}) 3 months**	**d (d_{corr}) 6 months**	**absolute 6 months**	**percentage 6 months**
PST	179	0.15* (0.27)	0.38* (0.56)	2.07 ± 4.0*	5.7*
C	50	-0.14	-0.20	-0.96 ± 2.4	-1
MST	60	0.50*	0.37*	1.72 ± 4.2*	6*

d_{corr}: net effect of PST compared with C, *significant change compared with start (0 months)



	physical	at work	sports	household	social life
PST	86.9%	60.2%	74.6%	38.8%	37.7%
MST	87.7%	61.3%	70.6%	49.5%	47.9%
C	22.5%	21.0%	17.4%	13.1%	17.1%

Diagram 9:
Percentage indicating a positive change in performance, cumulatively for categories «more or less applicable» and «totally applicable».



	In better phys. shape	relaxed in daily life	less stressed	quality of sleep
PST	89.0%	84.3%	57.7%	38.7%
MST	85.7%	80.0%	60.2%	36.7%
C	26.3%	29.0%	23.7%	15.8%

Diagram 10:
Percentage indicating a positive change in quality of life, cumulatively for categories «more or less applicable» and «totally applicable».

6.3 Self efficacy and satisfaction with life

There was no change in either construct.

The scale for general perceived self-efficacy is based on the concept of perceived self-efficacy as developed by Bandura. According to the authors, perceived self efficacy describes a person's constructive ability to cope with life[20] and measures the ability to master difficult situations (optimistic competence expectancy: maximum score = 40). The scores at the start of study were similar for all 3 groups: 31.16 (PST), 29.9 (MST) and 30.9 (C). After 6 months, only the PST group had achieved a significant improvement but to an insignificant extent ($p<0.05$, $d = 0.15$). The control group and the MST group recorded no change at the end of the study ($p>0.05$ and $d = 0.16$).

We used the «Satisfaction with Life Scale» (SWLS) in order to obtain data on general satisfaction with life. This scale provides a self assessment of a person's overall satisfaction with life[21]. The scores for each individual aspect are added together and the total score (max. 35 points) is then interpreted by sub-dividing it into 7 categories (35-31 = very satisfied, 26-30 = satisfied, 21-25 = more satisfied than dissatisfied, 20 = neutral, 15-19 = more dissatisfied than satisfied, 10-14 = dissatisfied, 5-9 = very dissatisfied). At the start of the study, all three groups were already recording the second highest category for average overall satisfaction with life and this did not change during the study.

6.4 Anthropometric measurements

The Body Mass Index for all three groups did not change during the study. Similarly, there were no changes to the thigh circumference.

7 Evaluation of other Parameters

7.1 Personal training goals

At the end of the study, participants in the MST and PST groups rated on a scale of 0 – 5 (0 = «not at all» and 5 = «in full») the extent to which they had achieved their original training goals after doing strength training for 6 months. These results were then transformed into a scale from 0-100 (see also Footnote 13, P. 3). The resultant percentage indicates the degree to which the participants had achieved their personal training goals. After six months, it was 73.6%. The following diagram shows the average extent to which both training groups had achieved their initial goals.

Overall, individual participants largely achieved their goals. The participants in the PST group achieved their goals by an average of 75%, somewhat higher than the equivalent percentage for the MST group, which was 69.4%. (Diagram 11).

20 JERUSALEM M. (1990): Persönliche Ressourcen, Vulnerabilität und Streßerleben. (Personal resources, vulnerability and stress) Göttingen: Hogrefe.
SCHWARZER R. (1994): Optimistische Kompetenzerwartung: Zur Erfassung einer personalen Bewältigungsressource. (Optimistic self-efficancy: Assessing a person's coping resource) Diagnostica, 40 (2), 105-123.

21 DIENER E., EMMONS RA., LARSEN RJ., GRIFFIN S. (1985): The Satisfaction with Life Scale. Journal of Personality Assessment, 49, 71-75. For this study, the following translation into German was used: SCHUMACHER J. (2003): SWLS – Satisfaction with Life Scale. In J. Schumacher, A. Klaiberg, E. Brähler (Hrsg), Diagnostische Verfahren zu Lebensqualität und Wohlbefinden (Diagnostic processes for quality of life and wellbeing) (Diagnostik für Klinik und Praxis, Volume 2) (Pages 305-309). Göttingen: Hogrefe.

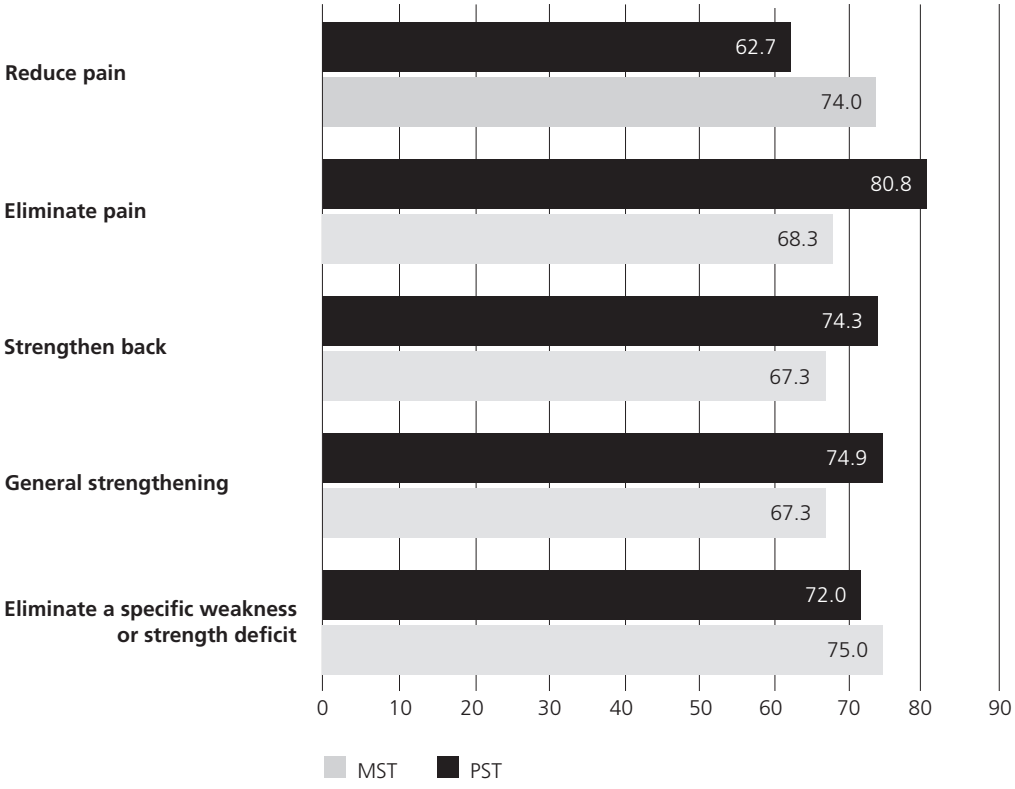


Diagram 11:
Extent to which training goals achieved based on initial training goals.

7.2 Level of satisfaction with service

The six months of the study provided active participants with an opportunity to put Kieser Training to the test. Participants in the MST and PST group were asked to rate their satisfaction with the service on a scale of 0 – 4 (0 = «poor» and 4 = «very good»).

- 72.4% of the PST group and 83.5% of the MST group rated the quality of service as «very good»; 27.2% of the PST group and 16.5% of the MST group as «good»[22].
- 61.1% of PST group and 61.5% of the MST group were very satisfied with the training; 36.9% of the PST group and 36.5% of the MST group were largely satisfied.
- 65.7% rated the quality of staff as «very good» and 29.7% as «good».
- 50.1% rated the ease of operation of machines and their external appearance as «very good» and 45.3% as «good».
- 62.8% rated the cleanliness of the facility as «very good» and 32.0% as «good».

22 Only 1 out of 371 individuals rated the service as «not so good».

8 Summary of results

Kieser Training works: The net effects obtained by comparing a training group and a control group in a randomised study provide clear evidence for this.
Physical wellbeing improved: this was evident from the reduction in both the prevalence of pain (pain index and pain severity) and the impact of that pain on daily life (effects of pain). As expected, strength training increased back strength, particularly in the group who did Medical Strengthening Therapy (MST). Muscle strength in this group was much weaker at the outset compared with those without back problems or pain. Anthropometric measurements did not change.
Mental wellbeing improved: there was an improvement in the subjective image participants had of their body, i.e. a decline in negative body image and an increase in vital body dynamics. In contrast, self-efficacy and satisfaction with life are apparently more stable traits as they did not change – this could also be the result of self-selection by persons in whom these traits were particularly marked. Only someone who believed that they could achieve something with training and reckoned they had a chance of being selected for the «Kieser Training works» study would have applied in the first place.

The study showed that strength training by both intervention groups achieved the following results:

- reduction in pain and its effects
- increase in strength
- improvement in personal body image and
- increase in overall performance and quality of life.

These positive changes were combined with a high percentage of participants achieving their personal training goals and a subjective perception that overall performance and quality of life had improved.

The following table provides a summary of aspects that resulted in relevant changes. As the study produced both positive and negative scores, the results are shown in a uniform way, i.e. a (+) change always means a positive change to that aspect, i.e. healthier, stronger, less pain, etc. In no group was there a deterioration in any aspect.

Table 12: Summary of results for «Kieser Training works» (PST: Group with preventive strength training, C: Control group, MST: Group with Medical Strengthening Therapy)

Category		PST	C	Net effect PST vs. C	MST
Pain	Index	+++	++	+	++
	Severity (PS)	+++	++	+	++
	Effect (EP)	++	+	+	++
Strength and mobility	Lumbar mobility				+
	Strength – lumbar extension	+		+	++
General physical health and wellbeing	Body Mass Index				
	Thigh circumference				
	Self efficacy				
	Satisfaction with life				
	Body image NBI	+		+	+
	Body image VBD	+		++	+

+++ significant change with major effect ≥ 0,8□
++ significant change with a moderate effect ≥ 0,5□
+ significant change with a small effect ≥ 0,2□

9 Abbreviations

C: Control group
d: Effect size
d_{corr}: Corrected effect size
EP: Effects of Pain scale from Medical Outcomes Study
LE: Therapy machine for lumbar extension
MST: Medical Strengthening Therapy
N: Number of persons in test
NBI: Negative body image
PS: Pain Severity scale from Medical Outcomes Study
PST: Preventive Strength Training
VBD: Vital body dynamics

10 Comments on effect sizes

The following extract is taken from a German-language publication by Fröhlich and Pieter (2009)[23] and explains using a specific example the concept of effect sizes.

«[…] The introduction of a new type of sprint training improved 100 m sprint times significantly by 0.05 s (pre-post difference = 0.05 s). How should we interpret this significant improvement in times resulting from the introduction of a new method of sprint training? Should we introduce the new training method? To answer these questions, below are two scenarios:

23 FRÖHLICH M., PIETER A. (2009): Cohen's Effektstärken als Mass der Bewertung von praktischer Relevanz – Implikationen für die Praxis. Schweizerische Zeitschrift für «Sportmedizin und Sporttraumatologie» (Cohen's effect sizes as a measure of evaluating practical relevance – practical implications. Swiss Journal for «Sports Medicine and Sports Traumatology») 57 (4), 139-142.

1. The average 100 m time of participants at a mass participation event (award of sports badges) was 13.04 s (SD = 2.02) and improved by 0.05 s. If we calculate the effect size «d», this gives 13.04 – 12.99/2.02 = 0.025.

2. The average time in the men's 100 m sprint final at the Olympic Games in 2008 was 9.92 s. (SD = 0.11) and improved by 0.05 s. The effect size «d» is, therefore 9.92 – 9.87/0.11 = 0.45.

From this example, we can see that the effect size «d» for sprinters in the final was considerable greater than for participants in the mass participation event. This means that the introduction of the new training method for world-class sprinters would have a practical relevance where for the average runner it would not […] This evaluation of practical significance – as the example of the 100 m sprint shows – has a long tradition and dates back to research done by Cohen (1969). The values normally used in the classification of effect sizes are as follows: small effect d = 0.20, medium effect d = 0.50 and large effect d = 0.80 (Cohen, 1969, Page 38; 1992, Page 157). […] As a result, the effect size – the effect of the new type of sprint training – would be interpreted as a trivial effect for amateur athletes and without any practical significance. However, for world-class sprinters the effect size is d = 0.45, and this would be a moderate or medium effect with medium practical importance or relevance […]»

11 Correlation between verifiable effects and sample size

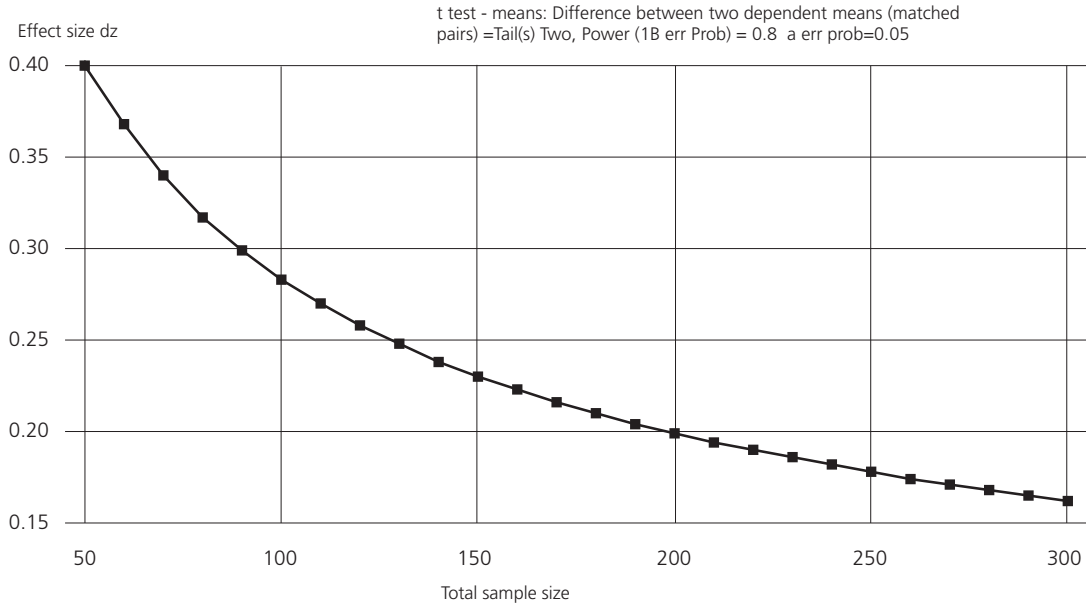


Diagram 12:
*Correlation between verifiable effects (Y axis) and sample size (X axis) Plotted with G*Power Version 3.1.2.*

Address for correspondence:

Dr. Sven Goebel
Head of Research & Development
Kieser Training AG
Kanzleistr. 126
CH-8026 Zurich

Email: sven.goebel@kieser-training.com
Phone: +41 44 296 17 17
Fax: +41 44 296 17 77