

Serie, Satz – Sieg!?



Welches ist die beste Kombination für Kraftzuwachs?

Für viele Methoden des Krafttrainings hat die Wissenschaft „signifikante Verbesserungen“ festgestellt. Aber: Bei vielen dieser Untersuchungen werden entscheidende Faktoren wie z.B. der Trainingszustand oder die Trainingserfahrung der Probanden nicht berücksichtigt. Prof. Dr. Theodor Stemper analysiert für F&G Meta-Studien, welche die sogenannte „Effektstärke“ mit einbeziehen und deshalb die Trainingseffekte verschiedener Trainingsmethoden für verschiedene Zielgruppen präzise ermitteln konnten.



Prof. Dr. Theodor Stemper
Sportwissenschaftler an der Bergischen Universität Wuppertal, 1. Stellvertretender Vorsitzender des Bundesverbandes Gesundheitsstudios Deutschland e.V. (BVGSD) und Ausbilderdirektor des DFAV e.V.

Wer kennt nicht die zahlreichen Empfehlungen zu einem „optimalen Krafttraining“? Sie reichen von „sanftes Krafttraining“ über „Ein-Satz-Training“ und „HIIT“ bis hin zu Intensivierungstechniken, wie „cheatings“, „Supersätze“ oder „forced reps“.

Alle Empfehlungen haben aber eines gemeinsam: Wenn überhaupt, dann treffen sie nur für bestimmte Trainingsgruppen zu und sind nicht universell gültig.

Wie kann man eine „optimale Methode“ ermitteln?

Leider wird der Erkenntnisweg aber nicht immer deutlich genug heraus gestellt. Und ebenfalls fehlt bei vielen Publikationen, besonders aus dem sog.

„Expertenbereich“, zumeist auch die Angabe wissenschaftlicher Kennzahlen. Vieles basiert lediglich auf Einzelfallstudien oder -beobachtungen, gelegentlich sogar nur persönlichen Meinungen bzw. Glaubenssätzen. Und wenn dann solche Kennzahlen doch einmal genannt werden, dann findet sich zumeist nur die Angabe der „Signifikanz“, also ein Hinweis, dass die vorgeschlagene Trainingsmethode zu „signifikanten Verbesserungen“ geführt hat.

Reicht das aber aus? Kann man so überhaupt eine „optimale Methode“ ermitteln und begründen? Oder bedarf es dazu anderer Angaben?

Im Folgenden sollen diese Fragen geklärt werden. Daran anschließend kann dann die aktuell gesicherte

Erkenntnislage zur Frage der sinnvollsten Dosis für die optimale, zielgruppenspezifische Wirkung eines Krafttrainings erläutert werden.

Dass die Klärung dieses Problems erforderlich ist, ergibt sich schlicht aus der Tatsache, dass nur die richtige „Dosis (dose)“ (Trainingsprogramm) auch zur gewünschten „Reaktion (response)“ (Trainingseffekt/-anpassung) führt. Andererseits kommt es entweder zu Übertraining bzw. Überbelastungen oder zur Unterforderung.

Signifikanz bedeutet nicht Relevanz

Was die wenigsten Laien – und erstaunlicher Weise auch manche Wissenschaftler – nicht kennen, ist die Tatsache, dass der alleinige Hinweis auf

„Signifikanz“ in der Regel nicht ausreichend, um die Effektivität einer Maßnahme ausreichend zu belegen. Denn, wenn ein Forschungsergebnis als „signifikant“ bezeichnet wird, dann drückt das nur aus, dass sich z.B. eine untersuchte Gruppe von einem früheren Trainingsstadium oder von einer anderen Untersuchungsgruppe oder im besten Falle auch von einer nicht trainierenden Kontrollgruppe so unterscheidet, dass das Ergebnis nur noch zu einem geringen Teil auf Zufall beruht. **Diese Restunsicherheit (Zufall) verdeutlicht man mit dem sogenannten „Signifikanzniveau“**, das in der Regel bei Trainingsuntersuchungen bei 5% angesetzt wird – bei sensibleren medizinischen Tests dagegen eher bei 1%, ggf. sogar auch bei 0,1%.

Die entsprechende Maßzahl ist ‚p‘. p steht dabei für „probability“ = (Irrtums-)Wahrscheinlichkeit. Zu lesen ist dann: $p < .05$ oder $p < .01$. Das bedeutet, dass z.B. für ein signifikantes Ergebnis von $p < .05$ noch eine 5%-ige Wahrscheinlichkeit gibt, dass es zufällig zustande gekommen ist. Andererseits besitzt es aber auch eine Sicherheit von 95%, dass das Ergebnis durch das eingesetzte Trainingsverfahren, oder ein Medikament oder eine Ernährungsweise usw. „erklärt“ werden kann.

So angesehen ‚ein signifikantes Ergebnis‘ aber nun auch sein mag – es sagt nur etwas über die Existenz eines Effekts aus, aber leider noch gar nichts über die praktische Bedeutsamkeit einer Maßnahme.

Denn erst eine zweite, dimensionslose Maßzahl, die als **„Effektstärke“ (ES)** bezeichnet wird und mit dem Buchstaben ‚d‘ dargestellt wird, gibt einen Hinweis auf die praktische Relevanz (vgl. Cohen, 1969, 1992).

Fehlt diese Angabe, hat ein Untersuchungsergebnis tatsächlich noch keine

große praktische Bedeutung, so signifikant es auch sein mag!

Dass die Effektstärke hier wichtiger als die Signifikanz ist, liegt zum einen daran, dass ‚signifikant‘ lediglich ‚überzufällig‘ bedeutet. Wenn also etwas „signifikant“ ist, wird das nur als Hinweis bzw. Voraussetzung dafür verstanden, daran anschließend die praktische Bedeutsamkeit (Relevanz) weiter zu untersuchen. Signifikanz drückt also erst einmal nur aus, dass sich eine weitergehende Analyse zum Untersuchungsbefund lohnt.

Wichtig zu wissen ist in diesem Zusammenhang auch, dass eine Signifikanz immer dann schnell errechnet wird, wenn Gruppen ausreichend groß oder recht homogen sind oder wenn die Irrtumswahrscheinlichkeit (p) „großzügig“ angesetzt wird. Besonders dann ist die ES umso wichtiger.



Effektstärke als Maß für Relevanz

Ähnlich wie die Signifikanz lässt sich auch die Effektstärke relativ leicht berechnen. Wenn es, wie in Studien zur Effektivität von Trainingsprogrammen üblich, um den Vergleich von Trainingsmethoden geht, dann werden zur Berechnung der ES entweder die Mittelwerte (MW) eines Vortests (prä) zum Nachtest (post) oder einer Versuchsgruppe (VG) zu einer Vergleichs-(Kontroll-)gruppe (KG) benötigt. Zusätzlich wird deren Streuung (s) herangezogen, d.h. die Verteilung der einzelnen Werte in den Gruppen.

Die Effektstärke ‚d‘ errechnet sich dann aus der Differenz der Mittelwerte, geteilt durch die Streuung, d.h. also $d = (MW1 - MW2)/s$.

Auf Basis der Effektstärken kann man wissenschaftlich fundierte Bewertungen zur Bedeutung der Resultate vornehmen.

Allerdings finden sich zu den Grenzwerten unterschiedliche Angaben. Konventionell gilt: Kein Effekt bei $d < .10$, kleiner Effekt bei d von .10 bis .30, mittlerer Effekt bei Werten von .30 bis .50 und großer Effekt bei $d > .50$ (vgl. Bortz & Döring, 2006). Nach Cohen (1969) liegen die Grenzen dagegen bei .20, .50 und .80.

Wissenschaftliche Analysen von Trainingsstudien

Aus den bisherigen Ausführungen sollte deutlich geworden sein, dass pauschale Empfehlungen, Erfahrungsberichte oder Expertenurteile nicht akzeptabel sind, um wissenschaftlich anerkannte Bewertungen von Trainingsstudien vornehmen zu können. Dazu bedarf es kontrollierter Studien mit Ausweis der Effektstärke d.

HAMMER LEASING

10er-Cardio-Bundle

- 3 x 240G-S1 Sitzfahrrad
- 2 x 250G-S1 Liegefahrrad
- 2 x 260G-S1 Laufband
- 3 x 280G-S1 Crosstrainer

ab 695€*



Highlights Ihres neuen Cardio-Bundles

- Top Cardio-Geräte mit modernem Design
- Über 30 verschiedene Trainingsprogramme
- 7" TFT-Display
- Stromlos betriebene Fahrräder und Crosstrainer
- Auf Wunsch mit individuellem Geräte-Branding (Wunsch-Logo)
- Niedrige Leasingraten für mehr finanzielle Freiheit auch für die Serien 2 und 3 möglich



Abbildungen ähnlich.

Jetzt beraten lassen!
☎ 0234. 436 23 | www.scotfit.de



* Die genannten Monatsraten verstehen sich zzgl. MwSt., gelten vorbehaltlich der Bonitätsprüfung und sind tages- sowie bonitätsabhängig. Laufzeit 54 Monate – 20% Restwert.

Doch daneben ist noch eine Reihe weiterer Bedingungen mit zu berücksichtigen, zu denen vor allem der Trainingszustand der Versuchsteilnehmer (zumindest Einteilung in trainiert/trainingserfahren vs. untrainiert/unerfahren) und die Dauer der Trainingsstudie (Wochen, Monate, Jahre) gehören. Und auch die Trainingsmethode bzw. Trainingsintervention selbst muss exakt dokumentiert werden, d. h. Intensität (z.B. in % des Maximums), Dauer (Wiederholungszahl), Umfang (Sätze oder Serien pro Muskelgruppe) und Häufigkeit pro Woche.

Nicht zuletzt ist es nachvollziehbar, dass wegen der bekannten Tatsache, dass Trainingseffekte bei Trainierten schwerer zu erzielen sind, ggf. auch die Grenzwerte der Effektstärken modifiziert werden sollten. Rhea (2004) schlägt daher vor, bei Untrainierten strengere (höhere) Grenzwerte anzusetzen, so dass mittlere Effekte erst bei 1.25 (bei Trainierten dagegen schon bei 0.50), große Effekte erst bei > 2.0 (Trainierte > 1.0) konstatiert werden sollten. Das sind deutlich strengere Maßstäbe als bei den oben genannten, üblichen Grenzwerten.

Dosis-Wirkungs-Beziehung des Krafttrainings

Rhea et al. stellten im Jahre 2003 eine Analyse von 140 Studien, die bis dahin zum Krafttraining vorlagen, mit insgesamt 1.433 Angaben zu Effektstärken vor. Sie analysierten bei jeder Studie im Vorher-Nachher-Vergleich die jeweilige Effektstärke, unterschieden dabei die Effekte nach der Trainingserfahrung und kamen abschließend zu folgenden Ergebnissen:

- Untrainierte erzielten den größten Kraftzuwachs, wenn die **Trainingsintensität** 60% des 1 RM (Einer-Wiederholungs-Maximum) betrug. (d.h. 12 RM). Bei Trainierten wurden dagegen 80% als optimal ermittelt (d. h. 8 RM).
- Auch bei der **Trainingshäufigkeit pro Woche** gab es Unterschiede, da Untrainierte bei ihrer Intensität jede Muskelgruppe optimaler Weise 3 mal pro Woche, Trainierte dagegen nur 2 mal pro Woche belasten sollten.
- Dagegen gab es beim **Umfang**, also der Anzahl der Sätze / Serien keine Unterschiede, da bei beiden Gruppen vier Sätze je Muskelgruppe optimal waren.

Allerdings unterschieden sich die „optimalen Belastungsnormative“ von den ‚benachbarten‘ Varianten nicht so dramatisch, dass nicht auch mit anderen Kombinationen deutliche Trainingserfolge zu erzielen gewesen wären.

So liegen z.B. bezüglich der Serien- bzw. Satzzahl die Effektstärken bei Trainierten bei den Satzzahlen 2 bis 5 mit den entsprechenden Effektstärken von 0.92, 1.00, 1.17 und 1.15 sehr eng beieinander. Nur beim Training mit einem Satz fällt die ES mit 0.47 erheblich geringer aus. Ähnlich verhält es sich bei Untrainierten, wo es bei 1 und 6 Sätzen nur schwache ES von 1,16 und 0.84 gab, aber bei 2 bis 4 Sätzen sehr nahe beieinander liegende Werte von 1,75, 1.94 und 2.28, bei 5 Sätzen dann jedoch wieder eine deutlich schlechtere ES von nur 1,34.



Senior © Monkey Business - Fotolia.com

Das heißt, dass bei allen Trainierenden 2 bis 4 Sätze beachtliche Kraftzuwächse bewirken. Die anfangs genannten 4 Sätze sind „optimal“.

Ähnliches gilt für die Intensität, hier aber nur bei Untrainierten. Dort war zwar die größte ES mit 2.8 bei 60 % zu verzeichnen, aber auch bei 40, 75 und 80 % gab es mit 2.1, 2.1 und 2.0 noch jeweils eine große ES. Anders bei Trainierten, wo bei 80 % mit einer ES von 1.8 das bei weitem beste Resultat vorlag. Dagegen gab es bei 70, 75 und 85 % lediglich eine ES von 0.70, 0.74 und 0.65.

Einschränkend bedenken sollte man bei der Bewertung der Studienergebnisse jedoch, dass zum Teil nur recht kleine Probandengruppen von 14 bis 38 Personen für die Ergebnisse verantwortlich waren.

Mehrsatz-Training langfristig mit größeren Effekten

Last not least ist bei der Frage nach der optimalen Trainingsbelastung auch noch die Laufzeit der dazu erfolgten Studien wichtig. So konnten z. B. Fröhlich et al. (2009) zeigen, dass sich ein Einsatz- und ein Mehrsatz-Training in den ersten 24 Wochen einer Trainingsintervention, wo die Effektstärken i. W. zwischen ca. 0.8 und 1.0 schwanken, kaum unterscheiden müssen, dass sich

dann aber nach etwa einem ½ Jahr in der Phase zwischen 25-30 Wochen mit Effektstärken von 1.24 zu 3.42 ein erheblicher Vorteil zugunsten des Mehrsatz-Trainings einstellt.

Dieser Befund kann jedoch nicht unbedingt überraschen, da bekanntlich gerade zu Beginn einer Trainingsmaßnahme ein erheblicher Kraftzuwachs koordinationsbedingt auftritt und der Muskelumbau sich vermehrt erst nach ca. 8-12 Wochen einstellt, wozu dann aber auch entsprechend höhere Belastungen erforderlich sind.

Fazit

Effekte des Krafttrainings lassen sich korrekt nur darstellen, wenn neben der Signifikanz, die das Vorhandensein eines Unterschieds anzeigt, auch die Effektstärke berechnet und publiziert wird.

Je nach Trainingsniveau sind hier unterschiedliche Grenzwerte sinnvoll, die ausweisen, ob eine Effektstärke ausreicht, um ihr eine kleine, mittlere oder sogar große Bedeutung bzw. Relevanz beizumessen. Verlautbarungen ohne diese Maßzahlen sind dagegen mit Vorsicht zu genießen, da sie eher im Bereich des Wunschdenkens anzusiedeln sind.

Prof. Dr. Theodor Stemper

Ausgewählte Literatur

- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Springer: Berlin, Heidelberg.
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. Academic Press: New York, London.
- Cohen, J. (1992). *A power primer*. *Quant. Meth. Psychol.* 112, 155-159.
- Fröhlich, M. & Pieter, A. (2009). *Cohen's Effektstärken als Maß (sic!) der Bewertung von praktischer Relevanz – Implikationen für die Praxis*. *Schw. Zschr. Sportmed. Sportraum*. 57 (4), 139-142.
- Rhea M.R. (2004). *Determining the magnitude of treatment effects in strength training research through the use of the effect size*. *J. Strength Cond. Res.* 18, 918-920.
- Rhea, M.R., Bernt, B.A., Burkett, L.N. & Ball S.D. (2003). *A Meta-analysis to Determine the Dose-Response for Strength Development*. *Med. Sci. Sports Exerc.* 35, 456-464.