

Zeitschrift: Jugend und Sport : Fachzeitschrift für Leibesübungen der Eidgenössischen Turn- und Sportschule Magglingen

Herausgeber: Eidgenössische Turn- und Sportschule Magglingen

Band: 28 (1971)

Heft: 7

Artikel: Überblick über die Untersuchungen zum Thema der Wertung von Turnwettkämpfen

Autor: Landers, Daniel M.

DOI: <https://doi.org/10.5169/seals-994670>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 14.03.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Überblick über die Untersuchungen zum Thema der Wertung von Turnwettkämpfen

von Daniel M. Landers (1)

Es bleibt zu hoffen, dass der Aufschwung im Mittelstreckenlauf der Frauen weiterhin anhält und recht viele Mädchen diese gesunde Sportart wählen. Es ist eine der wenigen Disziplinen, welche an der frischen Luft während des ganzen Jahres trainiert werden kann und dabei den Kreislauf belastet, die Durchblutung fördert und damit den Organismus stärkt.

Literatur: Leichtathletik / Autorenkollektiv/Sportverlag Berlin 1969 Biologie Prof. Dr. Nöcker, Sportverlag Berlin.

Seitdem es Wettkämpfe im Turnen gibt, haben Trainer und Beteiligte ihre Sorge hinsichtlich der Objektivität in der Bewertung dieser Wettkämpfe ausgesprochen. Im Turnen, wie auch in anderen Sportarten, die künstle-

(1) Dr. Landers ist ein ehemaliger Turner der Westküste und hat sowohl Big Ten- als auch NCAA-Turnerwettbewerbe bewertet. Zurzeit lehrt er auf dem Gebiet der Sozialpsychologie des Sports und ist Forschungsprofessor im Motor Performance and Play Research Laboratory, Children's Research Center, University of Illinois, Champaign, Ill.

rische Elemente enthalten, kommt man ohne menschliche Beobachter (nämlich Punktrichter) nicht aus, die den Leistungswert der Vorführung bestimmen. Es ist daher überraschend, dass ein Sport, der so vollkommen auf menschliche Beobachtung angewiesen ist, um das Ergebnis eines Wettbewerbs zu bestimmen, nicht mehr theoretisches Wissen hinsichtlich der Faktoren zeigt, die die Genauigkeit eines Richters beeinflussen können. Der Zweck dieses Artikels ist

- a) verschiedene Verfahren zu untersuchen, die die Beurteilung von Turnern objektivieren wollen, und
- b) einen Überblick über die Untersuchungen zu geben, die sich mit der Beurteilung von Turnvorführungen befassen.

Tabelle 2

Trainingsbeispiele aus zwei verschiedenen Perioden des 3. Wettkampfjahres:

Plan	Trainiert
18. bis 24. Januar 1970 (2. Aufbauperiode)	
Mo 60' Dauerlauf und Kraftprogramm	Mo 60' Dauerlauf und Kraftprogramm
Di 10 × kurze Steigung und Technik	Di 10 × kurze Steigung und Technik
Mi 30' Trab/10 Diagonalen und Kraft	Mi 30' Trab/10 Diagonalen und Kraft
Do 6 × 300 m lange Steigung und Technik	Do 6 × 300 m lange Steigung und Technik
Fr 40' Tempolauf	Fr 40' Tempolauf
Sa —	Sa —
Sa 50' Dauerlauf	Sa 50' Dauerlauf
8. bis 21. Juni 1970 (Wettkampfperiode)	
Mo 30' locker traben	30' locker traben
Di 2 × 1000 m in 3'22"	2 × 1000 m 3'12"/3'13"
Mi 4 × 400 m in 68 bis 70"	4 × 400 m 66,9"/68,2"/69,2"/68,2"
Do 30' locker traben	—
Fr —	25' locker traben
Sa 800 m in ungefähr 2'20"	800 m 2'15,4"
So —	—
Mo 45' Dauerlauf	50' Dauerlauf
Di 200 m/400 m/200 m/400 m 29" /65" /29" /65"	200 m/400 m/200 m/400 m 29" /65,3" /29,9" /66,5"
Mi 30' traben und 3 bis 6mal 600 m in 1'45"— 1'50"	1'50"/1'48"/1'44"/1'45"
Do 3 × 150 m in 21" 3 × 300 m in 45"	Regen. 30' Waldlauf
Fr —	—
Sa —	—
So SM Staffel in Zürich 3 × 800 m Ziel ungefähr 7' (Einzelzeit 2'16")	3 × 800 m 6'48" Schweizer Rekord Einzelzeit ungefähr 2'12"

Vorgeschlagene Verfahren zur Objektivierung von Turnvorführungen

Noch vor der Entwicklung des FIG-Punktesystems im Jahre 1949 wurden Beobachtungskategorien vorgeschlagen, die eine bessere quantitative Bestimmung und Objektivierung der Wertung im Wettkampf gestatten. 1934 stellte Zwarg²⁹ fest, dass in manchen Ordnungen eine Übung nach fünf Gesichtspunkten — Schwierigkeitsstufe, Aufbau, Ausführung, Einführung und Beendigung — beurteilt werden müsste, und zwar mit einer bestimmten Punkteanzahl für jedes Gebiet. Zwarg zeigte die Schwierigkeiten auf, denen sich die Richter bei dieser Methode gegenübersehen und schlug vor, dass Kürübungen nach den folgenden vier Gesichtspunkten beurteilt werden sollten:

- a) Korrektheit der Ausführung
- b) Schönheit der Ausführung
- c) Schwierigkeit der Teile
- d) Aufbau und Kombination.

Verschiedene Vorschläge wurden gemacht, ein Zweimannsystem für Turner auszuarbeiten, ähnlich wie es für Tennis²¹, Golf²⁷ oder Ringen⁴ benützt wird. Bei jeder dieser Methoden würde ein Turner in jeder Veranstaltung gegen einen bestimmten

Gegner des anderen Teams gestellt, und der jeweilige Sieger des Zweimannwettbewerbs festgestellt werden. Dieser Vorschlag würde dazu beitragen, die Beurteilung von Turnern zu objektivieren, da der Richter nur zwischen zwei Vorfürhungen anstatt zwischen sechs oder mehr zu entscheiden hätte.

Andere Vorschläge sind auf der Annahme begründet, dass die Genauigkeit und Zuverlässigkeit einer Beurteilung erhöht würde, wenn nur jeweils ein Teil der Übung bewertet würde. (Schwierigkeit, Ausführung oder Aufbau.) Bauer², Nooney und Warwick²⁰ und Welser²⁵ haben verschiedene solcher Spezialisationsysteme ausgearbeitet. Andere Autoren^{1/8/14/22} unterstützten die Idee, dass der Richter nur für ein Gebiet verantwortlich sein solle – entweder Schwierigkeit, Ausführung oder Aufbau. Bei dieser Methode werden die drei Wertungen – eine für Schwierigkeit, eine für Ausführung und eine für Aufbau – zum endgültigen Tabellenergebnis zusammengezählt. Neben dieser Trennung der Verantwortlichkeit des einzelnen Richters wurde ausserdem noch empfohlen, dass mehrere Richter (zwei oder mehr) für jedes Gebiet die Gültigkeit der Beurteilung erhöhen könnten und es dadurch möglich wäre, zwei Richter in verschiedenem Winkel zu dem Ausführenden aufstellen zu können^{19/22/26}. Ausser dieser Vereinfachung der Pflichten des Richters wurden Änderungen der FIG-Regeln vorgeschlagen, die die Beurteilung nach verschiedenen Leistungsstufen objektivieren sollen. Für Vorfürhungen der mittleren und unteren Leistungsstufe haben Bauer², Stebbins²², Weber und Koepke²⁴ sowie Wright²⁸ verschiedene Skalen der Schwierigkeitsstufe vorgeschlagen (verschiedene Abstufungen der bewerteten A-, B- und C-Teile), die eine realistischere Einschätzung der Anfänger und der mittleren Leistungsstufe ermöglichen würden als das gegenwärtige FIG-System. Für höhere Leistungsstufen haben Festa¹⁴ und Welser²⁶ vorgeschlagen, dass das Turnen das System übernehmen sollte, nach dem Kunstspringer beurteilt werden, nämlich Multiplikation der Bewertung für Schwierigkeit mit der Summe aus Ausführung plus Aufbau. Welser²⁶ ist

der Meinung, dass diese Methode zusammen mit erhöhten Wertungen der Schwierigkeit^{1/17} eine gerechtere Feststellung der Unterschiede zwischen Spitzeturnern ermöglichen würde, da sie einen weiteren Wertungsbereich vorsieht.

Wissenschaftliche Untersuchungen über die Beurteilung von Turnvorfürhungen

Auf der Basis der acht angegebenen Arbeiten, die sich mit der Beurteilung von Turnwettkämpfen beschäftigen, ergeben sich vier verschiedene, allgemeingängige Ansätze. Wissenschaftler, die die erste Methode benutzen^{12/16} haben untersucht, wie sich der Durchschnitt aller Wertungen der Richter (Bruttowertung) auf Änderungen in der Beurteilung des einzelnen auswirkt, im Gegensatz zu der herkömmlichen Methode, den Durchschnitt der mittleren Wertungen zu finden (Nettobewertung). Es stellte sich heraus, dass sich Unterschiede in der Endreihenfolge ergaben, wenn diese zwei Methoden benutzt wurden und dass der Durchschnitt der Wertungen von vier Richtern (Gesamtwertung) im wesentlichen das gleiche oder sogar ein besseres Resultat ergab als die Nettobewertung¹². Logischerweise sollten kaum Unterschiede zwischen der Bruttobewertung und der Nettobewertung bestehen, wenn alle Richter gleich kompetent sind und den gleichen Standard oder das gleiche Kriterium zum Vergleich benutzen. Der Hauptvorteil der Nettobewertung, speziell für unerfahrene Richter oder für Vorfürhungen, die schwer zu beurteilen sind, liegt im Schutz gegen eine übergrosse Beeinflussung der mittleren und extremen Bewertung eines Richters (die richtig oder auch unrichtig sein kann). Ohne einen äusseren objektiven Standard jedoch, der zum Vergleich herangezogen werden kann, sind Schlüsse, welche Methode besser ist, nicht möglich.

Ein zweiter Ansatz der häufig benutzt wird, um die Beurteilung von Turnern zu untersuchen, beruht auf der Annahme, dass die Qualität der Beurteilung umso geringer wird, je weiter man den Wertungsbereich für einen bestimmten Teilnehmer streut. Indem man den Durchschnitt oder das arithmetische Mittel der mittleren Wertungen der

Punktetafel als Bewertungsstandard^{6/7} oder den durchschnittlichen Unterschied zwischen hohen und niedrigen Wertungen¹⁵ nimmt, ergibt sich, dass gewisse Veranstaltungen und Treffen, die eher einen kleinen als einen grossen Unterschied zwischen den Bewertungen der Richter zeigen, «objektiver» und «besser» beurteilt werden als andere Veranstaltungen und Treffen¹⁵.

Die dritte und vorherrschende Methode, die zur Untersuchung von Wettkampfwertungen herangezogen wird, ist der Grad der Übereinstimmung zwischen zwei Gruppen von Bewertungen (Koeffizient der Wechselbeziehung) oder zwischen mehr als zwei Gruppen (Koeffizient der Übereinstimmung). Autoren, die diese Methode benutzen, haben die meisten Koeffizienten im allgemeinen hoch und signifikant gefunden. Sie waren ausserdem konstant für Veranstaltungen mit verschiedenen Schwierigkeitsstufen¹³. Daraus folgt, dass Punkterichter im allgemeinen den gleichen Standard anwenden^{5/11/12/16/23}. Es wurde ebenfalls festgestellt, dass die Übereinstimmung wenig konstant war bei Veranstaltungen, die schwache Leistungen aufwiesen^{5/12}, und bei der Bewertung von unbeendeten Vorfürhungen (zum Beispiel Sturz von den Geräten)¹¹. Diese Ergebnisse unterstützten die Argumente, dass Änderungen im gegenwärtigen FIG-System notwendig seien, sowohl für Spitzenvorfürhungen^{1/14/19/26}, als auch für Anfänger und mittlere Leistungsstufen, bei denen Stürze von den Geräten am häufigsten vorkommen^{2/22/24/28}. Andere Erkenntnisse, die ebenfalls die Daten der Wechselbeziehungen zur Grundlage haben, scheinen weniger konsequent zu sein. Wenn man zum Beispiel die Konsequenz der Ergebnisse einer Untersuchung mit anderen vergleicht, so scheint nur die Beurteilung von Salto- und Längspferdvorfürhungen^{11/16/23} konsequent geringe Übereinstimmung zu zeigen. Ausserdem zeigen die Richter wenig Übereinstimmung von einer Veranstaltung zur nächsten^{6/12}.

Bevor wir die vierte Methode beschreiben, ist eine kurze Diskussion einiger Unzulänglichkeiten der obigen Methoden der Turnerbewertung angebracht. Alles, was mit der Durchschnittsbeziehungsweise Korrelationsmethode

über einen Richter ausgesagt werden kann, der in seinen Beurteilungen von den anderen Richtern seines Ausschusses abweicht, ist die Tatsache, dass er nicht mit ihnen übereinstimmt. Artikel, die die «Objektivität» der Richterbewertung zum Thema haben ^{5/16}, zeigen eine unrichtige Anwendung dieses Begriffes ¹⁰. Calkin ⁶ und Sterling und Webb ²³ bemerken, dass «immer die Möglichkeit besteht, dass er (der Richter, der nicht mit den anderen übereinstimmt) sehr gut (objektiv) und dass die anderen drei schlecht (subjektiv) sind». Calkin ⁷ schlug vor, dass die Messmethoden verbessert werden könnten, wenn man die Wertungen der Richter eines Ausschusses mit denen anderer Richterausschüsse vergleicht, die den gleichen Turner oder die gleiche Mannschaft bewertet haben. Calkin ⁷ zeigt jedoch auch gewisse Probleme dieser Methode, da Mannschaften unter anderem gut und schlechte Tage haben, was einen Vergleich mit Wertungen anderer Richter schwierig, wenn nicht unmöglich macht.

Die vierte Methode, die eine gefilmte Übung in die Bewertung der Richter miteinbezieht wurde vorgeschlagen, um die Zuverlässigkeit und Genauigkeit der Richterbewertung besser überprüfen zu können. Diese Methodewürde ausserdem auch die Kontrolle vieler äusserlicher, manchmal sehr verwirrender Faktoren ermöglichen, die die Situation eines Treffens immer mit sich bringt. (Zum Beispiel Blickwinkel eines Richters ⁽²⁾, Beeinflussung durch die Zuschauer, Änderung der Bewertung durch die Richter nach einer Konferenz usw.) Indem die Wertungen von gefilmten Vorführungen benutzt werden und nicht die Abschlussergebnisse, können Hypothesen unter kontrollierten Umständen nachgeprüft werden und wirksamere Techniken (zum Beispiel parametrische und non-parametrische Analyse von Abweichungen) auf die Unterschiede zwischen den Bewertungen der Richter und den objektiven Bewertungen mit Hilfe einer Filmanalyse angewandt

werden. Dem Autor ist nur eine Untersuchung bekannt, die sich dieser letzten Methode bediente. Landers ¹⁸ prüfte den Einfluss von Richtern auf Einzel- und Gruppenbewertungen, die nur einzelne Bereiche beurteilten (Bauer Methode) und mit Richtern, die alle drei Bereiche beurteilten (FIG). Zwölf FIG-Richter, die die Nordkalifornische Turnrichterassoziation repräsentierten und zwölf Richter, die die Bauer Methode anwandten, wurden ausgewählt um die Meisterschaften der Big Ten-Universitäten zu beurteilen. Sie bewerteten 23 olympische Vorführungen verschiedener Leistungsstufen. Die zwölf Bewertungen nach der Bauer Methode (eine für Schwierigkeit, eine für Aufbau und eine für Ausführung) wurden wahllos vereinigt, um diese «Nominal»-Gesamtergebnisse mit den Gesamtergebnissen der FIG-Einzelrichter zu vergleichen.

Eine 24. Übung – ein Duplikat einer der anderen Übungen – wurde in den Film miteinbezogen, um die Zuverlässigkeit der Richter festzustellen. Eine Analyse des Films wurde mit Hilfe eines Filmeditors und den gedruckten Bewertungsnormen für jede Methode durchgeführt, um objektivere «absolute» Bewertungen für beide Beobachtungssysteme zu garantieren ⁽³⁾. Danach wurden zwei «absolute» Bewertungen, eine für jedes Beobachtungssystem und für jede der 23 Übungen gegeben. Zwei dieser «absoluten» Bewertungen der Übungen wurden vom Untersucher selbst bestimmt – eine für jedes Beobachtungssystem. Die anderen zwei «absoluten» Bewertungen wurden von anerkannten Autoritäten in jedem der beiden Beobachtungssysteme bestimmt. Die Bewertungen für Schwierigkeit, Aufbau und Ausführung wurden von den Personen, die «absolute» Bewertungen gaben, mehrere Male geprüft. Alle ausser einem (0,759) der daraus resultierenden Objektivitätskoeffizienten für Schwierigkeit, Aufbau und Ausführung waren grösser als

0,945. In den wenigen Fällen, in denen kleine Diskrepanzen (alle weniger als 0,04) bei den Bewertungen der Prüfer vorlagen, wurde der Durchschnitt aus den zwei «absoluten» Bewertungen genommen, so dass eine «absolute» Bewertung bestimmt werden konnte. Die Ergebnisse zeigten, dass die Richter der Bauer Methode, die nur ein Gebiet bewerteten, bedeutend weniger Abweichungen von den «absoluten» Bewertungen ($p < 0,01$), weniger interne Abweichungen von ihrem eigenen Mittel ⁽⁴⁾ ($p < 0,01$) und höhere Zuverlässigkeit (0,853) als die FIG-Richter (0,619) hatten. Als die «nominalen» Richter der Bauer Methode und des FIG-Systems in Gruppenbewertungen zusammengefasst wurden, zeigten sich keine bedeutenden Unterschiede der Methoden an Hand von Abweichungen von den «absoluten» Bewertungen; die Richter der Bauer Methode hatten jedoch bedeutend weniger interne Abweichungen ($p < 0,05$) als die FIG-Richter. Die Ergebnisse zeigen deutlich die grössere Wirksamkeit einer Teilung der Verantwortlichkeit des Einzelrichters (Bauer Methode) gegenüber Richtern, die alle Gebiete einer Vorführung beurteilen müssen (FIG). Ausserdem zeigte sich eine nicht signifikante Tendenz ($p < 0,05$, $< 0,10$) zum Vorteil für die Richter nach der Bauer Methode und deren grössere Genauigkeit in der Beurteilung gegenüber den Gruppenergebnissen der FIG-Richter, obwohl die Methode, die Bewertungen der Einzelrichter in beiden Methoden zusammenzufassen beabsichtigte, die Varianz des Einzelrichters herabzusetzen.

Zusammenfassung

Von allen Vorschlägen, die Beurteilung von Turnern weniger subjektiv zu gestalten, wurde nur eine Methode tatsächlich getestet – nämlich ob die Teilung der Verantwortlichkeit der Richter zu erhöhter Genauigkeit führt. Obwohl die Bauer Methode bei den Big Ten-Wettkämpfen nicht mehr benutzt wird, wurde ihre Überlegenheit über die FIG-Methode bewiesen,

(2) Man beachte, dass in der Untersuchung von Sterling und Webb ²³ die geringere Übereinstimmung der Richter bei Längspferdvorführungen, das Ergebnis der Blickwinkel der Richter gewesen sein könnte, und zwar in bezug auf den Flug vorher und nachher.

(3) Landers, Daniel M. A Comparison of Two Gymnastic Judging Methods (Ein Vergleich zweier Methoden der Turnerbeurteilung), widmet der Diskussion von Verschiedenheiten in der Regelinterpretation zwischen der Bauer Methode 1965 ³ und dem FIG-Punktesystem für Männer 1964, ein ganzes Kapitel.

(4) Interne Abweichung ist ein Mass der Standardabweichung des Richters über seine eigene mittlere Abweichung für alle 23 Vorführungen.

(zum Beispiel weniger Abweichungen von objektiveren «absoluten» Bewertungen). Die Abschaffung der Bauer Methode wurde teilweise durch

- a) niedrigere Bewertungen der Big Ten-Turner, verglichen mit FIG-Bewertungen veranlasst und
- b) durch die Möglichkeit, dass ein unfähiger oder unehrlicher Schwierigkeits- oder Aufbaurichter die ganze Veranstaltung über den Haufen werfen kann.

Die erste Kritik würde nur bestätigen, dass die Fähigkeit der Richter erhöht ist, jeden Fehler zu sehen und dementsprechend Punkte abzuziehen. Die Möglichkeit, dass ein Richter eine ganze Veranstaltung zunichte machen könnte, würde ausgeschaltet, wenn man sich strikte an den Vorschlag Welsers²⁶ und Stebbins²² hielt, zwei oder mehrere Richter für jedes Gebiet aufzustellen.

Zusammen mit all den einzelnen Auswirkungen des Problems einer Trennung der Verantwortlichkeit der Richter, verdienen die anderen Vorschläge zur Objektivierung der Turnerbeurteilung Berücksichtigung als wissenschaftliche Untersuchungen (zum Beispiel Bestimmung des Siegers in Zweimannkämpfen, verschiedene Bewertungssysteme für verschiedene Leistungsstufen...). Wie in den meisten solchen Bemühungen konnte jedoch die Theorie mit den praktischen Problemen nicht Schritt halten. Um viele der Irrtümer zu vermeiden, die bei einer tatsächlichen Veranstaltung auftreten, ist als Forschungsstrategie die Verwendung gefilmter Vorführungen von Anfang an in einer kontrollierten Situation zu empfehlen. Durch die Verwendung von Filmen der Vorführungen kann die Zuverlässigkeit der Richter bestimmt werden (wiederholtes Ansehen der Vorführung) und durch Filmanalyse kann ein objektivere «absolute» Bewertung aufgestellt werden, mit der die Bewertungen der Richter verglichen werden können. Sind diese Probleme dann unter kontrollierten Umständen getestet, und zeigt es sich, dass bestimmte Faktoren die Bewertung der Richter bedeutend beeinflussen, können diese Faktoren noch einmal unter semi-kontrollierten Umständen getestet werden (zum Beispiel ohne Zuschauer, ohne Änderung der Bewertung nach einer Kon-

ferenz, oftmaliges Prüfen des Blickwinkels usw.), und zwar an Hand der tatsächlichen und nicht der gefilmten Vorführung, um ihre generelle Bedeutung für die tatsächliche Wettkampfsituation festzustellen. Auf diese Weise könnten sich Änderungen und Verbesserungen des gegenwärtigen FIG-Punktesystems ergeben, die den «menschlichen Massstab» vervollkommen würden, der doch für alle Turnwettkämpfe notwendig ist.

- ¹ Bailie, Sam. We must have a new scoring system. [Ein neues Punktesystem ist notwendig.] *Modern Gymnast* 1955, 7, March, pp. 12–13.
- ² Bauer, George V. Let's modernize gymnastic judging. [Die Beurteilung von Turnern muss modernisiert werden.] *Modern Gymnast* 1962, 4, July-August, pp. 24–25.
- ³ Bauer, George V. Big Ten gymnastics judging rules for men. [Die Bewertungsregel der Big Ten-Turnerwettkämpfe für Männer.] Mimeographed paper, University of Wisconsin, Madison, 1963.
- ⁴ Beyer, Bud. A new concept in gymnastic meets. [Eine neue Auffassung von Turnveranstaltungen.] *Modern Gymnast* 1963, 5, September, p. 22.
- ⁵ Brown, James R. Objectivity of judging the 1963 Illinois state high school gymnastics meet. [Objektivität bei der Beurteilung des Gymnasialtreffens im Staate Illinois 1963.] Unpublished Master's thesis, Northern Illinois University, 1963.
- ⁶ Calkin, Gerald F. The computer takes a look at college judging. [Beurteilung der Universitätsturnveranstaltungen vom Computer aus gesehen.] *Modern Gymnast* 1968, 10, March, p. 24.
- ⁷ Calkin, Gerald F. The computer takes a new look at college judging. [Beurteilung der Universitätsturnveranstaltungen noch einmal vom Computer aus gesehen.] *Modern Gymnast* 1969, 12, January, p. 21.
- ⁸ Darling, Tom and Trainer, Ted. FIG judging? [Beurteilung nach dem FIG-System?] *Modern Gymnast* 1965, 7, March, p. 12.
- ⁹ Davis, Roy. Notes from a neurotic judge. [Notizen eines nervösen Richters.] *Modern Gymnast* 1966, 8, March, pp. 30–31.
- ¹⁰ Edwards, Allen L. Statistical methods for the behavioral sciences. [Statistische Methoden für Verhaltensforschung.] New York: Holt, Rinehart and Winston, 1963.
- ¹¹ Faulkner, John and Loken, Newt. Objectivity of judging at the national-collegiate meet: A ten-year follow-up study. [Objektivität bei der Beurteilung des nationalen Universitätstreffens.] *Research Quarterly* 1962, 33, October, pp. 485–486.
- ¹² Faulkner, John and Loken, Newt. The judging of the 1961 Western Conference and NCAA gymnastic meets. [Die Beurteilung des Western Conference und NCAA-Treffens 1961.] *Modern Gymnast* 1962, 4, January, pp. 26–27.
- ¹³ Faulkner, John and Loken, Newt. A further comment on gymnastic scores. [Noch ein Kommentar über Turnerbeurteilung.] *Modern Gymnast* 1962, 4, March, p. 25.
- ¹⁴ Festa, Angelo. A better system of judging gymnastics. [Ein besseres System für die Beurteilung von Turnern.] *Physical Educator* 1963, 20, March, pp. 19–20.
- ¹⁵ Hughes, Eric and Myerson, Gary. An evaluation of gymnastic judging. [Eine Beurteilung der Turnerbewertung.] *Modern Gymnast* 1959, 2, November–December, pp. 56–57.
- ¹⁶ Hunsicker, Paul and Loken, Newt. The objectivity of judging at the NCAA gymnastic meet. [Die Objektivität bei den NCAA Treffen.] *Research Quarterly* 1951, 22, December, pp. 423–426.
- ¹⁷ Landers, Daniel M. A comparison of two gymnastic judging methods. [Ein Vergleich zweier Beurteilungsmethoden für Turner.] Unpublished Master's thesis, University of Illinois, 1965 [University of Oregon microcard No. GV511].
- ¹⁸ Landers, Daniel M. Effect of the number of categories systematically observed on individual and group performance ratings. [Wirkung der Anzahl von systematisch beobachteten Kategorien auf die Bewertung von Einzel- oder Gruppenvorführungen.] *Perceptual and Motor Skills* 1969, 29, December, pp. 731–735.
- ¹⁹ Muzyczko, Ted. «Judges corner.» [Richterdiskussionen.] *Modern Gymnast* 1969, 11, November, p. 24.
- ²⁰ Nooney, J. and Warrick, A. Simplified approach to judging. [Vereinfachte Haltung der Beurteilung gegenüber.] *Modern Gymnast* 1965, 7, May-June, p. 36.
- ²¹ Ryser, Otto E. Proposed new method of scoring dual gymnastic meets. [Vorschlag für die Beurteilung von Doppel-Turnertreffen.] *Scholastic Coach* 1956, 26, September, pp. 65–69.
- ²² Stebbins, Clay. On objective judging [letter-to-the editor]. [Über objektive Beurteilung.] *Modern Gymnast* 1962, 4, December, p. 29.
- ²³ Sterling, Leroy F. and Webb, Roger W. Scoring behavior of gymnastics judges. [Richter und ihre Bewertungen.] *Modern Gymnast* 1969, 12, April, pp. 18–19.
- ²⁴ Weber, Chris and Keopke, Barry. A simple system of evaluating «hard to evaluate» low and intermediate routines. [Ein einfache⁹ System für die Beurteilung von schwer zu beurteilenden Vorführungen unteren und mittleren Grades.] *Modern Gymnast* 1968, 11, November–December, p. 41.
- ²⁵ Welsler, Lyle. Judging – as I see it. [Beurteilung – wie ich es sehe.] *Modern Gymnast* 1965, 7, Juli–August, pp. 34–36.
- ²⁶ Welsler, Lyle. Let's balance up – difficulty X execution. [Schwierigkeitsstufe X – wir beurteilen.] *Modern Gymnast* 1967, 9, October, p. 25.
- ²⁷ Wilson, Glenn G. Let's make judging easier for dual meets. [Die Beurteilung von Doppeltreffen soll leichter werden.] *Modern Gymnast* 1959, 2, May–June, p. 34.
- ²⁸ Wright, William. Judging high school gymnastics. [Die Beurteilung von Gymnasialturnern.] *Modern Gymnast* 1969, 12, August–September, p. 26.
- ²⁹ Zwarg, Leopold F. Judging and evaluation of competitive apparatus or gymnastic exercises. [Beurteilung und Bewertung von Geräte oder Bodenturnen bei Wettkämpfen.] *Mind and Body* 1934, 41, November, pp. 152–155; Also abstracted in the *Modern Gymnast* 1959, 2, April, p. 5.