

Nutzung multikriterieller und unscharfer (FUZZY-)Analysen zum Talentscreening im Sport

**Jochen Ester, Jochen Zinner, Till Utesch
& Dirk Büsch**

Informatik Spektrum

Organ der Gesellschaft für Informatik
e.V. und mit ihr assoziierter
Organisationen

ISSN 0170-6012

Informatik Spektrum

DOI 10.1007/s00287-020-01251-w



Your article is protected by copyright and all rights are held exclusively by Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



Nutzung multikriterieller und unscharfer (FUZZY-)Analysen zum Talentscreening im Sport

Jochen Ester¹ · Jochen Zinner² · Till Utesch³ · Dirk Büsch⁴

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2020

Zusammenfassung

Zum Talentscreening im Sport werden Testbatterien eingesetzt, die unterschiedliche sportmotorische Fertigkeiten sowie entwicklungsrelevante Parameter erheben. Standardauswertungsverfahren führen in der Regel zur Einordnung der Getesteten in zumeist fünfstufige Leistungsklassen. Insbesondere bei großen Stichproben werden allerdings Verfahrensweisen benötigt, die auch innerhalb der „besten Gruppe“ und „schwächsten Gruppe“ noch zuverlässig differenzieren. Dies gelingt z. B. im Projekt BERLIN HAT TALENT mit Verfahren, die auf multikriteriellen und unscharfen mathematischen Theorien und Methoden beruhen. Dabei ist hervorzuheben, dass subjektive Theorien und Wissensstände von Trainer*innen, Lehrer*innen und Sportwissenschaftler*innen explizit mit einbezogen werden können. Dies geschieht auf Basis diskursiver Validierungen und der Konstruktion von Bewertungsfunktionen, die auch als sogenannte Zugehörigkeitsfunktionen für Fuzzy-Sets interpretiert werden können. Die vorgestellten multiattributiven Entscheidungsmodelle zeigen unter Einbeziehung des Expertenwissens den Mehrwert dieses Vorgehens auf und können mit Hilfe einer interaktiven Software transparent nachvollzogen werden.

Abstract

In sports, motor assessments are often used in the context of talent screening. They consist of different motor skills as well as other relevant person parameters. Standard evaluation methods usually rely on the classification of the results mostly in five-point performance categories. Especially in large samples, however, procedures are needed that are able to provide reliable differentiations within the “best group” and the “worst group”. The solution, e.g., in the project BERLIN HAT TALENT, is the usage of multicriteria and fuzzy mathematical theories and methods. In this approach, subjective theories and knowledge of coaches, teachers and sports scientists can be explicitly integrated in the decision process. This is realized by discursive validations and the construction of evaluation functions, which can also be interpreted as so-called membership functions for fuzzy sets. The added value of the presented multi-attributive decision models is the possibility to integrate expert knowledge. Finally, the models are illustrated within an interactive software.

Problemstellung

Im Rahmen von bundesweiten Talentscreenings und Bewegungschecks besteht eine bedeutsame Zielsetzung darin, die Kinder zu identifizieren und zu fördern, die entweder eine gewisse Bewegungskompetenz bzw. sportliche Talente

oder gesundheitsrelevante Bewegungsdefizite aufweisen. Zur Identifizierung werden häufig Testbatterien, z. B. der Deutsche Motorik-Test 6–18 [4], eingesetzt, die unterschiedliche sportmotorische Fertigkeiten und Fähigkeiten adressieren, bei denen eine Bewertung über mehrstufige

Jochen Ester
jochen.ester@gmx.de

Jochen Zinner
jochen.zinner@dhgs-hochschule.de

Till Utesch
till.utesch@uni-muenster.de

✉ Dirk Büsch
dirk.buesch@uol.de

¹ Limbach-Oberfrohna, Deutschland

² Institut für Leistungssport und Trainerbildung, Deutsche Hochschule für Gesundheit & Sport, Berlin, Deutschland

³ Institut für Erziehungswissenschaft, Westfälische Wilhelms-Universität Münster, Münster, Deutschland

⁴ Institut für Sportwissenschaft, Fakultät IV – Human- und Gesellschaftswissenschaften, Carl von Ossietzky Universität Oldenburg, Oldenburg, Deutschland

Leistungskategorien unter Berücksichtigung des kalendarischen Alters und Geschlechts erfolgt [17, 36–38]. Auf der Grundlage eines Summenwerts oder eines Testprofils kann auf diese Weise zumindest eine zuverlässige kategoriale Einteilung, z. B. eine Dreiteilung von sportlich leistungsstarken („high performer“), differenziert leistungsstarken/-schwachen („average performer“) und leistungsschwachen Kindern („low performer“) vorgenommen werden. Für ein Talentscreening erweist sich die Kategorie leistungsstarker Kinder jedoch als zu unspezifisch, da eine Eignung für eine Sportart oder eine Sportartengruppe, z. B. Basketball oder Sportspiele, eine differenzierte Bewertung bzw. Gewichtung unterschiedlicher Faktoren der sportmotorischen Leistungsfähigkeit erfordert. Ebenso kann der individuelle Förderbedarf leistungsschwacher Kinder nicht aus einem Summenwert oder Leistungskategorien abgeleitet werden, sondern erfordert ein differenziertes Testprofil bzw. ein differenziertes Fertigungs- und Fähigkeitsprofil. Des Weiteren sind für eine valide Talent- oder Förderdiagnostik und -prognose neben den relevanten sportmotorischen Parametern weitere entwicklungsrelevante Parameter, z. B. die finale Körperhöhe bzw. das biologische Alter, das Trainingsalter, die elterliche Unterstützung usw. zu berücksichtigen [40].

Im Projekt BERLIN HAT TALENT [42] absolvieren jedes Jahr ca. 7000 und ab 2020 flächendeckend alle Kinder der dritten Schulklasse in Berlin den Deutschen Motorik-Test 6–18 [5], um anschließend Talent- bzw. Bewegungsfördergruppen zu bilden, deren Teilnehmer*innen über einen längeren Zeitraum, z. B. 6 oder 12 Monate, spezifisch gefördert werden. Schwerpunktmäßig in der Talentdiagnostik werden bereits seit einigen Jahren Verfahren erfolgreich eingesetzt, die auf multikriteriellen und unscharfen mathematischen Theorien und Methoden beruhen. Statt einer allgemeinen zusammenfassenden Kategorisierung in „Leistungsgruppen“ werden die Kinder innerhalb eines Ordnungsverfahrens auf der Grundlage gewichteter Personenmerkmale in eine „Rangfolge“ gebracht, die auch innerhalb der „besten Gruppe“, aber auch der „schlechtesten Gruppe“ noch zuverlässig differenziert. Als sehr vorteilhaft erweist sich bei diesen Verfahren, dass subjektive Theorien und Wissensstände von Trainer*innen, Lehrer*innen und Sportwissenschaftler*innen explizit einbezogen werden können [40]. Dieses Wissen wird mithilfe diskursiver Validierungen erschlossen und in Bewertungsfunktionen abgebildet, die in der FUZZY-Theorie als sogenannte Zugehörigkeitsfunktionen definiert bzw. interpretiert werden können. Die auf diese Weise jedem Probanden zugeordneten Werte bilden die Grundlage zur Berechnung einer Rangfolge: Das Kind respektive das Talent, welches in dieser Rangfolge an erster Stelle steht, ist bezüglich der vergebenen Bewertungen (Gewichtungen, Substitutionsraten, Kompensationsgrade, Fehlerangaben usw.) der/die Beste, weil die bei ihm/ihr festgestellten Merkmale (Leis-

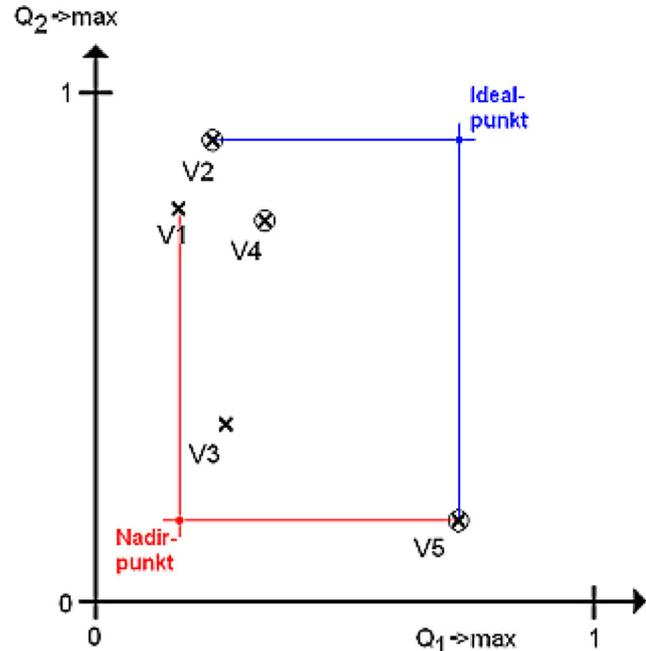


Abb. 1 Multiple-Attribute-Decision-Making (MADM)-Problem

tungsmerkmale, Körperbaumerkmale, biologisches Alter usw.) den „Gütemodellen“ am besten entsprechen [12, 40, 42]. Die konkrete Vorgehensweise, mit der Talentscreenings und Bewegungsschecks durchgeführt werden können, sodass Kinder mit einer möglichen Eignung für eine Sportart oder Sportartengruppe (in unserem Beispiel die Spielsportarten) zu identifizieren sind, wird in diesem Beitrag erläutert.

Die Methodik des multikriteriellen und unscharfen Variantenvergleichs

Für die Bewertung von verschiedenen Varianten/Probanden (hier zum Beispiel den Kindern der dritten Schulklassen in Berlin), die jeweils durch eine endliche Anzahl von Merkmalen, sogenannten Attributen, charakterisiert sind, erscheint ein multikriterieller Variantenvergleich (Multiple Attribute Decision Making, MADM), d. h. ein multiattributives Entscheidungsverfahren, als ein geeigneter Zugang [2, 21, 24]. Ziel ist dabei entweder das Finden der besten Variante, der klassische MADM-Fall, oder die Ermittlung einer Rangfolge der Varianten hinsichtlich einer optimalen „Gesamtgüte“ (hier des Talents für eine Sportart oder eine Sportartengruppe). Seit den 70er-Jahren des vorigen Jahrhunderts beschäftigt sich die Wissenschaft mit dieser Fragestellung und es wurden eine Vielzahl von Grundalgorithmen entwickelt, die bis heute in zahlreichen Varianten weiterentwickelt wurden (siehe z. B. [1, 13, 20, 35]). Eine methodologisch hervorzuhebende Lösung für multikriteri-

elle Entscheidungs- und Optimierungsprobleme ergab sich infolge der Einführung von „unscharfen Mengen“, sogenannte FUZZY-Sets, deren Grundlagen von Zadeh in den 60–70er-Jahren des vorigen Jahrhunderts eingeführt wurden [39].

Die Grundproblematik kann am folgenden einfachen Beispiel beschrieben werden. Falls beispielsweise 5 Varianten ($n=5$: V_1, V_2, V_3, V_4, V_5) durch zwei Kriterien ($m=2$: Q_1, Q_2) beschrieben werden und der Erfüllungsgrad der Kriterien quantifiziert werden kann, z. B. mit den Werten zwischen 0 (unzureichende Ausprägung) und 1 (herausragende Ausprägung), so ist dieses Problem mit der Grafik in Abb. 1 vollständig beschrieben. In einem Koordinatensystem mit den zwei Kriterien kann jede der Varianten ausreichend gut beschrieben werden. Es gibt demnach einen besten Punkt (Idealpunkt) sowie einen schlechtesten Punkt (Nadirpunkt).

Unter der Prämisse, dass mit den verwendeten Merkmalen alle güterelevanten Eigenschaften erfasst wurden, kann zunächst keine vollständige Ordnung der Varianten erzeugt werden. Im Beispiel (siehe Abb. 1) ist die Variante V_1 schlechter als die Variante V_2 und die Variante V_3 schlechter als V_4 . Diese teilweise Ordnung wird als (Vektor-)Halbordnung bezeichnet. Alle Varianten, die in diesem Sinne nicht durch andere Varianten dominiert werden, bilden die sogenannte Pareto-Menge (auch effiziente Menge). Im vorliegenden Beispiel sind das die Varianten V_2, V_4 und V_5 . Welche von diesen wiederum die beste ist, kann ohne zusätzliche Informationen nicht entschieden werden.

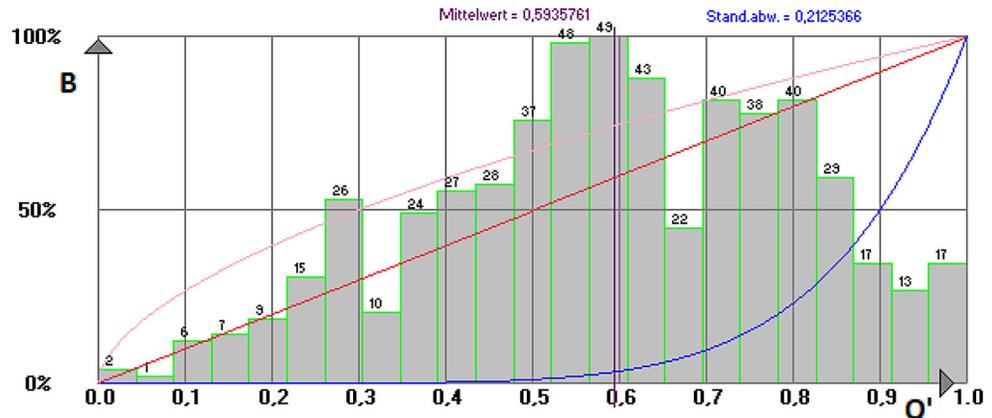
Besteht die Fragestellung darin, *die beste* Variante zu finden, so können alle nicht effizienten Varianten ausgesondert werden, denn die beste Variante muss pareto-optimal sein. Soll jedoch eine *Rangfolge* erzeugt werden, müssen zusätzliche Informationen berücksichtigt werden, z. B.:

- Gewichtungsfaktoren (Gewichtungen, Gewichte): Positive Zahlen zwischen Null und Eins, die die Bedeutsamkeit eines Kriteriums in Bezug auf die Gesamtgüte anzeigen.
- Zufriedenheitsgrade: Bewertung der Ausprägung eines Kriteriums hinsichtlich der angestrebten Zielerreichung.
- Substitutionsraten: Positive Verhältniszahlen, die angeben, um wie viele Einheiten ein Kriterium verbessert werden muss, wenn sich ein anderes Kriterium um eine gewisse Anzahl von Einheiten verschlechtert hat und die weiteren Kriterien konstant geblieben sind.
- Fehler- oder Toleranzangaben: Messfehler, die aus der Problemstellung resultieren.
- Empfindlichkeits- und Schwellenwerte: Die Angabe einer Differenz in einer Merkmalsausprägung, die groß genug ist, dass eine Variante gegenüber einer anderen Variante eindeutig präferiert werden kann.
- Zielstrukturen: Die Angabe einer Zielstruktur, anhand derer Unterattribute strukturiert werden können.

- Kompensationsgrade: Kann ein Verlust bei einem Kriterium durch den Zugewinn bei einem anderen Kriterium überhaupt kompensiert werden?
- Wunschwerte, Wunschintervalle, Mindestausprägungen u. a. m.

Notwendige zusätzliche Informationen lassen sich durch *diskursive Validierungen* mit Verfahrensbeteiligten oder Expert*innen (in unserem Falle also insbesondere Trainer*innen, Lehrer*innen und Sportwissenschaftler*innen) generieren. Die damit zwingend erforderliche *Auseinandersetzung mit den Daten* begünstigt einerseits eine fundierte und bewusste Position des „Untersuchenden“ sowohl zu den Analyseinhalten als auch zu den Analyseergebnissen. Die unterschiedliche Art von Zusatzinformationen führt andererseits dazu, dass sich mehrere, jeweils plausible und zweckmäßige, aber doch deutlich unterschiedliche Vorgehensweisen und Methoden entwickeln lassen, die auch zu veränderten Analyseergebnissen führen. Das liegt insbesondere daran, dass sich die Wirkung der einzelnen Parameter auf eine Gesamteinschätzung zur sportmotorischen Begabung bzw. zum sportlichen Talent beispielsweise nicht durch relative Häufigkeiten, sondern insbesondere durch unscharfe, vage, subjektive Abschätzungen beurteilen lassen. Es lohnt sich deshalb nicht, zwischen diesen Methoden a priori auszuwählen und sich auf eine einzige beschränken zu wollen. Ziel sollte es vielmehr sein, die Vielfältigkeit der Methoden zu nutzen, um das (komplexe) Problem aus verschiedenen Perspektiven zu beurteilen. Das kann ein Vorteil gegenüber der Nutzung von herkömmlichen statistischen Verfahren sein, weil im Vergleich eine gewisse „Entfremdung“ eintritt, da die Ergebnisse vieler statistischer Analysen meist nur angenommen oder verworfen werden können. In den letzten Jahren wurden dazu in Berlin entsprechende Verfahren entwickelt und in der Software MAOE (Modellieren, Analysieren, Optimieren, Entscheiden [12]) umgesetzt, die u. a. im Rahmen des Projekts BERLIN HAT TALENT [32] systematisch eingesetzt wird. Während in der Literatur immer wieder über bemerkenswerte Anwendungsfälle multikriterieller, unscharfer Optimierungsverfahren auf den unterschiedlichsten Gebieten berichtet wird, z. B. auf dem Gebiet der Planung robuster Flugrouten [19], der Meteorologie [8], der Lösung von Logistikproblemen [14] sowie der integrierten Technikbewertung [15, 16], und sich auch vielfältige Workshops mit diesem Gegenstand beschäftigen (zum Beispiel der Workshop der GOR-Arbeitsgruppe „Entscheidungstheorie und -praxis“ 2017 in Wien), beschränken sich vergleichbare Ergebnisse innerhalb der deutschen Sportlandschaft insbesondere auf Arbeiten im Umfeld von Jürgen Perl [23, 25] und Joachim Mester [33]. Die über viele Jahre systematisch weiterentwickelte Software MAOE und die integrierten Möglichkeiten zu deren interaktiven Nutzung

Abb. 2 Lineare (rot, $A=1$), pessimistische (lila, $A<1$) und optimistische (blau, $A>1$) Bewertung $B=f(Q',A)$



sowie die inhaltliche Zusammenarbeit der Autoren (und Institutionen) dieses Beitrags vor allem zur Abbildung der diskursiven Analyse auf verschiedene Typen von Zugehörigkeitsfunktionen, Hierarchien und Gewichtungen zeigen darüber hinaus die übergreifende Nützlichkeit dieser Herangehensweise für ein theoriegeleitetes und systematisches Talentscreening. Ein weiterer bedeutsamer Anwendungskontext, auch im Sport, ergibt sich, wenn bei (leistungs-)diagnostischen Untersuchungen differenzierte, komplexe Problemstellungen in (oftmals „aufgeblähte“) Datenmatrizen überführt werden, bei der Analyse subjektive Bewertungen zu berücksichtigen sind und integrative Ergebnisse (auch interaktiv) entwickelt werden sollen.

Vorbereitende Verfahrensschritte

Für die Methoden des multikriteriellen Variantenvergleichs sind zwei vorbereitende Schritte angezeigt:

- **Normierung:**
Jedes der verwendeten Attribute/Merkmale wird in seinem „Erfüllungsgrad“ bei jeder Variante mit einer Maßzahl aus einer Kardinalskala, d. h. metrischen Skala, ausgedrückt. Diese natürliche Bewertung, z. B. einer Länge kann in m, mm, km usw. erfolgen. Um den Einfluss einer solchen willkürlich gewählten Maßeinheit auf eine Vergleichsrechnung mit anderen Kriterien zu korrigieren, ist es notwendig, die Bewertung jedes einzelnen Kriteriums zu „entdimensionieren“, d. h. zu normieren. Im Idealfall wird die jeweils denkbar beste Ausprägung (Q_b) und die jeweils denkbare schlechteste Ausprägung (Q_w) mit dazwischen liegendem linearen Anstieg genutzt, um die Variantenbewertung auf das Intervall $[0,1]$ zu normieren. Folglich bezeichnet im Weiteren Q das Kriterium mit einer natürlichen Skala sowie Q' das normierte Kriterium.
- **Bewertungsfunktionen:**
Im Rahmen der Nutzentheorie [21] wurde u. a. der Gedanke entwickelt, dass die Bewertung einer Kriteriums-ausprägung nicht linear sein muss (vgl. rote Linie in

Abb. 2). Häufig ist davon auszugehen, dass eine Verbesserung bei kleinen Ausprägungen stärker bewertet wird. In diesem Fall wird die Bewertung (lila Kurve) konvex von oben gekrümmt sein (sog. „pessimistische Bewertung“). Der umgekehrte Fall (blaue Kurve) wird als „optimistische Bewertung“ bezeichnet.

Die Bewertungen kann auch weiter differenziert werden, z. B. Sprungfunktionen, Stufenfunktionen, Schwellenfunktionen u. a. Eine Bewertungsfunktion kann als Zugehörigkeitsfunktion zur Menge der besten Lösungen in Bezug auf das Merkmal Q gewertet und behandelt werden.

Methoden des multikriteriellen Variantenvergleichs

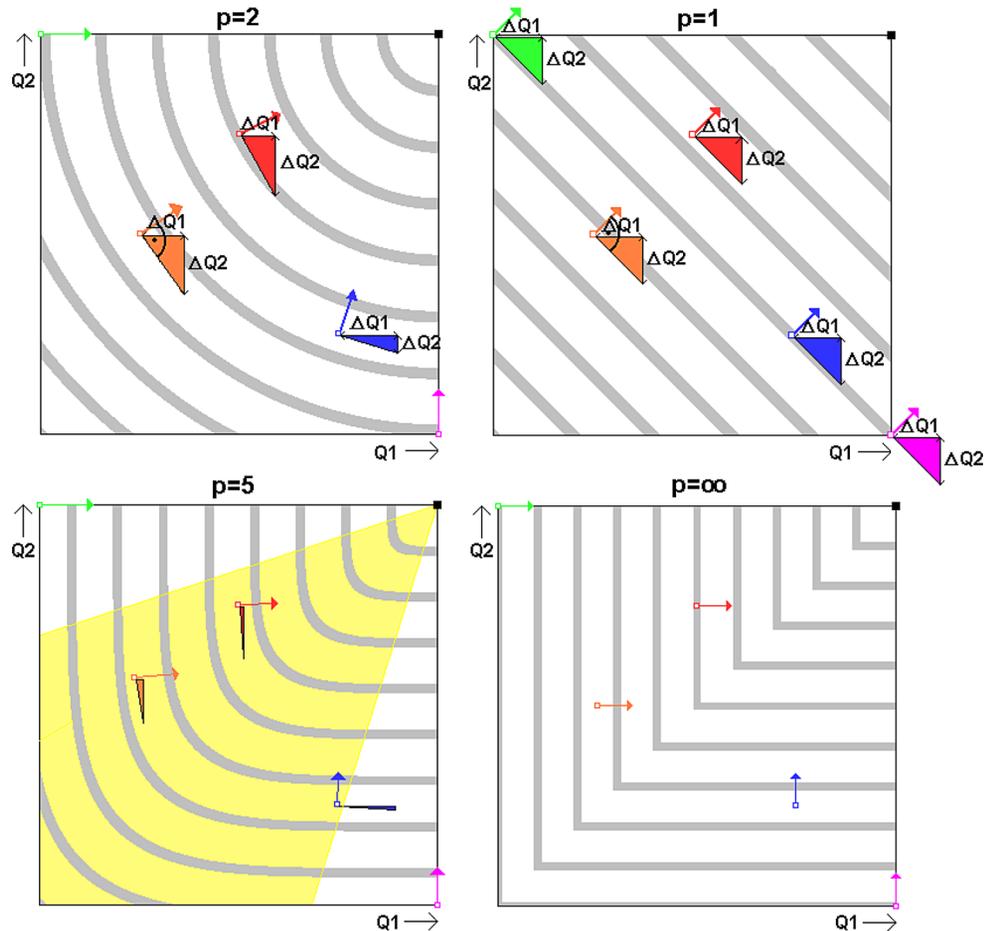
Rückwärtsfilterung

Die Grundidee der Rückwärtsfilterung [10, 34] lässt sich wie folgt formulieren:

Derjenige ist in der Rangfolge vorn, der am „nächsten“ am Ideal liegt, wobei die Kompensation zwischen den Merkmalen mit dem verwendeten Abstandsmaß gesteuert wird.

Die Idee klingt einfach. Alle Varianten/Probanden werden nach ihrem Abstand zu einem Idealpunkt im Zielraum geordnet. Wer am nächsten zum Ideal liegt, ist der Beste, und wer am weitesten weg liegt, ist der Schlechteste. Kann man also einen „Wunsch“-Punkt (Q^{REF}) vorgeben oder wählt man in Ermangelung eines Wunschpunkts den Idealpunkt (Q^{Ideal}), so entsteht die Frage nach der Wahl des Abstandsmaßes zu diesem Referenzpunkt. Das muss nämlich nicht der meist intuitiv gewählte „normale“ euklidische Abstand sein. Mit dieser Wahl legt man dann auch den Kompensationsgrad und die Substitutionsraten fest.

Abb. 3 Wirkung des Parameters p auf die Form der Abstandsfunktion (die Niveaulinien der Abstandsfunktion sind grau dargestellt und der Referenzpunkt befindet sich in der oberen rechten Ecke)



Ein hinreichend allgemeines Abstandsmaß A ist die sogenannte gewichtete Lp-Norm

$$A = |\underline{Q}, \underline{Q}^{\text{REF}}| = \left[\sum_i |g_i (Q_i^{\text{REF}} - Q_i)|^p \right]^{1/p} \quad (1)$$

$\infty > p \geq 1, \quad g_i \geq 0, \quad \sum_i g_i = 1$

mit dem zu bewertenden Merkmalsvektor

- $\underline{Q} = (Q_1, \dots, Q_m)$

und dem Referenzpunkt $\underline{Q}^{\text{REF}} \geq \underline{Q}^{\text{ideal}}$ (im Sinne der Vektorhalbordnung)

- $\underline{Q}^{\text{REF}} = (Q_1^{\text{REF}}, \dots, Q_m^{\text{REF}})$.

In Abb. 3 sind für $m=2$ und $g_1=g_2=0,5$ die Niveaulinien der Lp-Norm für unterschiedliche Werte von p dargestellt. Dabei sind insbesondere die folgenden Sonderfälle zu verzeichnen:

- $p=2$: Euklidische Norm
- $p=1$: City-Block-Norm
- $p=\infty$: Tschebyscheff-Norm

Es sind fünf Varianten zusammen mit der jeweiligen Richtung des steilsten Abfalls eingetragen. Bei der braunen Variante ist gezeigt, wie das sogenannte Substitutionsdreieck entsteht. Im rechten Winkel zur Richtung des steilsten Anstiegs liegt die Tangente der Niveaulinie, deren Anstieg bzw. Abfall das Substitutionsdreieck festlegt. Die jeweilige Substitutionsrate ist das Verhältnis der Δ -Werte. Bei $p=1$ ist dieses Verhältnis ortsunabhängig. Bei $p=2$ und darüber hinaus erkennt man eine Ortsabhängigkeit: Bei großen Werten von Q_1 und kleinen Werten von Q_2 ist das Dreieck so beschaffen, dass eine geringe Verschlechterung von Q_2 nur durch eine große Verbesserung von Q_1 kompensiert werden kann. Das entspricht einem „vernünftigen“ Entscheidungsverhalten. Andererseits heißt das aber auch, dass in den „Randgebieten“ die Kompensierbarkeit immer mehr abnimmt, womit das Substitutionsdreieck immer flacher wird. Als bewährtes Beispiel für dieses Entscheidungsverhalten dient die „Hunger-Durst-Problematik“: Wer Hun-

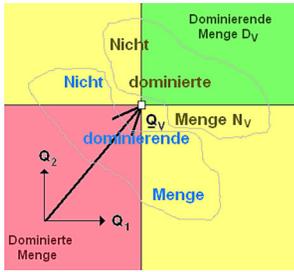


Abb. 4 Dominanzmengen der Vektorhalbordnung für eine Variante \underline{Q}_v

ger leidet, tauscht viel Wasser für wenig Brot, wer an Durst leidet, tauscht viel Brot gegen wenig Wasser.

Mit p wird auf diese Weise also die Kompensierbarkeit zwischen den Kriterien „gesteuert“: Erhöht man p , so verschlechtert sich diese, bei $p=1$ (City-Block-Norm) ist vollständige Kompensierbarkeit gegeben, bei $p=5$ kann man von merklicher Kompensierbarkeit nur noch innerhalb des gelben Kegels sprechen, bei $p=\infty$ (Tschebyscheff-Norm) ist nirgends mehr eine Kompensierbarkeit vorhanden und die Bewertung richtet sich immer nach der jeweils schlechtesten Ausprägung; unabhängig davon, wie gut oder schlecht die anderen Kriterien ausgeprägt sind.

In Bezug auf unser Sportbeispiel ergibt sich daraus:

Bei $p=1$ ist der beste Mehrkämpfer und bei $p=\infty$ der mit der besten schlechtesten Disziplin vorn („Eine Kette ist so stark, wie ihr schwächstes Glied!“).

Mithilfe des totalen bzw. vollständigen Differenzials lässt sich zeigen, dass folgender Zusammenhang für die Substitution von i -Werten durch j -Werte gilt:

$$S_{ij} = \left(\frac{g_j}{g_i}\right)^p \cdot \left(\frac{Q_j^{\text{REF}} - Q_j}{Q_i^{\text{REF}} - Q_i}\right)^{p-1} = \frac{1}{S_{ji}} \quad (2)$$

Für $p=1$ ergibt sich eine konstante Substitutionsrate als Verhältnis der beiden Gewichtungsfaktoren, wodurch in diesem Fall beide Informationen ineinander umgerechnet werden können. Außerdem erkennt man, dass an den Stellen, wo eine Güteausprägung den Referenzwert hat ($Q_i^{\text{REF}} = Q_i$), die dazugehörigen Substitutionsraten 0 oder ∞ sind, d. h. eine Substitution nicht möglich ist.

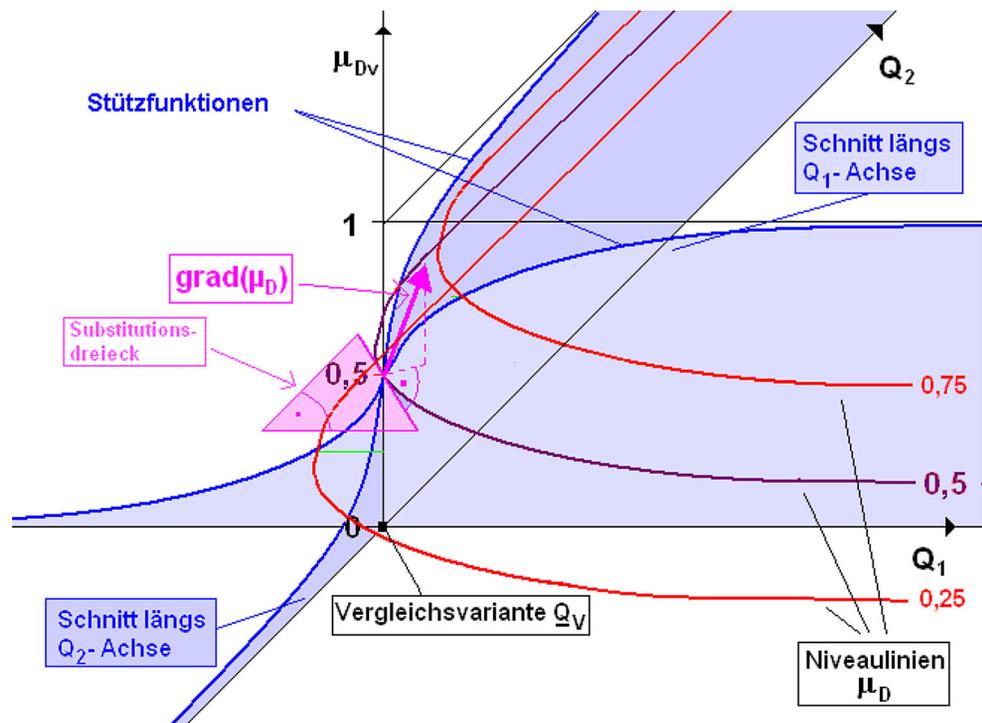
Unschärfe Dominanzmengen

Die Grundidee der unscharfen Dominanzmengen [10] drückt sich wie folgt aus:

„Derjenige ist in der Rangfolge vorn, der die meisten anderen dominiert und von den wenigsten dominiert wird; mit anderen Worten der ‚bescheidene‘ Gewinner!“

Für jede Variante bzw. jeden Probanden gibt es eine Menge im Merkmalsraum, die in Bezug auf eine „Gesamtgüte“ besser als diese ist, d. h. die dominierende Menge, und eine Menge im gleichen Merkmalsraum, die in Bezug auf diese „Gesamtgüte“ von dieser Variante dominiert wird und als

Abb. 5 Zugehörigkeitsfunktion („membership function“) der Menge dominierender Varianten einer Vergleichsvariante \underline{Q}_v (3D-Darstellung der Zugehörigkeitsfunktion über zwei Merkmale)



dominierte Menge bezeichnet wird. Da es sich um eine Halbordnung handelt, gibt es aber auch Bereiche, die weder zur dominierenden noch zur dominierten Menge gehören (siehe Abb. 4).

Eine Variante ist dann am besten, wenn sie für alle anderen Varianten zur dominierenden Menge D gehört und zu den dazugehörigen nicht dominierten Mengen N . Man muss für jede Variante diese Mengen ermitteln. In Anbetracht der Tatsache, dass alle Vorstellungen zu einem Gütemodell vage und unscharf sind, liegt der Gedanke nahe, diese Mengen gleich als unscharfe Mengen, d.h. FUZZY-Sets, mit dazugehörigen Zugehörigkeitsfunktionen („membership functions“) zu modellieren.

Als Basisinformationen zur Konstruktion einer solchen Zugehörigkeitsfunktion werden die Substitutionsraten S_i und die geschätzten absoluten Messfehler F_i für jede Variante i genutzt. Die geforderten Eigenschaften für eine Zugehörigkeitsfunktion $\mu_{Dv}(Q_1, Q_2)$, welche die Menge der dominierenden Varianten (rechte obere Ecke in Abb. 4) einer beliebigen Vergleichsvariante im Kriterienraum beschreibt, werden in Abb. 5 dargestellt.

Eigenschaften:

1. Der Wert der Zugehörigkeitsfunktion muss mit zunehmender Ausprägung gleichfalls steigen (asymptotisch von Null auf asymptotisch Eins, siehe die Schnitte längs der Kriterien-Achsen).
2. Am Punkt der Lage der Vergleichsvariante Q_v muss die Zugehörigkeit 0,5 sein (eine andere Variante mit exakt den gleichen Ausprägungen kann chancengleich besser oder schlechter sein).
3. Der im gleichen Punkt vorhandene Gradient der Zugehörigkeitsfunktion wird in seiner Richtung durch das vorgegebene Substitutionsdreieck bestimmt und in seinem Betrag (Steilheit) von den geschätzten „Mess“-Fehlern.

Die angegebenen Stützfunktionen werden wie folgt berechnet [8]:

$$SF_i = \frac{1}{\pi} \arctan \left(C_i (Q_i - Q_{vi}) + \frac{\pi}{2} \right), \quad i = 1(1)m$$

$$\mu_{Dv}(Q) = \left(\prod_{i=1}^m SF_i \right)^{1/m} \tag{3}$$

Die m freien Parameter C_i werden so berechnet, dass die drei oben genannten Eigenschaften erfüllt sind.

Unschärfe Güte

Bei der Grundidee der unscharfen Güte [10] ist zu beachten, dass die Vorgehensweise der Methode der unscharfen Dominanzmengen ähnelt, jedoch die verwendeten Substitutionsraten von der „Lage“ des Probanden im Koordinatensystem abhängig sind (siehe „Hunger-Durst-Problematik“ bei der Methode der Rückwärtsfilterung).

Ziel ist die Konstruktion einer Zugehörigkeitsfunktion zum denkbar besten Element (Idealpunkt) und die anschließende Ordnung der Varianten nach dem Wert dieser Zugehörigkeit. Als Basisinformationen werden hier die Substitutionsraten für alle Varianten benötigt. Bei diesem Verfahren wird auch die Gl. 1 für die L_p -Norm verwendet, wobei die Gewichtungsfaktoren und der Exponent nicht vorgegeben, sondern aus den angegebenen Substitutionsraten für jede Variante berechnet werden. Dabei ist es möglich, dass auch Werte kleiner als Eins, evtl. Null oder kleiner als Null, für den Exponenten entstehen. Über einen nicht linearen Regressionsansatz werden die freien Parameter in der Formel für die L_p -Norm gesucht, die die Summe der quadrierten Winkelabweichungen der Substitutionsdreiecke jeder Variante über alle Varianten minimieren.

Betrachtet man die Gl. 1, benutzt als Referenzpunkt den Nadirpunkt und arbeitet mit normierten Ausprägungen, so ergibt sich als Bewertungsfunktion μ :

$$\mu = \left[\sum_i (g_i |Q_i|)^p \right]^{1/p}, \quad g_i \geq 0.$$

$$\sum_i g_i = 1, \quad \infty \geq p \geq -\infty \tag{4}$$

Setzt man hier formal $p=0$, so ergibt sich ein in dieser Form nicht berechenbarer Ausdruck. Man kann aber mithilfe einer Grenzwertbetrachtung [11] zeigen, dass

$$\lim_{p \rightarrow 0} \left[\sum_i (g_i |Q_i|)^p \right]^{1/p} = \prod_i |Q_i|^{g_i} \tag{5}$$

Für $p=1$ erhält man die Formel für einen gewichteten arithmetischen Mittelwert (Summenbewertung) und für $p=0$ ergibt die Gl. 5 einen gewichteten geometrischen Mittelwert (faktorielle Bewertung). Abb. 6 zeigt, dass mit einem Referenzpunkt (Nadir-Punkt) p jeden reellen Wert annehmen kann.

Zwei weitere Grenzwertbetrachtungen führen zu den beiden Extremfällen, Maximum- und Minimumnorm.

Für $p=\infty$ erhält man die Maximumnorm (Bewertung entspricht der besten Ausprägung)

$$\mu = \max_i (g_i Q_i), \quad g_i \geq 0, \quad \sum_i g_i = 1 \tag{6}$$

und für $p=-\infty$ erhält man die Minimumnorm (Bewertung entspricht der schlechtesten Ausprägung).

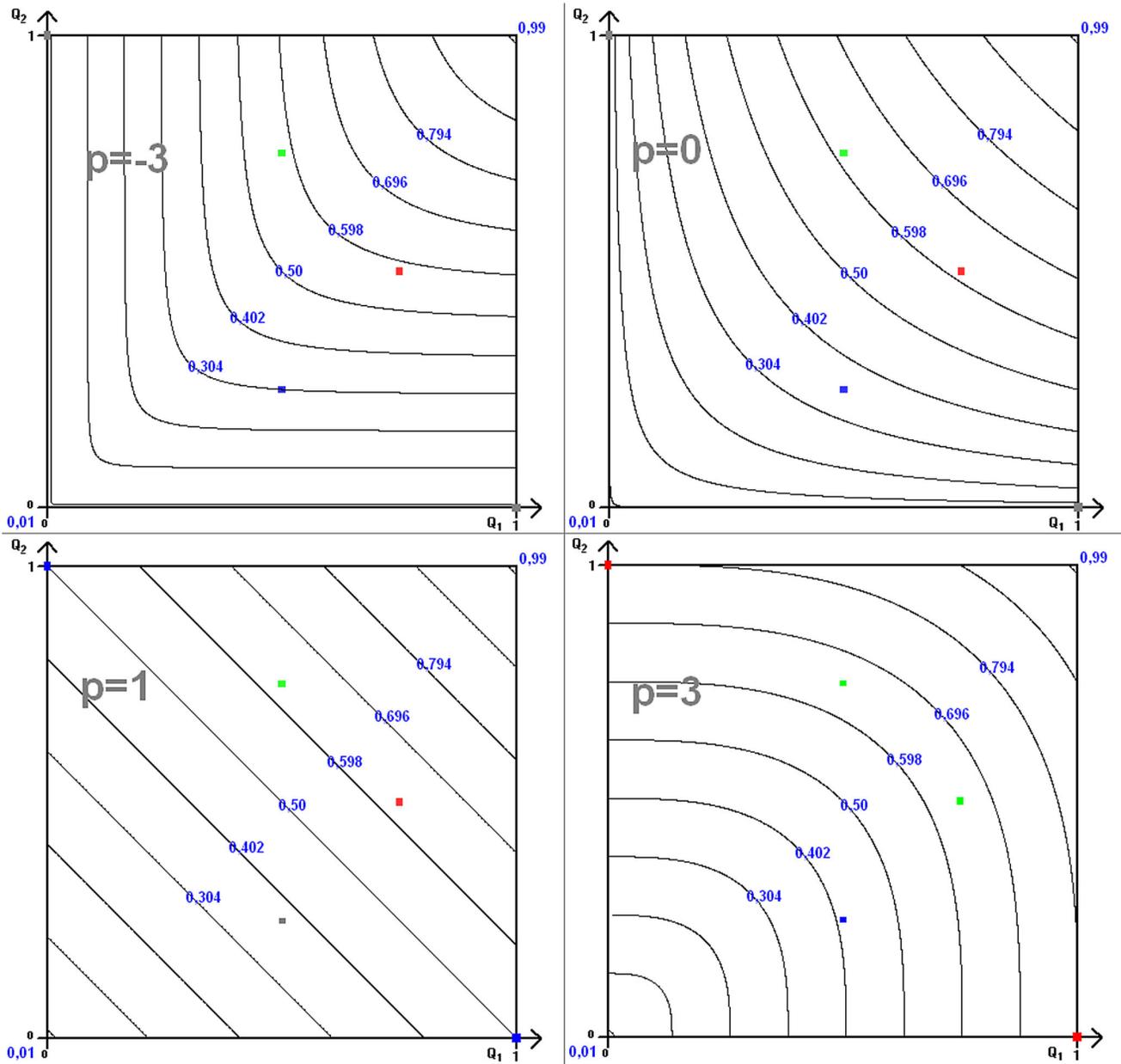


Abb. 6 Wirkung des Exponenten p auf die Niveaulinien der Zugehörigkeit (siehe zum Vergleich Abb. 3)

Gl. 4 beschreibt für $p < 1$ streng genommen auch keine Norm mehr.

$$\mu = \min_i (g_i Q_i), \quad g_i \geq 0, \quad \sum_i g_i = 1 \quad (7)$$

In diesem Verfahren werden die Substitutionsraten für jede Variante (individuell) verwendet. Sie hängen also von der Lage der Varianten im Koordinatensystem ab. Mit Bezug auf die „Hunger-Durst-Problematik“ ergibt sich die Analogie: Zwei Wanderer in der Wüste haben unterschiedliche Rationen von Essen und Trinken. Im Extremfall gibt ein

Verdurstender für einen Schluck Wasser seine gesamte Essensration und ein Verhungender für einen Bissen Brot seinen gesamten Wasservorrat. Könnte man die Substitutionsraten für jede Variante exakt vorgeben, wäre dieses Verfahren sehr vorteilhaft und faktisch fehlerlos. Da das aber in der Regel nicht möglich ist, kann man das beispielsweise durch die Veränderung von p simulieren, um zu prüfen, ob die entstehenden Substitutionsraten den praktischen Gegebenheiten entsprechen.

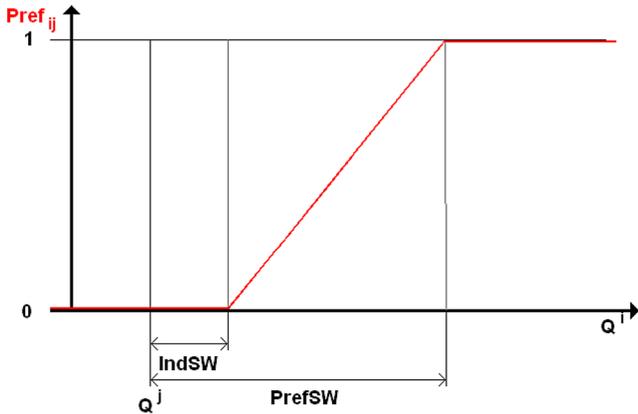


Abb. 7 Indifferenzschwelle (IndSW) und Präferenzschwelle (PrefSW) und dazugehöriger Verlauf eines Präferenzwertes

PROMETHEE-Verfahren

Die Grundidee des PROMETHEE-Verfahrens [3, 6, 7, 35] lässt sich wie folgt beschreiben:

Derjenige ist in der Rangfolge vorn, der die meisten dominiert, egal wie oft er selbst dominiert wird. Es ist derjenige im Turnier, der nicht nur die meisten Spiele gewinnt, sondern dabei noch die höchste Trefferquote erreicht, d. h. der „narzisstische“ Gewinner.

Hier handelt es sich um ein Outranking-Verfahren, wie sie insbesondere unter den Namen ELECTRE [Elimination Et Choix Traduisant la Réalité] eingeführt und in vielen Varianten weiterentwickelt wurden [26–28]. Im Folgenden werden nur die Varianten I (Kerngrad) und II („net-flow“) dargestellt, für die Informationen über Schwellenwerte und Gewichtungsfaktoren benötigt werden (Abb. 7).

Die Indifferenzschwelle hat den Charakter einer Toleranz, innerhalb derer eine Präferenz mit Sicherheit nicht ausgesprochen werden kann und die Präferenzschwelle gibt die Wertedifferenz an, ab der mit Sicherheit die Vergleichsvariante in Bezug auf das einzelne Kriterium Q^j dominiert wird. Bei konstantem Kriteriumswert Q^j einer Variante gibt

die rote Kurve den Präferenzwert $Pref_{ij}$ dafür an, dass eine Variante mit der Ausprägung Q^i bei diesem Kriterium gegenüber j zu präferieren ist. Das entspricht der Überführung der binären „>“ Relation in eine graduelle Relation, die auch als Zugehörigkeitsfunktion betrachtet werden kann. Der Verlauf zwischen den Schwellen muss nicht notwendigerweise linear, jedoch streng monoton sein. Die Schwellenwerte können auch von den Ausprägungen der Vergleichsvariante abhängig oder jeder Variante direkt zugeordnet sein. Jede Variante erhält bei jedem Kriterium von jeder anderen Variante einen Präferenzwert: Ist dieser Wert positiv, so wird das „Pluskonto“ mit diesem Wert, der mit dem Gewichtungsfaktor multipliziert wird, erhöht, und die andere Variante erhöht um den gleichen Wert ihr „Minuskonto“. Am Ende dieser Berechnung hat also jede Variante ein Siegeskonto und ein Niederlagenkonto. Wenn beim Vergleich zweier Varianten eine dieser Varianten ein höheres Siegeskonto und ein niedrigeres Niederlagenkonto hat, so wird die andere als „ausgemustert“ bezeichnet. Die Berechnung der Rangfolge erfolgt für PROMETHEE I (Kerngrad) dahingehend, dass der Beste die meisten ausmustert und der Schlechteste die wenigsten ausmustert, wohingegen für PROMETHEE II („net-flow“) bei jeder Variante das Niederlagenkonto vom Siegeskonto abgezogen und anschließend nach der Größe dieser Zahl geordnet wird.

Zielstrukturen und hierarchische Methoden

Bei der Grundidee der Zielstrukturen und hierarchischen Methoden [9, 18, 29–32] wird davon ausgegangen:

Derjenige ist vorn, der jeweils in den Strukturknoten am „nächsten“ an der idealen Ausprägung ist und damit den „strukturiertesten“ Gewinner darstellt.

Die bekannteste Methode, die Zielstrukturen verwendet, ist der sogenannte „analytische Hierarchieprozess“ [29], der seine Weiterentwicklung im „analytischen Netzwerkprozess“ [32] gefunden hat. Eine Bewertungshierarchie der

Abb. 8 Beispiel für eine Bewertungshierarchie zur körperlichen Leistungsfähigkeit

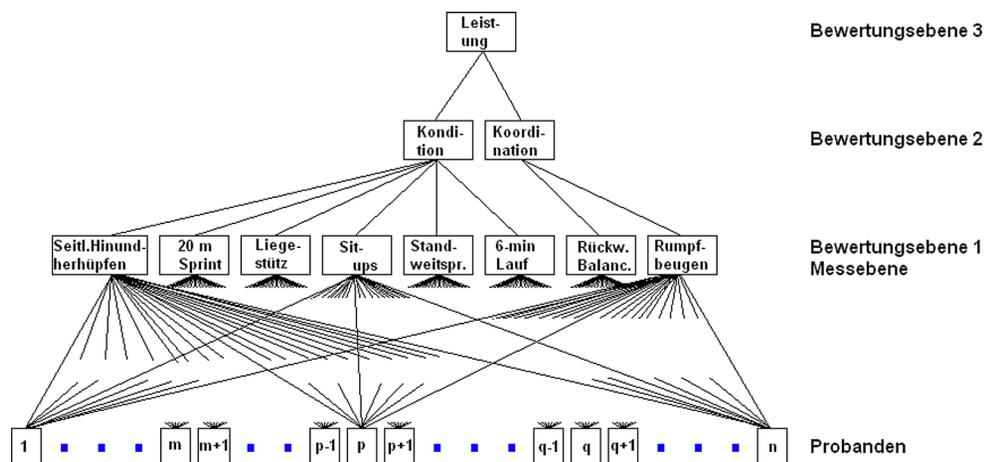
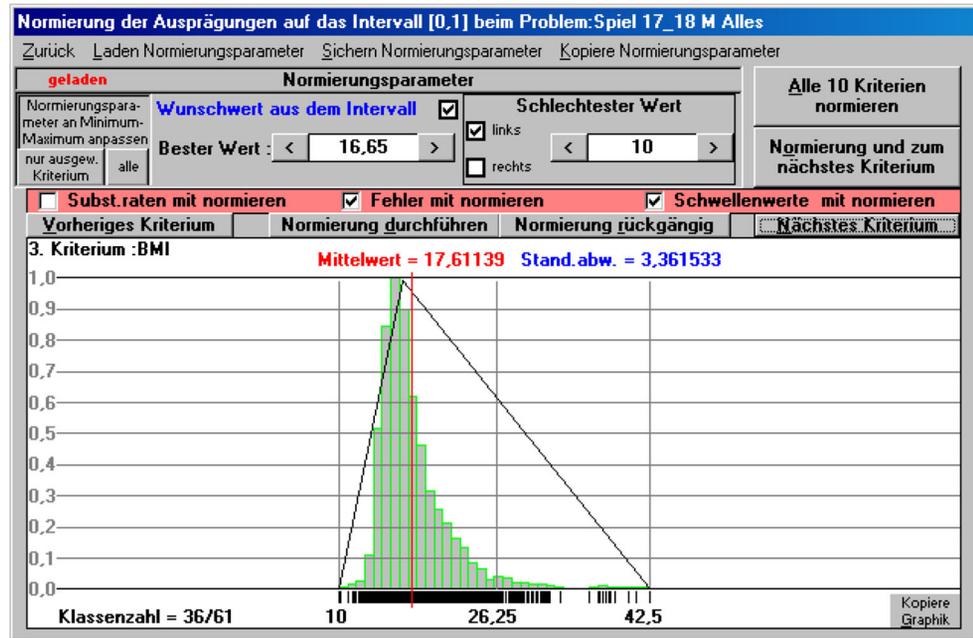


Abb. 10 Normierung am Beispiel des Parameters BMI (Wunschwert: 16,65)



Eine weitere Möglichkeit, diese Parameter zu bestimmen, besteht darin, durch die Vorgabe einer Anzahl n von „Muster“-Varianten und deren Bewertung nach dem Gesamtkriterium (Notenvergabe) diese freien Parameter mittels Fehlerminimierung zu bestimmen (11), um sie dann auf das gesamte Entscheidungsproblem anzuwenden (Auffassung der Bewertungshierarchie als künstliches neuronales Netz).

$$\min \left\{ F = \sum_{j=1}^n \left[B_j - Q^j (\underline{g}, \underline{p}) \right]^2 \right\} \quad (10)$$

Dabei stellt B_j die vergebene Benotung und Q^j die berechnete Bewertung für die j -te Mustervariante dar. Der Vektor \underline{g} bezeichnet alle Gewichtungen in allen Knoten und \underline{p} bezeichnet die Exponenten in allen Knoten. Algorithmisch berechnet man Q^j in der Hierarchie aufsteigend ab Bewertungsebene 2 so, dass immer alle untergeordneten Bewertungen vorliegen. Da bei jedem Knoten die Normierung der Gewichtungsfaktoren berücksichtigt werden muss und auch für g_i keine negativen Werte zugelassen sind, stellt (10) eine Aufgabe der nicht linearen Optimierung mit Gleichungs- und Ungleichungsbeschränkungen dar. Wegen der schwierigen Bestimmung der Ableitungen von F sollte zur Lösung von (10) ein gradientenfreies Suchverfahren mit numerisch gewonnenen Anstiegsinformationen verwendet werden.

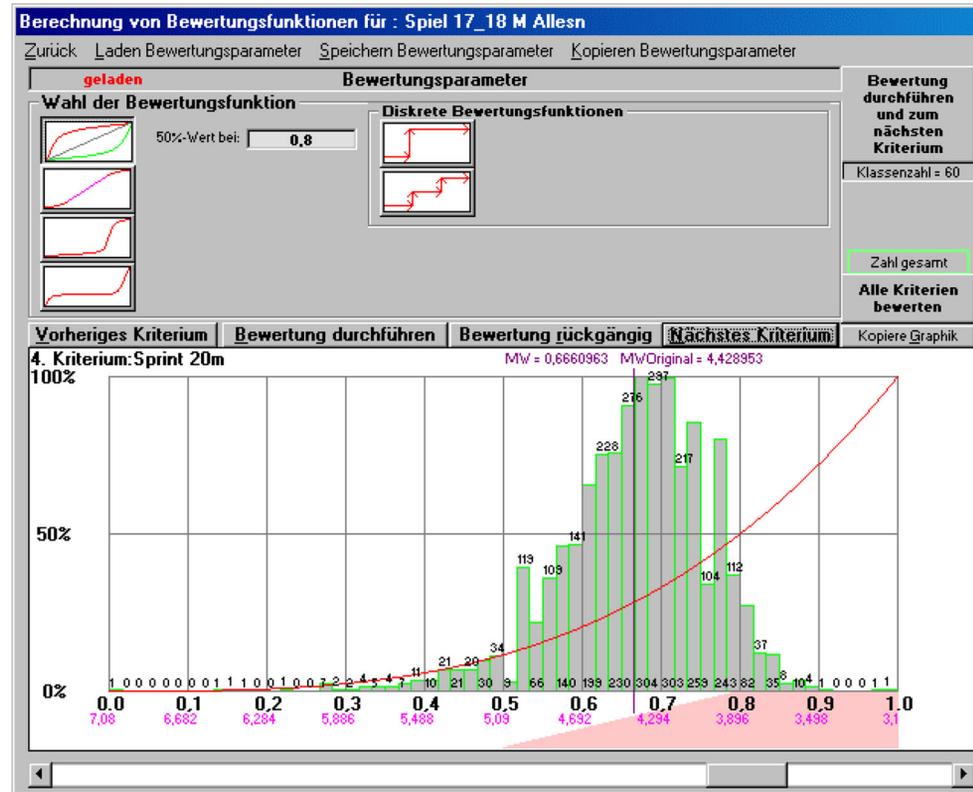
Kombination von Auswahlverfahren und Auswertung unterschiedlicher Rangfolgen

Die Anzahl der Methoden zur Erzeugung einer Rangfolge bei multikriteriellen Aufgaben der Variantenauswahl ist hoch. Allein in den vorliegenden Ausführungen werden bereits mehr als 5 Möglichkeiten dargestellt, die z.B. mittels des Exponenten p auch noch parametrisiert werden können. Die dargestellten multikriteriellen Verfahren beruhen im Wesentlichen auf analytischer Geometrie und sind in diesem Sinne „scharfe“ Verfahren. Die „Unschärfe“ ergibt sich insbesondere durch die Einbeziehung subjektiver Theorien der involvierten Experten. Des Weiteren verarbeiten die Verfahren unterschiedliche Informationen, d.h. Gewichtungen, Substitutionsraten, Fehlerangaben, Schwellenwerte usw. Doch allen ist gemein, dass vermutlich keines der Verfahren den tatsächlichen Beitrag eines Einzelmerkmals zur Erreichung einer Gesamtzielstellung in realiter korrekt nachbilden kann¹.

Insofern sind alle Methoden zuvorderst gleichberechtigte Näherungsmodelle. Der nächste Schritt besteht deshalb darin, die von den verschiedenen Verfahren mit unterschiedlichen Ausgangsinformationen berechneten Rangfolgen zu vergleichen und daraus verallgemeinernde Schlussfolgerungen zu ziehen. In diesem Schritt können Beurteilungen über die Eignung der unterschiedlichen Verfahren bzw. die un-

¹ „Wenn die Komplexität eines Systems zunimmt, wird unsere Fähigkeit, präzise und dennoch signifikante Aussagen über dieses System zu machen, entsprechend geringer, bis eine Grenze erreicht wird, von der an Präzision und Signifikanz (oder Relevanz) zu sich ausschließenden Merkmalen werden“ [17, S. 28].

Abb. 11 Bewertung am Beispiel des Parameters 20-m-Sprint. Es „interessiert“ praktisch erst eine Zeit unter 5 s, ab dann werden alle Varianten entsprechend einer exponentiellen Zugehörigkeitsfunktion bewertet. Ab einer Zeit von vier Sekunden wird die Leistung als ausreichend eingeschätzt, die Bewertung liegt bei 50%. Die erreichten Sprintzeiten sind auf der x-Achse sowohl als normierte Werte sowie – auszugsweise – auch als absolute Werte (von 7,08 s bis zu 3,1 s) aufgetragen



terschiedlichen Grundideen abgeleitet und in diesem Zusammenhang Gewichtungen für die Rangfolgen von den „Experten“ vergeben werden. In einem weiteren Verfahrensschritt lassen sich dann die parallel ermittelten Rangfolgen zu einem Gesamtergebnis kombinieren, beispielsweise dahingehend, dass die Einzelergebnisse bzw. Rangfolgen entsprechend gewichtet und gemittelt werden. Somit kann man die Rangfolgen der Verfahren, die auf der Basis von Ersatzzielfunktionen beruhen, stärker gewichten, als jene Rangfolgen, die auf Ordnungsprinzipien beruhen und deshalb instabiler sind.

Für jede Variante bzw. jeden Probanden kann final ein (gewichteter) mittlerer Platz berechnet und eine Gesamtrangfolge der mittleren Platzierungen festgehalten werden. Dieses Vorgehen kann auch als kompensatorisches Modell verstanden werden, da Vor- und Nachteile der verschiedenen Verfahren ausgeglichen werden. Sollten sich jedoch Informationen hinsichtlich der besseren Eignung eines bestimmten Verfahrens ergeben, so kann man diese durch eine Ungleich- bzw. Höhergewichtung von Rangfolgen bei der Berechnung der mittleren Platzierung der Varianten/Probanden berücksichtigen.

Die Anwendung des multikriteriellen und FUZZY-Vorgehens auf das Talentscreening im Projekt BERLIN HAT TALENT

Im Schuljahr 2017/18 wurden 3718 Jungen aus den dritten Schulklassen im Alter von 8–10 Jahren in Berlin mit dem Deutschen Motorik-Test 6–18 (DMT [5]) untersucht. Mithilfe eines Screenings sollten die Kinder gesichtet werden, die ein gewisses Talent für die Sportsportarten haben könnten.

Im ersten Schritt der diskursiven Validierung wurden mit verschiedenen Sportwissenschaftlern, Trainern und Lehrern zunächst die aus deren Sicht für die Sportsportarten wichtigsten Parameter aus dem DMT ausgewählt, die um die sportsportrelevanten Parameter Körperhöhe (KH), geschätzte finale Körperhöhe (FKH), biologisches Alter (BA), kalendarisches Alter (AM) und den Body-Maß-Index (BMI) ergänzt wurden. Für jeden dieser Parameter wurde nun eine Fehlerbreite, die auch biologisch verursacht sein könnte, ermittelt, z. B. für die Körperhöhe das Maß von 1 cm, und eine prinzipielle Optimierungsrichtung abgeleitet, z. B. für die Körperhöhe (KH) das Maximum, d. h. je größer, desto besser. Auf dieser Basis erfolgte die Normierung der Parameter. Alle Parameter wurden linear normiert, jedoch wurde für den BMI die Normierung in Anlehnung an Kromeyer-Hauschildt et al. [22] so gewählt, dass der Mittelwert für deutsche Kinder dieses Alters und Geschlechts (16,65) als

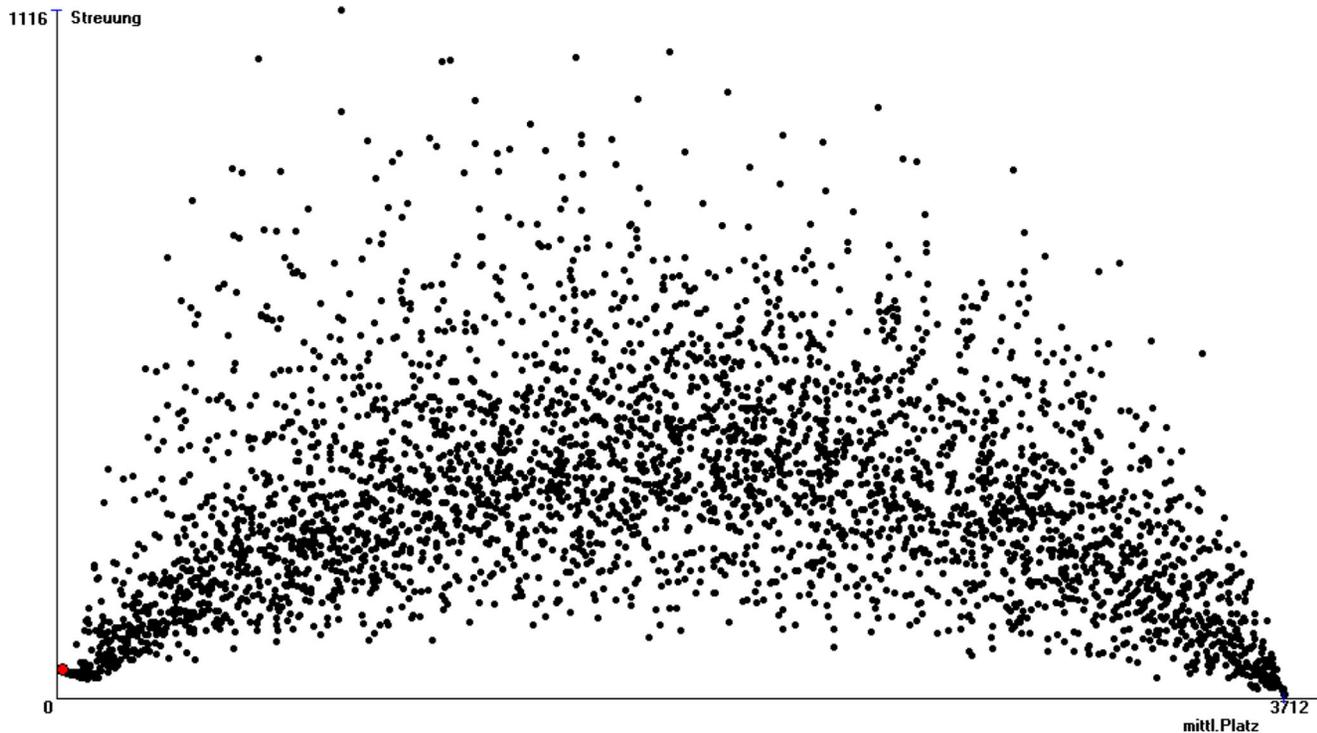


Abb. 12 Differenzierung und Diskriminierung im Bereich der „am meisten talentierten“ (*links*) und der „am wenigsten talentierten“ Kinder (*rechts*) in den Spielsportarten

Wunschwert gilt (Bewertung 1) und Abweichungen nach oben und unten gleichermaßen schlechter gewertet wurden.

Im nächsten Schritt wurde dann ein Strukturdiagramm entwickelt und das Gewicht jedes Parameters in Bezug auf den jeweils zugehörigen Knoten geschätzt. Beispielsweise konnte man sich darauf einigen, im „Knoten“ der Kondition die Parameter 20-m-Sprint (20m), Standweitsprung (SW), Liegestütz (LS) und den 6-min-Lauf (6min) anzulegen und dabei dem Parameter 20m mit 100% die höchste Priorität und LS mit 85% die niedrigste Priorität zuzuordnen (siehe Abb. 9). Abschließend wurden noch die Präferenzen für die Parameter vergeben (Abb. 10) und Bewertungsfunktionen vereinbart (Abb. 11). Beispielsweise wurden Kinder nur „in die Wertung“ genommen, wenn unter der Annahme einer exponentiellen Zugehörigkeitsfunktion im Parameter 20-m-Sprint eine Leistung von schneller als 5s erreicht wurde. Darüber hinaus wurde erst ein Ergebnis von besser als 4s als ausreichend bewertet. Auf diese Weise gelingt es, auch linguistische Informationen, die eben durch Ungenauigkeiten gekennzeichnet sind, in berechenbare Problemstellungen umzuwandeln [8].

Mithilfe des Softwaresystems MAOE, das speziell zur Lösung von multikriteriellen Aufgaben entwickelt wurde, konnten daraus die notwendigen Informationen wie Gewichtungen, Kompensationsgrade, Schwellenwerte usw. ermittelt und mit den vorgestellten Verfahren verifiziert und realisiert werden. So wurden acht Rangfolgen der Proban-

den hinsichtlich ihres Talents für die Spielsportarten bestimmt, die zu einer Gesamtrangfolge der mehr als 3500 untersuchten Schüler aggregiert wurden. Diese Vorgehensweise führte zu einer guten Differenzierung und Diskriminierung insbesondere im Bereich der „am meisten talentierten“ bzw. der „am wenigsten talentierten“ Kinder in den Spielsportarten (Abb. 12) – auch in Abgrenzung z.B. zur allgemeinen Fitness [37]. Ob die „am meisten talentierten Kinder“ nun aber ihr Potenzial langfristig tatsächlich ausschöpfen können und im Sport erfolgreich sein werden, hängt von vielen weiteren Einflussgrößen ab. Der (prognostische) Wert der ermittelten Rangfolge wird sich letztlich erst in der Zukunft zeigen. Die Bedeutung für zu treffende Förderentscheidungen ist dennoch im Spiegel der oftmals divergierenden Meinungen von Trainer*innen und Lehrer*innen bei der Zusammenfindung von Talentsichtungsgruppen unstrittig, da weder motorisch leistungsstarke noch motorisch leistungsschwache Kinder übersehen, sondern gleichermaßen mit unterschiedlichen Programmen gefördert werden sollten (zur Plausibilität der Rangfolge siehe [41]).

Fazit

Zusammenfassend ist festzuhalten, dass die in einem integrativen Team von Sportwissenschaftler*innen, Trainer*innen und Lehrer*innen erarbeiteten Annahmen re-

spektive Hypothesen hinsichtlich der Bewertung von Merkmalsausprägungen im Programmsystem MAOE immer interaktiv angepasst und geprüft werden können. Expert*innen – und nicht nur im und aus dem Sport – haben damit die Möglichkeit, auf der Basis quantifizierter Ergebnisse subjektive Urteile zu modifizieren und damit neue, weiterführende Ergebnisse zu begründen. Auf diese Weise ist das hier beschriebene komplexe multikriterielle FUZZY-Vorgehen eine verlässliche und vielversprechende Methode zur Ergebnisgenerierung unter Unsicherheit: Die Untersuchenden treten in den Dialog mit der Lösung, sie können zielgerichtet in der „Pareto-Menge wandern“ bzw. Rangfolgen beurteilen. Insgesamt verdeutlichen Plausibilitätsbetrachtungen die Unterschiede in der subjektiven Einschätzung verschiedener Expert*innen und es lässt sich empirisch begründet über deren Vor- und Nachteile „streiten“, was in der Regel zu einem tieferen Verständnis der Problemsituation führt.

Die innovative Idee dieser Methodik ist das Zusammenführen der Resultate verschiedenartiger, an sich bekannter Entscheidungsmodelle, bei einer sehr benutzerfreundlichen Einbindung unterschiedlicher Informationen von Experten. Die Vorgehensweise ist allgemein auf alle MADM-Probleme anwendbar. Die entwickelte Software MAOE [12] gestattet ein effektives und schnelles, interaktives Arbeiten zur Berechnung von Gesamtrangfolgen und deren differenzierte Auswertung, was bei der Komplexität und dem Umfang der notwendigen Daten einen theoretisch-inhaltlichen und praktisch relevanten Mehrwert hat.

Literatur

- Anteneh S (1994) Zur Lösung komplexer mehrkriterieller Entscheidungsprobleme mittels Decision Support Systemen. Peter Lang, Frankfurt
- Bamberg G, Coenenberg AG (2008) Betriebswirtschaftliche Entscheidungslehre, 14. Aufl. Vahlen, München
- Behzadian M, Kazemzadeh RB, Albadvi A, Aghdasi M (2010) PROMETHEE: a comprehensive literature review on methodologies and applications. *Eur J Oper Res* 200(1):198–215
- Bös K, Albrecht C, Schlenker L (2016) Der Deutsche Motorik-Test (DM 6–18). In: Zinner J, Poller T (Hrsg) Berlin hat Talent. Hochschule für Gesundheit & Sport, Technik & Kunst, Berlin, S 24–34
- Bös K et al (2016) Deutscher Motorik-Test (DMT 6–18), 2. Aufl. Schriftenreihe der Deutschen Vereinigung für Sportwissenschaft. Czwalina, Hamburg
- Brans JP, Marshal B (2005) PROMETHEE Methods. In: Figueira J, Greco S, Ehrgott M (Hrsg) Multiple criteria decision analysis: state of the art surveys. Springer, New York, S 163–186
- Brans JP, Vincke P, Marshal B (1986) How to select and how to rank projects: the PROMETHEE method. *Eur J Oper Res* 24:228–238
- D'Onofrio S, Portmann E (2015) Von fuzzy-sets zu computing-with-words. *Informatik-Spektrum* 38(6):543–549
- Eisenführ F, Weber M (1986) Zielstrukturierung: Ein kritischer Schritt im Entscheidungsprozess. *Z Betriebswirtsch Forsch* 38: 907–929
- Ester J (1987) Systemanalyse und mehrkriterielle Entscheidung. VEB Verlag Technik, Berlin
- Ester J, Tröltzsch F (1986) On generalized notions of efficiency in multicriteria decision making. *Syst Anal Model Simul* 3(2):153–161
- Ester J, Zinner J (2019) Handbuch MAOE – Modellieren, Analysieren, Optimieren, Entscheiden. Multikriterielle und Fuzzy-Entscheidungsverfahren (nicht nur) im Leistungssport. Deutsche Hochschule für Gesundheit & Sport, Institut für Leistungssport und Trainerbildung (ITLt), Berlin
- Figueira J, Greco S, Ehrgott M (2005) Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, New York
- Geiger MJ (2017) Mehrkriterielle Optimierung in taktischen Logistikproblemen. Workshop der GOR Arbeitsgruppe „Entscheidungstheorie und -praxis, 27.–28.02.2017 Universität Wien, Wien
- Geldermann J (2006) Mehrzielentscheidungen in der industriellen Produktion. Verlag Karlsruhe, Karlsruhe
- Geldermann J, Lerche N (2014) Leitfaden zur Anwendung von Methoden der multikriteriellen Entscheidungsunterstützung. Methode: Promethee. Georg-August-Universität Göttingen, Göttingen
- Holzweg M, Ketelhut K, Brandt S (2012) Der Einfluss des gewählten Bewertungsverfahrens auf die Interpretation der Ergebnisse des Deutschen Motorik Tests (DMT 6–18). *Sportunterricht* 61(8):239–243
- Homburg C (1998) Hierarchical multi-objective decision making. *Eur J Oper Res* 105(1):155–161
- Ide J, Schöbel A (2016) Robustness for uncertain multi-objective optimization: a survey and analysis of different concepts. *OR Spectr* 38(1):235–271
- Ishizaka A, Nemery P (2013) Multicriteria decision aid: methods and software. Wiley, Chichester
- Keeney RL, Raiffa H (1976) Decision analysis with multiple conflicting objectives. Wiley & Sons, New York
- Kromeyer-Hauschild K et al (2001) Perzentile für den Body-Mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschr Kinderheilkd* 149(8):807–818
- Lames M, Perl J (2006) Prozessmodellierung von Handballspielen – stochastische Beschreibung und antagonistische Modellierung. In: Raab M et al (Hrsg) Zukunft der Sportspiele: fördern, fordern, forschen, 2. Aufl. University Press, Flensburg, S 186–190
- Laux H (2003) Entscheidungstheorie, 5. Aufl. Springer, Berlin
- Perl J (2008) Modellbildung und Simulation von Prozessen im Sport. *Inform Spektrum* 31(4):292–300
- Rangel LAD, Gomes LFAM, Moreira RA (2009) Decision theory with multiple criteria: an application of ELECTRE IV and TODIM to SEBRAE/RJ. *Pesqui Operacional* 29:577–590
- Roy B (1980) Selektieren, Sortieren und Ordnen mit Hilfe von Präferenzrelationen. *Z Betriebswirtsch Forsch* 32:465–497
- Roy B, Bouyssou D (1993) Aide multicritère à la décision. Economica, Paris
- Saaty TL (1980) The analytic hierarchy process. McGraw-Hill, New York
- Saaty TL (1990) How to make a decision: the analytic hierarchy process. *Eur J Oper Res* 48(1):9–26
- Saaty TL (1990) Multicriteria decision making—the analytic hierarchy process. Planning, priority setting, resource allocation, 2. Aufl. RWS Publishing, Pittsburgh
- Saaty TL (1996) The analytic network process. RWS Publications, Pittsburgh
- Seifriz F, Mester J (2012) Fuzzy-Logik zur Auswertung und Interpretation leistungsdiagnostischer Kraftparameter. In: Byshko R et al (Hrsg) 19. Symposium der dvs-Sektion Sportinformatik 12.–14.09.2012 Universität Konstanz, Konstanz, S 52–54
- Steuer RE (1986) Multiple criteria optimization. Wiley & Sons, Chichester
- Gwo-Hshiung T, Jih-Jeng H (2011) Multiple attribute decision making. CRC Press, Boca Raton

36. Utesch T, Tietjens M, Büsch D, Ghanbari M-C, Seidel I (2015) Die Überprüfung der Konstruktvalidität des Deutschen Motorik-Tests 6–18 für 9- bis 10-Jährige. *Z Sportpsychol* 22(2):77–90
37. Utesch T, Dreiskämper D, Strauss B, Naul R (2018) The development of the physical fitness construct across childhood. *Scand J Med Sci Sports* 28(1):212–219
38. Utesch T, Zinner J, Büsch D (2018) Stabilität der physischen Fitness im Kindesalter. *Ger J Exerc Sport Res* 48(3):404–414
39. Zahdeh LA (1973) Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans Syst Man Cybern SMC-3*(1):28–44
40. Zinner J, Büsch D, Ester J (2017) Individuelle Leistungseinschätzung im Deutschen Motorik-Test. Zur Individualisierung von DMT-Auswertungen mit Hilfe von multiattributiven FUZZY-Analysen. *Leistungssport* 47(2):4–11
41. Zinner J, Büsch D, Ester J, Becker M, Utesch T (2019) Mit diskursiver Validierung zu plausiblen Förderentscheidungen im Projekt BERLIN HAT TALENT – Forschungsbericht der DHGS. Fakultät Sport. Deutsche Hochschule für Gesundheit und Sport, Berlin
42. Zinner J, Büsch D, Poller T, Bartko R (2015) Berlin hat Talent – ein Zwischenbericht für die Jahre 2012 und 2013. *Leistungssport* 45(5):12–18