

Die clusteranalytische Auswertung eines Musik-Video-Experiments

Die Clusteranalyse (CA) kennt – wie die Faktorenanalyse – zwei Grund-Versionen, nämlich R- und Q-Technik. Bei der R-Technik der CA sprechen wir kurz von einer Variablen-Cluster-Analyse (VCA), bei der Q-Technik von einer Personen- oder Profil-Cluster-Analyse (PCA). Mit der VCA werden Merkmale bzw. Variablen, mit der PCA Personen bzw. Merkmals-Profile nach Ähnlichkeit gruppiert. Die dabei resultierenden Gruppierungen werden Cluster genannt. Diese Cluster sollen möglichst groß (ökonomischer Aspekt), möglichst homogen (typologischer Aspekt) sein und dabei die Gesamtheit der M Variablen (VCA) bzw. der N Personen (PCA) durch ein Minimum von Clustern möglichst vollständig repräsentieren (Repräsentativitäts-Aspekt). Diese Bedingungen determinieren zugleich auch eine markante, empirisch fundierte Typologie. Die resultierenden Cluster sind inhaltlich zu spezifizieren durch Cluster-Ladungen (VCA) bzw. mittlere, sog. »charakteristische« Merkmals-Profile (PCA).

Diese methodologische Zielsetzung kann in Abhängigkeit vom gewählten clusteranalytischen Algorithmus mehr oder weniger optimal realisiert werden. Die betreffende methodologische Problematik soll hier nicht weiter ausgebreitet werden (s. dazu die einschlägige Literatur). Der potentielle Anwender kann jedoch – insbesondere in Anlehnung an Milligan (1980, 1981) bzw. dessen grundlegende methodologische Studien – davon ausgehen, im Rahmen sozialwissenschaftlicher, typologischer Fragestellungen dieses zu erwartende Optimum vorzugsweise mit dem Algorithmus »Group Average« (Sokal & Mitchener 1958) mit hoher Wahrscheinlichkeit zu erreichen.

Bei diesem »agglomerativen« Verfahren (»allokative« und »divisive« Verfahren sollte man meiden!) werden Untergruppen von einander hochgradig ähnlichen Variablen (VCA) bzw. Personen (PCA) auf der Basis eines Abstandsmaßes nach Maßgabe der o.g. Kriterien gebildet. Dieser Vorgang wird auf verschiedenen Ähnlichkeits- bzw. Homogenitäts-Niveaus ausgeführt. Die auf den verschiedenen Ähnlichkeits-Niveaus resul-

tierenden Cluster bilden jeweils eine Partition der clusteranalytisch zu unterteilenden Gesamtheit. Dem Analytiker fällt dabei die nicht immer leichte Aufgabe zu, aus dem Angebot der verschiedenen Partitionen, nach Möglichkeit unter Zuhilfenahme kritischer Zusatz-Informationen, die »optimale« Partition bzw. typologisch interessante und bedeutsame Cluster auszuwählen.

Dieses analytische Vorgehen soll jedoch nun nicht weiter in abstrakter Form behandelt werden, sondern am konkreten Beispiel des Video-Experiments die einzelnen Analyseschritte dargestellt und erläutert werden. Die 93 Versuchspersonen dieses Experiments (Musik-Studenten, Musik-Lehrer, Laien) sollten jeweils zwei Videopaare im Vergleich beurteilen. Im ersten Fall spielten *Birgit* und *Constantin* den *As-Dur-Abschiedswalzer* von *F. Chopin*, im zweiten Fall sah man *Claudia* und *Christian* bei der Interpretation des *Capriccio d-moll* (op. 116, Nr. 7) von *J. Brahms*. Tatsächlich handelte es sich in drei Fällen – nur *Christian* spielte original – um gedoppelte Aufnahmen, die Tonspuren der beiden zu beurteilenden Paare waren also jeweils identisch! Wenn jemand das Spiel von *Birgit* – im Vergleich – für *ausdrucksvoller* oder jenes von *Christian* für *präziser* hielt, so muß man dies als eine Auswirkung der visuellen Ebene betrachten. Neben den jeweils acht (komparativen) Skalen für die Chopin- bzw. Brahms-Interpretation wurde u.a. auch gefragt, ob man gewohnheitsmäßig (AVTYP) oder in der konkreten Versuchssituation (AVEINST) eher Auge- oder eher Ohr-orientiert gewesen sei.

Für die Analyse werden nun zunächst die Meßwerte der 16 Beurteilungs-Skalen (CHOPIN-sicherer bis BRAHMS-überzeugender) zu Standard-z-Werten normiert – wie im Rahmen multivariater Verfahren allgemein verbindlich. Im Gegensatz zu konventionellen korrelations-statistischen Methoden (z.B. Faktorenanalyse, Multiple Regression u.a.m.) werden nun nicht mittlere Produkte je zweier normierter Meßwertreihen (d.h. Variablen [VCA] bzw. Merkmals-Profile [PCA]) berechnet, sondern mittlere quadratische Differenzen, deren Quadratwurzel die bekannte *euklidische Distanz* ergibt. Solche Distanzen haben einen statistischen Erwartungswert (E) von 1.41 (d.i. der Mittelwert aller je zu bestimmenden euklidischen Distanzen). Mit diesem Erwartungswert können nun empirische Distanzen (D) nach der einfachen Beziehung

$$\ddot{A} = (E - D)/E$$

zu Ähnlichkeitswerten (\ddot{A}) transformiert werden, die wie Korrelationskoeffizienten in den Grenzen $(-1, +1)$ variieren. Hohe Ähnlichkeitswerte entstehen zwischen Meßwertreihen mit relativ gering voneinander abweichenden Profilverläufen, unabhängig von deren Verlaufsform. Hierdurch ist zugleich auch der Begriff »Homogenität« zu definieren. Korrelationskoeffizienten sind unter dieser Bedingung noch weitgehend unbestimmt und daher für Gruppierungs-Verfahren unter Homogenitäts-Erwartungen untauglich. Im clusteranalytischen Agglomerations-Prozeß werden nun auf verschiedenen Ähnlichkeits-Niveaus (z.B. .80, .75, .70, ...) die jeweiligen Partitionen ermittelt und, versehen mit mehr oder weniger umfangreichen Zusatzinformationen bzw. kritischen Kennwerten, zu Papier gebracht. Es dürfte unmittelbar einsichtig sein, daß mit fallendem Ähnlichkeits-Niveau die Cluster-Größe zunimmt und daher die Anzahl der Cluster pro Partition abnimmt. Auf einem sich ergebenden niedrigsten Ähnlichkeits-Niveau bilden schließlich sämtliche Variablen (VCA) bzw. Personen (PCA) ein einziges Cluster. Diese letzte Partition ist wie diejenige auf dem Ähnlichkeits-Niveau 1.00 trivial. Interessante Lösungen liegen dazwischen – wo genau, wird der Analytiker – oder dessen Kritiker – kraft fundierter und überzeugender Argumente entscheiden müssen. Die CA ist also kein Ergebnis produzierender *deus-ex machina* – wie von naiven Methoden-Anwendern immer wieder erwartet – sondern ein dialog-orientiertes, argumentativ unterstützendes Analyse-Verfahren in einem offenen empirischen Erkundungsprozeß.

Nun zu den Ergebnissen. Das wichtigste Resultat einer VCA ist das Dendrogramm der Variablen. Es gibt den gesamten clusteranalytischen Prozeß in einer einzigen überschaubar gegliederten und hinsichtlich der Beziehungs-Struktur der Variablen nachvollziehbaren Graphik wieder (s. Abb. 1). Die einzelnen (punktierten) Linienführungen dieser Graphik kennzeichnen in Verbindung mit den jeweiligen Ähnlichkeits-Niveaus die clusteranalytische Agglomeration (um nicht zu sagen Aggregation) als hierarchisches System von Variablen und Variablen-Clustern (es sind auch nicht-hierarchische Systeme verifizierbar und ggf. sinnvoll, worauf hier jedoch nicht näher eingegangen werden soll). Wie man deutlich erkennt, bilden die Beurteilungs-Skalen bei BRAHMS ein kompaktes Cluster auf

THEMA: VIDEOEXPERIMENT HANNOVER * CHOPIN - BRAHMS *
 ANSATZ: VARIABLEN-CLUSTER-ANALYSE
 METHODE: GROUP AVERAGE (SOKAL/MITCHENER 1958)

DENDROGRAMM DER CLUSTER-STRUKTUR-EBENEN BZW. PARTITIONEN 1 - 15
 ZAHLENWERTE IM DENDROGRAMM: AEHNLICHKEITS-NIVEAU (OHNE +/-)

17 MAENNER	17.....20	
20 LEHRER	20.....02	*
9 BRAHMS SICHERER	9.....34	*
11 BRAHMS PRAEZISER	11.....30	*
13 BRAHMS VIRTUOSER	13.....32	*
12 BRAHMS DRAMATISCHER	12.....38	*
15 BRAHMS DYNAMISCHER	15.....40	*
10 BRAHMS KLANGVOLLER	10..48	*
14 BRAHMS AUSDRUCKSVOLLER	14	*
16 BRAHMS UEBERZEUGENDER	16.....18	*
3 CHOPIN PRAEZISER	3.....28	*
1 CHOPIN SICHERER	1.....32	*
5 CHOPIN VIRTUOSER	5.....36	*
7 CHOPIN DYNAMISCHER	7...44	*
6 CHOPIN AUSDRUCKSVOLLER	6	*
8 CHOPIN UEBERZEUGENDER	8.....30	*
4 CHOPIN DRAMATISCHER	4.....34	*
2 CHOPIN KLANGVOLLER	2.....40	*
21 LAIEN	21.....06	*
22 AVTYP	22.....18	*
23 AVEINST	23.....00	*
18 FRAUEN	18.....10	*
19 STUDENTEN	19.....00	*

Abb. 1: Dendrogramm der Variablenclusteranalyse

dem Ähnlichkeits-Niveau .30. Die CHOPIN-Skalen clustern gemeinsam auf dem Ähnlichkeits-Niveau .28, jedoch nur unter Einschluß des Status-Merkmals *Laie*. Das bedeutet inhaltlich, daß die Laien beim CHOPIN-Urteil mehrheitlich dem einen Pol (positivere Beurteilung der Interpreta-

tion *Constantins*) zuneigten. Die übrigen Persönlichkeitsmerkmale clustern erst in der Nähe von .00 bzw. nicht mehr signifikant und sind daher als unabhängige Variablen zu erachten. Weiterhin läßt sich feststellen, daß CHOPIN etwas differenzierter beurteilt wird als BRAHMS. Insgesamt läßt dieses Dendrogramm die Schlußfolgerung zu, daß die CHOPIN-Interpretationen einerseits und die BRAHMS-Interpretationen andererseits relativ einheitlich, jedoch unabhängig voneinander und offenbar auch relativ unabhängig von den Persönlichkeitsmerkmalen *Geschlecht*, *Status*, *audiovisuelle Wahrnehmung* und *Einstellung* beurteilt wurden.

Ob und ggf. in welcher (nicht-linearen) Weise die Persönlichkeitsmerkmale die Beurteilung der musikalischen Darbietungen doch beeinflussen, läßt sich also nicht auf dem Wege einer VCA klären, jedoch mit den weitergreifenden Möglichkeiten der PCA, deren Resultate im folgenden wiedergegeben und eingehender dargestellt werden sollen.

Das wichtigste Ergebnis der PCA sind die sog. *charakteristischen Profile* der resultierenden Personen-Cluster. Diese sind nichts anderes als mittlere Merkmals-Profile der den jeweiligen Clustern angehörenden Personen. Die *charakteristischen Profile* der hier ausgewählten Cluster sind in Abb. 2 tabellarisch und in Abb. 3 graphisch repräsentiert. Bei den Variablen 17–21 sind Prozentwerte/100 mitgeteilt, alle übrigen Skalen sind fünfstufig. Die Profile lassen interessante Beziehungen zwischen Persönlichkeitsmerkmalen und Beurteilungsverhalten erkennen, die nun im einzelnen dargelegt und kommentiert werden sollen. Die hier nur knapp referierte Auswertung ist ausführlicher bei Behne (1990) dargestellt.

In den Abb. 3 a–g sind für jeden der sieben Cluster die normierten Daten der clusterbildenden Variablen (1 bis 16) sowie der unabhängigen Variablen 17 bis 23 veranschaulicht (Variablennamen in Abb. 2). Cluster 1 (Abb. 3a, n = 21) ist insofern »unauffällig«, als sich das Profil nur in der »Mitte« bewegt, d.h. hier überwiegen Personen, die sowohl beim CHOPIN- (1–8) wie auch beim BRAHMS-Urteil (9–16) keiner der Interpretationen den Vorzug gaben, also ein sehr zurückhaltendes Votum artikuliert haben. Dieses Votum war geschlechts- und gruppenunabhängig. Cluster 2 (3b, n = 21) tendiert in der Graphik durchweg nach unten, d.h. inhaltlich, daß die beiden *weiblichen* Interpretationen (*Birgit* und *Claudia*) durchweg positiver beurteilt wurden. Die Graphik gibt auch einen Hinweis auf die auffällige

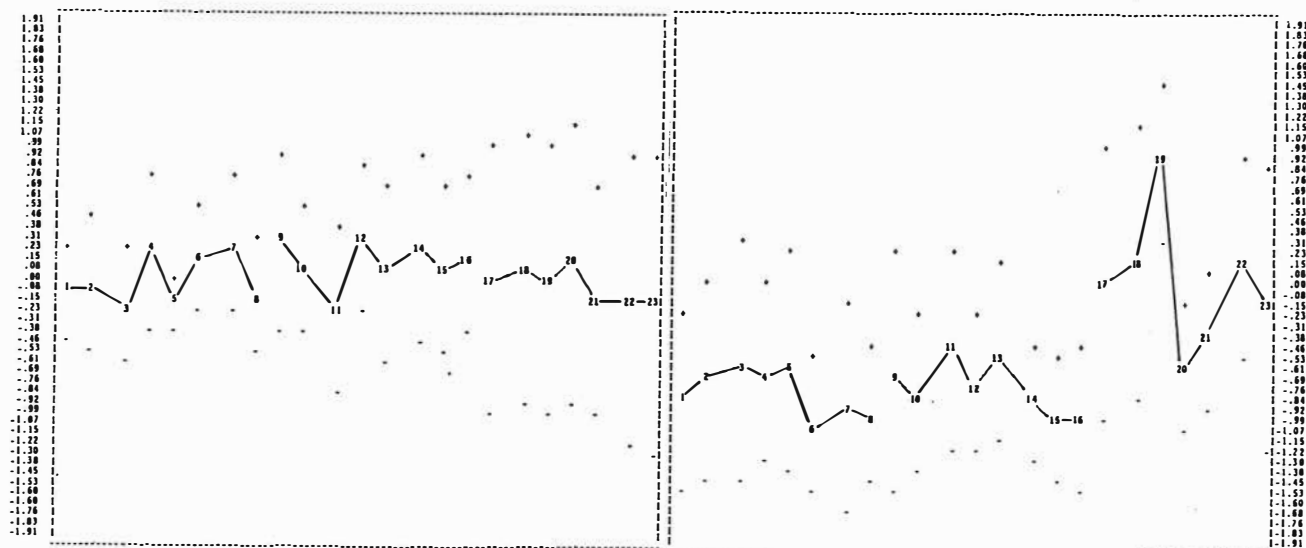
CHARAKTERISTISCHE PROFILE: MITTELWERTE (ROH-DATEN)								-----			
	1	2	3	4	5	6	7			CLUSTER-NR.	CLUSTER-GR.
	21	21	15	10	6	4	4				
1	3.10	2.38	4.00	4.20	2.83	2.75	3.25	*	1	CHOPIN	SICHERER
2	3.05	2.38	3.93	4.70	2.83	2.00	1.75	*	2	CHOPIN	KLANGVOLLER
3	3.14	2.81	3.80	4.10	3.17	2.50	3.50	*	3	CHOPIN	PRAEZISER
4	3.10	2.14	3.87	4.10	2.50	2.00	2.00	*	4	CHOPIN	DRAMATISCHER
5	2.95	2.57	4.07	4.00	3.67	2.75	1.75	*	5	CHOPIN	VIRTUOSER
6	3.14	1.62	4.13	4.50	2.83	1.75	1.75	*	6	CHOPIN	AUSDRUCKSVOLLER
7	3.29	1.95	3.67	4.30	3.00	2.75	2.00	*	7	CHOPIN	DYNAMISCHER
8	2.95	1.90	4.40	4.30	3.33	1.75	2.25	*	8	CHOPIN	UEBERZEUGENDER
9	3.10	2.19	2.47	3.50	4.33	3.25	2.00	*	9	BRAHMS	SICHERER
10	3.00	1.90	2.13	4.60	4.83	3.25	1.75	*	10	BRAHMS	KLANGVOLLER
11	3.10	2.86	2.80	4.10	4.83	3.25	3.75	*	11	BRAHMS	PRAEZISER
12	2.90	1.62	2.00	3.00	4.83	2.75	1.75	*	12	BRAHMS	DRAMATISCHER
13	2.86	2.29	2.40	4.10	4.17	3.00	2.00	*	13	BRAHMS	VIRTUOSER
14	3.00	1.57	2.00	4.90	4.50	2.75	1.75	*	14	BRAHMS	AUSDRUCKSVOLLER
15	2.90	1.57	2.13	4.00	4.83	3.75	3.50	*	15	BRAHMS	DYNAMISCHER
16	2.95	1.43	2.33	4.00	5.00	3.75	2.00	*	16	BRAHMS	UEBERZEUGENDER
17	.43	.43	.53	.10	.67	.50	.25		17	MAENNER	
18	.48	.52	.40	.60	.17	.50	.50		18	FRAUEN	
19	.48	.90	.20	.00	.17	.50	1.00		19	STUDENTEN	
20	.38	.05	.40	.00	.83	.50	.00		20	LEHRER	
21	.14	.05	.40	1.00	.00	.00	.00		21	LAIEN	
22	2.83	3.19	3.08	3.57	3.20	2.50	2.75		22	AVTYP	
23	2.89	2.90	3.10	3.83	3.75	2.50	3.75		23	AVEINST	

Abb. 2 : Mittelwerte für sieben ausgewählte charakteristische Profile

Zusammensetzung dieses Clusters (überwiegend Studenten, wenig Lehrer), die (in diesem Fall vorhandene) Signifikanz ist jedoch nicht unmittelbar den Abbildungen zu entnehmen, sondern muß separat (t-Test oder Chi-Quadrat-Test) ermittelt werden. Die durchgehende Bevorzugung weiblicher Interpreten erfolgte unabhängig vom Geschlecht der Beurteiler! Ein *gemischtes* Votum ergab sich für das dritte Cluster (3c, n = 15). Hier wurde bei CHOPIN der männliche, aber bei BRAHMS durchgehend die weibliche Interpretin positiver bewertet. Die Graphik zeigt auch, daß diese Urteilskonfiguration bei Studenten deutlich unterrepräsentiert ist. Besonders auffällig ist das Urteilsverhalten des vierten Clusters (3d, n = 10), in dem durchgängig die *männlichen* Interpretationen wesentlich positiver erlebt wurden. Ungewöhnlich ist dieser Cluster vor allem dadurch, daß er sich ausschließlich aus *einer* Untergruppe, den (musikinteressierten) Laien rekrutiert.

Diese vier Cluster dominierten in der untersuchten Stichprobe, wobei sich generell feststellen läßt:

- das zurückhaltende (»richtige«) Votum des Clusters 1 erfolgte unabhängig von musikalischer Vorbildung;



»Profilschlauch«: 1. Charakteristisches Profil

1–23 = Cluster-Mittelwerte (Aktive Variablen)

+ = Cluster-Mittelwerte + Stand.-Abw.

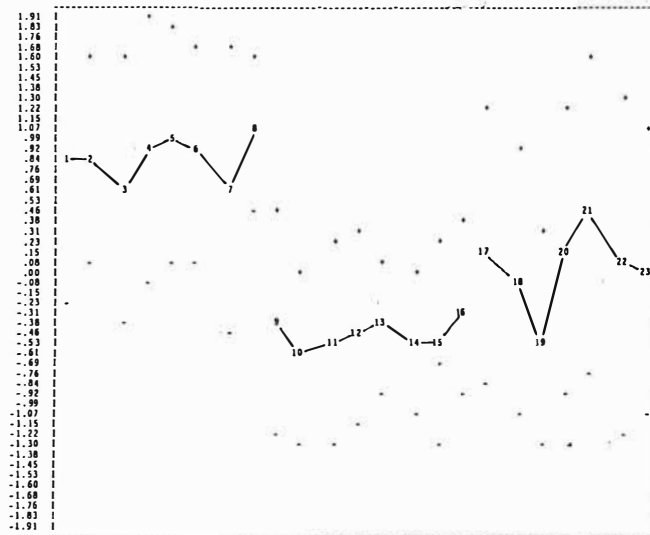
– = Cluster-Mittelwerte – Stand.-Abw.

»Profilschlauch«: 2. Charakteristisches Profil

1–23 = Cluster-Mittelwerte (Aktive Variablen)

+ = Cluster-Mittelwerte + Stand.-Abw.

– = Cluster-Mittelwerte – Stand.-Abw.

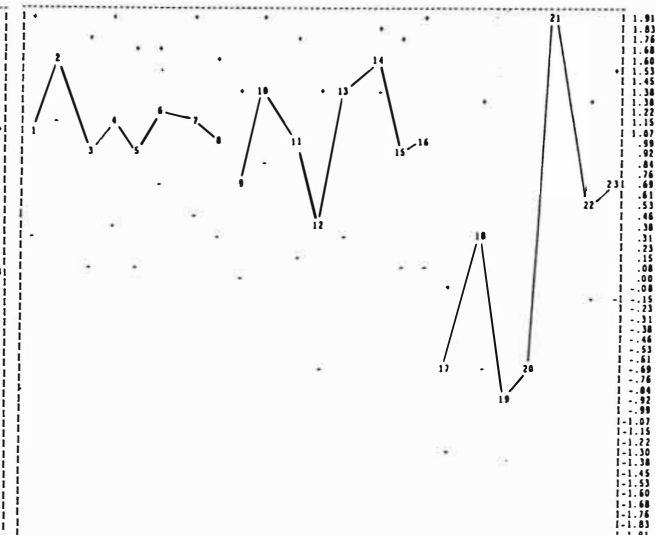


»Profilschlauch«: 3. Charakteristisches Profil

1–23 = Cluster-Mittelwerte (Aktive Variablen)

+ = Cluster-Mittelwerte + Stand.-Abw.

– = Cluster-Mittelwerte – Stand.-Abw.

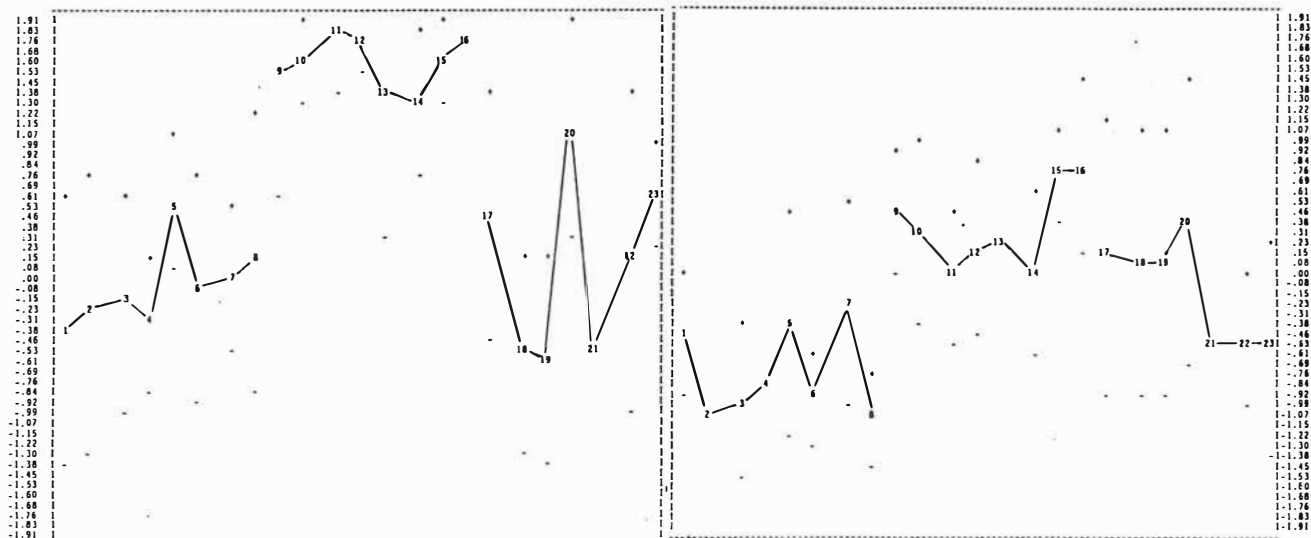


»Profilschlauch«: 4. Charakteristisches Profil

1–23 = Cluster-Mittelwerte (Aktive Variablen)

+ = Cluster-Mittelwerte + Stand.-Abw.

– = Cluster-Mittelwerte – Stand.-Abw.



»Profilschlauch«: 5. Charakteristisches Profil

1-23 = Cluster-Mittelwerte (Aktive Variablen)

+ = Cluster-Mittelwerte + Stand.-Abw.

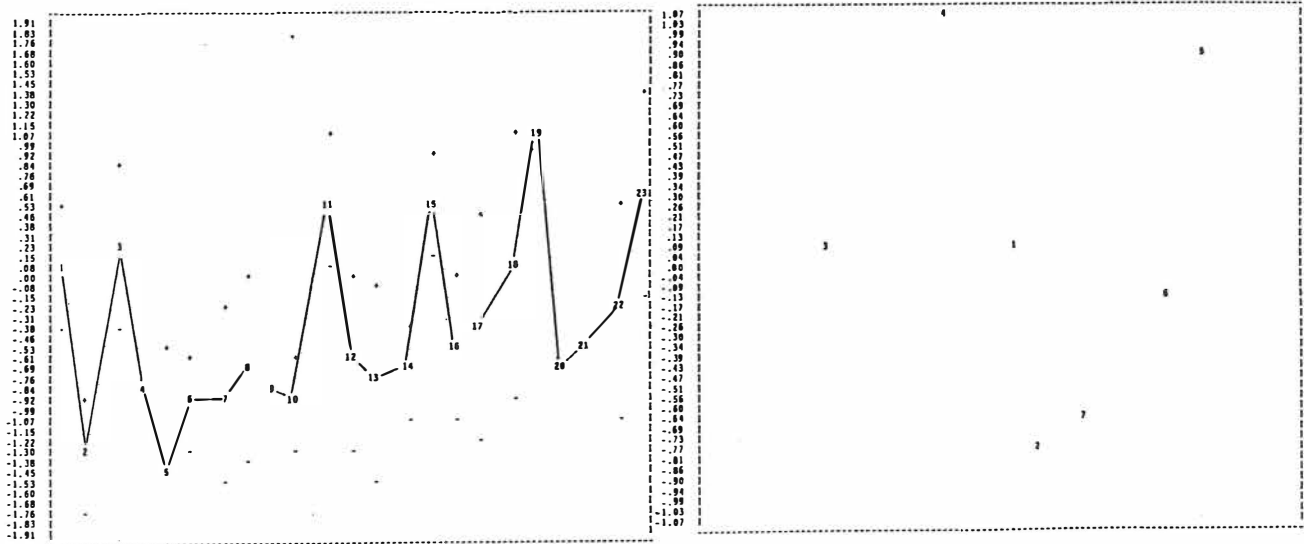
- = Cluster-Mittelwerte - Stand.-Abw.

»Profilschlauch«: 6. Charakteristisches Profil

1-23 = Cluster-Mittelwerte (Aktive Variablen)

+ = Cluster-Mittelwerte + Stand.-Abw.

- = Cluster-Mittelwerte - Stand.-Abw.



»Profilschlauch«: 7. Charakteristisches Profil

1–23 = Cluster-Mittelwerte (Aktive Variablen)

+ = Cluster-Mittelwerte + Stand.-Abw.

– = Cluster-Mittelwerte – Stand.-Abw.

Diagramm der Haupt-Achsen 1 (vertikal) und 2 (horizontal)

Zahlen = Cluster-Nummern

Abb. 3: Graphische Veranschaulichung der charakteristischen Profile

- es besteht eine auffällige Tendenz, sich global für *eine* der beiden Interpretationen zu entscheiden und *sowohl-als-auch*-Urteile (z.B. »Birgit spielte ausdrucksvoller, Constantin aber präziser«) zu vermeiden;
- das Geschlecht der Interpreten spielte eine große Rolle, das der Urteiler aber nur eine untergeordnete Rolle.

Die Cluster 5 bis 7 (3e, f, g) setzen sich jeweils nur aus fünf bzw. vier Urteilern zusammen und sind deshalb weniger quantitativ als qualitativ bedeutsam. In ihnen manifestieren sich Minderheitsvoten, die bei normalen Auswertungsstrategien überhaupt nicht berücksichtigt würden. Die Konstellation aller sieben Cluster ist in Abb. 3h – ähnlich wie bei einer Faktorenanalyse – räumlich dargestellt. Hier sieht man recht gut die »Mittenposition« des »enthaltamen« Clusters 1 sowie in der Kreisanordnung der übrigen sechs Cluster interessante »Oppositionen« bzw. »Nachbarschaften«. Von Interesse ist schließlich auch die Gruppe derjenigen Urteiler, die keinem der sieben Cluster angehören: es sind überwiegend Lehrer mit ausgesprochen ideosynkratischem Urteilsverhalten, die deshalb singulär blieben.

Die Clusteranalyse ist von einer Reihe von Entscheidungsprozessen begleitet, die hier nicht im Detail ausgebreitet werden konnten. So ist jeweils neu festzulegen, welche Variablen clusterbildend sein sollen und welche nicht. In diesem konkreten Beispiel hätte man auch die CHOPIN- und BRAHMS-Daten getrennt analysieren können. Die Entscheidung darüber, an welchen Positionen bzw. Partitionen des hierarchischen Clusterungsprozesses man die zu interpretierenden Cluster auswählt, ist zwar teilweise durch Berechnung kritischer Kennwerte objektivierbar, erfordert aber stets ein abwägendes Urteil des Analysierenden, das um so sinnvoller ausfallen wird, je intensiver man sich mit dem »Innenleben« einer Datenstruktur auseinandersetzt.

Literatur

- K.-E. Behne, 1990 – »Blicken Sie auf die Pianisten?!« Zur bildbeeinflussten Beurteilung von Klaviermusik im Fernsehen. Medienpsychologie (i.D.).
- T. Eckes & H. Roßbach, 1980 – *Clusteranalysen*. Stuttgart: Kohlhammer.
- G.W. Milligan, 1980 – *An Examination of the Effect of six Types of Error Perturbation in fifteen Clustering Algorithms*. Psychometrika 45, 325–342.

- G.W. Milligan, 1981 – *A Review of Monte Carlo Tests of Cluster Analysis*. Multiv. Behav. Res. 16, 379–407.
- R.R. Sokal & C.D. Mitchener, 1958 – *A statistical Method for evaluating systematic Relationships*. Univ. of Kansas Science Bull. 38, 1409–1438.
- D. Steinhausen & K. Langer, 1977 – *Clusteranalyse – Einführung in Mathematik und Verfahren der automatischen Klassifikation*. Berlin: de Gruyter.