

Zur Methode der Clusteranalyse

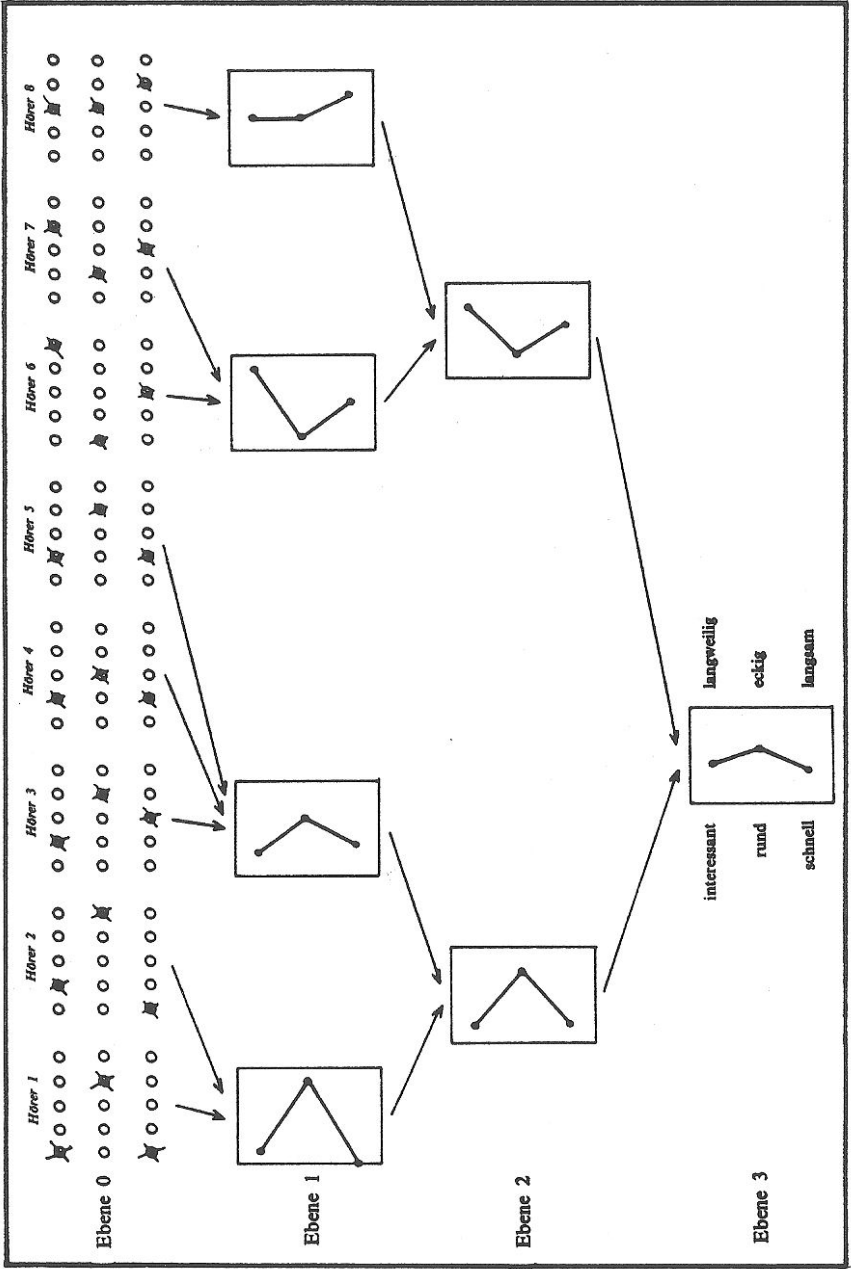
KLAUS-ERNST BEHNE

*Hans Günther Bastian (Hg.): Umgang mit Musik. - Laaber: Laaber 1985.
(Musikpädagogische Forschung. Band 6)*

Obwohl die Clusteranalyse als ein besonderes Auswertungsverfahren in der deutschsprachigen Musikpsychologie seit 1974, wenngleich bisher nur vereinzelt, eingesetzt worden ist¹, so ist sie doch noch weitgehend unbekannt. Dies mag z. T. darin begründet sein, daß manche, angesichts der unübersehbaren Fülle statistischer Auswertungsverfahren, „konservativ“ auf das Bewährte zurückgreifen, andererseits darin, daß auch in der Allgemeinen Psychologie und in den empirischen Sozialwissenschaften die Clusteranalyse noch keineswegs den Bekanntheits- und Nutzungsgrad hat, den sie verdient. Letzteres ist aber wiederum darauf zurückzuführen, daß bisher kaum ein Rechenzentrum die notwendigen Programme in befriedigender Qualität zur Verfügung stellt.² Die folgende, denkbar knappe Einführung richtet sich vor allem an Musikwissenschaftler und Musikpädagogen, die mit diesem Verfahren noch gar nicht vertraut sind.

Die Clusteranalyse ist ein *einordnendes* Verfahren, bei dem Personen, die sich gleich oder ähnlich verhalten haben, zu Clustern (d. h. zu Gruppen von Personen) zusammengefaßt werden. Diese Definition, die sich auf die häufigste Anwendung (Personenclusteranalyse, PCA) bezieht, macht deutlich, daß die Grundidee der Clusteranalyse im Prinzip sehr einfach und überschaubar ist. Sie soll an einem fiktiven Beispiel erläutert werden.

Gegeben sei ein Musikstück, das von 8 Personen gehört und erlebt wird und von jedem Hörer auf drei 5-stufigen Polaritäten beurteilt werden soll (s. Abb. S. 266 oben). Die 8 Hörer wurden auf „Ebene 0“ bereits so angeordnet, daß offensichtlich ähnlich Urteilende benachbart stehen und eine ganz bestimmte Gruppierung nahelegen: Die ersten beiden Hörer wird man (auf Ebene 1) zu einem Zweiercluster zusammenfassen, da beide die Musik als ausgesprochen *schnell*, *interessant* und *eckig* empfinden. Die Hörer 3, 4 und 5 zeigen eine ganz ähnliche Urteilstendenz, vermeiden aber extreme Urteile. Sie sind auf Ebene 1 zu einem Dreiercluster zusammengefaßt, der dem ersten Cluster relativ ähnlich ist. In offenkundigem Kontrast hierzu steht das Urteil der Hörer 6 und 7, die das Stück als ausgesprochen *langweilig* und *rund* sowie als *mittelschnell* eingestuft haben. Auch sie bilden auf Ebene 1 einen Cluster, dessen Profil auch graphisch den Gegensatz zu den beiden ersten Clustern deutlich macht. Es bleibt der 8. Hörer, der das Stück zwar als eher *lang-*



sam empfindet, in seinen sonstigen Urteilen aber indifferent oder unentschlossen bleibt und deshalb — auf dieser Ebene — noch keinem anderen Cluster zugeordnet werden sollte.

Bei einer entsprechend großen Stichprobe könnte es bereits sinnvoll sein, diese vier Cluster als „die vier Arten des Musikerlebens bei Musikstück X“ zu betrachten und als Ergebnis mitzuteilen. Man könnte jedoch auch argumentieren, daß die Unterschiede zwischen Cluster 1 und 2 nur graduell sind und beide deshalb — auf Ebene 2 — zu einem Fünfercluster zusammengefaßt werden sollten. Entsprechend könnte man, wenngleich nicht ganz so zwingend, die Hörer 6, 7 und 8 zu einem Dreiercluster auf Ebene 2 bündeln. Auf dieser Ebene würde man als Ergebnis also zwei Erlebnisprofile mitteilen, eine Mehrheitsauffassung als *ziemlich interessant, schnell und eckig*, sowie eine konträre, aber nicht ganz so ausgeprägte Auffassung bei nur drei Hörern.

Wollte man jedoch wissen, welches die allgemeinen, übergreifenden Urteilstendenzen bei diesem Stück sind, so würde man abschließend auf Ebene 3 den Fünfer- und den Dreiercluster von Ebene 2 zu einem Achtercluster zusammenfassen, somit nur noch die Mittelwerte der gesamten Gruppe betrachten und zu der Feststellung gelangen, daß das Stück generell eher *schnell, interessant* und *eckig* eingestuft wird.

Damit sind die wesentlichen Schritte einer Clusteranalyse an einem sehr überschaubaren Beispiel vereinfachend dargestellt. Die Ergebnisse auf allen Ebenen sind im Prinzip richtig, wenngleich unterschiedlich exakt bzw. nivellierend, die Frage ist nur, ob sie auch sinnvoll sind. Am exaktesten ist es zweifellos, die Rohdaten (Ebene 0) als Ergebnis mitzuteilen, dies wäre aber kaum sinnvoll! Am elegantesten (und häufig am verführerischsten und bequemsten) wären die Gesamtmittelwerte der Ebene 3, aber sie würden ganz offensichtlich wichtige Informationen nivellieren. Ein sinnvolles Ergebnis müßte also irgendwo zwischen Ebene 0 (Rohdaten) und Ebene 3 (Gesamtmittelwerte) liegen.

Das Hauptproblem der Clusteranalyse besteht für den Benutzer nun in der Festlegung der zu interpretierenden Cluster, in unserem Beispiel auf Ebene 1 oder 2, wobei es grundsätzlich keine „richtigen“ und „falschen“ Cluster gibt. Je mehr man sich der Ebene der Rohdaten nähert, um so genauer ist das Ergebnis, aber um so schwieriger, weil umständlicher ist es zu beschreiben. Je mehr man sich hingegen der letzten Ebene nähert, um so eleganter kann die Ergebnisdarstellung gelingen, um so größer ist aber die Gefahr, daß wichtige Informationen über Urteilsstrukturen bei Minderheiten verloren gehen könnten. Bei dieser Wahl der zu interpretierenden Cluster kann man

sich von statistischen Kriterien leiten lassen, aber die Entscheidung muß letztlich durch den Benutzer erfolgen.

Ein naheliegendes und vernünftiges Kriterium ist die Tatsache, ob zwei (oder mehr) ähnliche Cluster, bevor sie im nächsten Schritt der Analyse zusammenfallen, sich hinsichtlich eines oder mehrerer Merkmale signifikant unterscheiden. Vor allem bei einer größeren Zahl von clusterbildenden Variablen stellt sich jedoch die Frage, ob schon signifikante Unterschiede bei nur einer Variablen ausreichen, um die betreffenden Cluster isoliert zu betrachten, oder ob hierfür nicht Unterschiede auf zwei (oder mehr) Variablen nötig wären, um unnötig „feine“ Lösungen — d. h. zu viele Cluster — zu vermeiden. In unserem fiktiven Beispiel wäre es z. B. denkbar, daß sich die beiden profilähnlichen (linken) Cluster auf Ebene 1 zwar bei ein oder zwei Variablen signifikant unterscheiden, daß es aber unabhängig von statistischen Kriterien psychologische Argumente gäbe, die beiden Cluster zusammenzufassen. Wenn man nämlich unterstellt, daß die gleichgerichteten, aber verschiedenen ausgeprägten Urteile nur darauf zurückzuführen wären, daß individuell verschiedene Bezugspunkte wirksam gewesen sind oder daß die Verbalisierung der fünf Skalenstufen unterschiedlich verstanden wurde, so läge es durchaus nahe, die beiden Gruppen zu einem Fünfercluster auf der nächsten Ebene zusammenzufassen. Da ähnliche Argumente sich für die beiden rechten Cluster auf Ebene 1 nicht aufdrängen, könnte es sinnvoll sein, den großen linken Cluster auf Ebene 2 und die beiden kleinen rechten Cluster auf Ebene 1 als Endergebnis einer PCA zu interpretieren.³

Auf die verschiedenen methodischen Probleme der Clusteranalyse, so die Unterschiede zwischen verschiedenen Ähnlichkeitsmaßen sowie die Entscheidung für ein bestimmtes Verfahren, kann hier nicht eingegangen werden; der interessierte Leser sei deshalb auf die einschlägige Literatur (Schlosser 1976) hingewiesen. Man kann jedoch unterstellen, daß für die verschiedenen Verfahren der Clusteranalyse ähnliches gilt wie für jene der Faktorenanalyse, bei denen Revenstorf (1978) feststellte, daß sie i. W. zu den gleichen Ergebnissen führten, eine mehrfache Analyse mit verschiedenen Verfahren also letztlich nur ein unsinniger Aufwand ist.

Der Hauptvorteil der Clusteranalyse ist vor allem darin zu sehen, daß sie die Daten fast voraussetzungslos analysiert⁴ und deshalb keine Artefakte, sondern nur diejenigen Strukturen sichtbar machen kann, die für den betreffenden Datensatz jeweils kennzeichnend sind. Sie garantiert deshalb dort prägnante Ergebnisse, wo diese auch wirklich in den Daten *schlummern*. Die Ergebnisse werden u. a. deshalb prägnanter, weil zwei Gruppen von *un-*

erwünschten Versuchspersonen erkannt und von der weiteren Analyse ausgeschlossen werden können. Das sind zum einen diejenigen, die — aus welchen Gründen auch immer — ihre Urteile wahllos und zufällig abgeben und sich deshalb im Verlauf der Clusteranalyse erst sehr spät einem bestehenden Cluster zuordnen lassen. Da es für diejenigen, die sich einer ehrlichen Beantwortung eines Fragebogens verweigern wollen, durchaus schwierig sein kann, ihre Kreuze rein zufällig aufs Papier zu setzen, kann es auch den Nonsensurteiler geben, der bestimmte „Strickmuster“ kreuzt, also z. B. abwechselnd links und rechts. Treten solche Urteiler in größerem Umfang auf, so bilden sie eigene Cluster, die man sinnvollerweise nicht weiter berücksichtigen wird; treten sie seltener auf, so werden sie, wie die Zufallsurteiler, erst spät an vorhandene Cluster angebinden. Clusterlösungen auf mittleren Ebenen, die häufig nur 80 — 95 % der Befragten berücksichtigen, klammern deshalb vor allem diejenigen Fälle aus, die für die Irrtumsvarianz verantwortlich sind. Diese Verringerung der Irrtumsvarianz ist ein selten gewürdigter Nebeneffekt der Clusteranalyse.

Anmerkungen

- 1 Die vier ersten Untersuchungen, die — relativ unabhängig voneinander die Clusteranalyse auf musikpsychologische Probleme anwendeten, erschienen 1976 (Batel, 2xBehne, Jost), es folgten die Arbeiten von Schaffrath (1978) und Bastian (1980).
- 2 Die Verwendung von CLUSTAN ist bei größeren Stichproben, und zwar schon bei $N > 100$ (!), umständlich und kompliziert, die angekündigte SPSS-Version war zum Zeitpunkt unserer Untersuchung noch nicht verfügbar.
- 3 Dieses Beispiel spricht dagegen, ausschließlich anhand statistischer Kriterien nach einer optimalen oder richtigen Ebene zu suchen. Auch hier ist die vielleicht eleganter anmutende Lösung (auf Ebene I) nicht jene, die auch zugleich die in den Daten vorhandenen Strukturen optimal, d. h. möglichst wenig verfälschend sichtbar macht.
- 4 Intervallskalenqualität ist wünschenswert, leichte Abweichungen dürften aber letztlich nicht sonderlich ins Gewicht fallen; grundsätzlich ist auch die Analyse von Daten mit Nominalskalenqualität möglich.

Literatur

- Bastian, H. G.: Neue Musik im Schülerurteil. Eine empirische Untersuchung zum Einfluß von Musikunterricht, Mainz 1980.
- Batel, G.: Komponenten musikalischen Erlebens. Eine experimentalpsychologische Untersuchung, Diss. Göttingen, Kassel 1976.

- Behne, K.-E.: „Zeitmaße“ — Zur Psychologie des musikalischen Tempoempfindens, in: Die Musikforschung 29, 1976, S. 155-164.
- Ders.: Zur Struktur und Veränderbarkeit musikalischer Präferenzen, in: Zeitschrift für Musikpädagogik 1, 1976, S. 139-146.
- Jost, E.: Sozialpsychologische Faktoren der Popmusik-Rezeption, Mainz 1976.
- Revenstorff, D.: Vom unsinnigen Aufwand, in: Archiv f. Psychologie, 130, 1978, S. 1-36.
- Schaffrath, H.: Der Einfluß von Information auf das Musurteil. Eine Kontextstudie am Beispiel 15-jähriger Gymnasiasten, Herrenberg 1978.
- Schlosser, O.: Einführung in die sozialwissenschaftliche Zusammenhanganalyse, Reinbek (Hamburg) 1976.

Prof. Dr. Klaus-Ernst Behne
Meisenweg 7
Schloß Ricklingen
D-3008 Garbsen 5