

# Einstellungsmessung

FRANZ PETERMANN

Günter Kleinen (Hg.): *Außerschulische Musikerziehung*. - Laaber: Laaber 1987.  
(Musikpädagogische Forschung. Band 8)

## 1. Einleitung

Die Einstellungsmessung hat in der Sozialpsychologie eine knapp 60jährige Tradition und geht auf Arbeiten von US-amerikanischen Sozialpsychologen zurück, die ökonomische Verfahren zur Erfassung von Einstellungen entwickeln wollten (Thurstone Chave 1929). Messung bedeutet in diesem Zusammenhang die quantitative Erfassung der Ausprägung einer Meinung, Einstellung u. ä. zu einem Objekt. Der Meßbegriff läßt sich in verschiedener Weise präzisieren, wobei hierauf nicht eingegangen wird; dafür soll auf die einführende Arbeit von Gigerenzer (1985) verwiesen werden.

Einstellungen im Bereich der Musik können unterschiedlich konkret gefaßt werden. So kann man die Haltung gegenüber der Musik, gegenüber bestimmten Musikrichtungen oder einer konkreten Musikdarbietung analysieren; man könnte auch Details der Musikdarbietung, die Interpretation oder den musikalischen Rahmen bewerten lassen. Darüber hinaus könnte man überprüfen, ob sich allgemein oder bei bestimmten Musikrezipienten über die Zeit spezifische Neigungen herausbilden. Da sehr viele und völlig verschiedenartige Fragestellungen denkbar sind, sollen die weiteren Ausführungen eher allgemein gehalten werden. Es soll jedoch soviel Basiswissen vermittelt werden, daß eine gezielte Suche nach den angemessenen methodischen Zugängen dem Leser möglich ist.

## 2. Der Einstellungsbegriff

Eine der umfassendsten Analysen des Einstellungsbegriffes lieferte schon 1935 Allport. Allport ging davon aus, daß man eine Einstellung als „*Reaktionsbereitschaft*“ auffassen kann, die das menschliche Verhalten bestimmt. In diesem weiten Verständnis sind Einstellungen sowohl aktuelle Erwartungen als auch umfassende stabile Dispositionen. Solche Dispositionen wären etwa weltanschauliche Überzeugungen oder stark verinnerlichte Werte bzw. Wertvorstellungen. Allports Einstellungsbegriff besitzt für den „*Einstellungsmesser*“ seine Tücken! Seine weitgefaßte Konzeption von Einstellung verstellt

ihm den Blick dafür, daß Einstellungen an ein Objekt gekoppelt sein müssen, um sie bestimmen zu können. Diese Fassung vertraten Ende der zwanziger Jahre Thurstone und seine Mitarbeiter. Für Thurstone waren Einstellungen eindimensionale Gebilde, die sich auf einem Kontinuum von Pro und Contra im Hinblick auf ein spezifisches Einstellungsobjekt abbilden lassen. Die historische Einschränkung auf eindimensionale Einstellungen wurde bald zugunsten mehrdimensionaler Modelle aufgehoben (vgl. Togerson 1958; Petermann 1980). Heute dominieren jedoch im Forschungsalltag immer noch einfache, eindimensionale Ansätze wie das Vorgehen von Likert, das den meisten Einstellungsfragebögen zugrunde liegt. Auf diese eindimensionalen Ansätze zur Einstellungsmessung wollen wir deshalb hier ausführlicher eingehen. Hierbei wird kein Bericht technischer Details vermittelt, sondern lediglich die Logik der Skalenkonstruktion vertieft.

### *3. Indirekte Ansätze zur Einstellungsmessung*

Wir können die Ansätze zur Einstellungsmessung nach indirekten und direkten untergliedern. Direkt bezeichnet man ein Vorgehen dann, wenn man eine Versuchsperson unmittelbar nach ihrer Einstellung befragt. Ein indirektes Vorgehen zur Einstellungsmessung liegt dann vor, wenn die Befragten nicht erkennen, daß sie an einer Erhebung teilnehmen, in der sie eindeutige Äußerungen zu einem Einstellungsobjekt aussprechen sollen. Indirekte Ansätze liegen in zweifacher Weise vor: (1) getarnte Verfahren und (2) nicht-reaktive Verfahren.

Bei den getarnten Verfahren erfährt die Versuchsperson nicht den realen Zweck der Erhebung, und man täuscht sie bewußt, indem man unstrukturierte Vorlagen vorgibt, bei denen man die Absichten nicht erkennen kann. Ein anderer Weg besteht darin, falsche Untersuchungsziele der Versuchsperson vorzuspiegeln (vgl. Webb et al. 1975; Petermann & Noack 1984). Solche Verfahren dürften jedoch nur geringe Verbreitungschancen besitzen, da die Ergebnisse solcher Studien nur sehr schwer zu interpretieren sind.

Schon interessanter sind nicht-reaktive Verfahren, bei denen die Versuchsperson nicht weiß, daß sie untersucht wird; man zieht hierzu folgende Verfahren heran:

- *Analyse physikalischer Spuren* (z. B. das Ausmaß der zerstörten Bestuhlung bei einem Rock-Konzert als Hinweis für die Emotionalisierung durch Pop-Musik)

- *nicht-reaktive Verhaltensbeobachtung* (z. B. als „normaler“ Konzertbesucher oder mit Hilfe einer nicht-erkennbaren Video-Kamera beobachtet der Feldforscher, welche Publikumsreaktionen ein bürgerliches Publikum auf avantgardistische Musikdarbietungen zeigt)
- *nicht-reaktive Feldexperimente* (z. B. beobachtet man die Publikumsreaktionen bei avantgardistischen Musikdarbietungen, wobei das Ausmaß der Erklärungen über die dargebotenen Inhalte variiert wird)
- *Inhaltsanalyse vorhandener Dokumente* (z. B. schriftliche — historische — Äußerungen zu bestimmten Musikrichtungen/-stilen)

Die Auswertung nicht-reaktiver Maße bzw. Daten kann sehr mühsam sein, wenn man etwa an den Kategorisierungsaufwand bei einer systematischen Verhaltensbeobachtung oder einer Inhaltsanalyse denkt. Der uneingeschränkte „Datenstrom“ bei nicht-reaktiven Maßen wird auf diese Weise für diesen Ansatz zum Problem. Hinzu kommt, daß man nur eine beschränkte Aussage über die Validität der Befunde anstellen kann. Unter Validität versteht man in diesem Zusammenhang, daß man aufgrund der beobachteten Reaktion eines Zuschauers im Konzertsaal nicht auf seine Einstellung zu der aktuellen Musikdarbietung schließen kann. Seine denkbaren Unmutsäußerungen können auch dadurch bedingt sein, daß er trotz des erbaulichen Musikvergnügens nicht in der Lage ist abzuschalten und ihm zeitweise seine Ehekrise bewußt wird und er darüber verzweifelt ist.

#### *4. Direkte Ansätze zur Einstellungsmessung*

Die hierfür erforderlichen Vorlagen können als standardisiert bezeichnet werden, da sie immer in der gleichen Weise vorgegeben werden. Man kann weiter in Verfahren untergliedern, die aufgrund eines Skalenmodells aufwendig konstruiert werden, und solchen, die diese statistische Absicherung nicht erfahren haben. Wir wollen zunächst mit Verfahren *ohne* Skalenmodell beginnen — also solche ohne spezifischen Konstruktionsaufwand kennenlernen.

##### *4.1 Verfahren ohne Skalenmodell*

In der Regel sind diese Verfahren nicht im Hinblick auf ein spezielles Einstellungsobjekt hin konstruiert worden. Kennzeichnend für diese Ansätze ist,

daß sie auf dem Prinzip der Eigenschaftszuschreibung beruhen und so Aufschluß über die Einstellung zum jeweiligen Gegenstand geben.

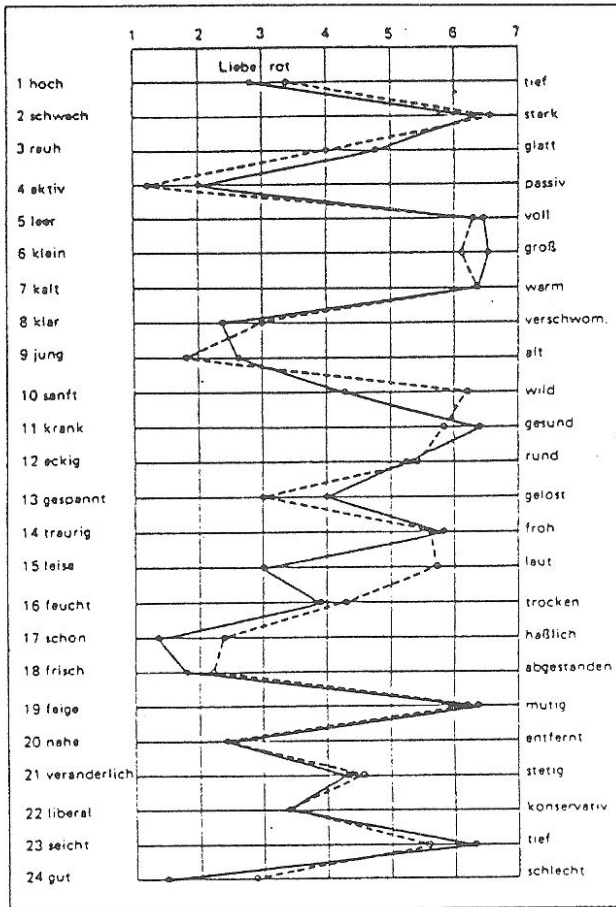
Das einfachste Vorgehen ohne Skalenmodell bildet der Einsatz einer Prüfliste (Checkliste). Eine solche Liste enthält ausgewählte Eigenschaften oder Verhaltensweisen, von denen die Versuchsperson diejenigen ankreuzen soll, die ihrer Meinung nach das Einstellungsobjekt kennzeichnen. Wenn die Prüfliste umfänglich genug ist und damit den zentralen Eigenschaftsraum abdeckt, können durch diese Verfahren einstellungsrelevante Informationen eingeholt werden. Hierzu ein Beispiel:

Das dargebotene Musikstück ist . . .

- abwechslungsreich \_\_\_\_\_
- aggressiv \_\_\_\_\_
- aktivierend \_\_\_\_\_
- bedächtig \_\_\_\_\_
- chaotisch \_\_\_\_\_
- erregend \_\_\_\_\_
- harmonisch \_\_\_\_\_
- langweilig \_\_\_\_\_
- laut \_\_\_\_\_
- melancholisch \_\_\_\_\_
- romantisch \_\_\_\_\_
- witzig \_\_\_\_\_

Um eine Vergleichbarkeit der Antworten zu gewährleisten, werden die Befragten gebeten, nur eine bestimmte Anzahl von Urteilen abzugeben.

Ein zweites, weit verbreitetes Vorgehen bezeichnet man als Polaritätenprofil (auch Eindrucksdifferential oder Semantisches Differential genannt); dieses Verfahren ist vielfältig einsetzbar und wurde von Hofstätter in den fünfziger Jahren eingeführt sowie von Bergler (1975) systematisch erforscht. Dieses Verfahren ist vielseitig einsetzbar, wobei durch eine größere Anzahl von Adjektivpaaren der Aussagegehalt bzw. die Bedeutung eines Begriffes bzw. Objektes beschrieben werden kann. Man verwendet hierbei z. B. die folgende Skala (vgl. Hofstätter 1970, S. 30):



**Abbildung 1:** Polaritätenprofil nach Hofstätter (1970, S. 30)

In der graphischen Darstellung erkennt man 24 Adjektivpaare, die entgegengesetzte Endpunkte einer Skala verdeutlichen sollen. Beispielhaft werden in der Darstellung zwei Begriffe (Liebe, rot) in ihrem *Bedeutungsprofil* gekennzeichnet. Wir erkennen in dem Beispiel von Hofstätter, wie ähnlich die Einschätzungen der beiden Begriffe „Liebe“ und „rot“ sind. Diese Ähnlichkeiten können auf gemeinsame, zugrundeliegende Beurteilungsgesichtspunkte zurückgeführt werden. Das Polaritätenprofil kann schon als weitgehend standardisiert angesehen werden, da es in der vorliegenden Form in vielen Bereichen zur Anwendung gelangen kann.

## 4.2 Verfahren mit Skalenmodell

Wir kommen jetzt zur eigentlichen Einstellungsmessung, bei der man mit Hilfe einer speziell konstruierten Skala (= Einstellungsfragebogen) Daten sammelt, die eine Information über eine Person oder Personengruppe hinsichtlich ihrer Position zu verschiedenen Objekten oder ebenfalls Personen beinhaltet. In der Regel werden den Antworten der befragten Personen Zahlenwerte zugeordnet, die die Position des Einzelnen auf einem Kontinuum widerspiegeln. Wir werden kurz auf verschiedene, einfache Verfahren zur Konstruktion einer Einstellungsskala eingehen (vgl. ausführlicher Stosberg 1980).

Allen Bemühungen im Rahmen der Einstellungsmessung ist gemeinsam, daß man die Variation in den Reaktionen einer Person auf eine vorgegebene Aussage bestimmen möchte. Die Aussagen können alle nur denkbaren Bereiche umfassen wie etwa die Beurteilung von Komponisten, die subjektive Erlebnisqualität im Hinblick auf verschiedene Musikstücke oder die Beurteilung von Verkaufschancen neuer Produkte aus dem Bereich der Popmusik.

Um die Variation in den Antworten von Personen bestimmen zu können, muß man sich die Quellen vor Augen halten, die zur Varianz im Antwortverhalten beitragen können. Man kann in Anlehnung an Togerson (1958) drei Varianzquellen unterscheiden:

- die reagierende Person selbst,
- die unterschiedliche Bedeutung der Aussagen (Fragen), die der Person vorgelegt werden,
- die Wechselwirkung zwischen dem Verhalten der reagierenden Person und der unterschiedlichen Bedeutung der Fragen.

Liegt die Varianz in der Person begründet, dann wäre es in der Einstellungsmessung konsequent, sich ausschließlich dieser Varianzquelle zuzuwenden und diese aufzuklären. Wir nennen diesen Ansatz deshalb auch personenorientierte Messung, der durch das Verfahren von Likert repräsentiert wird. Geht man jedoch davon aus, daß die Varianz durch die vorgegebenen Aussagen (Fragen) bedingt ist, dann liegt eine indikatororientierte Messung vor, die durch die Methoden von Thurstone repräsentiert wird. Die Vorschläge von Thurstone umfassen den Paarvergleich und die Methode der gleicherscheinenden Intervalle. Eine kombinierte Betrachtung beider Varianzquellen nennt man reaktionsorientierte Messung, die von Guttman in der nach ihm benannten Skalogrammenanalyse vorliegt.

Bei der Konstruktion einer Einstellungsskala geht man von der Definition der zu erfassenden Einstellung aus und sammelt aufgrund von theoretischen Vorgaben oder auch Plausibilitätsüberlegungen Aussagen, die diese Einstellung erfassen. In der Sozialpsychologie nennt man die Aspekte der zu erfassenden Einstellung auch Dimensionen, und man unterscheidet zwischen ein- und mehrdimensionalen Einstellungen, je nachdem wie viele Aspekte erfaßt werden sollen (vgl. Petermann 1980). Der Dimensionsbegriff unterstellt, daß eine homogene bzw. in homogene Aspekte zerlegbare Einstellung gegeben ist.

Im nächsten Schritt muß geprüft werden, ob die gesammelten Aussagen in ihrem Bedeutungsgehalt eindeutig sind und sich jeweils nur auf einen einzigen Aspekt beziehen. So wäre ein Beispiel für eine Aussage, die sich nur auf einen einzigen Aspekt bezieht der Satz: „Die Beatles haben auf die moderne Musikgeschichte einen großen Einfluß ausgeübt.“ Ein Beispiel für eine Aussage, die mehrere Aspekte gleichzeitig umfaßt, wäre etwa: „Alles in allem glaube ich, daß jede Zeit ihre Musik hat.“ Es ist leicht einzusehen, daß solch eine vieldeutige Aussage schwer zu interpretieren ist. Aus diesem Grund ist man bemüht, möglichst eindimensionale Aussagen zu finden.

Im nächsten Schritt muß der Forscher prüfen, inwieweit ein bestimmtes Skalenmodell (z. B. der Likert- oder Thurstone-Ansatz) für die Fragestellung angemessen ist. Diese Wahl wird wieder von der Annahme abhängen, durch welche Bedingungen die Varianz im Antwortverhalten bestimmt wird (vgl. die einleitenden Überlegungen in diesem Abschnitt).

Von grundlegender Bedeutung ist die Überprüfung, inwieweit die Aussage über die Zeit hinweg die gleiche Bedeutung beibehält, d. h. stabil ist. So sind etwa Fragen nach politischen Einstellungen starkem Wandel unterworfen. Darüber hinaus muß eine Aussage bei unterschiedlicher Plazierung im Fragebogen die gleiche Bedeutung besitzen.

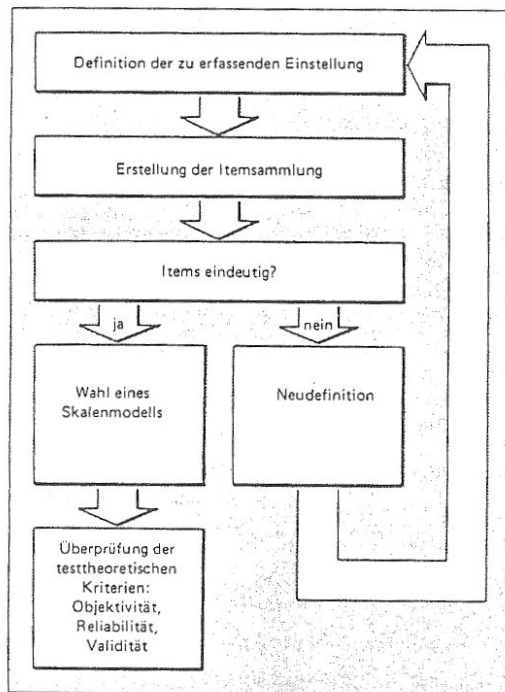
Weiterhin müssen die grundlegenden Kriterien der Testkonstruktion (vgl. Lienert 1969) auch an die Skalenkonstruktion angelegt werden, was bedeutet, daß eine Skala objektiv, reliabel und valide sein soll. Die Schritte der Datenerhebung müssen objektiv sein, und daher dürfen die Bedingungen der Datenerhebung das Ergebnis nicht systematisch beeinflussen (z. B. unterschiedliche Versuchsleiter, unterschiedliche Anordnung der Aussagen, unterschiedliche Umgebung bei der Beantwortung zu Hause, am Arbeitsplatz usw.). Weiterhin darf die Auswertung und Interpretation der Ergebnisse der Befragung nicht vom Versuchsleiter beeinflußt werden.

Eine Voraussetzung für die Reliabilität ist, daß ein eindimensionales Merk-

mal durch eine ausreichende Anzahl von Aussagen (vereinfacht: je mehr Aussagen, desto zuverlässiger) beschrieben wird. Weiterhin sollen die Aussagen trennscharf sein. Dies bedeutet, es muß die Möglichkeit bestehen, daß Personen mit unterschiedlich ausgeprägten Einstellungen differenziert auf vorgegebene Aussagen reagieren. Diese Eigenschaft einer Skala nennt man „Trennschärfe“, und diese muß für den gesamten Bereich der Skala, also nicht nur für deren Extrembereich gewährleistet sein.

Ein schwer prüfbares Kriterium einer Skala ist ihre Validität. Die Validität gibt den Grad der Genauigkeit an, mit dem diese Skala dasjenige Verhalten, das sie messen soll oder zu messen vorgibt, auch tatsächlich mißt.

Die abschließende Abbildung versucht nochmals, den Ablauf der Skalenerstellung zu verdeutlichen. Ein Überblick über das technische Vorgehen bei der Erstellung einer eindimensionalen Einstellungsskala, auf die sich die bisherigen allgemeinen Aussagen beziehen, gibt Stosberg (1980).



**Abbildung 2:** Grober Ablaufplan der Konstruktion einer Einstellungsskala

## 5. Neuere Ansätze zur Einstellungsmessung

Die letzten beiden Jahrzehnte standen im Blickpunkt mehrdimensionaler und/oder probabilistischer Modelle, die es gestatten, mehrere Einstellungsaspekte simultan zu erfassen (vgl. Liepmann 1983; Petermann 1980). Von ganz besonderem Interesse waren die probabilistischen Ansätze, die aber von einem anderen Einstellungskonzept ausgehen. So wird unterstellt, daß die gemessene (manifeste) Einstellung nicht mit der latenten (realen) Einstellung deckungsgleich ist. Demnach sind die beobachteten Einschätzungen einer Person lediglich Indikatoren für eine nicht unmittelbar erkennbare Einstellung (Wakenhut 1980). Bei den probabilistischen Modellen sind latente und manifeste Einstellung nicht nach einer 1:1-Relation miteinander verbunden, so daß diese Modelle Fehler und Zufallseinflüsse mitberücksichtigen können. Die Verknüpfung zwischen latenter und manifester Dimension erfolgt aufgrund von Wahrscheinlichkeitsbeziehungen.

Im einzelnen haben sich verschiedene probabilistische Ansätze herauskristallisiert wie die Analyse latenter Klassen nach Lazarsfeld, die Analyse nach Rasch und Mokken (vgl. Wakenhut 1980). Die Kritikpunkte an diesen neueren Verfahren liegen in ihren Restriktionen bei der Konstruktion einer Einstellungsskala, so daß oft nur wenige Fragen mit den mathematisch komplexen Vorannahmen übereinstimmen. Es bleibt nach gut 20 Jahren Erfahrung mit probabilistischen Modellen zur Einstellungsmessung völlig offen, ob diese mathematisch befriedigenderen, aber praktisch aufwendigeren Ansätze sich jemals durchsetzen werden. Heute dominiert in der Forschungspraxis die eindimensionale Einstellungsskala, die ökonomisch auswertbar ist. In den Fällen, in denen die Vielgestaltigkeit eines Einstellungsbereichs untersucht wird, kombiniert man vielfach mehrere Skalen miteinander und stellt die einzelnen Kennwerte gegenüber.

### Literatur

- Allport, G. W.: Attitudes. In: C. Murchinson (ed.): A handbook of social psychology. Vol. 2. Worcester, Mass. 1935.
- Bergler, R. (Hg.): Das Eindrucksdifferential. Bern 1975.
- Gigerenzer, G.: Messung und Modellbildung. In: H. Bruhn/R. Oerter/H. Rösing (Hg.): Musikpsychologie. München 1985.

- Hofstätter, P. R.: Psychologie. Frankfurt 1970.
- Lienert, G. A.: Testaufbau und Testanalyse. Weinheim <sup>3</sup>1969.
- Liepmann, D.: Einstellungsmessung. In: D. Frey/S. Greif (Hg.): Sozialpsychologie. München 1983.
- Petermann, F. (Hg.): Einstellungsmessung — Einstellungsforschung. Göttingen 1980.
- Petermann, F./Noack, H.: Nicht-reaktive Meßverfahren. In: E. Roth (Hg.): Sozialwissenschaftliche Methoden. München 1984.
- Stosberg, M.: Klassische Ansätze in der Einstellungsmessung. In: F. Petermann (Hg.): Einstellungsmessung — Einstellungsforschung. Göttingen 1980.
- Thurstone, L. L./Chave, E. J. The measurement of attitude. Chicago 1929.
- Togerson, W. S.: Theory and methods of scaling. New York 1958.
- Wakenhut, R.: Probabilistische Modelle in der Einstellungsmessung. In: F. Petermann (Hg.): Einstellungsmessung — Einstellungsforschung. Göttingen 1980.
- Webb, E. J./Campbell, D. T./Schwartz, R. D./Sechrest, L.: Nicht-reaktive Meßverfahren. Weinheim 1975.

Prof. Dr. Franz Petermann  
Psychologisches Institut der Universität Bonn  
Römerstraße 164  
5300 Bonn 1