

3.4 Individuelle Unterschiede in psychophysischen Aufgaben – einige Probleme und Modelle –

Alf Zimmer

(Universität Oldenburg, Fach Psychologie, z. Zt. Stanford University)

Zusammenfassung

Individuelle Unterschiede im Bereich der Psychophysik können vor allem an den folgenden Punkten auftreten

- bei der Beziehung zwischen der Empfindungs- und der Urteilsskala
- bei der Beziehung zwischen der individuellen Empfindungsskala und dem allgemeinen psychophysischen Gesetz
- bei der Relevanz unterschiedlicher Beurteilungsdimensionen in mehrdimensionalen Reizkontinua (z. B. Lärm)

Es wird gezeigt, mit welchen Modellen individuelle Unterschiede in diesen Bereichen untersucht werden können und welche Annahmen dabei gemacht werden müssen.

Meßtheoretische Analysen der Psychophysik, insbesondere des Schalls (siehe z. B. FALMAGNE 1974, LEVINE 1974 und ZAUS 1980, in diesem Band) nehmen an, daß situative und interindividuelle Unterschiede genauso wie Ungenauigkeiten bei der physikalischen Realisierung zu einem weiter nicht analysierbaren Fehlerterm zusammengefaßt werden können. In diesem Beitrag werden – ausgehend von der Annahme, daß die individuellen Unterschiede systematischer Natur sind – eine Reihe von Modellen skizziert, die eine Analyse solcher Unterschiede ermöglichen. Dabei muß allerdings festgehalten werden, daß die meßtheoretischen Grundlagen dieser Modelle bei weitem nicht so gut untersucht sind, wie bei den obengenannten. Ihre Anwendung hat – was den wissenschaftstheoretischen Aspekt angeht – also eher heuristischen Charakter.

3.4.1 Individuelle Unterschiede in der Beziehung von Empfindung und Beurteilung bei Diskriminationsaufgaben

Gibt man den Versuchspersonen in einem Diskriminationsexperiment die Anweisung, auf einer Ratingskala anzugeben, wie sicher sie sich hinsichtlich der Lautheitsunterschiede zweier Reize sind (z. B. „bin sehr sicher, daß Reiz 1 lauter ist als Reiz 2“, „bin sicher, daß Reiz 1 lauter ist“, „die Reize 1 und 2 sind nicht unterscheidbar hinsichtlich Lautheit“, „bin sicher, daß Reiz 2 lauter ist“, „bin sehr sicher, daß Reiz 2 lauter ist“), dann lassen sich aufgrund dieser Urteile Punkte auf einer Receiver-Operating-Characteristic-Kurve

entsprechend den Annahmen der Signal-Entdeckungstheorie (GREEN & SWETS 1966; EGAN 1975; speziell für die Anwendung der Rating-Methode: LEE 1971) bestimmen (siehe Abbildung 47 und 48).

In Abbildung 47 und 48 ist der Fall dargestellt, bei dem entweder die Rating-Urteile einer Person analysiert werden oder alle Personen das Ra-

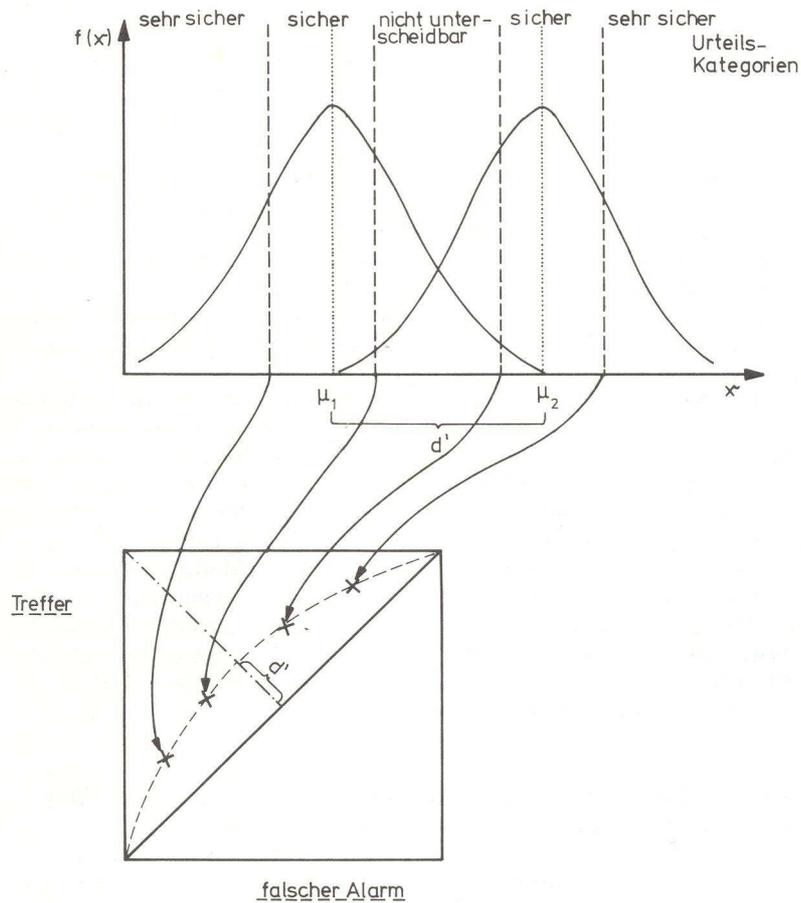


Abbildung 47: Verteilung von Signalstärke und Rauschstärke (normalverteilt) und entsprechende Kategoriengrenzen.

Abbildung 48: ROC-Kurve (interpoliert) für die Signal-Rauschen-Verteilung in Abbildung 47. Die Punkte auf der Kurve (x) entsprechen Kategoriengrenzen.

tingssystem in gleicher Weise anwenden. In den Abbildungen 49 und 50 wird dargestellt, welche Konsequenzen es für das Modell hat, wenn zwei Versuchspersonen das Ratingsystem unterschiedlich anwenden, d. h. ihre Empfindungen unterschiedlich in Rating-Urteile übersetzen.

Aus Abbildung 50 wird deutlich, daß die unterschiedlichen Kategoriengrenzen zur gleichen ROC-Kurve führen: die Vpn unterscheiden sich nur hinsichtlich des Urteilsverhaltens.

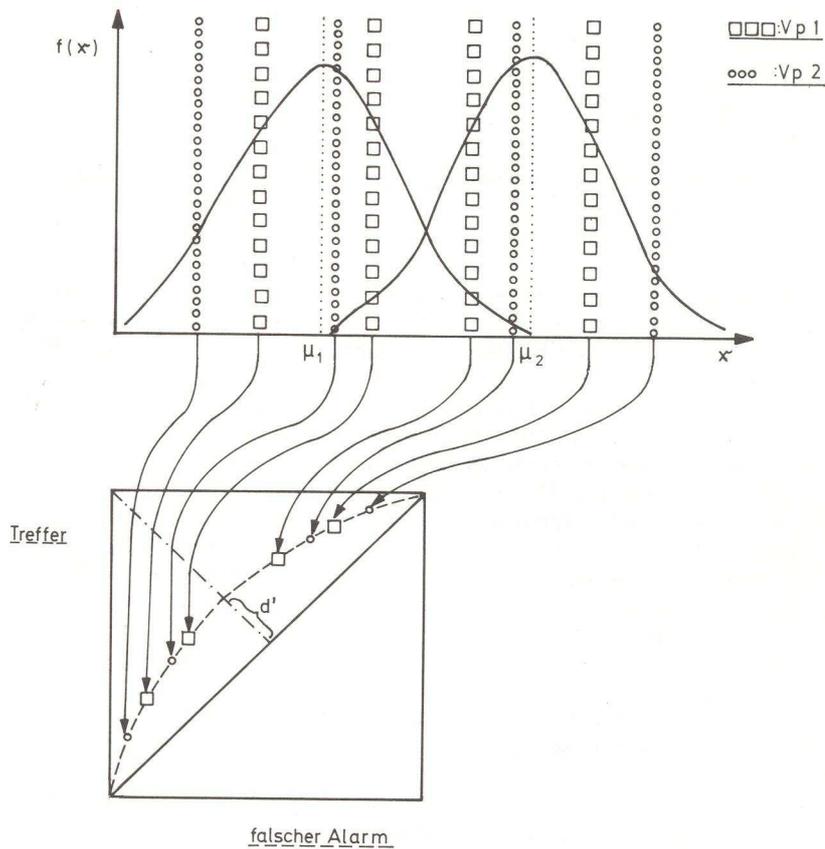


Abbildung 49: Signal-Rauschen-Verteilung mit individuell unterschiedlichen Kategoriengrenzen.

Abbildung 50: ROC-Kurve (interpoliert) für die Signal-Rauschen-Verteilung in Abbildung 49. Die Punkte auf der Kurve (□ und ●) entsprechen den individuellen Kategoriengrenzen.

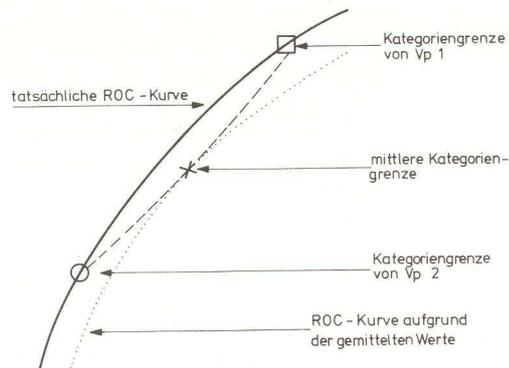


Abbildung 51: Auswirkung der Mittelung individueller Kategoriengrenzen auf den Verlauf der ROC-Kurve.

Welche Konsequenzen hat es nun, wenn diese individuellen Unterschiede nicht berücksichtigt werden? In Abbildung 51 werden die Auswirkungen dieses Mittelungsprozesses für den Übergang der ersten zur zweiten Urteils-Kategorie verdeutlicht.

Dieses Ergebnis läßt sich verallgemeinern: wenn sich die Versuchspersonen hinsichtlich der Anwendung des Rating-Systems unterscheiden, dann führt eine Zusammenfassung der Urteile immer zu einer Verringerung von d' , d. h. die tatsächliche Unterscheidbarkeit wird unterschätzt.

Zusammenfassend läßt sich sagen, daß die Anwendung des Signal-Entdeckungs-Modells individuelle Unterschiede in der Anwendung von Urteils-systemen bei Diskriminationsaufgaben berücksichtigen kann, wenn sich diese Unterschiede ausschließlich auf die Breite der Bereiche bezieht, die von den einzelnen Urteilkategorien abgedeckt wird.

3.4.2 Individuelle Unterschiede bei Absolutbeurteilungen von akustischen Reizen

Bei empirischen Untersuchungen, speziell Felduntersuchungen von Phänomenen, wie Lärm, Lautheit oder akustische Belästigungen, ist häufig die absolute Beurteilung von Reizen die einzig durchführbare Methode der Datenerhebung. Herkömmliche Auswertungsmethoden (s. z. B. EDWARDS 1951) nehmen dabei an, daß die Beurteiler sich bei der Anwendung ihres Urteilssystems weder hinsichtlich der Lageparameter (was ist ein normaler Geräuschpegel?) noch hinsichtlich der Dispersionsparameter (welchen Umfang hat die Kategorie „ziemlich laut“?) unterscheiden. Formal ausgedrückt

führt dies für den vereinfachten Fall des additiven Zusammenwirkens aller Faktoren zu dem folgenden Fehlermodell:

$$(1) \quad u_g = (p_g \Theta + e_g + e_i) e_d$$

u_g : mittlere Beurteilung des Reizes g
 p_g : Korrelation der Beurteilungsskala u mit der Empfindungsskala
 Θ : Empfindungsskala
 e_g : Fehler bei der Herstellung des Reizes g
 e_i : individuelle Variation der Lage des Beurteilungssystems zu
 e_d : individuelle Variation der Dispersion der Kategoriumfänge

Bei der Bestimmung von u_g durch die Berechnung des Mittelwertes oder Medians wird von e_g , e_i und e_d angenommen, daß sie nicht korrelieren.

Weiterentwicklungen der Latent-trait-Theorie durch ANDERSON (1973), SAMEJIMA (1972), BOCK (1972), BEJAR (1977) und ANDRICH (1978) analysieren die Wahrscheinlichkeit, daß auf die Frage oder eine Aufgabe g mit der Kategorie x geantwortet wird, wenn ein bestimmtes Θ vorliegt. Wenn man annimmt, daß Unterschiede auf der Θ -Skala den individuellen Unterschieden hinsichtlich des Lageparameters entsprechen, dann lassen sich die Latent-trait-Modelle folgendermaßen auf die Absolutbeurteilungen in psychophysischen Aufgaben übertragen:

Die Wahrscheinlichkeit, daß eine Versuchsperson i mit einem bestimmten Lageparameter Θ_i auf den Reiz g mit der Kategorie x antwortet, ist eine Funktion des Abstandes der Reizstärke b_{xg} von dem ‚Normal‘-Reiz der Versuchsperson ($b_{xg} - \Theta_i$) und der von der Versuchsperson i verwendeten Kategoriebreite (a_i).

Bei Anwendung des Modells von SAMEJIMA (1972) erhält man

$$(2) \quad p(x_g | \Theta) = \frac{1}{\sqrt{2\pi}} \int_{a_i(\Theta - b_{x_{g+1}})}^{a_i(\Theta - b_{x_g})} e^{-t^2/2} dt$$

Die gesuchten Parameter a_i und b_{xg} lassen sich unter den Annahmen der Normalverteilung und Eindimensionalität einfach schätzen.

$$(3) \quad \hat{a}_i = \frac{\hat{Q}_i}{\sqrt{1 - \hat{Q}_i^2}}$$

Dabei ist \hat{Q}_i die Korrelation der Empfindungsskala von Versuchsperson i und der Reiz-Skala,

und

$$(4) \quad \hat{b}_{xg} = \Phi^{-1}(\hat{p}_{xg}) / \hat{Q}_g$$

Dabei ist Φ die inverse Normalverteilungsfunktion, \hat{p}_{x_g} ist die relative Anzahl von Versuchspersonen, die dem Reiz g die Kategorie x zuordnen, und \hat{q}_g ist die Korrelation zwischen der Reiz- und der Kategorienskala für g .

Für den speziellen Fall mnestisch stabilisierter Bezugssysteme im Sinne von WITTE (1960) läßt sich das von ZIMMER (1980) entwickelte Modell anwenden, in denen Kategorien als unscharfe Mengen im Sinne von ZADEH (1965) interpretiert werden; dieses Modell ermöglicht eine schärfere Prüfung der Voraussetzungen und impliziert nicht die Normalverteilungsannahme.

Wenn man die Operationen von Versuchspersonen im Herstellungsverfahren als kontinuierliche Antwortreaktion (z_g) interpretiert und in das Intervall $[0,1]$ transformiert, so lassen sich diese Operationen ebenfalls durch ein Latent-trait-Modell analysieren (SAMEJIMA 1972).

3.4.3 Individuelle Unterschiede bei der Beurteilung komplexer Reize

Die alltägliche Beobachtung, daß Schallereignisse von verschiedenen Beurteilern ganz unterschiedlich eingeschätzt werden, bzw. daß die subjektive Lästigkeit zweier Schallereignisse mit gleichem Dauerschallpegel (z. B. Autobahn und Eisenbahn) stark differiert, macht das Konzept einer eindimensionalen Lärmskala (wie sie z. B. der Bestimmung von „noy“ zugrundeliegt) fraglich. Um die Dimensionalität des komplexen Ereignisses Lärm zu untersuchen, ist es daher sinnvoll, Lärm nicht absolut zu beurteilen, sondern Schallereignisse paarweise hinsichtlich ihrer Ähnlichkeit einzuschätzen. Unter der Voraussetzung, daß der generelle Eindruck von Lärm sich additiv aus verschiedenen Komponenten (jede für sich eindimensional, aber nicht notwendig physikalisch interpretierbar und voneinander unabhängig) zusammengesetzt, kommt man zu einer Klasse von Modellen, die annehmen, daß die Ähnlichkeit zweier Reize ($s(x,y)$) eine monoton fallende Funktion ihrer Distanz ($d(x,y)$) in einem Eigenschaftsraum ist:

$$(5) \quad f[s(x,y)] = d(x,y) \\ = h^{-1} \left[\sum_i^k h(|x_i - y_i|) \right]$$

dabei ist h eine Funktion und h^{-1} die entsprechende Inverse. x_i und y_i sind die Ausprägungen der Reize in der i -ten Dimension. Die Dimensionalität des Eigenschaftsraumes ist k . BEALS, KRANTZ und TVERSKY (1968) haben bewiesen, daß h eine Potenzfunktion ist.

$$(6) \quad h(x) = x^r$$

Ein besonders interessanter Fall ist das Modell mit $r = 2$ (die sog. euklidische Metrik), da in ihr die Distanzen invariant gegenüber Rotation des Achsen-

systems sind. Untersuchungen des Lärms mit dem Modell weisen darauf hin, daß dem komplexen Reiz ‚Lärm‘ die Eigenschaften Lautheit, Vorhersagbarkeit, spektrale Zusammensetzung (Überwiegen der hohen Frequenzen) und Disharmonie zugrundeliegen; allerdings sind diese Ergebnisse nicht eindeutig. Ursache für die mangelnde Eindeutigkeit kann sein, daß individuelle Unterschiede bezüglich der Lärmwahrnehmung nicht berücksichtigt worden sind. Das einfachste Modell für die Berücksichtigung individueller Unterschiede in diesem Fall geht davon aus, daß alle Beurteiler zwar die gleichen Eigenschaften ihrem Ähnlichkeitsurteil zugrundelegen, aber die Eigenschaften unterschiedliches Gewicht haben. So mag die „Vorhersagbarkeit“ für eine Versuchsperson kaum Bedeutung haben, während sie für eine andere von entscheidender Bedeutung ist. Im Modell wird dies durch unterschiedliche Gewichtung der Dimensionen berücksichtigt. In Abbildung 52 wird

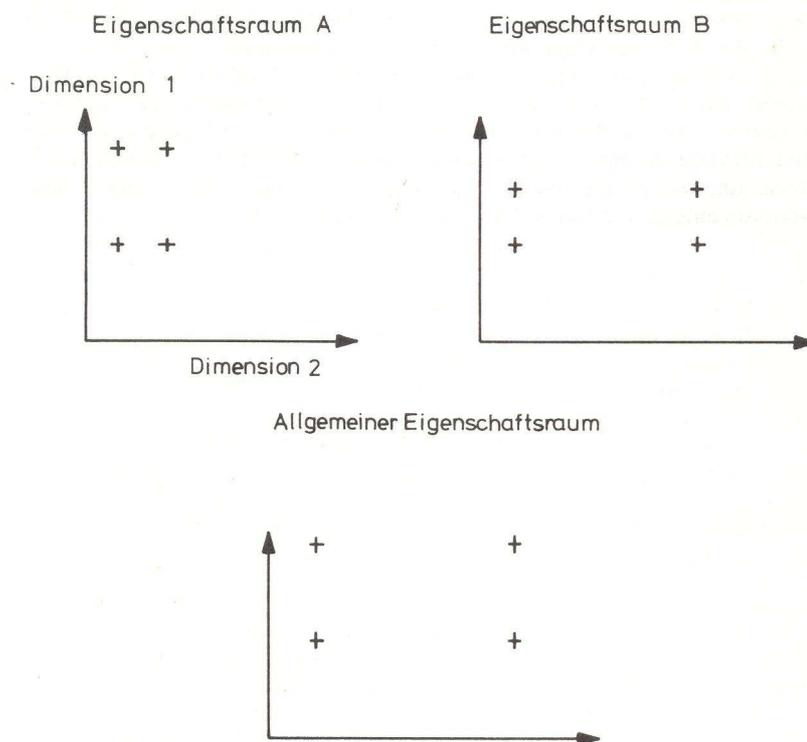


Abbildung 52: Räumliche Konfigurationen für Reize bei Beurteilern, die Dimensionen unterschiedlich gewichten.

gezeigt, wie sich die Eigenschaftsräume zweier Versuchspersonen unterscheiden, die den jeweiligen Eigenschaften entgegengesetzte Bedeutung zumessen.

Das diesen Veranschaulichungen zugrundeliegende Distanzmodell gewichtet jede Dimension i mit dem für die Versuchsperson j entsprechenden Faktor $0 \leq w_{ij} \leq 1$ (CARROLL & CHANG 1970).

$$(7) \quad d_j(x,y) = \left[\sum_i^k w_{ij}(x_i - y_i)^2 \right]^{1/2}$$

Ein Problem bei der Anwendung dieses Modells ist insbesondere die Interpretierbarkeit der Ergebnisse, da die Gewichte nicht invariant gegenüber der Rotation der Achsen sind, und die so gefundenen Dimensionen häufig in keiner sinnvollen Beziehung zu davon unabhängig definierbaren Variablen (z. B. physikalischen Charakteristika von Lärm) stehen.

Der Verfasser arbeitet z. Zt. an einem Modell der individuellen Unterschiede bei Ähnlichkeitsurteilen, das auf der Axiomatik der Ähnlichkeit von TVERSKY (1977) aufbaut und nicht die Annahme der Zerlegung in additive dimensionale Komponenten macht. Dieses Modell führt zu individuellen Beziehungsnetzen, die einen interindividuellen Vergleich dann ermöglichen, wenn sie eine hierarchische Baumstruktur aufweisen.