

4. Kapitel

Semantische Differential Technik

Bernd Schäfer

1. Einleitung

Die Semantische Differential Technik ist eine Methode zur Analyse der Bedeutung von Zeichen. Ein Semantisches Differential (SD) besteht aus einer (nicht verbindlich festgelegten) Anzahl von bipolaren (meist siebenstufigen) Rating-skalen, deren Endpunkte in der Regel durch Adjektive gekennzeichnet sind. Im deutschen Sprachraum werden u. a. auch die Bezeichnungen ‚Eindrucksdifferential‘ und ‚Polaritätsprofil‘ verwendet. Die SD-Technik wurde von Osgood und Mitarbeitern (Osgood 1952; Osgood & Suci 1955; Osgood et al. 1957) zur Analyse der dem sprachlichen Bedeutungsverhalten zugrunde liegenden Dimensionalität entwickelt und wird seitdem auch außerhalb der psycholinguistischen Problemstellung mit großer Häufigkeit in der empirischen Sozialforschung und in nahezu allen Bereichen der psychologischen Forschung eingesetzt. Darstellungen theoretischer, methodischer und technischer Art geben Osgood, Suci & Tannenbaum (1957), Heise (1969), Snider & Osgood (1969), Bergler (1975) und Osgood, May & Miron (1975).

1.1 Zugrundeliegende Modelle

Die SD-Technik wird von Hörmann (1976, 92) im Hinblick auf die Komplexität ihrer theoretischen Begründung als „eine glänzende Leistung des ‚aufgeklärten Neobehaviorismus‘“ gewürdigt. Als Methode zur Erfassung bedeutungsspezifischer Reaktionen auf Zeichen ist sie durch eine Verhaltenstheorie der Bedeutung von Zeichen fundiert, die mit einem Mess- und einem Raummodell in einen Korrespondenz-Zusammenhang eingebettet ist. Zwar sind diese drei Modelle logisch voneinander unabhängig, ihre Verbindung kann aber als charakteristisch für die SD-Technik gelten.

1.1.1 Verhaltensmodell (*representational mediation theory*)

Die Bedeutung von Zeichen wird vom beobachtbaren Verhalten gegenüber den bezeichneten Dingen hergeleitet. Die Genese von Zeichen und ihrer Bedeutung wird von Osgood (1971, 11) folgendermaßen beschrieben:

„a stimulus pattern (S') which is not the same physical event as the thing signified (S) will become a sign of that significate when it becomes conditioned to a mediation process, this process: (a) being some distinctive representation of the total behavior (R_T) produced by the significate, and (b) serving to mediate overt behaviors (R_X) to the sign which are appropriate to („take account of“) the significate.

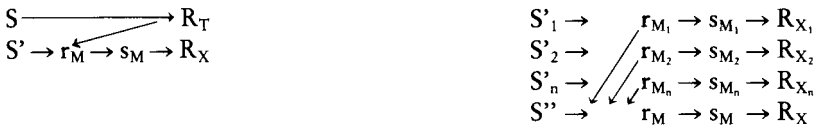


Abb. 1: (Nach Osgood et al. 1957, 7)

a) Entwicklung primärer Zeichen

b) Entwicklung sekundärer Zeichen

Das potentielle Zeichen S' , z.B. ein Wort, löst also nach raum-zeitlicher Verbindung mit einem Ding S einen (reduzierten) Teil des Gesamtverhaltens auf S aus, der - bei alleiniger Darbietung des Zeichens - als repräsentationale Response r_M die Funktion hat, ein Autostimulationsmuster s_M zu vermitteln, das die zeichenspezifischen, dem bezeichneten Sachverhalt S rechnungstragenden, offenen Verhaltensweisen R_X auslöst. Die Bedeutung eines Zeichens ist nach dieser Konzeption durch einen spezifischen repräsentationalen Vermittlungsprozeß bestimmt.

Die Mehrzahl aller Zeichen hat ihre Bedeutung durch Verbindung mit anderen Zeichen und nicht unmittelbar mit den bezeichneten Dingen erhalten (sekundäre Zeichen, vgl. Abb. 1b). Die meisten Sechsjährigen, von denen die wenigsten jemals einem Zebra begegnet sind, verstehen das Wort ‚Zebra‘: sie haben Bilder von Zebras gesehen, gehört, daß Zebras gestreift sind, wie Pferde laufen und gewöhnlich wild leben (Osgood et al. 1957, 8). Das Reizmuster ‚Zebra‘ (S'') erhält Teile derjenigen Mediations-Responses r_M , die bereits von den primären Zeichen ausgelöst werden.

Für die bedeutungsspezifischen r_M gilt, daß sie aus Komponenten (r) bestehen; die Eigenart der r_M wird jeweils durch die spezifische Kombination der r_m bestimmt.

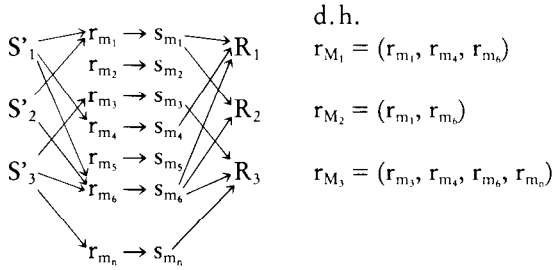


Abb. 2: (Nach Osgood 1971, 12) Komponenten zeichenspezifischer Vermittlungsprozesse.

Eine Diskussion der Osgoodschen Bedeutungskonzeption findet sich bei Fuchs (1975a).

1.1.2 Meßmodell

Um die Bedeutung von Zeichen zu erfassen, müssen der Mediationskonzeption gemäß die repräsentationalen Vermittlungsprozesse ($r_M \rightarrow s_M$) zugänglich gemacht werden. Diese äußern sich, abhängig von der Stärke der Verbindung mit den Autostimulationsprozessen und situativen Bedingungen, in den offenen Verhaltensweisen R_x . Die von Osgood und Mitarbeitern vorgeschlagene Lösung ist die SD-Technik: eine Kollektion von siebenstufigen bipolaren Ratingskalen, deren Merkmale eine repräsentative Stichprobe bedeutungsspezifischen Urteilsverhaltens darstellen und dazu dienen sollen, Zeichen quantitativ zu qualifizieren, z.B. Hans ist (außerordentlich) klug, (sehr) mächtig, (etwas) alt usw.

HANS

dumm	—	:	—	:	—	:	—	:	—	:	—	:	x	klug
mächtig	—	:	x	:	—	:	—	:	—	:	—	:	—	machtlos
jung	—	:	—	:	—	:	—	:	x	:	—	:	—	alt
	-3		-2		-1		0		1		2		3	

1.1.3 Raummodell

In der Analogie einer räumlichen Darstellung von Bedeutung lassen sich Zeichen als Punkte in einem geometrischen Raum lokalisieren. Die Bedeutung eines Zeichens wird durch einen Vektor repräsentiert, der vom Schnittpunkt der Achsen, dem Punkt vollständiger Bedeutungslosigkeit, ausgeht: die Länge des Vektors entspricht der Bedeutsamkeit (Intensität, Sättigung), seine Rich-

tung der ‚semantischen Qualität‘ des Zeichens. Die Bedeutungsähnlichkeit von Zeichen wird durch die Größe der Distanzen zwischen den Punkten abgebildet.

Als Achsen eines derartigen Bedeutungsraumes könnten die SD-Ratingskalen aufgefaßt werden. Da sie einen Raum konstituieren würden, dessen Ordnung beliebig bliebe, wird die empirische Analyse der dimensionalen Struktur zum zentralen Problem der Entwicklung des Bedeutungsraumes.

1.2 Integration der Modelle

Osgood (1971) legt Wert auf die Darlegung, daß diese drei Modelle streng aufeinander bezogen sind.

1. Die Komponenten r_m der repräsentationalen Responses r_M werden mit den empirisch gewonnenen Hauptachsen des Bedeutungsraumes, den Dimensionen E(valuation), P(otency) und A(ctivity) identifiziert. In ihrer Verschiedenheit repräsentieren die Bedeutungs-Komponenten E, P und A nach Osgoods Auffassung solche Aspekte des Verhaltens gegenüber Dingen, die aufgrund unterschiedlicher Anpassungsfunktion differentiell verstärkt werden. In dieser Betrachtungsweise sind Zeichen seit den Zeiten des Neandertalers dadurch bedeutsam, d.h. verhaltensrelevant, daß sie vor allem spezifizieren, in welchem Maße die bezeichneten Dinge ‚gut‘ oder ‚schlecht‘ (E), ‚stark‘ oder ‚schwach‘ (P), ‚aktiv‘ oder ‚passiv‘ (A) sind.
2. Da die offenen Reaktionsweisen, die von den Vermittlungsprozessen repräsentiert werden, nach einem reziprok-antagonistischen Muster organisiert seien, folge, daß die Komponenten r_m in eben dieser Weise funktionierten. Da die offenen Reaktionsweisen im Hinblick auf ihre Intensität variieren, wird auch ihren repräsentationalen Vermittlungsprozessen diese Eigenschaft zugeschrieben. Aufgrund der Unvereinbarkeit gleichzeitiger Tendenzen in Richtung auf antagonistische r_m (also Tendenzen z.B. in Richtung E+ und E-) wird angenommen, daß sich beide, wenn sie auftreten, in Richtung auf Neutralität oder Bedeutungslosigkeit aufheben.

Dem reziprok-antagonistischen Charakter der bedeutungsspezifischen Mediatoren und ihrer Intensitätsvariation wird durch die Verwendung bipolarer, quantitativ abgestufter Beurteilungsskalen Rechnung getragen.

3. Schließlich werden die Punkte im Raum, die die Bedeutung von Zeichen repräsentieren, mit den r_M insgesamt, wie sie durch Zeichen hervorgerufen werden, identifiziert.

Osgood bezeichnet die Beziehung der in Abb. 3 wiedergegebenen Modelle als „isomorph“.

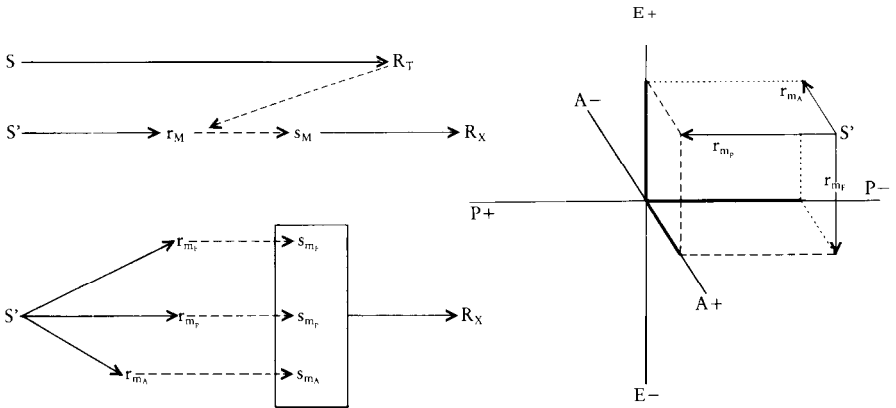


Abb. 3: (Nach Osgood 1962, 11 und 1971, 17) Integration der Modelle („Isomorphismus“)

2. Ordnung von SD-Daten: Architektur eines universellen Bedeutungsraumes

Die Bedeutung eines Zeichens ist - der Konzeption Osgoods gemäß - durch seine Lage im Bedeutungsraum bestimmt. Das zentrale Problem der SD-Forschung betrifft die Identifizierung der Dimensionalität dieses Raumes; die dazu vorliegende Lösung ist die EPA-Struktur. Aufgrund von Belegen ihrer Stabilität gegenüber der Variation von Beurteilern und Konzepten (sprachlichen Zeichen, meist Nomen) und zwar sowohl innerhalb wie zwischen Sprachen/Kulturen gilt sie als ein generelles bzw. universelles Sprachmerkmal. Der anwendungsorientierten Forschung dient sie als ein allgemeines Bezugssystem für quantitative Vergleiche der Bedeutung von Zeichen.

Es verdient an dieser Stelle hervorgehoben zu werden, daß die SD-Technik nur in Verbindung mit dem Modell eines Semantischen Raumes als eine eigenständige Forschungstechnik anzusehen ist. Abgelöst vom Raummodell stellt ein SD lediglich eine Ansammlung von Ratingskalen dar. Allerdings ist die SD-Technik keineswegs an die EPA-Lösung der Struktur von Zeichen-Bedeutung gebunden. Im folgenden sollen zunächst die grundlegenden Befunde zur Geltung dieses semantischen Faktorsystems skizziert, sowie Bedingungen erörtert werden, die das Auftreten der EPA-Struktur beeinflussen.

Wichtigste Erkenntnisgrundlage sind die Forschungsberichte ‚The Measurement of Meaning‘ (Osgood et al. 1957) und ‚Cross-Cultural Universals of Affective Meaning‘ (Osgood et al. 1975), in denen ein beispielloses ehrgeiziges, mehrphasiges semantisches ‚Raumfahrt‘-Programm beschrieben wird, das Osgood und Mitarbeiter zusammen mit ‚Copiloten‘, computerorientiertem ‚Bodenpersonal‘ und ‚Beobachtungsstationen‘ rund

um die Erde durchgeführt haben. Nur die kritische Würdigung der Befunde zur Konzeption des Semantischen Raumes trägt zur Klärung der praktisch bedeutsamen Frage bei, inwieweit das SD als eine wiederverwendungsfähige semantische ‚Raumfähre‘ eingesetzt werden kann und wann die Exploration semantischer Räume spezifischer SDs bedarf.

Die folgende Darstellung geht von den für SD-Erhebungen typischen Daten aus. SD-Urteile werden von Beurteilern auf Skalen für Konzepte abgegeben; die Daten lassen sich also in einer dreimodalen Matrix anordnen (vgl. Abb.4).

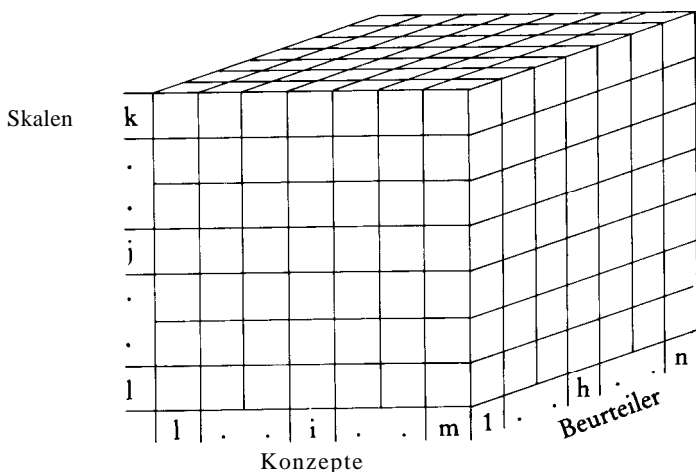


Abb. 4: Dreimodaler SD-Datenkubus

2.1 Skalen-Kovariation: Generalität der EPA-Struktur

2.1.1 Grundlegende Befunde (*The Measurement of Meaning: Osgood et al. 1957*)

In einer ersten Analyse wurden zu 40 Nomen (der Kent-Rosanoff-Liste) bei Studenten Primärassoziationen (Adjektive) gesammelt. Die fünfzig häufigsten wurden zusammen mit ihren Antonymen in der Standardform 100 Studenten zur Beurteilung von 20 (anderen) Konzepten vorgelegt. Die Skalenkorrelationen (über Beurteiler und Konzepte) wurden nach der Centroid-Methode faktorisiert und orthogonal rotiert. Als Ergebnis wurde die klassische EPA-Lösung gefunden: Evaluation (33,8 % ges. Varianz) mit hohen Ladungen auf den Skalen ‚good - bad‘, ‚beautiful - ugly‘, ‚sweet - sour‘; Potency (7,6% ges. Varianz) durch ‚large - small‘, ‚strong - weak‘, ‚heavy - light‘; Activity (6,2% ges. Varianz) mit ‚fast - slow‘, ‚active - passive‘ und ‚hot - cold‘.

Um die Unabhängigkeit dieser Lösung von der Auswahl der Konzepte zu überprüfen, wurden in einer zweiten Analyse alle fünfzig Adjektivpaare einer Stichprobe von 40 Beurteilern aus der gleichen Population paarweise vorgelegt. Ein Adjektiv des ersten Paares diente als ‚Konzept‘; das zu ihm am besten passende Adjektiv des anderen Paares sollte ausgewählt werden. Die Prozentwerte der Wahlhäufigkeiten in einer Skalen x Skalen-Matrix wurden in einer für Rohdaten modifizierten Form der Diagonalmethode nach Thurstone faktorisiert. Die Ähnlichkeitsbeziehungen der Skalen, wie sie sich in dieser Prozentwerte-Matrix ausdrücken, setzen jeweils perfekte Bipolarität voraus - eine Annahme, die angesichts der ad-hoc-Erzeugung der Antonyme und der inzwischen vorliegenden Befunde (vgl. Abschn. 3.1) problematisch erscheint. Die Ergebnisse wiesen befriedigende Übereinstimmung mit denen der ersten Analyse auf; die Autoren konstatieren, daß die EPA-Struktur damit gegenüber einem Wechsel der Beurteilerstichprobe und der Datenerhebungsmethode stabil geblieben ist.

Als Beitrag zur Begründung der Generalität der EPA-Struktur werden in den frühen Arbeiten des Measurement of Meaning die Ergebnisse einer dritten Studie interpretiert. Um die Abhängigkeit dieser Lösung von der vorher getroffenen Skalenauswahl einzuschätzen und mögliche weitere Dimensionen des Semantischen Raumes zu identifizieren, wurden die Beschreibungsmerkmale nun systematisch aus Roget's Thesaurus ausgewählt. Wiederum 100 Studenten beurteilten mit Hilfe von 76 Skalen (Kapazitätsgrenze des Computers) 2.0 Konzepte unterschiedlicher Kategorienzugehörigkeit. Da die unrotierte Centroid-Lösung die EPA-Struktur erkennen ließ, wurde sie durch Pivot-Skalen fixiert; für die Residualvarianz wurden fünf weitere Faktoren tentativ interpretiert. Der erheblich geringere Anteil der durch EPA erklärten Varianz (28 % gegenüber 48 % bzw. 44 % der ges. Varianz), der im wesentlichen zu Lasten des E-Faktors ging, wird von den Autoren auf die Art des Skalensamplings zurückgeführt.

Ergänzende Analysen, die die EPA-Struktur auch für die Beurteilung von nicht-verbalen Konzepten (Ortungssignale, Gemälde) belegen, werden berichtet.

Die frühen Arbeiten deuten auf eine spezifische, durch EPA beschriebene Struktur des Urteilsverhaltens hin, die gegenüber einer gewissen Variation der Datenanalysetechnik (im Bereich der Faktorenanalyse), dem Kriterium der Auswahl ‚repräsentativer‘ Skalen, der Auswahl von Konzepten und der Wahl der Stichprobe von Beurteilern aus einer Population (Studenten) zumindest qualitativ stabil bleibt.

Eine systematische Untersuchung der EPA-Generalität hat die Art des Samplings in allen drei Varianzquellen, sowie die Aufklärung der gesamten, in den Daten enthaltenen Varianz zu berücksichtigen.

2.1.2 Berücksichtigung der verfügbaren Varianz von SD-Daten

Die den drei Modalitäten entsprechenden Varianzquellen sind für die Analyse der Dimensionalität des Semantischen Raumes in unterschiedlicher Weise relevant: Der Semantische Raum wird durch Achsen bestimmt, deren Definition auf der Kovariation zwischen den Skalen basiert. Gesucht wird die geringste Anzahl von Achsen, die Komponenten r_m der r_M des mediationstheoretischen Modells, die das bedeutungsspezifische Verhalten hinreichend beschreiben. Für die Autoren des Measurement of Meaning (Osgood et al. 1957), wie die Autoren von Cross-Cultural Universals of Affective Meaning (Osgood et al. 1975) ist die Faktorenanalyse die Methode der Wahl zur Lösung dieses Problems gewesen und geblieben.

2.1.2.1 Daten-Reduktionstechniken

Die während der frühen Arbeiten verfügbaren faktorenanalytischen Modelle waren nur für zweimodale Datenmatrizen anwendbar. Für die Überführung der dreimodalen Ausgangsmatrix in eine Korrelationsmatrix diskutieren Miron & Osgood (1966) drei Reduktionstechniken:

„stringing out“

Bei dieser, häufig und auch von Osgood et al. (1957) verwendeten Technik werden die Korrelationen zwischen den Skalen für die ‚aufgereihten‘ Markierungen aller Beurteiler bei allen Konzepten berechnet. Miron & Osgood (1966; vgl. auch Osgood et al. 1975, 45f und insbes. Miron 1972, 315ff) bemängeln, daß hier die Struktur der Skalenvariation nicht nur im Hinblick auf die Konzepte, sondern auch gleichzeitig durch Beurteilervarianz bestimmt wird. Bei geringer Konzeptvarianz determiniere diese u.U. entscheidend die Skalenvariation.

Summation

Die Beurteilervarianz wird durch Summierung über die Beurteiler innerhalb der Konzepte reduziert. Die Skalen-Korrelationen werden für die Summenwerte bei den Konzepten bestimmt. Diese Methode wurde im Rahmen des kulturvergleichenden Projekts (Osgood et al. 1975) bevorzugt.

Durchschnittliche Korrelation

Die für alle Beurteiler gesondert berechneten Skalen-Korrelationsmatrizen (für alle Konzepte) werden über die Beurteiler gemittelt.

Während also im ersten Falle die gesamte Beurteiler- und Konzeptvarianz bei der Berechnung der durchschnittlichen Skalenvariationen berücksichtigt wird, wird die Beurteilervarianz in den beiden anderen Verfahren als Fehlervarianz betrachtet und durch Mittelung reduziert. Insbesondere ein der ‚Summation‘ entsprechendes Vorgehen muß schon deshalb empirisch gut begründet werden, weil es - statt einer strengen Prüfung - die generelle (d.h. u.a. auch beurteiler-unabhängige) Geltung der EPA-Struktur als eine Folge der metho-

dischen Manipulation implizieren kann: Beurteilervarianz wird hier gänzlich ignoriert. Dafür, daß sie zu Recht vernachlässigbar ist, werden folgende Argumente vorgebracht (Osgood et al. 1957; Miron & Osgood 1966; Osgood et al. 1975):

- Faktorenanalysen aufgrund von Skalenkorrelationen nach der ‚stringing-out‘ und der Summations-Methode führen (in einem Beispielfall) zu sehr ähnlichen Ergebnissen.
- Faktorenanalysen von Beurteiler x Beurteiler-Matrizen mit den Korrelationen der beurteilerspezifischen Skalenkorrelationen führen (in zwei Referenzfällen: Ware 1958; Tanaka & Osgood 1965) zu Lösungen, die nur einen Beurteilerfaktor bedeutsam erscheinen lassen.
- Die Ergebnisse von seit neuerem vorliegenden dreimodalen Faktorenanalysen lassen erkennen, daß der ‚Löwenanteil‘ der Beurteilervarianz jeweils durch nur einen Faktor erklärt wird.

Die Analyse interindividueller Differenzen wird in Abschn. 2.1.5 noch ausführlicher erörtert.

2.1.2.2 Konzeptvarianz

Während die Autoren des Measurement of Meaning durch Verwendung der stringing-out-Technik Beurteiler- und Konzept-Varianz noch gleichermaßen in die Datenanalyse einbeziehen

(„since our long run purpose was to set up a semantic measuring instrument which would be applicable to people and concepts in general, we wished to obtain that matrix of intercorrelations among scales which would be most representative or typical“, Osgood et al. 1957, 35),

gehen die späteren Arbeiten (insbes. Osgood et al. 1975) davon aus, daß die Beurteilervarianz Vernachlässigbar ist. Demgegenüber erwies sich die Replizierbarkeit der EPA-Struktur nicht nur von einer bevorzugten Berücksichtigung der Konzeptvarianz in den Skaleninterkorrelationen abhängig, sondern vor allem auch von der Organisation der Konzeptvarianz durch Art und Anzahl der zur Beurteilung vorgelegten Konzepte. In einer großen Zahl von Untersuchungen werden von der EPA-Struktur abweichende Lösungen berichtet, wenn weniger umfangreiche oder spezifische Konzeptklassen analysiert werden. Derartige Abweichungen lassen sich selbst dann beobachten, wenn die EPA-repräsentativen Beurteilungsskalen verwendet werden. Wenn, wie man vermuten kann, die Skalenkorrelationen angesichts der meist sehr begrenzten Umfänge der Konzeptstichproben in vielen Fällen nach der stringing-out Methode berechnet wurden, könnte der Einwand von Miron (1972) zutreffend sein, daß in diesen Fällen die Beurteilervarianz eine geringe Kon-

zeptvarianz ‚überschwemmt‘ habe. Allerdings ist dieser Einwand problematisch, weil er die Abweichung durch eine Ursache erklärt, die die Annahme substantiell bedeutungsloser Beurteilervarianz gefährdet.

Gravierender als diese Vermutung erscheint die Kritik, die Carroll schon 1959 gegen die Ergebnisse von Osgood et al. (1957) vorgebracht hat: Im explorativen Zusammenhang muß die verwendete Stichprobe von Konzepten gewährleisten, daß die Skalen entsprechend der gesuchten ‚wahren‘ dimensional Struktur kovariieren. Der gesamte Raum muß entsprechend durch eine angemessene Stichprobe von Konzepten repräsentiert werden; zur Begründung von m unabhängigen Dimensionen fordert Carroll, daß mindestens $2m + 2^m$ Konzeptbeurteilungen zugrundegelegt werden. Dieser Forderung entsprechen die Analysen des Measurement of Meaning offenkundig nicht hinreichend, da die Verwendung von 20 Konzepten danach allenfalls zur Interpretation von vier unabhängigen Faktoren führen kann. Aus diesem Grunde haben Osgood et al. (1975) die Zahl der Konzepte auf 100 erhöht.

Soweit die gesuchte Urteils-Struktur nicht schon als eine inhärente Ordnung der Skalenvariation angesehen und entsprechend - konzeptfrei - begründet wird (vgl. Analyse 2 des Measurement of Meaning), kann sie - anhand von Konzeptbeurteilungen - nur dann identifiziert und repliziert werden, wenn eine nach Umfang und Art hinreichende Stichprobe von Konzepten zugrunde gelegt wird. Lösungen für einzelne Konzepte oder Konzeptklassen lassen Rückschlüsse auf eine generelle Struktur des Urteilsverhaltens nicht zu.

In welcher Weise beschreibt aber eine konzeptunabhängige, generelle EPA-Struktur die Bedeutung der einzelnen Konzepte, wenn deren Bedeutung beim gleichen methodischen Zugriff durch andere als die EPA-Dimensionen bestimmt wird?

Methodisch entspricht diese Differenz jener zwischen gemeinsamer und spezifischer Kovarianz der Skalen für die Konzepte. EPA repräsentiert die allen (verbalen) Zeichen gemeinsamen Bedeutungsaspekte. Die in sie nicht eingehende, ausgesonderte spezifische Urteilsvarianz führt bei der Analyse einzelner Konzepte/Konzeptklassen zu entsprechend spezifischen Lösungen. Der empirische Gehalt der EPA-Konzeption bemißt sich nicht zuletzt danach, in welchem Ausmaß die auf EPA entfallende gemeinsame Varianz im Verhältnis zur konzeptspezifischen an der Varianz des Urteilsverhaltens beteiligt ist.

In der Tradition der SD-Forschung ist allerdings die Frage, wie gemeinsame und spezifische Bedeutung begrifflich zu fassen sind, von wesentlich größerem Interesse gewesen. Die wichtigste, keineswegs völlig entsprechende Unterscheidung ist die zwischen konnotativer und denotativer Bedeutung.

2.1.2.3 EXKURS: Affektive (konnotative) und denotative Bedeutung

Osgood selbst hat immer wieder nachdrücklich betont, daß EPA Bedeutung nicht in erschöpfender Weise repräsentiert. Ihre Eigenart wird begrifflich als konnotative (emotive, metaphorische) von denotativer (designativer, referentieller) Bedeutung abgegrenzt (Osgood et al. 1957, 321). Die Verwendung dieses in Sprachphilosophie, Linguistik und Psychologie vielfältig variierten Begriffsdualismus durch Osgood hat insbes. von linguistischer Seite heftige Kritik erfahren und zwar sowohl im Hinblick auf ihre begriffliche Klarheit wie auch ihren empirischen Gehalt (vgl. Nordenstreng 1969).

Osgood hatte zunächst erwartet, die zeichenspezifischen Vermittlungsprozesse r_m als Dimensionen sensorischer Art, wie z.B. (visuell) Helligkeit, Farbton und -Sättigungsfaktoren, (auditiv) Lautstärke und Tonhöhe, sowie olfaktorische Faktoren identifizieren zu können. Er bekennt 1964, daß diese Erwartung seiner eigenen mediationstheoretischen Konzeption nicht entsprach, in welcher Bedeutung durch response-spezifische Vermittlungsprozesse definiert wird. Der reaktive Charakter der wiederholt beobachteten EPA-Dimensionen ließ Osgood nicht nur die bessere Übereinstimmung mit seiner Theorie erkennen, sondern auch die affektive Natur dieser Bedeutungsprozesse deutlicher hervortreten (vgl. dazu auch Ertel 1964). Als Reaktion auf einen Beitrag von Kuusinen (1969) tendiert er 1969 dazu, die durch philosophische Sprachverwendung belastete Unterscheidung von Konnotation und Denotation zu verwerfen und von ‚affektiven‘ und ‚anderen‘ Bedeutungskomponenten zu sprechen.

Diese neue Kennzeichnung von Bedeutungsarten erfolgt auf empirischer Grundlage: Die beobachteten EPA-Faktoren weisen bemerkenswerte Ähnlichkeit mit den Wundtschen Hauptrichtungen der Gefühle (Lust/Unlust, Erregung/Beruhigung, Lösung/Spannung) und den - allerdings davon nicht unabhängig formulierten - Dimensionen des mimischen Ausdrucks nach Schlosberg (1954) auf. Über diese begriffliche Analogie hinaus legt Osgood (z.B. 1969, 1971) seiner Interpretation der EPA-Bedeutung Beobachtungen der Dominanz und der Universalität der EPA-Dimensionen zugrunde. Insbesondere aufgrund ihres universellen, kulturunabhängigen Auftretens wird die EPA-Struktur mit neurophysiologischen Spekulationen als angeboren interpretiert; der Dominanz entspreche die fundamentale, ursprüngliche Bedeutung von affektiven Reaktionen für das menschliche Verhalten und zwar sowohl im phylo- wie im ontogenetischen Sinne. In funktionalistischer Betrachtungsweise werden die durch EPA charakterisierten Bedeutungsdimensionen als zentrale Modi der Umweltbewältigung aufgefaßt.

Die operationale Trennung von EPA- und anderen Faktoren, sowie der Hinweis auf eine bemerkenswerte Ähnlichkeit der ersteren mit Emotionsdimen-

sionen, kann eine begriffliche Unterscheidung ‚affektiv-konnotativer‘ gegenüber ‚denotativ-anderen‘ Bedeutungsarten nicht ersetzen. Osgood et al. (1975, 393) akzeptieren die Kritik von Nordenstreng (1969; vgl. auch Miron 1969), insoweit er eine klare, nicht-operationale Definition für das verlangt, was mit einem SD erfaßt wird.

Einen älteren Vorschlag Osgoods (1964) aufgreifend, werden ‚konnotativ‘ solche Urteile genannt, in denen die Bedeutung von Zeichen durch Bezug auf ihre (übergeordnete) Klassenzugehörigkeit spezifiziert wird: ein BABY ist klein, und zwar aufgrund des Vergleichs mit Menschen im allgemeinen; als ‚denotativ‘ werden solche Urteile bezeichnet, die relativ zu einem Standard auf der gleichen hierarchischen Ebene abgegeben werden: NINA ist klein, d.h. im Vergleich zu anderen Babies und nicht als Mensch.

Da beim semantischen Differenzieren, wie die Autoren vermuten, Urteile über ein Konzept relativ zu einem übergeordneten Konzept abgegeben werden (s.o.), würden denotative Urteilsweisen zurückgedrängt. Leider begründen die Autoren die angemessene Verwendung dieser linguistischen Unterscheidung im Hinblick auf das SD-Urteilsverhalten nicht. An anderer Stelle (Osgood et al. 1975, 400) wird das Hervortreten der affektiv-konnotativen EPA-Struktur weiter dadurch erklärt, daß durch die Beurteilung der verschiedensten Konzepte mit den gleichen Beurteilungsmerkmalen eine metaphorische Sprachverwendung begünstigt wird (vgl. Osgood 1969, 1971).

Angesichts der begrifflichen Unschärfe der SD-relevanten Bedeutungsarten beschränken sich operationale Bestimmungen affektiver und denotativer Bedeutung bei homogenen Konzeptklassen auf die Trennung von EPA-spezifischen Bedeutungsanteilen als ‚Affekt‘ von der übrigen, interpretierbaren Skalen-Kovarianz als ‚Denotation‘. Kuusinen (1969) identifiziert affektive und denotative Strukturen von Persönlichkeitsbeurteilungen, indem er aus den Interkorrelationen der Skalenmittelwerte die Korrelationen mit den EPA-Markierskalen herauspartialisiert und sowohl die ursprüngliche, wie die Partialmatrix faktorisiert. Das als denotativ bezeichnete Restsystem ist aber nach diesem Verfahren nicht ein von den EPA-Affekt-Dimensionen gesäubertes System, sondern es ist die gesamte Kovarianz der einzelnen EPA-Markierskalen eliminiert. Tzeng (1975, 1977; Tzeng & May 1975) hat deshalb vorgeschlagen, die Trennung durch Konstruktion von orthogonalen Subräumen für Affekt und Denotation im Bedeutungsraum vorzunehmen: Transformations-Matrizen der Rotation von Markierskalen zur (reinen) EPA-Lösung und zur Einfachstruktur werden dabei auch zur Rotation von Ladungsmustern für die übrigen Skalen verwendet; EPA- und ‚sonstige‘ Bedeutungsanteile der einzelnen Skalen sind in den jeweiligen Subräumen feststellbar. Im Verhältnis zur Elaboriertheit des Verfahrens tritt die bereits konstatierte begriffliche Vagheit besonders deutlich hervor: als Kriterium der Unterscheidung von affektiver und

denotativer Bedeutung fungiert Ähnlichkeit bzw. Nicht-Ähnlichkeit mit der EPA-Struktur.

Die Bemühungen um eine Präzisierung der SD-Bedeutungsprozesse haben bislang einschlägige Beiträge der ‚imagery‘-Forschung kaum berücksichtigt. Paivio (1969; 1971), ihr bedeutendster Vertreter, faßt ‚imagery‘ als ein non-verbales assoziatives Vermittlungssystem zwischen verbalen Reizen und Responses auf, das neben und unabhängig von einem (teilweise mit ihm verbundenen) verbalen, assoziativen System existiert. Beide Kodierungs-Systeme leisten die kognitive Repräsentation unserer Welt. Die Verfügbarkeit des imaginalen Kodierungs-Systems hängt nach Paivio entscheidend davon ab, in welchem Ausmaß ein verbaler Reiz konkret oder abstrakt ist. d.h. sich auf sinnlich erfahrbare Merkmale von Sachverhalten bezieht:

„The hypothesis is that concrete terms such as „house“ derive their meaning through association with concrete objects and events as well as through association with other words, and thereby acquire the capacity to evoke both nonverbal images and verbal processes as associative (meaning) reactions, which could function as alternative coding systems affecting mediation and memory. Abstract terms such as „truth“, on the other hand, derive their meaning largely through intraverbal experiences and more effectively arouse verbal associative than imaginal processes“ (Paivio 1969, 248).

Denotative Bedeutung bezeichnet die Verknüpfung von Vorstellungsbild (Image) und sprachlichem Zeichen. (Mit Bezug auf Staats, 1968, wird denotative Bedeutung als der auf ein sprachliches Zeichen konditionierbare Teil der sensorischen ‚Reaktion‘ auf das (bezeichnete) Objekt aufgefaßt.)

Insofern besteht zwischen der Dimension abstrakt/konkret und der denotativen Bedeutung von Zeichen ein enger Zusammenhang: abstrakte Wörter, d.h. solche ohne bezeichnete Objekte im Unterschied zu konkreten Wörtern rufen keine Vorstellungsbilder hervor und weisen mithin keine denotative Bedeutung auf.

Godfrey & Natalicio (1970) haben im Anschluß an Paivio die Abstraktheit/Konkretheits-Dimension, deren Rolle von Osgood et al. (1975) mehrfach gering veranschlagt wird (z.B. p. 401, vgl. allerdings p. 187), auf die Evaluationsdimension bezogen und ihren Beitrag durch den Titel ‚Evaluation on SD equals abstraction plus error‘ gekennzeichnet. Diesem Befund liegen z.T. sehr hohe Korrelationskoeffizienten für den Zusammenhang von Urteilsvarianz auf E-Skalen relativ zu NonE-Skalen mit Abstraktheits/Konkretheits-Rangordnungen zugrunde: Bei abstrakten Konzepten tritt mehr E gegenüber NonE als bei konkreten Konzepten auf und vice versa. Lohr (1976) hat gezeigt, daß imaginal-denotative Bedeutung (sensorischer Art) und evaluative Bedeutung (emotional sensu Osgood) als distinkte Vermittlungsprozesse voneinander unabhängig und konkurrierend konditionierbar sind.

KostiE & Das (1971) haben versucht, die Art der mit dem SD erfaßten Bedeutung begrifflich durch Ausschluß von nicht erfaßten Bedeutungsaspekten zu präzisieren:

- Die durch die EPA-Faktoren definierte Bedeutung ist nicht-lexikalischer Art, obwohl die Differentialurteile auf lexikalischen Bedeutungen basieren. Osgood selbst hat wiederholt Beispiele für EPA-bedeutungsgleiche aber lexikalisch verschieden bedeutsame Konzepte gegeben.
- Die mit dem SD erfaßte Bedeutung ist begrenzt auf verbal abstrahierbare Bedeutungsaspekte und zwar solche, die vielen Konzepten gemeinsam sind. Idiosynkratische, spezielle und einzigartige Bedeutungsaspekte werden nicht berücksichtigt.
- Die Definition von Bedeutung durch EPA ist eine allgemeine, ohne Einschränkung im Hinblick auf die individuelle oder spezifische Situation. Im sprachlichen Kommunikationsverhalten wird die Bedeutung von Konzepten durch Adjektive spezifiziert; durch Eingrenzung oder Reduktion vom allgemeinen auf das weniger allgemeine erfüllen Adjektive ihre kommunikative Funktion. EPA leistet eine solche Bedeutungsspezifizierung nicht.

2.1.3 Variationen des Modus der Dimensionsanalyse

Bereits Osgood et al. (1957, 31f) haben die Frage gestellt, ob die dominante EPA-Lösung außer von der Art des samplings in den drei Varianzquellen - Personen, Skalen, Konzepte - auch von der Faktorisierungsmethode abhängig ist. Sie vergleichen die Ergebnisse von Centroid- und Rohwertanalysen und konstatieren befriedigende Übereinstimmung (p. 42ff). Orlik (1965; 1967) bemängelt, daß in Rohwertanalysen dieser Art artifizielle Varianz zwischen den Skalen eingeht und in der Regel einen zusätzlichen, Pseudo-Faktor konstituiert. Für mittelwertszentrierte Rohwert-Produktsummen (Kovarianzen) wird die Möglichkeit erwartungstreuer Abbildung psychologischer Merkmalsräume belegt (vgl. dazu auch Revenstorff, 1973a).

Die Ergebnisse des kulturvergleichenden Forschungsprojekts (Osgood et al. 1975), mit dem die Generalität/Universalität der EPA-Struktur begründet wird, basieren auf Hauptkomponenten-Analysen, die zu einer Lösung führen, die der nach der Centroid-Methode entspricht (Ertel 1965 b; vgl. Harman 1970, 174). Die EPA-Struktur erweist sich als relativ stabil, wenn über die Urteiler gemittelte SD-Ratings faktorisiert werden (vgl. Heise 1969, 415) - ‚repräsentative‘ Stichproben von Konzepten und Skalen vorausgesetzt. Auf die Berücksichtigung interindividueller Differenzen wird noch besonders eingegangen.

Die Validität der EPA-Struktur des semantischen Raumes kann nicht überzeugend durch Variationen der faktorenanalytischen Technik belegt werden, wenn das Urteilsverhalten in der Form von SD-Ratings konstant gehalten wird. Osgood et al. (1957, 143ff) berichten über erste Versuche, die EPA-

Faktoren als die zentralen Dimensionen des bedeutungsspezifischen Urteilsverhaltens über Ähnlichkeitsurteile für Konzepte zu validieren. Anderson (1970) hat diesen Gedanken aufgegriffen und zwölf Adjektive (jeweils zwei für EPA charakteristische Paare) im Paarvergleich auf Ähnlichkeit beurteilen lassen. Aufgrund der stress-Werte der MDS nach Kruskal erwies sich die Dreidimensionalität auch des Ähnlichkeitsraumes als begründet, bei der allerdings - wie auch gelegentlich für SD-Faktoren berichtet wird - A und P zusammenfallen. Diesen, von Osgood als ‚Dynamism‘ bezeichneten komplexen Faktor fand auch Arnold (1971) aufgrund von Kruskal-MDS-Analysen, wobei jeweils vier Dimensionen (nicht-euklidischer Metrik) der Unähnlichkeit für eine Nomen-, eine Adjektiv- und eine Verbliste angemessen erschienen. Diese Unähnlichkeits-Dimensionen wiesen zwar substantielle Korrelationen mit den in Hinblick auf E, P und A auch unidimensional skalierten Listen auf, ohne daß jedoch eine dimensionale Korrespondenz auffindbar wurde. Im Unterschied zu den Ergebnissen von Anderson (1970) basieren die Unähnlichkeitsräume bei Arnold (1971) allerdings auf Wortlisten, deren Geeignetheit zur Reproduktion von EPA zweifelhaft ist. Aber auch angesichts weiterer vorliegender Befunde zur Korrespondenz von SD und MDS-Lösungen für Ähnlichkeits- oder Präferenzräume (Nordenstreng 1968; Green et al. 1969; Magnusson & Ekman 1970; Everett 1973; Shikiar et al. 1974; Gärling 1976) kann eine Entsprechung - zumal im Hinblick auf EPA - nicht hinreichend klar festgestellt werden.

2.1.4 Transkulturelle *Stabilität*

Nachdem sich die EPA-Struktur gegenüber Variationen der Skalen-, Konzept- und Personen-Stichproben - soweit sie hinreichend repräsentativ waren - resistent erwiesen hatte und bereits einige Befunde vorlagen, denen zufolge sie auch über Sprachen/Kulturen Geltung zu haben schien (Kumata & Schramm 1956; Kumata 1957; Triandis & Osgood 1958; Suci 1960) initiierte Osgood 1959 ein gigantisches Forschungsprojekt, bei dem die Generalität der EPA-Struktur über Kulturen und Sprachen überprüft werden sollte. Im Unterschied zu den vorliegenden Befunden wurde auf die Verwendung übersetzungsäquivalenter Beurteilungsmerkmale verzichtet, um sprachlich-kulturellen Eigenarten verschiedener semantischer Systeme Rechnung zu tragen.

Insgesamt gingen Daten aus 25 Sprach/Kultur-Gemeinschaften in diese Untersuchung ein; trotz eines Übergewichtes indo-europäischer Sprachen war versucht worden, linguistische und kulturelle Differenzen zu maximieren.

In einem mehrphasigen Auswahlprozeß wurden 100 Substantive als Beurteilungsgegenstände gesammelt, die ein hohes Maß kultureller Allgemeinheit gewährleisten sollten. Diese zunächst amerikanisch-englische Liste wurde über-

setzt und die Nomen wurden an allen Erhebungsorten jeweils 100 Schülern/Studenten zur Charakterisierung durch jeweils ein Adjektiv vorgelegt. In regulärer grammatikalischer und orthografischer (ggf. transkribierter) Form wurden diese Nennungen von Osgood und Mitarbeiter in Illinois ‚blind‘ im Hinblick auf ‚productivity‘ (Aufretenshäufigkeit und Verteilung über die Konzepte, vgl. Abschn. 4.1), sowie ‚Unabhängigkeit‘ von anderen Beurteilungsmerkmalen geordnet. Auf diese Weise wurden für jede Sprach/Kultur-Gemeinschaft die für sie relevanten Beurteilungsmerkmale ausgewählt. Die örtlichen Forschungsgruppen erhielten eine bis zu 70 Adjektive umfassende Liste zurück, die gemäß diesen Kriterien die höchsten Rangplätze einnahmen. Mit Hilfe von 10 unabhängigen Experten wurden zu ihnen Antonyme erhoben und die danach verbleibenden 50 Adjektivpaare (vorläufig) um 10 Paare ergänzt, die - aufgrund der automatischen Selektion ausgesondert - am Erhebungsort für wichtig gehalten wurden. In einem weiteren Schritt wurden sodann von 200 neu-rekrutierten Vpn (ebenfalls männlichen Schülern/Studenten) alle 100 Konzepte mit diesen Merkmalspaaren in der Form von Rating-Skalen beurteilt. Für die Skalen-Interkorrelationen wurden in Illinois sowohl für Sprache/Kultur spezifische, wie auch pankulturelle Faktorenanalysen gerechnet. In eindrucksvoller Weise wird in nahezu allen Analysen, insbes. der pankulturellen, gezeigt, daß EPA nicht nur durchgängig als dominierende Dreier-Struktur auftritt, sondern ihre semantische Ähnlichkeit zwischen Sprachen/Kulturen teilweise bis in übersetzungsäquivalente Beurteilungsmerkmale reicht. Im theoretischen Zusammenhang ist damit die zentrale Hypothese bekräftigt: „... regardless of language or culture, human beings utilize the same qualifying (descriptive) framework in allocating the affective meanings of concepts“ (Osgood et al. 1975, 6). Dieser Sachverhalt ermöglicht nach Meinung der Autoren für alle Sprachen/Kulturen SDS zu entwickeln, die Unterschiede in subjektiven Kulturen vergleichbar machen, wenn ihre Items EPA repräsentieren.

2.1.5 Interindividuelle Unterschiede

Osgood et al. (1975, 364) stellen zu Recht fest, daß die Osgoodsche Bedeutungskonzeption nicht im Widerspruch zur Möglichkeit individueller Unterschiede bei Bedeutungssystemen steht, die auf unterschiedlichen Erfahrungen beim Lernen von Zeichen oder in Unterschieden hinsichtlich Emotionalität, Intelligenz usw. basieren könnten. Die Universalität der EPA-Struktur ist insoweit nicht theoretisch begründet: Die Annahme ihrer Geltung über Personen geht vielmehr auf Beobachtungen zurück, denen zufolge EPA bei Stichproben verschiedener Personen-Kategorien auftritt, so z.B. im Hinblick auf Alter, Geschlecht, Intelligenz, politische Orientierung, Normalität (vgl. die Übersicht bei Osgood et al. 1975, 58ff; Rosenbaum et al. 1971). Die Befunde zur Generalität über sehr unterschiedliche kulturelle Gruppen haben - auch wenn dabei jeweils nur Stichproben von männlichen Schülern/Studenten be-

rücksichtigt wurden - die Vorstellung bekräftigt, daß alle Menschen ein gemeinsames affektives semantisches Bezugssystem teilen.

In einigen Arbeiten (Williams 1972; Denmark et al. 1972) werden zwar Einschränkungen aufgrund sozio-ökonomischer Klassifizierung von Urteilern nahegelegt; es fällt jedoch schwer, die gefundenen Mängel an Übereinstimmung zwischen den sozialen Gruppen zu interpretieren, da sie gering sind oder durch die Auswahl der beurteilten Konzepte begründet sein können. Hinweise auf kulturspezifische Variationen der Dimensionslösungen werden von Tanaka & Osgood (1965) und Tanaka et al. (1963) bei übersetzungsäquivalenten Skalen in der Weise interpretiert, daß bestimmte Beurteilungsskalen entweder im Hinblick auf Konzepte und/oder Personen faktoriell instabil sind (Konzept-Skalen- und Person-Skalen-Interaktionseffekte).

Den Vergleichen inter- und intrakultureller Stichproben von Personen liegt in der Regel der von Osgood und Mitarbeitern bevorzugte Aggregierungsmodus der dreimodalen Ausgangsdaten 'Summation' d.h. Mittelwertbildung über Personen bei den Konzepten zugrunde.

Wiggins & Fishbein (1969) bezweifeln, daß Befunde wie die hier angeführten geeignet sind, die Frage nach der universellen Geltung der EPA-Struktur hinreichend zu beantworten. Wenn diese Frage nicht nur die Bedeutung habe, ob es ein gemeinsames semantisches Bezugssystem gebe, das die intraindividuelle Struktur einer 'gemittelten pankulturellen Person' widerspiegeln, sondern auch, ob dieses Bezugssystem repräsentativ für (intrakulturelle und) individuelle Strukturen innerhalb der Kultur sei, dann müßte eine auf gemittelten Maßen basierende Struktur auch die beste Repräsentation der semantischen Struktur von Individuen innerhalb einer Kultur sein.

Osgood et al. (1975) rechtfertigen die Verwendung von Gruppenmittelwerten nicht nur mit dem Hinweis, daß für ihre Untersuchungen geeignete dreimodale Analyseverfahren nicht verfügbar waren, sondern sie vertreten weiterhin die Meinung, daß diese Datenreduktion im Hinblick auf interindividuelle Varianz angemessen ist. Erkenntnisgrundlagen dafür sind:

- Faktorlösungen für Datenmatrizen nach Summation- und stringing-out-Reduktion zeigen keine bemerkenswerten Unterschiede. (Bei der stringing-out-Prozedur gehen die einzelnen Personen ein, allerdings ist auch hier Person- und Konzept-Varianz konfundiert.)
- Skaleninterkorrelationen über die Konzepte für jede einzelne Person korrelieren untereinander so hoch, daß die Matrix der Korrelationen über die korrespondierenden Zellen (Personen-Matrix) zu Faktorlösungen führt, die als einfaktoriell anzusehen sind (Ware 1958; Tanaka & Osgood 1965).
- Die Eindimensionalität der Person-Varianz erscheint durch die Ergebnisse

neuerdings vorliegender dreimodaler Faktorenanalysen bekräftigt (Levin 1965; Litt 1966; Snyder 1967; Tzeng 1975).

Gegenüber dem zuletzt genannten Punkt verweisen Wiggins & Fishbein (1969) darauf, daß gerade die Ergebnisse der dreimodalen Faktorenanalyse von Levin die Möglichkeit zur Interpretation mehrerer Person-Faktoren aufweisen. Sie selbst führen eine Tucker-Messick-MDS von Ähnlichkeitsurteilen für 15 charakteristische EPA-Skalen durch und erhalten 3 Person-Faktoren. Während sich die Dimensionalität der Skalen beim ersten Person-Faktor („group average space“) bemerkenswert klar durch EPA reproduzieren ließ, ergab die Einlagerung der Personen in den dreidimensionalen Person-Raum kreisförmige Arrangements um den zweiten und dritten Faktor, deren Repräsentation durch 10 kegelförmig angeordnete Vektoren (Idealisierte Personen) angemessen erschien. Die Ähnlichkeitsurteile dieser 10 homogenen Personen-Gruppen ergaben jeweils Faktorenlösungen („viewpoints“), die zwischen 2 und 4 Skalen-Dimensionen (mit jeweils mehr als 90% Anteil an der gesamten Varianz) aufwiesen. Dabei war nicht nur der Beitrag der einzelnen Skalen zur Definition der semantischen Dimensionen unterschiedlich, sondern es konnte auch beobachtet werden, daß die Dimensionszugehörigkeit der Skalen vom Gruppendurchschnitt zu den idealisierten Personen und zwischen den idealisierten Personen keineswegs stabil war. Dabei ist zu berücksichtigen, daß nur Skalen mit ‚klarer‘ EPA-Indikatorfunktion verwendet wurden.

Bei personspezifischen Faktorenlösungen (Meßwiederholungen über Konzepte) war Ertel (1965 b) zwar zum Ergebnis einer personunabhängigen stabilen EPA-Struktur mit invarianter faktorieller Struktur der Skalen gelangt. Allerdings lagen diesen Analysen Erhebungen bei nur vier Personen zugrunde. Crockett & Nidorf (1967) fanden demgegenüber bei zwölf Vpn eine EPA-Lösung nur für fünf Personen möglich, für die übrigen waren zweidimensionale Lösungen angemessen. Keine Vp zeigte die erwartete Gruppierung der EPA-Skalen zu separaten EPA-Faktoren. Lediglich der E-Faktor war bei allen Personen klar identifizierbar. Hinweise auf interindividuelle Differenzen im Hinblick auf die Struktur des semantischen Raumes finden sich auch bei den Ähnlichkeitsdaten Andersons (1970) und Q-Faktorenanalysen für einzelne Urteilkonzepte durch Revenstorff (1973 a).

Nun wird man die Feststellung von interindividuellen Differenzen im Prinzip als trivial zur Kenntnis nehmen können, solange ihr Ausmaß nicht spezifiziert ist. Der varianzanalytisch feststellbare geringe Varianzanteil zu Lasten der Beurteiler relativ zu dem aufgrund der Konzepte und Skalen (Fuchs 1973; Revenstorff 1973 a, Schäfer 1975 a) läßt keine hinreichenden Rückschlüsse auf die Stabilität der EPA-Struktur über Personen zu. Auch die Interaktions-Varianzen explizieren nicht den strukturellen Aspekt interindividueller Differenzen. Die aufgrund dreimodaler Faktorenanalysen im Hinblick auf die Personen beobachteten mehrdimensionalen Lösungen weisen, soweit Eigenwert-

verlauf und/oder Varianzanteile mitgeteilt werden, für den ersten (Gruppen-durchschnitts-)Faktor den Löwenanteil erklärter Personenvarianz aus. Sie lasen aber kaum mehr erkennen, als daß die EPA-Struktur nicht die beste Repräsentation aller individuellen semantischen Strukturen darstellt (Tzeng 1975, 1977; Snyder & Wiggins 1970; Muthen et al. 1977). Shikiar et al. (1974) konnten zwar die Befunde von Wiggins & Fishbein (1969) bekräftigen, die Nützlichkeit der Berücksichtigung interindividueller Differenzen erwies sich für die Vorhersage politischer Präferenzen allerdings als gering. Idiosynkratische- und Gruppendurchschnitts-E-Maße korrelierten nicht nur hoch untereinander, für keine von fünf idealisierten Individuen war die Korrelation des idiosynkratischen E-Maßes mit einem Maß der Wahlpräferenz höher als die zwischen Durchschnitts-E-Maß und Wahlpräferenz. Selbst wenn man die Voraussetzung eines perfekten Zusammenhangs zwischen evaluativen und Präferenz-Urteilen akzeptiert, trägt auch dieser Befund nur sehr vorläufig zur Würdigung individueller Variation im Hinblick auf die Bedeutungsstruktur bei: Interindividuelle Differenzen sind nur im Hinblick auf die E-Dimension berücksichtigt, und die E-Maße konnten nur im Rahmen von SD-Skalen, die aufgrund von ‚Durchschnitts-Analysen‘ für EPA charakteristisch sind, ‚idiosynkratisch‘ sein.

Osgood et al. (1975, 346) konstatieren, daß die affektiven semantischen Systeme von Individuen innerhalb von Kulturen keineswegs völlig homogen sind. Darüberhinaus kann angesichts der vorliegenden Befunde als gesichert gelten, daß EPA weder kulturspezifisch, noch ein Artefakt der Durchschnittsbildung über Personen ist.

2.2 Interaktionsvarianz: Konzept-Skalen-Interaktion

Bedenken gegen die Annahme einer generellen EPA-Struktur sind seit den grundlegenden Arbeiten mit Beobachtungen begründet worden, über die bereits Osgood et al. (1957) berichten: ‚the meanings of scales and their relations to other scales vary considerably with the concept being judged‘ (p. 187). Während eine Person-Skalen- und/oder eine Person(gruppen)-Konzept-Skalen-Interaktion seltener berichtet wird (Krieger 1963; Tanaka et al. 1963; Nordenstreng 1970; Snyder & Wiggins 1970) - was Wunder angesichts der üblichen Durchschnittsbildung der Urteilsmaße über Personen -, ist die Konzept-Skalen-Interaktion ein bevorzugtes Thema der Kritik an der SD-Technik. Allerdings sind die Schlußfolgerungen aus der Analyse der Konzept-Skalen-Interaktion höchst unterschiedlich: sie reichen von der Auffassung, es handle sich um ein ‚Scheinproblem‘ oder methodisches Artefakt (z.B. Ertel 1965a; Kahneman 1963) bis zum Vorschlag, das SD als Datenerhebungstechnik von seiner bedeutungstheoretischen Grundlage abzulösen (z.B. Darnell 1970). Aufgrund des dominierenden Interesses an der Frage der dimensionalen Struk-

tur des semantischen Differenzierens basieren die Interpretationen dieser konzeptspezifischen Einflüsse - wie auch schon für Personeneinflüsse (interindividuelle Differenzen) festgestellt - in aller Regel auf den Ergebnissen von Dimensionsanalysen und selten auf einer varianzanalytischen Erkenntnisgrundlage. Das Ausmaß dieser ‚Anomalie‘ ist deshalb kaum quantitativ präzisierbar und auch nicht im Verhältnis zur Zufallsvariation testbar.

Konzept-Skalen-Interaktionseffekte manifestieren sich in einer erheblichen Variation der Korrelationen zwischen Skalen bei verschiedenen Konzepten (vgl. z.B. Presly 1969; Burns 1976). Entsprechend werden für die einzelnen Konzepte und Klassen von Konzepten unterschiedliche Ladungsmuster der Skalen, einschließlich verschiedener Faktorenstrukturen berichtet (z.B. Osgood et al. 1957; Osgood 1962, Tanaka et al. 1963; Tanaka & Osgood 1965; Darnell 1966; Kubiniec & Farr 1971; Bynner & Romney 1972; Heskin et al. 1973; Klemnack & Ballweg 1973, Burns 1976).

Es entspricht unterschiedlichen Fragestellungen der SD-Forschung, wenn das Auftreten von Konzept-Skalen-Interaktionseffekten einerseits in seiner Relevanz für die Geltung einer generellen EPA-Struktur und andererseits in bezug auf seine Konsequenzen für die Entwicklung eines generell verwendbaren Instruments zur Bedeutungsdifferenzierung gesehen wird.

Osgood und Mitarbeiter haben den Sachverhalt, daß Konzept-Skalen-Interaktionen (im Unterschied zu Person-Skalen-Interaktionen) auftreten, nicht negiert, sondern betont und durch konzeptspezifische Bedeutungsverschiebungen von Skalen zu erklären versucht: In Übereinstimmung mit dem Kongruenzprinzip tendierten im menschlichen Urteilsprozeß alle Skalen zu Bedeutungsverschiebungen in Richtung auf Parallelismus mit dem dominanten Bedeutungsattribut des Urteiskonzeptes. Evaluative Skalen seien für derartige Verschiebungen in besonderer Weise anfällig, SD-Skalen generell als Funktion der ‚evaluativness‘ von Konzepten. Osgood et al. (1957, 188 u. 326f) vertreten deshalb - was häufig übersehen wird - die Auffassung, man werde zur Erfassung der Bedeutung von Konzepten solche Skalen verwenden müssen, die EPA in konzept(klassen)-spezifischer Weise repräsentierten. Osgood et al. (1975, 351) stellen fest, daß Bedeutungsmaße für bestimmte Konzepte zwischen verschiedenen Kulturen aufgrund von Konzept-Skalen-Interaktionen „cum grano salis“ zu interpretieren seien.

Ertel (1965 a) hat das Auftreten von Konzept-Skalen-Interaktionen einer methodologischen Kritik unterzogen. Er geht davon aus, daß konzeptspezifische Korrelationen keine hinreichende Begründung für Schlußfolgerungen auf eine zugrundeliegende allgemeine Dimensionsstruktur liefern können. Vielmehr könne sich die dimensionale Unabhängigkeit der Urteilsfaktoren nur durch gezielte Variation der Urteilsgegenstände nach dem Kriterium ‚dimensionaler Repräsentativität‘ der Konzeptstichprobe erweisen. Bei einer derartigen Stich-

proben-Organisation würden begriffsspezifische Korrelationen ‚intersituativ‘ verschwinden. Die konzeptspezifische Kovariation der Skalen wird danach verursacht durch irrelevante und störende Bedeutungsaspekte von Konzepten und/oder Skalen. Ähnlich äußert sich Revenstorff (1973a), der das Interaktionsproblem für trivial hält, da man durch eine nicht-repräsentative Auswahl von empirischen Objekten in jeder Korrelationsellipse Ausschnitte wählen kann, die die Korrelation beliebig variieren lassen.

Nun wird man zwar den frühen Arbeiten von Osgood und Mitarbeitern vorhalten können, daß die Anzahl der berücksichtigten Konzepte zu gering war und die Konzepte auch nicht dimensional repräsentativ ausgewählt waren (vgl. die Kritik Carrolls 1959). Immerhin haben sie zu sehr ähnlichen Ergebnissen geführt wie die Ertelschen Arbeiten, in denen die Bedeutungsstruktur aufgrund eines Prozesses sukzessiver Approximation der Auffindung dimensional reiner Skalen und Konzepte als eine mit Erregung, Valenz und Potenz bezeichnete Konfiguration resultierte. Die Forderung nach einer breiteren Urteilsbasis, auch i.S. eines ‚Wechsels der Situation‘, haben Osgood et al. (1975) im Cross-Cultural Projekt erfüllt. Konzept-Skalen-Interaktionseffekte lassen sich dennoch in der beschriebenen Weise beobachten. Ertel selbst hat die Annahme ihres Verschwindens bei dimensional-repräsentativer Konzept- und Skalenauswahl empirisch nicht überprüft. Die Forderung nach repräsentativer Konzeptauswahl ist im übrigen für die meisten Anwendungsfälle der SD-Technik nicht vertretbar.

Während der Zugang Ertels eher das Problem der Identifizierung der Struktur des semantischen Differenzierens und der ihr zugeordneten Skalen betrifft als die Lösung des Problems der Konzept-Skalen-Interaktion, sind andere methodisch begründete Argumente vorgetragen worden, die diese Effekte als ein Artefakt erscheinen lassen. Kahneman (1963) hält die psychologische Erklärung, insbesondere die Annahme konzeptspezifischer Bedeutungsverschiebungen der Skalen für unnötig. Er geht davon aus, daß jedes Rating s_{ijk} bei Konzept j durch Person i auf Skala k als Summe dreier Komponenten aufgefaßt werden kann: dem ‚wahren‘ Wert (Mittelwert der Personen-Population) auf Skala k bei Konzept j , der konstanten Abweichung der Person i auf Skala k und einer spezifischen Abweichung von Person i auf Skala k bei Konzept j , die ihrerseits eine konsistent-idiosynkratische und eine Fehlerkomponente aufweisen. Alle diese Komponenten sind nach Kahnemans Befund korreliert, einschließlich der ‚Fehler‘-Abweichungen, und weisen eine ähnliche Struktur auf. Kahneman kann die in seinen Daten (allerdings in nur geringem Ausmaß) beobachteten Konzept-Skalen-Interaktionseffekte im wesentlichen durch die idiosynkratische Komponente der spezifischen Abweichung in Form einer konstanten Überschätzungs/Unterschätzungs-Tendenz des wahren Wertes erklären: unterschiedliche Korrelationen (Vorzeichen) zwischen Skalen bei einzelnen Konzepten können infolge dieser interindividuellen Unterschiede er-

wartet werden, je nachdem, ob die wahren Werte auf der gleichen oder verschiedenen Seite des Skalenneutralpunktes liegen. Unterschiedliche Skalen-Korrelationen (der wahren Werte) bei verschiedenen Konzeptklassen seien demgegenüber nicht auf kognitive Strukturen der Urteiler, sondern auf reale ‚ökologische‘ Beziehungen zwischen den Konzepten zurückzuführen. Auch für diese Art der Konzept-Skalen-Interaktion wird die Annahme von Bedeutungsverschiebungen der Skalen zurückgewiesen.

Heise (1969) betont, daß Beobachtungen von Konzept-Skalen-Interaktionen in der Regel bei Analysen auf der Basis von Personwerten und nicht von Gruppenmitteln berichtet werden. Daß hierbei person-spezifische Varianz die diskutierte Interaktion zu beeinflussen scheint, wird durch die Arbeit von Snyder & Wiggins (1970) nahegelegt: die Autoren weisen aufgrund einer dreimodalen Faktorenanalyse nicht nur auf eine „interaktionale Beziehung zwischen idealisierten Personen, Konzepten und Skalen“ (p. 466) hin, sie charakterisieren die idealisierten Personen durch Extremisierungstendenzen, die allerdings abhängig von Urteilsdimensionen und Skalen erscheinen. Nordenstreng (1969) hebt hervor, daß die Kahnemansche Unterscheidung des kognitiven und des ökologischen Systems immerhin deutlich macht, daß sowohl Personen wie Konzepte zu Konzept-Skalen-Interaktionseffekten beitragen. Die Feststellung einer psychometrischen Beziehung könne ihre konzeptuelle Interpretation aber nicht ersetzen. Er unterscheidet vier Typen (A-D):

„. . . different correlations between corresponding scales indicate in type **A** the extent to which a set of concepts is related differently to two individuals, in type **B** the extent to which an individual is related differently to two sets of concepts, in type **C** the extent to which a concept is related differently to two sets of individuals, and in type **D** the extent to which a set of individuals is related differently to two concepts“ (p. 13).

Heise (1969) spricht von ‚wahrer‘ Konzept-Skalen-Interaktion in Abgrenzung von methodisch artifizieller, wie sie durch unangemessene Konzeptauswahl, Verwendung irrelevanter Skalen und Polarisierungsfehler bedingt sei. Wahre Interaktionseffekte könnten durch unterschiedliche Relevanz der Skalen für verschiedene Konzepte und durch Bedeutungsverschiebungen in den Skalen im Hinblick auf Klassen von Konzepten zustandekommen.

Das Merkmal ‚süß‘ - ‚sauer‘ mag hoch relevant zur Beurteilung von (bestimmten) Nahrungsmitteln, mäßig relevant zur Beurteilung von Mitmenschen und kaum relevant zur Beurteilung abstrakter Ideen sein. Entsprechend wird bedeutsame Urteilsvarianz auf dieser Skala geringer werden. Insoweit die Kovariation mit anderen Skalen dadurch beeinflußt wird, variieren - konzeptspezifisch - die Ladungsmuster von Faktorenlösungen.

Eng verbunden mit der konzeptspezifisch (und mutmaßlich auch person-spezifisch) differentiellen Relevanz von SD-Skalen scheint die auf Bedeutungsarten bezogene Erklärung von Verschiebungen der Skalen-‚Bedeutungen‘ durch

Osgood (1962) zu sein. Durch die Integration des affektiv-energetischen und des sensorisch-motorischen Diskriminierungssystems mit den ihnen zugeordneten konnotativen bzw. denotativen Bedeutungsreaktionen in gleichen Systemen des Sprachverhaltens trete in den auf die Erfassung von affektiver (konnotativer) Bedeutung gerichteten Skalen ‚denotative Kontamination‘ auf. Einzelne Konzepte schränken die Skalenbedeutungen in selektiver Weise ein: so werde durch das Konzept LAVA für die Skala ‚heiß - kalt‘ deren Denotation hervorgerufen, während deren (activity) Konnotation durch Konzepte wie JAZZ und FESTIVAL betroffen werde.

Welche Folgerungen lassen sich aufgrund der vorliegenden Befunde aus dem Sachverhalt der Konzept-Skalen-Interaktion im Hinblick auf die Frage nach der Generalität der EPA-Struktur einerseits und der generellen Verwendbarkeit eines SD-Instruments andererseits ziehen?

Zunächst operieren derartige konzept- wie auch personspezifischen Einflüsse gegen die Begründung eines generellen/universellen affektiven Bedeutungssystems. Ihre Auswirkungen können aber angesichts der vorliegenden Befunde (vgl. auch die Ergebnisse der ‚konzeptfreien‘ Skalenstruktur-Analysen bei Osgood et al. 1957; Wiggins & Fishbein 1969) die Feststellung einer stabilen EPA-Kernstruktur des semantischen Differenzierens nicht erschüttern; abgesehen von Meßfehleranteilen verweisen sie vielmehr auf spezifische Bedeutungsaspekte, die außerhalb des Bedeutungshorizonts der EPA-Konfiguration für Konzepte (und Personen) charakteristisch sind. Offenkundig verfügt keine der von Osgood und Mitarbeitern im Cross-Cultural-Projekt berücksichtigten Sprachen/Kulturen über „reine“, konzeptunabhängige EPA-Indikatoren. Die Beurteilungsskalen transportieren jeweils Bedeutungsaspekte über die durch EPA definierten hinaus und zwar in interindividuell und zwischen Konzepten varianter Weise.

Es erscheint gerechtfertigt, diese spezifischen Abweichungen bei der Suche nach einer *allgemeinen* Struktur im Kognitionsverhalten als Fehlerkomponente aufzufassen. Sie scheinen jedoch von hinreichend systematischer Art zu sein, um *spezifische* Bedeutungsstrukturen identifizierbar zu machen (vgl. Kuusinen 1969; Tzeng 1975, 1977; Tzeng & May 1975).

Die Frage nach den Konsequenzen der Konzept-Skalen-Interaktion für Konstruktion und Anwendung eines SD außerhalb des Kontextes der Suche nach einer allgemeinen Bedeutungsstruktur ist weniger eindeutig zu beantworten. Einerseits ist der Wert der SD-Technik wesentlich von der Bezugsmöglichkeit auf die allgemeinen Bedeutungsdimensionen EPA abhängig. Andererseits kann kaum bezweifelt werden, daß konzept- und personspezifische Einflüsse die Bedeutungsdimensionierung in unbestimmter Weise verzerren, wenn sog. generelle Skalen verwendet werden; die dimensionale Repräsentativität auch von ‚typischen‘ EPA-Skalen variiert in Abhängigkeit von diesen Einflüssen.

Aus der Generalität der EPA-Struktur folgt allerdings nicht, daß diese Dimensionen für alle Konzepte und Personen(gruppen) durch die gleichen Skalen optimal repräsentiert werden. Im Gegenteil zeigen die Interaktionseffekte, daß dies nicht der Fall ist. Die sprach/kulturspezifischen Lösungen des Cross-Cultural-Projekts zeigen vielmehr, daß EPA (bei Sprach/Kulturgruppen) durchaus auch im Gewande verschiedener Skalensets auftritt. Eine Lösung des Dilemmas könnte also darin bestehen, für eine bestimmte homogene Klasse von Konzepten und eine bestimmte Population von Personen spezifische EPA-Indikatoren zu suchen und zu verwenden.

Die üblicherweise zur Behandlung von Konzept-Skalen-Interaktionseffekten vorgeschlagenen ‚Lösungsmöglichkeiten‘ betreffen die Gewichtung der einzelnen Skalen für die Berechnung von Person-Maßen für die einzelnen Dimensionen (Faktoren-Scores), die - häufig als Bedeutungs- und Einstellungsmaße verwendet - in besonderer Weise von instabilen Faktorstrukturen und Ladungsmustern beeinträchtigt werden. Presly (1969) hält mit Bezug auf Konzept-Skalen-Interaktionseffekte eine Gewichtung von Skalen-Werten nur auf der Grundlage konzeptspezifischer Faktorenladungen für vertretbar. Bynner & Romney (1972) wollen wenigstens die damit (aufgrund unterschiedlicher Faktorenlösungen) preisgegebene Vergleichbarkeit der Faktoren-Scores in einem gemeinsamen Faktoren-System retten und empfehlen Faktorenanalysen sowohl für die einzelnen wie über alle Konzepte. Soweit Faktoren der Analyse über die Konzepte auch in den konzeptspezifischen Analysen auftreten, sollten für diese Faktoren-Scores nach Maßgabe der Gewichte der Analyse über die Konzepte berechnet werden. Levy (1972) hebt hervor, daß die Identifizierung gemeinsamer und konzeptspezifischer Faktoren erhebliche Probleme aufwerfe und schlägt andere Lösungen vor: Faktorisierung der (Konzepte/Skalen) x (Konzepte/Skalen) Korrelationsmatrix (vgl. auch Kubiniec & Farr 1971; Klemmack & Ballweg 1973); oder: Projektion der Faktorenstrukturen der untersuchten Konzeptklasse in einen Bedeutungsraum für Standard-Konzepte, etwa den EPA-Raum; oder: Verwendung einer dreimodalen Faktorenanalyse-Prozedur. Zur Berücksichtigung von Variation in Mittelwerten und Standardabweichungen der Skalen über die Konzepte wird die Analyse von Kreuzprodukt-Rohwerten und Kovarianzen empfohlen.

Datenanalyseverfahren können Konzept-Skalen-Interaktionseffekte deutlich und lokalisierbar machen. Sie können dazu beitragen, der Konzept-Skalen-Interaktion als einer Anomalie Rechnung zu tragen. Sie können aber weder das Problem störender konzept- und personspezifischer Einflüsse „lösen“, noch ihr Auftreten beeinflussen. Letzteres kann allerdings durch angemessene Auswahl von Skalen für bestimmte Konzepte (bei bestimmten Personen) erreicht werden (vgl. Abschn. 4).

2.3 „Fehlervarianz“

2.3.1 Systematische Urteilsfehler

2.3.1.1 Extremisierung

Bei der Inspektion von SD-Datenlisten sind in der Regel bei einzelnen Personen auftretende Bevorzugungen bestimmter Skalenpositionen, und zwar der Mittel- und der Extrempositionen auffällig. Derartige Urteilstendenzen sind für Rating-Skalen häufig beschrieben worden (z.B. Peabody 1962; Hamilton 1968). Wir beschränken uns hier auf einige Beobachtungen und Erklärungen dieses Sachverhalts bei der Analyse von SD-Daten. Dabei steht die ‚Extremisierungstendenz‘ im Vordergrund. Die Bevorzugung der mittleren Antwortkategorie scheint nach den vorliegenden Befunden (Mitsos 1961; Orlik 1965; Mikula & Schulter 1970; Grimm et al. 1973) von der Auswahl der Skalen für die Konzeptbeurteilungen abzuhängen.

Herrmann (1962) berichtet über zwei voneinander unabhängige Tendenzen der Präferenz von Skalenstufen: „Urteilsnuanciertheit“ als Wahl der Stufen 1 und 7 gegenüber 3 und 5 und „Fraktionierung des Bezugssystems“ als Wahl von 4 gegenüber 2 und 6 bei der siebenstufigen Skala. Diese Faktoren des semantischen Differenzierens werden als genuine Aspekte von Bedeutungsurteilen und nicht als Fehlerkomponenten aufgefaßt. Die Diskussion Herrmanns legt eher eine Abhängigkeit von der Art der beurteilten Konzepte als eine differentiell-psychologische Interpretation nahe. Demgegenüber resümiert Peabody (1962, 73):

‚Individual differences in average extremeness show wide generality across different extremeness scores. This generality extends to responses in opposing directions - suggesting that the differences primarily represent response sets, and only to a secondary degree actual differences in intensity.‘

Wenn man Gruppenmittel als ‚wahre Werte‘ betrachtet, lassen sich - wie Kahneman (1963) gezeigt hat - die individuellen Abweichungen bei SD-Daten als Über-/Unterschätzungstendenzen interpretieren. Osgood et al. (1957, 226 ff) diskutieren sehr ausführlich Befunde, derartige Urteilstendenzen als persönlichkeitspezifische Stile aufzufassen. Die in neueren Untersuchungen berichteten Beziehungen zu Merkmalen wie Alter und Geschlecht (Maltz 1963; Light et al. 1965; Long et al. 1968; Schludermann & Schludermann 1969; Washington 1975), IQ (Neuringer 1963; Brod et al. 1964; Light et al. 1965; Stricker & Zax 1966; Long et al. 1968), Ambiguitätstoleranz (Mogar 1960; Rydell 1966; Chen 1977), psychischen Störungen (Neuringer 1963; Zax et al. 1964; Marks 1965; Arthur 1966; Priest 1971) sind allerdings uneinheitlich. Washington (1975) stellt fest, daß aufgrund unterschiedlicher methodischer Ansätze die Vergleichbarkeit der Befunde im Hinblick auf verschiedene Kon-

zepte und Skalen kaum gegeben sei. Er empfiehlt, entsprechende Analysen für einzelne Konzepte und dimensionsgleiche Skalen (auch) getrennt durchzuführen. Zwar hat Arthur (1966) bei einer studentischen Vpn-Stichprobe hohe Korrelationen für Häufigkeiten extremer Markierungen über ein Vierwochen-Intervall und über verschiedene Konzepte gefunden. Die Stabilität der Extremisierungstendenz scheint nach dem Befund von Snyder & Wiggins (1970) aber sowohl im Hinblick auf Skalen als auch Konzepte eingeschränkt: die Differenzierung zweier Personentypen (als Ergebnis einer dreimodalen Faktorenanalyse) basierte auf Überbewertungen bei den E-Skalen durch die erste idealisierte Person, während der zweite Personentyp durch ähnliche Tendenzen bei den P-A-Skalen charakterisiert war. Allerdings waren diese allgemeinen Tendenzen von Konzept zu Konzept verschieden. Long et al. (1968) zufolge verwenden Schüler bei Selbst-Ratings für E- und P-Skalen unterschiedlich stark Extremkategorien, und zwar Mädchen stärker bei E und Jungen stärker bei P.

Personen(gruppen) unterscheiden sich im Ausmaß der Verwendung extremer Skalenpositionen. Obwohl diese Tendenz systematisch zu sein scheint, lassen sich ihr keine spezifischen Persondispositionen zuordnen. Sie scheint von Konzept-Skalenauswahl abhängig zu sein. Heise (1969) schlägt vor, ihr (in Experimenten) durch randomisierte Zuordnung der Vpn oder durch statistische Kontrolle Rechnung zu tragen. Johnson & Wall (1969) kritisieren die Anwendung von ‚Ausfilterungs-‘ und anderen statistischen Korrekturtechniken: Wenn die Extremheit eines Urteils zwar nicht eindeutig als Intensitätsmaß aufgefaßt werden kann, andererseits aber nicht unabhängig von der Bedeutung eines beurteilten Konzeptes variiert, wird durch derartige Techniken die Bedeutungseinschätzung eines Konzeptes mitverändert.

2.3.1.2 Soziale Erwünschtheit

Die Kontamination von Urteilsverschiebung und Bedeutungsaspekten wird in einer systematischen Urteilstendenz deutlich, die empirisch nicht immer ohne nähere Analyse von der Extremisierungstendenz zu unterscheiden ist (vgl. die zitierte Untersuchung von Long et al. 1968) und als Interpretationsalternative bei Bedeutungsmaßen die Validität einer SD-Bedeutungsinterpretation in Frage stellen kann: Die Extremheit der Urteile hängt auch davon ab, wie sehr die Konzepte und Skalen sozial normierte Beurteilungssachverhalte und -gesichtspunkte betreffen. Sozial positiv bewertete Wörter weisen höher polarisierte Bedeutungsmaße auf als negative, wobei Polarisierung und Evaluation konfundiert erscheinen (Howe 1965). Nickels & Shaw (1964) stellten fest, daß die Korrelation zwischen E-Faktoren-Scores (als Einstellungsmaßen) und Maßen von Thurstone-Skalen je nach dem Grad der Bedeutsamkeit (saliency) der beurteilten Konzepte für Beurteiler variierte. Personen seien sozialen Normie-

rungen des Urteilsverhaltens bei salienten Konzepten stärker ausgesetzt und die wirkten sich beim transparenteren Einstellungsmaß, dem SD, stärker aus.

Krieger (1963) hat die soziale Erwünschtheit des Urteilsmediums ‚Skalen‘ kontrolliert; sie fand für balancierte Skalen geringere Beiträge zur E-Dimension. Ford & Meisels (1965) beziffern das Ausmaß der Korrelation zwischen Maßen sozialer Erwünschtheit von SD-Adjektivpaaren und Ladungen auf der E-Dimension mit .88 bis .92, bei Ladungen auf P und A dagegen mit nur .13 bis .22. Nun belegen derartige Befunde lediglich, daß für die Bedeutungsdifferenzierung auf der E-Dimension (im Unterschied zu P-A) Merkmale konstitutiv sind, die Sachverhalte nicht unwesentlich durch die unterschiedliche Zuordnung (sozial) erwünschter Eigenschaften charakterisieren. Diese empirische Korrespondenz impliziert noch keineswegs die Feststellung systematischer Urteilsverzerrungen, wie Heise anzunehmen scheint. (Revenstorff (1971, 196) hält gar - konsequenterweise - das Vorkommen eines evaluativen Faktors als nachteilig für den Vergleich von Persönlichkeitsbeschreibungen.) Als Tendenzfehler ist die (differentielle) Abhängigkeit des Urteilsverhaltens von der sozialen Erwünschtheit der Konzepte und/oder der Skalen zu belegen.

Meisels & Ford (1969) haben in einer weiteren Arbeit personspezifische Urteilstendenzen i. S. sozialer Erwünschtheit nachzuweisen versucht. Die berichteten Korrelationskoeffizienten zwischen verschiedenen derartigen Tendenzmaßen und Maßen für EPA bestätigen die Unabhängigkeit von P und A und lassen - entgegen der Interpretation der Autoren - allenfalls eine schwache Beziehung zu den E-Maßen erkennen.

Wenngleich ungewöhnliche personspezifische Tendenzen, Urteile gemäß ihrer sozialen Erwünschtheit abzugeben, für die SD-Technik nicht belegt sind, wird man deshalb eine entsprechende Fehlerquelle nicht negieren können (vgl. dazu Voyce & Jackson 1977). Insbesondere bedarf der Zusammenhang von sozialer Erwünschtheit mit der E-Dimension einer näheren Begründung.

Für die Erklärung eines möglichen Tendenzfehlers der beschriebenen Art könnte der Bezug auf die in der Tradition der Einstellungsforschung und der Urteilstheorien untersuchten Polarisierungseffekte nützlich sein. Beurteiler, für die ein Urteilsgegenstand ausgeprägten Wertbezug aufweist, tendieren dazu, die (Un)Günstigkeit von Feststellungen über diesen Sachverhalt polarisierter zu beurteilen (Hovland & Sherif 1952; Zavalloni & Cook 1965; Eiser & Stroebe 1972; Eiser 1971a, 1971 b; Eiser & Mower White 1974). Jones (1969) hat gezeigt, daß die Bedeutung des am meisten akzeptierten Beurteilungsstatements erheblich stärker durch die E-Komponente charakterisiert ist als die Bedeutung des am meisten abgelehnten - ein Hinweis, daß positiver Wertbezug sich wie social desirability und evtl. auch saliency durch Polarisierung auf der E-Dimension auswirkt. Im Unterschied dazu stellte Jones signifikant höhere Ratings bei den abgelehnten Statements auf der P-Dimension fest.

Auch die Theorien zur Wahrnehmungs- und sozialen Akzentuierung (vgl. Irle 1975; Lilli 1975), so die Reizklassifikationstheorie Tajfels (1959, 1975), betonen die Polarisierung von Urteilen in Abhängigkeit von einem Wertbezug.

2.3.2 Zufallsfehler - Reliabilität von SD-Urteilen

Osgood et al. (1957, 126ff) haben mit großer Sorgfalt die Reproduzierbarkeit der wichtigsten Bedeutungsmaße aufgrund von Meßwiederholungen untersucht. (Wegen der z.T. geringen Streuungen der Urteilsmaße über die Personen halten sie die Korrelationstechnik *zur* Reliabilitätseinschätzung für ungeeignet.) Sie unterscheiden die Aspekte der Reproduzierbarkeit auf den Ebenen der Skaleneinstufungen und der Faktoren-Scores, sowie der von letzteren abhängigen Konzept-Bedeutungsmaße im semantischen Raum. Die berichteten 5%-Signifikanzgrenzen für Differenzen zwischen je 2 Erhebungen liegen für die einzelnen Personen und Skalen (7 Stufen) bei mehr als 2 Skaleneinheiten, für Faktoren-Scores einzelner Personen zwischen 1 und 1,5 Skaleneinheiten (je nach Faktor; E am stabilsten) und für gruppenspezifische Faktoren-Scores (Gruppenmittel) bei etwa 0,5 Skaleneinheiten. Die mittleren absoluten Abweichungen zwischen 2 Erhebungen bei den verschiedenen Skalen liegen im Durchschnitt bei etwa $3/4$ Skaleneinheiten (vgl. Osgood et al. 1957; Piaggio 1968).

Da der Spielraum für Diskrepanzen zwischen den Meßwiederholungen von der Extremität der ersten Messung abhängig ist, bezieht Norman (1959) die beobachteten Diskrepanzen auf die maximal möglichen. Die auch schon von Osgood et al. berichteten geringen Konsistenzen der einzelnen Ratings lassen sich nach Norman durch Faktoren-Scores leicht verbessern, wenn mindestens drei Skalen zugrunde gelegt werden. (Dieser Befund basiert allerdings auf Daten, in denen die Skalenzahl nicht unabhängig von der Art der Bedeutungsdimensionen variiert wurde. Im übrigen läßt sich der Meßfehleranteil natürlich durch ‚Test‘verlängerung reduzieren (Piaggio 1969). Bei dimensional repräsentativ ausgewählten Skalen sind allerdings in der Regel nur wenige geeignete Indikatoren verfügbar; auch dürfte der Gewinn gering sein.) Während die D(istanz)-Maße für einzelne Skalen und Personen zwischen Konzepten extrem geringe Stabilität aufwiesen, korrelierten D-Maße, gemittelt über die Personen, sowie gemittelte Skalenmarkierungen zwischen den beiden Messungen hoch. Die Stabilität von Ratings ist nach Norman sowohl bei verschiedenen Konzepten wie bei verschiedenen Personen unterschiedlich.

DiVesta & Dick (1966) haben eine umfangreiche Untersuchung zur Reliabilität von SD-Daten bei Schulkindern durchgeführt, mit zwei Messungen in unmittelbarem zeitlichem Zusammenhang bzw. im Abstand von etwa vier Wochen. Unter der Bedingung des 4-Wochen-Intervalls betrug die durchschnittlichen Korrelationskoeffizienten für die einzelnen Skalen über die Schulklassen .27

bis .56 und für einzelnen Klassen aufsteigend von .33 (in der 2. Klasse) bis .55 (7. Klasse). Die Höhe der Koeffizienten stieg bei Addition von jeweils zwei dimensionsgleichen Skalen-Werten zu Faktoren-Scores, nunmehr zwischen 35 und .86 variierend, wobei ab der 4. Klasse eine deutliche Steigerung zu beobachten war. Korrelationen für Maße der Bedeutungs sättigung im semantischen Raum lagen etwa im Bereich von .50 bis .70. Für diese insgesamt unbefriedigenden Ergebnisse werden von den Autoren Bedeutungsänderungen der Konzepte zwischen den Meßzeitpunkten geltend gemacht: Die Koeffizienten waren für Erhebungen mit einem Zeitintervall von einem bis zwei Tagen erheblich höher. Für die untersuchten Klassenstufen 3, 5 und 7 betrug die durchschnittlichen Korrelationen über die Skalen .56, .56 bzw. .67, für die einzelnen Skalen .42 bis .77. Faktoren-Scores für Personen korrelierten bei den EPA-Faktoren zwischen .62 und .84 (am höchsten für E) und für Konzepte gemittelt über Personen zwischen .73 und .94. Entsprechend erhöht sind auch die Korrelationen für die Bedeutungs-Distanzmaße. Die Autoren bewerten diese Ergebnisse als Beleg für akzeptable Stabilität von SD-Daten, eine Würdigung, die allenfalls angesichts des Alters der untersuchten Personen akzeptabel erscheinen mag.

Eher befriedigende (Alpha)Koeffizienten werden von Oles (1973) berichtet: sie liegen für jeweils sechs E-Skalen über neun Konzepte bei Dritt- bis Fünftkläßlern bei .86 bis .92 und summiert über die Skalen innerhalb der Konzepte bei .54 bis .72. Als Stabilität über sieben Monate sind - angesichts der untersuchten Altersgruppen immerhin noch - Koeffizienten von 35 bis .44 für Konzeptsummen bei bedeutungsstabilen Konzepten ausgewiesen.

Miron (1961) betont zu Recht, daß Einschätzungen von Eigenschaften eines SD dem Umstand Rechnung tragen müssen, daß ein SD nicht als ein spezifischer Test, sondern als eine Technik anzusehen sei. Zu den variablen Merkmalen dieser Technik gehören auch die Instruktionsbedingungen. Er variiert systematisch die Faktoren Schnelligkeit der Bearbeitung und Rekapitulation der Markierungen des ersten Durchgangs (jeweils in zwei Stufen) im unmittelbar darauffolgenden Retest. Faktoren-Scores über drei bis fünf Skalen für EPA-Faktoren, gemittelt über die Vpn, korrelieren über 20 Konzepte unter allen Bedingungen mit mindestens .97. Eine Varianzanalyse der Test-Retestabweichungen für die gemittelten Konzept-Scores weist einen signifikanten Haupteffekt zu Lasten der Rückruf-Bedingung: die Vpn konnten, wenn sie entsprechend instruiert wurden, Markierungen erinnern. Dieser Befund läßt die Bevorzugung der unmittelbar erfolgten Retest-Messung bei der Interpretation der Reliabilität von SD-Messungen durch DiVesta & Dick (1966) als problematisch erscheinen und belegt die Berechtigung der von Gulliksen (1958) bereits formulierten Forderung, Parallelversionen zu verwenden; in dieser Hinsicht ist die Arbeit von Coyne & Holzman (1966) eine Ausnahme geblieben.

Abweichungen der Skalen von klarer Bipolarität (fragwürdige Gegensätzlichkeit der polaren Adjektive) und Konzept-Skalen-Kombinationen, die das Auftreten von Interaktions-Effekten begünstigen, scheinen das SD-Urteilsverhalten kaum zu beeinflussen. Vidali (1976) fand unter diesen Bedingungen keine nennenswerten Unterschiede bei Reliabilitätsmaßen für einzelne Rater (um .50) und für Gruppen von Urteilern (um .97). Maruyama (1971) zeigte, daß die mittlere Antwortkategorie der Skalen Reliabilitäts- und Stabilitätsmaße beeinflussen kann: Die von ihm berechneten Koeffizienten waren höher, wenn o-stufige Skalen (ohne Mittelpunkt) verwendet wurden, im Vergleich zu 7-stufigen Skalen.

Auf Reliabilität im Sinne hoher Stabilität und Reproduzierbarkeit der Faktorladungen für die charakteristischen EPA-Skalen, und zwar für Korrelationen individueller wie über Personen gemittelter Ratings bei 25 Konzepten und einer repräsentativen amerikanischen Stichprobe, verweist Tzeng (1975). Daß hohe Skalen-Homogenität, als Reliabilitätsmaß ausgedrückt durch einen Generalisierbarkeits-Koeffizienten (vgl. Gleser et al. 1965), erreichbar ist, belegen die konzept- und personspezifisch konstruierten SDs von Fuchs (1973) und Schäfer (1975a, 10).

Es ist Fuchs (1975, 84f) zuzustimmen, der die Ergebnisse der einschlägigen Arbeiten folgendermaßen zusammenfaßt.

- (1) „Geht man von den einzelnen Beurteilern aus, ist die Reliabilität unter allen Aspekten - obwohl deutlich besser als nach dem Zufall zu erwarten wäre - nicht zufriedenstellend. Geht man dagegen von Gruppenmittelwerten aus, erhält man - verglichen mit anderen subjektiven Tests - sehr zufriedenstellende Reliabilitätswerte.
- (2) Es gibt deutliche Unterschiede zwischen Beurteilern (z.B. urteilen ältere Kinder konsistenter als jüngere), Skalen (z.B. werden Skalen der Bewertungsdimension konsistenter verwendet als andere) und Konzepten (z.B. werden ‚objektive Konzepte‘ konsistenter eingeschätzt als ‚subjektive‘, ‚nicht-neutrale‘ konsistenter als ‚neutrale‘).
- (3) Dimensionswerte, d.h. über alle auf einen bestimmten Faktor hoch und möglichst rein ladenden Skalen - evtl. unter Berücksichtigung der Ladungshöhe - gebildete Konzeptmeßwerte, sind zuverlässiger als Item-Einstufungen . . .
- (4) Die Reliabilität nimmt offensichtlich ab mit dem zeitlichen Abstand zwischen Test und Retest.“

3. *Metrische Eigenschaften von SD-Skalen: 'Statik' des semantischen Raumes*

Es ist bereits erwähnt worden, daß die SD-Technik meßtheoretisch zunächst keinen anderen Status beanspruchen kann als andere Rating-Verfahren: Es handelt sich um Messungen ‚per fiat‘, numerische Indizierung ohne Lösung des Repräsentationsproblems. Im Hinblick auf die Vertretbarkeit von Transformationen im Prozeß der Analyse von SD-Daten (einen Überblick über die gängigen SD-spezifischen Techniken geben Diehl & Schäfer 1975) sind bereits von Osgood et al. (1957) einige metrische Eigenschaften der Skalen diskutiert worden, die für den Architekten eines semantischen Raumes als Probleme der Statik gelten können. Zu den grundlegenden metrischen Annahmen, die als Merkmale der Skalen Eigenschaften des semantischen Raumes betreffen, gehören Bipolarität, Intervallgleichheit und Nullpunktlage gemäß der numerischen Kodierung.

3.1 Bipolarität

Mit Bezug auf die Charakterisierung der bedeutungsspezifischen Vermittlungsreaktionen als reziprok-antagonistisch ist für die Bedeutungsdimensionen und die sie konstituierenden Skalen zu fordern, daß ihre Pole Gegensätze auf eindimensionalen Kontinua repräsentieren.

Kjeldergaard & Higa (1962) konnten zeigen, daß das Wiedererkennen von Wörtern durch den Grad ihrer Polarisierung im semantischen Raum begünstigt wird. Aufgrund von Analysen der Enkodierung und Speicherung von Wörtern im Kurzzeitgedächtnis sowie des Reproduzierens lassen sich die Pole der EPA-Dimensionen als verschiedene Klassen der Kodierung von Wörtern auffassen (Markel et al. 1966; Wickens & Clark 1968; Wickens 1970; Kroes & Libby 1971). Turvey et al. (1969) und Turvey & Fertig (1970) konnten zeigen, daß die Unähnlichkeit von Wörtern zwischen und die Ähnlichkeit in diesen Klassen auf die Polarität der EPA-Dimensionen zurückgeführt werden kann. Nach Befunden von Haygood (1966) und Taylor & Haygood (1968) wurden semantische Konzepte um so schneller gemäß den EPA-Dimensionen kategorisiert, je stärker die Kategorien polarisiert waren.

Derartige Befunde stimmen zwar mit der Annahme überein, daß die beim semantischen Differenzieren verwendeten Dimensionen als bipolar zu konzipieren sind. Sie erübrigen aber nicht die Prüfung, ob, in welchem Ausmaß und unter welchen Bedingungen Bipolarität als ein Merkmal von SD-Skalen und -Dimensionen gilt.

Aufgrund von Analysen der Assoziationen für die häufigsten englischen und die Standard-SD-Adjektive kommt Deese (1964) zu dem Ergebnis, daß das

Schema polarer Gegensätze - wenn auch nicht für alle Adjektive - durchaus im Sprachverhalten begründet ist. Die Ergebnisse weisen die SD-Polaritäten überwiegend als linguistische Kontraste aus. Carter et al. (1969), die ihre Vpn auch aufforderten, das Antonym zur nur einseitig markierten Skala einzusetzen, fanden für die 15 am höchsten ladenden SD-Skalen in der Mehrzahl die von Osgood et al. (1957) verwendeten wieder, in anderen Fällen aber bemerkenswerte Abweichungen. ROSS & Levy (1960) bezweifeln, daß Adjektive polarer Anordnung im Hinblick auf ihre semantische Eindeutigkeit gleich und entgegengesetzt sind (vgl. Terwilliger 1962). Im Unterschied zur nominalen Antonymie fordert Mordkoff (1963, 1965), daß polare Gegensatzpaare auch funktional antonym sein müßten, derart, daß sich SD-Beurteilungen der ‚bipolaren‘ Adjektive (Konzepte) als symmetrische und äquidistante Profile zum Nullpunkt - zumindest für die einzelnen Skalen, wenn schon nicht über die Skalen hinweg - darstellen ließen. Die mittels Hotellings T^2 geprüften Abweichungen von dieser Bedingung erwiesen sich in einer Vielzahl von Fällen als signifikant, nominell antonyme Adjektivpaare also keineswegs gesichert auch als funktional gegensätzlich. Bei systematischer Variation der konzeptspezifischen Information bilanzierten Malmstrom & French (1963) im Hinblick auf die Symmetrie von evaluativen SD-Skalen günstiger: Die Polarität der Urteile korrespondierte eng mit der Polarität der auf E-Skalen gegebenen Konzept-Information.

Auch Andersons (1970) Ähnlichkeits-Ratings von 12 EPA-Adjektiven bestätigen - trotz bemerkenswerter Variation bei Polaritäten und Personen - im wesentlichen die Angemessenheit der Bipolaritäts-Annahme, und zwar sowohl im Hinblick auf die Größe der Distanzen zwischen bipolaren Adjektiven, wie auch auf die Äquidistanz vom Ursprung.

Green & Goldfried (1965) argumentieren, daß die vorgegebene bipolare Etikettierung der Rating-Skalen die Überprüfung einer dementsprechenden Bipolarität der Urteilsdimensionen nicht erlaubt. Die Bipolarität der EPA-Struktur sei durch die Anordnung der Adjektivskalen erzwungen. Sie selbst präsentierten ihren Vpn die Antonyme jeweils einzeln (unipolar) zur Beurteilung von Konzepten und konstatierten erhebliche Abweichungen von funktionaler Bipolarität bei EPA-Skalen, die sich in Null- oder positiven Korrelationen zwischen ‚gegensätzlichen‘ Adjektiven äußern. Demgegenüber hatte Ertel (1964) aufgrund einer ähnlichen Vorgehensweise der Korrelierung von Daten unipolarer Skalen diejenigen als faktoriell gegensätzlich gefunden, die bereits intuitiv als Gegensätze angesehen worden waren. Nach Green & Goldfried (1965) sind alle Tendenzen von gegensätzlichen Adjektiven, Pole einer einzigen Skala zu bilden, konzept- oder konzeptklassenabhängig. Ausgehend von der größeren Anfälligkeit von unipolaren Rating-Skalen für Zustimmungstendenzen berichtet Bentler (1969) für Adjektiv-Zuordnungen zu den Polen der EPA-Dimensionen ähnliche Korrelationskoeffizienten wie Green & Goldfried

(1965), die sich jedoch dramatisch zur Übereinstimmung mit der Bipolaritäts-Annahme verändern, wenn Zustimmungstendenz als die Gesamtzahl der Adjektive, die über alle sechs Pole verwendet wurden, kontrolliert wurde. Bipolarität von Skalen, die die Bipolarität der EPA-Dimensionen repräsentieren sollen, haben Atwood & Falkenberg (1971) für unipolare SD-Ratings nach Auspartialisierung von Zustimmungstendenz als die Summe aller Skalenmarkierungen über alle Konzepte für jeden Beurteiler weniger deutlich und allenfalls für konkrete Konzepte (Makkaroni, Akkordeon, Automobile) im Unterschied zu abstrakten (Glück, Notwendigkeit, Theorie) gefunden.

Gegenüber Erklärungen der beobachtbaren Abweichungen von der funktionalen Bipolaritätsannahme (i.S. von Skalensymmetrie) durch Person- oder Konzept-Merkmale hat Gilpin (1973) lexikalische Markierungseffekte, eine strukturelle Asymmetrie der Skalen, als Bedingung verantwortlich gemacht. Er bezieht sich auf das Prinzip lexikalischer Markierung (Clark 1969). Danach werden solche Adjektive als „unmarkiert“ bezeichnet, die in zweifacher Weise verwendet werden: ‚nominal‘, insofern sie das gesamte Kontinuum einer bipolaren Skala bezeichnen (z.B. ‚günstig‘ das Günstigkeitskontinuum von extrem günstig bis extrem ungünstig) und ‚kontrastiv‘, insofern sie eine Spezifizierung im Hinblick auf einen Standard oder Kontrast implizieren (z.B. günstig im Unterschied zu ungünstig). Markierte Adjektive weisen demgegenüber nur eine, nämlich kontrastive Bedeutung auf (z.B. ungünstig). Auf die Frage wie gut, interessant oder wichtig z.B. ein neues Produkt ist, wird die Antwort einer nominalen oder kontrastiven Verwendung dieser Adjektive entsprechen; der Befragte ist auf die kontrastive Bedeutung festgelegt, wenn die Frage lautet wie schlecht, uninteressant oder unwichtig das Produkt ist. Differenzen der Urteilsmaße auf unmarkierten und markierten unipolaren Skalen zu denen auf ihren bipolaren Skalen waren für die markierten Adjektive signifikant größer, mithin einen Effekt der Asymmetrie erzeugend. Aufgrund des Umstandes, daß unmarkierte Adjektive überwiegend evaluativ positive Bedeutung haben und markierte negative (Hamilton & Deese 1971) ist dieser Befund auch für eine entsprechend alternative Interpretation offen.

Kaplan (1972) kritisiert, daß aufgrund der Definition der Mittelkategorie durch Osgood et al. (1957, 29 und 83) Indifferenz und Ambivalenz konfundiert sind. Osgood habe zwar die Gegensätzlichkeit der Adjektivpaare als Repräsentanten der reziprok-antagonistischen Tendenzen konzipiert, den Vpn würde jedoch die Mittelkategorie als ‚weder-noch‘ und als ‚sowohl-als auch‘-Kategorie erläutert. Auf diese Weise würden ambivalente Urteilstendenzen, statt als Markierungen auf beiden Seiten der Skala aufzutreten (was mit der Konzeption reziprok-antagonistischer Prozesse nicht vereinbar ist), in der Neutralkategorie aufgefangen. (Dazu ist zu bemerken, daß Osgood diesen Sachverhalt nicht nur gesehen, sondern theoretisch postuliert hat, s.o., S. 157). Mit Bezug auf die evaluativen SD-Skalen schlägt Kaplan eine Trennung

der Antonyme als orthogonale (liking-disliking) Komponenten vor, die jeweils durch unipolare Ratings (0 - 3 bzw. -3 - 0) zu erfassen sind und - zusätzlich zu den üblichen bipolaren SD-Daten - eine unabhängige Einschätzung des Ausmaßes von Ambivalenz erlauben sollen.

Die vorliegenden Befunde zur Bipolaritäts-Annahme sind überaus uneinheitlich. Dieser Sachverhalt läßt sich kaum zureichend mit den verwendeten, unterschiedlichen Operationalisierungen der Bipolarität begründen, da auch ähnliche Operationalisierungen zu unterschiedlichen Ergebnissen führen und verschiedene Operationalisierungen zu ähnlichen Ergebnissen. Vielmehr scheinen Varianten der Kombination von Skalen mit Konzepten dafür verantwortlich zu sein. Der Gegensatz von ‚gut‘ ist nicht invariant ‚schlecht‘, sondern u.U. auch ‚böse‘ (Brandt 1972). Da die Angemessenheit des polar-gegensätzlichen Schemas nicht grundsätzlich in Zweifel steht, sondern konkrete Formulierungen die Zweifel an der Geltung der Bipolaritäts-Annahme begründen, bedarf die Wahl von Antonymen in jedem Falle einer empirischen Begründung, und zwar unter Berücksichtigung der spezifischen Urteiler- und Konzeptpopulation.

3.2 Intervallgleichheit

Bei der Verarbeitung von SD-Daten wird - sowohl von Osgood & Cie, wie von anderen SD-Raumkonstrukteuren und -Anwendern - in aller Regel davon ausgegangen, daß die Antwortkategorien der (meist 7-stufigen) Skalen das bipolare Kontinuum nach gleichen Intervallbreiten aufteilen.

Soweit Adverbien die Kategorien auf den Beurteilungs-Skalen definieren, können für die von Osgood et al. (1957, 1975) verwendeten (slightly, quite, extremely) etwa gleiche Intensitätszuwächse angenommen werden (Cliff 1959; vgl. auch Howe 1962, 1966 a, 1966 b). Vergleiche der durch die numerische Kodierung bestimmten Kategoriengrenzen mit solchen, die aufgrund einer Skalierung nach dem Gesetz des kategorialen Urteils bestimmt sind, sind von Messick (1957), Revenstorff (1973a) und Fuchs (1974) durchgeführt worden. Diese Arbeiten unterscheiden sich im Hinblick auf die Anzahl der zur Beurteilung gegebenen Kategorien (7,10 bzw. 9) und die Skalen- (und Konzept-) Auswahlen (Standard Osgood, Standard Ertel, bzw. konzeptspezifisch). Die Ergebnisse stimmen in den wesentlichen Punkten überein:

- Die Abweichungen von den vorgegebenen gleicherscheinenden Intervallen sind beträchtlich: Die Intervallbreiten verengen sich zur Skalenmitte hin und erscheinen besonders auf der positiv markierten Seite der Skalen gedehnt. Bei Revenstorff (Skala ohne Mittelpunkt) und Fuchs erscheint auch die mittlere Kategorie gedehnt. Fuchs berichtet - im Unterschied zu Messick - über erhebliche Variabilität zwischen den Skalen.

- Die Regression der skalierten auf die angenommenen Intervallgrenzen ist generell linear: Die entsprechenden Korrelationskoeffizienten liegen bei allen Autoren um .97 und darüber.

Nach diesen Befunden erscheinen insbesondere die auf dem euklidischen Distanzmodell basierenden D-Maße (Osgood & Suci 1952; Cronbach & Gleser 1953; Osgood et al. 1957) problematisch, da sie Intervallgleichheit auf den Skalen und über die Skalen voraussetzen (vgl. Diehl & Schäfer 1975).

Die von Revenstorff (1973a, 125) geäußerte Erwartung, „daß man in einer Faktorenanalyse der skalierten Durchschnittsprofile sehr ähnliche Aussagen über den konnotativen Raum gewinnen würde, wie bei der Faktorisierung der unskalierten Profile“ wird durch einen entsprechenden Befund von Fuchs (1974) auch empirisch belegt. Allerdings geht in die Erwartung die Voraussetzung ein, daß die verwendeten Skalen in einem gemeinsamen Nullpunkt symmetrisch sind.

3.3 Nullpunktlage

Die Arbeiten von Messick (1957) und Fuchs (1974) enthalten auch Hinweise, daß der subjektive Nullpunkt der Skalen nicht mit dem Skalenmittelpunkt zusammenfällt, sondern leicht zur Seite der ‚positiven‘ Skalenmarkierung verschoben ist. Dem entsprechen im Zusammenhang mit der Bipolaritäts-Annahme berichtete Befunde der Skalen-Asymmetrie. Anderson (1970) hat die Winkel bestimmt, die die bipolaren Adjektiv-Punkte mit dem Nullpunkt bilden: Sie sind nach seinen Befunden befriedigend zum Nullpunkt distanz-symmetrisch. Die Gemeinsamkeit des Nullpunktes setzt im Ursprung Cosinus-Werte von -1.00 voraus, d.h. die Geraden zwischen den polaren Punkten bilden im Nullpunkt einen Winkel von 180°. Aufgrund der beobachteten Abweichung konstatiert Anderson, daß die Ergebnisse allenfalls als ein schwacher Beleg für die Annahme eines gemeinsamen Nullpunktes angesehen werden können.

Wenn Dawes (1977) demonstriert, daß die korrelations-analytische Behandlung von gemittelten Rating-Urteilen, einschließlich solcher auf einer SD-Skala, bei Größenschätzungen zu Ergebnissen führt, die Messungen auf der Basis einer Repräsentationstechnik entsprechen, so mag dies als Beleg dafür gelten, daß Ratingmaße nützlich verwendet werden können. Für die SD-Technik ist dieser Anwendungsfall aber zumindest uncharakteristisch; der schon von Heise (1969) beklagte Mangel an Forschung zu Problemen der Metrik von SD-Skalen wird durch derartige Befunde kaum geringer.

4. Wahl von SD-Skalen zur Exploration von Bedeutungs- Räumen: Konstruktion von Semantischen Differentialen

Bei der Verwendung eines SDS als Instrument zur Erfassung der Bedeutung von Konzepten kommt der Antwort auf die beiden folgenden Fragen erhebliche Bedeutung zu:

- Wird die EPA-Struktur als ein angemessenes Bezugssystem zur Bedeutungsdimensionierung akzeptiert?
- Lassen sich die relevanten Bedeutungsdimensionen durch einen Standardsatz von SD-Skalen repräsentieren?

Wenn - wie bei einem Großteil der SD-Verwendungen - beide Fragen bejaht werden, wird der Forscher von aufwendigen, technischen Konstruktionsarbeiten entlastet, und die Wahl von SD-Skalen wird zu einer Kompositions-Aufgabe: Das Material bilden die Ladungsmuster der Skalen aufgrund vorliegender Faktorenlösungen. Als semantische ‚Raum-Fähre‘ ist das SD - u.U. bei Modifikation der äußeren Verkleidung - jederzeit wiederverwendungsfähig.

Angesichts der Beobachtungen von Konzept-Skalen-Interaktionseffekten und ihrer Erklärung durch Bedeutungsverschiebungen bei den Skalen empfehlen Osgood et al. (1957, 78 ff) beim Einsatz des SDS zur Analyse der Bedeutung spezifischer Konzepte die Wahl der Skalen auch unter dem Gesichtspunkt der Relevanz, der semantischen Stabilität und der Bipolarität im Hinblick auf die untersuchten Konzepte vorzunehmen. Durch diese Zusatzkriterien wird allerdings nicht nur die Generalität der Indikatorfunktion von SD-Skalen, sondern auch die Verfügbarkeit einer Standardliste von Beurteilungsmerkmalen, die zur Lösung der Kompositions-Aufgabe herangezogen werden können, in Frage gestellt: Die Erfüllung derartiger Kriterien kann nicht impressionistisch, sondern nur auf der Grundlage systematisch-empirischer Prüfung gewährleistet werden. Vorliegende Merkmalslisten bieten dafür keine hinreichende Grundlage.

Aufgrund seiner - abweichenden - Einschätzung der Konzept-Skalen-Interaktionsproblematik konnte Ertel (1964, 1965a, 1965 b) eine Standardform des SDS vorschlagen, die den Anspruch genereller Verwendbarkeit erhebt, mit der Einschränkung allerdings, daß diejenigen Skalen aus der Merkmalsliste eliminiert werden müßten, die die Bedeutung eines Konzeptes ‚im eigentlichen Sinne‘, d.h. nicht-metaphorisch, denotativ spezifizieren. (Diese Einschränkung steht im Zusammenhang mit Ertels Bemühen, die Art der mit dem Eindrucksdifferential zu erfassenden Bedeutungsprozesse als „emotionale“ zu präzisieren.) Abgesehen von der unregelmäßigen Handhabung derartiger Ausschlüsse zieht Ertel auf diese Weise offenbar nur eine restriktive Konsequenz aus dem von Osgood - wie vorher erörtert - mit ‚denotativer Kontamina-

tion' bezeichneten Problem, das zur Erklärung von Konzept-Skalen-Interaktionseffekten herangezogen wird und zur Abkehr vom Ideal eines universell - über alle Konzepte und Personen - verwendbaren Standard-Instruments führte.

Osgood et al. (1957, 76) betonen, daß es unangemessen sei, das SD als ein bestimmtes Instrument (als eine Art Test) zu apostrophieren. Vielmehr handele es sich um eine verallgemeinerbare Forschungstechnik, die jedem Forschungszweck anzupassen sei. Diese vorsichtige Einschätzung wird von den Autoren allerdings nicht durchgängig vertreten - so z.B. anders bei der Diskussion des SD als verallgemeinerter Einstellungs-Skala (p. 189) - und von Anwendern häufig unbeachtet gelassen. Durch den Bezug auf spezifische Untersuchungsbedingungen wird im übrigen das EPA-System nicht obsolet. Es ist auch nicht als eine problematische oder überflüssige Zugabe anzusehen, wie es - explizit - bei Darnell (1970) bzw. - implizit - bei Hofstätter (1955, 1959) geschieht. Vielmehr wird man - Miron (1969) zustimmend, wenn er behauptet, daß die EPA-Struktur des semantischen Raumes viel häufiger unter verschiedenen Bedingungen repliziert worden ist als die meisten anderen nicht unmittelbar beobachtbaren „Fakten“ unserer Wissenschaft - diese auch als Strukturmerkmal von SD-Skalen anzusehen haben. Abgesehen von den Einschränkungen bei der Reproduktion der EPA-Struktur für spezifische Konzepte/Konzeptklassen, die methodisch bedingt sein können, bedeutet die Akzeptierung der EPA-Struktur weder, daß diese Urteilsdimensionen alle Konzepte/Konzeptklassen in gleicher Weise charakterisieren, noch, daß sie bei allen Konzepten gleichermaßen durch identische Merkmale optimal repräsentierbar sind. Der EPA-Raum ist aufgrund des gewählten Analyseansatzes ein ‚Durchschnitts‘-Raum, der diejenigen Bedeutungsaspekte repräsentiert, die Beurteilern im Hinblick auf Konzepte gemeinsam sind. Die korrespondierenden Bedeutungsprozesse, die diesen Raum konstituieren, manifestieren sich im Urteilsverhalten als „gut“ reactions‘ (Osgood 1971, 37), d.h. abgesehen von der ‚ausgemittelten‘ Variation, die durch sprachliche Äußerungen von Personen im Hinblick auf Konzepte verursacht wird. Darüber hinaus stehen auch Abweichungen bei der Dimensionalität für Personen(gruppen) und Konzepte/Konzeptklassen nicht notwendig im Widerspruch zur ‚allgemeinen‘ EPA-Struktur; es ist bereits darauf hingewiesen worden, daß derartige spezifische Bedeutungsaspekte (bei Personen und für Konzepte) als Ergänzungen zur EPA-Durchschnittsstruktur verstanden werden können.

Im folgenden werden die wichtigsten Schritte der Konstruktion eines SD skizziert, wobei - je nach Forschungsintention - sowohl der Orientierung an der allgemeinen EPA-Struktur wie auch an spezifischen Bedeutungsstrukturen Raum gegeben wird. Das Verfahren berücksichtigt weitgehend die Vorgehensweise, die Osgood et al. (1975) zur Begründung der pankulturellen EPA-Struktur gewählt haben.

4.1 Merkmals-Relevanz

Bizarre Kombinationen von Konzepten und Skalen hatten schon in einer der ersten Würdigungen des ‚Measurement of Meaning‘ die Skepsis des Rezensenten hervorgerufen (Brown 1958: ‚Is a boulder sweet or sour?’). Osgood et al. (1957, 78f) hatten dieses Problem durchaus erkannt und die Relevanz der verwendeten Skalen für die zu beurteilenden Konzepte als Kriterium formuliert. Für den Fall der Verwendung irrelevanter Skalen hatten sie eine Tendenz zu uncharakteristischen ‚Neutral‘-Urteilen festgestellt. Oetting (1967) hat diesen Sachverhalt aufgrund entsprechender Beobachtungen bestätigt. An dieser Stelle ist auch ein Befund von Mitsos (1961) zu erwähnen, der für die jeweils drei persönlich bedeutsamsten von je sieben typischen EPA-Skalen größere Distanzen zwischen (7) Konzepten und größere Distanzen der Konzeptpunkte vom Bedeutungsnullpunkt fand als für die übrigen Skalen. Reduzierte Variabilität der Konzeptbeurteilungen, die in unmittelbarem Zusammenhang mit Validitätseinbußen steht, wird von Orlik (1965) für subjektiv ‚sachlich nicht einschlägige‘ und von Grimm et al. (1973) für ‚allgemeine‘ gegenüber ‚inhaltsorientierten‘ SD-Skalen berichtet. Mikula & Schulter (1970, 383) bekräftigen diesen Befund und spezifizieren, daß „annähernd 45% der Gesamtvarianz der Einstufungen durch die ‚Geeignetheit‘ der verwendeten Polarität determiniert sind“. Dabei wird eine Tendenz zur Verwendung extremerer Skalenkategorien bei verbal begabteren Versuchspersonen festgestellt. Diese Befunde entsprechen der ‚meaningful-polarization“-Hypothese von O’Donovan (1965), wonach Reaktionen auf bedeutsame stimuli polarisiert werden, während Urteile auf bedeutungslose stimuli in Richtung auf die Indifferenz-Kategorie tendieren. Nur geringe Ähnlichkeit/Unterschiede zwischen den Konzeptbeurteilungen anhand subjektiv sachlich einschlägiger und weniger einschlägiger SD-Skalen berichtet dagegen Schick (1968). Eher wirkte sich objektive (d.h. inferenz-statistisch definierte) Trennschärfe auf die Höhe der einzelnen Konzeptähnlichkeiten aus; die Struktur der Ähnlichkeitsbeziehung zwischen den Konzepten für die beiden nach der ‚objektiven Trennschärfe‘ verschiedenen Skalensätze war allerdings wieder sehr ähnlich. Daß für die jeweils fünf in den Analysen von Osgood et al. (1957) am höchsten ladenden EPA-Indikatoren Relevanz für die Beurteilung anderer als der von diesen Autoren berücksichtigten Konzepte keineswegs gewährleistet ist, belegt ein Befund von Carter et al. (1969). Den Beurteilern war Gelegenheit gegeben, bei allen Skalen-Konzeptkombinationen (bei selbstgewählten Skalen-Antonymen) eine ‚wouldn’t use‘-Kategorie zu markieren. Von dieser Möglichkeit wurde in insgesamt 44% der Fälle, bei einigen Kombinationen zu mehr als 90%, Gebrauch gemacht.

Auf eine formale Bedingung der Angemessenheit von Konzept-Skalenkombinationen weisen Smith & Nichols (1973) hin: Konzepte wie Skalen sollten im Hinblick auf eher ‚intensionale‘ oder eher ‚extensionale‘ Bedeutung unterschieden werden. Durch diese Termini, die der philosophischen Sprachtradi-

tion entlehnt sind (Inhalt/Umfang), wird von diesen Autoren eher subjektive, ausschließlich konnotative Bedeutung abgegrenzt von objektiver, sowohl denotative wie konnotative Aspekte einschließende Bedeutung (z.B. idealistisch - realistisch bzw. sauber - schmutzig). Die faktorielle Instabilität als Ausdruck von Konzept-Skalen-Interaktionseffekten war reduziert, wenn Konzepte und Skalen der gleichen Bedeutungsart kombiniert waren.

In einer weiteren Gruppe von Untersuchungen wird gefragt, ob die mit erheblichem Aufwand verbundene Auswahl ‚inhaltsorientierter‘ oder ‚konzeptspezifischer‘ Skalen aufgrund der resultierenden Ergebnisse gerechtfertigt ist. Grimm et al. (1973) vertreten aufgrund ihrer Befunde die Ansicht, daß ein enger Problembezug der Skalen sich über die höhere Diskriminationsleistung der Skalen in einer deutlichen Validitätssteigerung auswirkt. Die unzureichende Begründung des verwendeten Validitätskriteriums in dieser Untersuchung mindert allerdings den Wert dieser Interpretation. Flade (1968) hält eine konzeptspezifisch zusammengestellte Merkmalsliste (Franke 1976; vgl. dazu Franke & Bortz 1972; Bortz 1972) einer allgemeinen, unspezifischen (Hofstätter 1971) gegenüber deshalb nicht für „geeigneter“, weil in beiden Fällen je drei Faktoren einen etwa gleichen Anteil der Konzeptvarianz erklären. Allerdings kann bezweifelt werden, daß die Spezifität der Merkmalsliste von Franke für die untersuchten Konzepte als adäquat gelten kann. Vor allem erscheint aber die Höhe des auf eine bestimmte Faktorenlösung entfallenden Varianzanteils als ‚Effizienz‘- und Präferenz-Kriterium fragwürdig, insbesondere da die Interpretation dieser Faktoren für die beiden Listen verschieden und ein Validitätskriterium nicht verfügbar ist.

Techniken, bei denen die Geeignetheit von Skalen in einer vorgeschalteten Erhebung durch Beurteiler eingestuft wird (Mills 1970; vgl. auch Mitsos 1961, Schick 1968, Mikula & Schulter 1970) setzen voraus, daß das relevante Beurteilungsrepertoire bereits bekannt ist. Sie sind deshalb für die Begründung der SD-Skalen nur sehr eingeschränkt tauglich.

Ein interessanter Ansatz, der es ermöglichen würde, ‚konzeptadäquat‘ (Bergler 1975) individuelle Bedeutungsstrukturen zu explorieren, ist von Micko (1962; vgl. auch Triandis 1959 a, 1959 b, 1960) in Anlehnung an Kellys ‚Role-Construct-Repertory-Test‘ vorgeschlagen worden: Die Personen werden gebeten, aus je drei aller Konzepte das gemeinsame Merkmal der beiden ähnlichsten und das Unterscheidungsmerkmal zum dritten Konzept zu benennen. Auf diese Weise wird für jede Person eine Liste von Beurteilungsmerkmalen gefunden, die zu einem individuellen SD zusammengestellt werden können. Zwar könnten die so gefundenen Merkmale auch entsprechend dem von Osgood und Mitarbeiter beim Cross-Cultural-Projekt angewendeten Verfahren (s.o.) behandelt werden, ein individualisierendes Vorgehen würde aber einen angemessenen Zugang zum Problem der Beurteilung interindividueller Differenzen im Hinblick auf Bedeutungsstrukturen ermöglichen; die berichteten

Lösungen von MDS-Analysen und dreimodalen Faktorenanalysen berücksichtigen zwar auch Personvarianz, setzen aber die Angemessenheit der allgemeinen Beurteilungsmerkmale (Skalen) für das Urteilsverhalten der untersuchten Personen voraus. In der Einstellungsforschung sind inzwischen Techniken erprobt, die die Analyse individueller Strukturen, sowie deren Aggregation zu Strukturtypen erlauben (Feger 1974, 1975). Eine solche Vorgehensweise ist aber angesichts des erheblichen Aufwandes der individuellen Erhebungen für die meisten der typischen SD-Anwendungsfälle nicht geeignet.

Das konventionelle, von Osgood et al. (1975; vgl. auch Fuchs & Schäfer 1972) entwickelte Verfahren, das im Bezugssystem des Durchschnitt-Bedeutungsraumes personengruppen- und konzept(klassen)-spezifischer Variation Rechnung tragen kann, dürfte für die meisten Fragestellungen zu brauchbaren Lösungen führen.

Für eine nach einem Repräsentativitätskriterium bestimmte Stichprobe von Konzepten aus der Population der zu untersuchenden, oder - soweit möglich - für alle zu untersuchenden Konzepte, werden (adjektivische) Qualifikatoren gesucht. Osgood und Mitarbeiter verwendeten das Frageschema „(Konzept) ist -“ und „(Das) - (Konzept)“. (Diese Erhebung sollte bei Personen durchgeführt werden, die die Personpopulation repräsentieren, die für die spätere Untersuchung in Betracht genommen ist.)

Die erhaltenden Adjektive werden sodann nach drei Kriterien geordnet: ‚saliency‘ (Verwendungshäufigkeit über alle Konzepte), ‚diversity‘ (Zahl der verschiedenen Konzepte, für die die Adjektive verwendet wurden), ‚independence‘ (Ausmaß der Korreliertheit über die Konzepte). Die beiden ersten Kriterien können kombiniert als ‚productivity‘ durch Shannons H-Maß indiziert werden.

Die Adaptation lautet:

$$\text{index } H_j = - \sum_i p_{ij} \log_2 p_j(i) \quad (\text{productivity})$$

wobei i das Konzept- und j das Adjektiv-System bezeichnet, $p(ij)$ und $p_j(i)$ die Wahrscheinlichkeit des Auftretens eines Adjektivs bei allen bzw. den einzelnen Konzepten:

$$p_{ij} = f_{ij}/N_T \quad \text{und} \quad p_j(i) = f_{ij}/N_j$$

Bezogen auf die absoluten Häufigkeiten läßt sich für die einzelnen Adjektive auch schreiben:

$$H = 1/N_T \left[\sum_i (f_{ij} \log_2 f_{ij}) - f_j \log_2 f_j \right]$$

H steigt mit der Gesamthäufigkeit eines Adjektivs und mit der Häufigkeit der Konzepte, für die es genannt wurde an. H wird = Null, wenn ein Adjektiv nur

bei einem einzigen Konzept verwendet wurde, unabhängig von der Häufigkeit der Verwendung. Das Maximum von H würde erreicht, wenn alle Personen für alle Konzepte dasselbe Adjektiv nennen würden.

Zur Reduzierung semantischer Redundanz verwenden Osgood und Mitarbeiter die Phi-Statistik als Index der Unabhängigkeit („Quasi-Synonymität“) von Merkmalen. Für jedes Adjektiv werden die Fälle von gemeinsamem und nicht-gemeinsamem Vorkommen (und Nicht-Vorkommen) mit jedem anderen, in der Rangordnung der H-Maße folgenden Adjektiv bei allen Konzepten gezählt und die Summen in die Berechnungsformel eingesetzt. Zur Vermeidung von Typ I-Fehlern (Ausschluß wegen angenommener Gleichheit trotz vorhandener Unterschiedlichkeit) wird eine hohe Signifikanzgrenze als Selektionskriterium verwendet.

Es kann zweifelhaft erscheinen, daß ein derartiges Maß stochastischer (Un-)Abhängigkeit semantische Synonymität angemessen operationalisiert. Überdies ist nicht geklärt, welche Auswirkungen diese ‚Säuberung‘ auf die weiteren Konstruktionsschritte, die Auswahlmöglichkeit von bipolaren Skalen und die Analyse der dimensionalen Struktur des Urteilsverhaltens hat. Eine so bestimmte Unabhängigkeit der Merkmale ist u.E. problematisch und entbehrlich.

4.2 Merkmals-Polarität

Die in Abschnitt 3.1 erörterten Argumente und Befunde lassen es notwendig erscheinen, die Bipolarität von SD-Skalen empirisch-systematisch und nicht bloß intuitiv zu begründen.

Osgood et al. (1975) haben Antonyme in jeder Sprache/Kultur-Gruppe durch jeweils ca. 10 kompetente Sprecher dieser Gruppen erhoben. Da aufgrund der einschlägigen Untersuchungen konzept- und personspezifische Variationen zu erwarten sind, erscheint für Forschung, die sich nicht unmittelbar auf die Analyse einer allgemeinen Bedeutungsstruktur bezieht, die Berücksichtigung entsprechender Besonderheiten nicht unwesentlich. Bei einer Stichprobe von Beurteilern aus der vorgesehenen Population von Untersuchungspersonen sollten demnach Antonyme für die einzelnen Adjektive im Hinblick auf die zu beurteilenden Konzepte erhoben werden: ‚Das Gegenteil von einem (Adjektiv) (Konzept) ist ein _____ (Konzept)‘.

Es gibt bislang weder systematische noch konventionelle Kriterien, die das Maß der noch akzeptablen Beurteiler-Nichtübereinstimmung spezifizieren. Fuchs & Schäfer (1972) weisen auf einen bei eindeutigen Gegensätzen zu erwartenden Sprung in der Häufigkeitskurve hin. Wenn zwei, jeweils relativ häufig verwendete Antonyme auftreten, dürfte es zweckmäßig sein, zunächst

beide beizubehalten, da sie möglicherweise verschiedene Urteilstkontinua repräsentieren. In jedem Falle sollte überprüft werden, ob sich Abweichungen auf wenige Konzepte konzentrieren; ggf. ist zu erwägen, solche Konzepte als untypisch zu eliminieren.

4.3 Dimensionale Repräsentativität

Es ist festgestellt worden, daß die Repräsentation von Beurteilungs-Dimensionen ein charakteristisches Merkmal der SD-Technik ist. Mit der EPA-Struktur wird dafür ein allgemeines Bezugssystem zur Bedeutungs-differenzierung bereitgestellt. Soweit das Forschungsinteresse darauf gerichtet ist, die Bedeutung sehr verschiedenartiger Konzepte für Beurteiler vergleichbar zu machen, die nicht näher spezifiziert sind als durch ihre Zugehörigkeit zu einer Sprach/Kultur-Gemeinschaft, ist das Angebot, dafür den EPA-Raum zu wählen, konkurrenzlos. Die Reproduktion dieser Struktur durch Dimensionsanalysen von SD-Urteilen ist, wie mehrfach betont, von einer entsprechend allgemeinen, breit gefächerten Konzeptauswahl, die bereits im Prozeß der Merkmalsfindung zugrunde gelegt wird, abhängig. Wichtiges Material für die von Osgood et al. (1975) untersuchten Sprach/Kultur-Gruppen findet sich hierzu in dem der Veröffentlichung beigefügten ‚Semantischen Atlas‘.

Je spezifischer die untersuchte Konzeptklasse und Personengruppe ist, um so wichtiger wird die Frage, ob die EPA-Struktur als Vergleichsstandard dienen soll und kann, d.h. ob die aufgrund der Konstruktionsarbeit erhältliche spezifische Information ausgeschöpft werden oder eine Anpassung an die wohlbegründete, wenngleich unspezifische Struktur versucht werden soll. Durch eine Konzeptauswahl, die - ggf. mit Hilfe von Atlas-Daten - die Kombinationsmöglichkeiten der Oktanten des SD-Raumes abdeckt, durch Ergänzung der Skalenliste um EPA-Markierskalen, sowie durch entsprechende Rotation der Faktorenstruktur wird die letztere Lösung begünstigt. Heise (1969) ist der Auffassung, daß in einem solchen Falle mindestens 40 Konzepte (je 5 pro Oktanten) verwendet werden sollten, schon um die Skalenkorrelationen auf der Basis von Konzept-Mittelwerten bestimmen zu können. Wenn eine geringere Anzahl von Konzepten zugrunde gelegt wird, sollten die Skalenkorrelationen über $m \times n$, d.h. über alle Personen bei allen Konzepten berechnet und zusätzlich auftretende Faktoren ignoriert werden.

Ein unbedingtes Festhalten an der Verbindlichkeit der EPA-Struktur ist allerdings nicht begründbar. Es entspricht auch nicht der Konzeption Osgoods, der immer wieder betont hat, daß EPA zwar die zentralen Dimensionen affektiver Bedeutung von Zeichen repräsentiert, den Bedeutungsraum aber keineswegs im Hinblick auf Person- und Konzept-Variation erschöpfend beschreibt.

Die hier beschriebene, personengruppen- und konzeptklassen-spezifische Auswahl von SD-Skalen steht der Identifizierung von EPA-Dimensionen nicht im Wege. Sie erlaubt zwar keinen Vergleich und keine Typisierung individueller konzeptklassen-spezifischer Bedeutungsstrukturen, aber die Exploration weiterer und/oder konzeptspezifischer Bedeutungsdimensionen für bestimmte Personengruppen, vorausgesetzt, daß die untersuchte Konzeptklasse in der Konstruktionsphase hinreichend repräsentiert war.

4.4 Variationen der Präsentationsweise

4.4.1 Reihenfolge der Konzept-Skalenkombination

In der üblichen Form der Anwendung erhalten die Beurteiler Antwortbögen, auf denen die Skalen unterhalb der Nennung des jeweils zu beurteilenden Konzeptes in dimensional gemischter Reihenfolge und balancierter Polung der Bewertungsrichtung aufgeführt sind. In den früheren Arbeiten haben Osgood und Mitarbeiter eine Variante verwendet, bei der das Urteilskonzept zu jeder Skala neu festgelegt wird, wodurch eine Permutation der Konzept-Skalenkombinationen ermöglicht wird. Durch die Wahl dieser Form sollte die Wahrscheinlichkeit des Auftretens von Halo-Effekten verringert werden. Andererseits ist nicht auszuschließen, daß die Bedeutung der Konzepte durch die unterschiedlichen Kontextbedingungen stärker variiert. Im direkten Vergleich der Ergebnisse beider Verfahrensweisen fanden Osgood et al. (1957, 82) keine nennenswerten Unterschiede zwischen den Skalenmittelwerten. Die von ihnen des weiteren berichtete Resistenz von SD-Urteilen gegenüber Kontext-Ankereffekten wird von Sommer (1965) bekräftigt.

Mögliche kontextbedingte Fehlervarianz aufgrund einer Standardreihenfolge der Konzeptbeurteilungen wird häufig durch die technisch einfache Variation der Konzeptabfolge zu reduzieren versucht. Kane (1969) hat ein Computer-Programm entwickelt, mit dem sowohl die Reihenfolge der Konzepte, wie auch die Reihenfolge und Polung der Skalen systematisch permutiert werden können; in einer weiteren Arbeit (Kane 1971) berichtet er, daß eine Standardabfolge gegenüber verschiedenen Anordnungsvariationen weder im Hinblick auf die Faktorenstruktur und Faktoren-Scores, noch hinsichtlich der Markierungskonsistenz bei benachbarten Skalen zu unterschiedlichen Ergebnissen führt.

Osgood et al. (1975, 118 f) überprüften Ermüdungseffekte, soweit sie sich in einer geringeren oder gesteigerten Polarisierung der Skalenurteile auswirkten. Für umgekehrte Konzeptreihenfolgen wurden keine systematischen Effekte entdeckt (wobei allerdings Durchschnittswerte über die Personen zugrunde liegen).

4.4.2 Verankerung der Skalen

Die Uneinheitlichkeit der im Zusammenhang der Bipolarität von SD-Skalen berichteten Befunde kann - wie hier vorgeschlagen - zur Forderung einer systematischen Begründung der Antonyme von Merkmalen führen, die in bipolarer Anordnung als SD-Skalen verwendet werden. Wenn das bipolare Schema grundsätzlich in Frage gestellt wird, wird - allerdings nicht ohne Konsequenzen für die Konzeption der bedeutungsspezifischen Vermittlungsprozesse als reziprok-antagonistisch - die Verwendung unipolarer Skalen in Erwägung gezogen (z.B. Green & Goldfried 1965; Kaplan 1972).

Aus nicht näher beschriebenen Gründen haben Vidali (1973) und Vidali & Holeway (1975) eine Abwandlung der Green & Goldfried-Technik als ‚Stapel-Skalen‘ vorgeschlagen. Dabei wird die Skalenmitte durch ein Adjektiv gekennzeichnet, von dem die Ziffern 1 bis 3 auf- (+) und absteigen (-). Die Feststellung, daß die Ergebnisse insgesamt nicht signifikant verschieden von jenen ausfallen, die mit Hilfe von bipolaren SD-Skalen zu erhalten sind, kann allerdings nicht als eine hinreichende Begründung für den Vorzug oder eine ‚alternative‘ Verwendung von ‚Stapel-Skalen‘ akzeptiert werden. Im übrigen dürfte diese Präsentationsweise auch empfindlich für Effekte aufgrund ‚lexikalischer Markierung‘ sein, wonach die einzelnen Adjektive Urteilskontinua in unterschiedlicher Weise repräsentieren (vgl. S. 186).

Es werden auch Abweichungen vom adjektivischen Modus der Urteilsmerkmale vertreten. Ertel (1965a; vgl. auch Fuchs 1973) zieht die substantivische Form vor, weil er vermutet, daß diese die Beurteilung eher als Ähnlichkeitsvergleich fördert und nicht als Anheften von Attributen auffassen läßt. Er verspricht sich auch eine geringere Tendenz zu Verzerrungen im Sinne sozialer Erwünschtheit. Der Vergleich der Urteile auf einer adjektivischen und einer substantivischen Liste läßt gewisse Unterschiede zwischen diesen in der erwarteten Richtung erkennen (Ertel 1965 b).

Eine andere Abwandlung von der adjektivischen Skalenetikettierung wird von Mindak (196 1) vorgeschlagen und in der Markt- und Meinungsforschung häufiger verwendet: Die Beschreibungsmerkmale werden unabhängig von einer grammatikalischen Regel erhoben und verwendet, z.B. ‚really modern - sort of old-fashioned‘. Es scheint, daß auf diese Weise dem person- und konzeptspezifischen Urteilsverhalten sehr weitgehend Rechnung getragen werden kann, um den Preis allerdings, daß die Bestimmung verbindlicher Antonyme und bereits insoweit die Begründung eindimensionaler Kontinua erschwert ist.

4.4.3 Zahl der Antwortkategorien

Für die im Hinblick auf Rating-Skalen häufig diskutierte Frage nach der optimalen Anzahl der Abstufungskategorien (vgl. Miller 1956; McKelvie 1978) hat Gulliksen (1958) in bezug auf die SD-Technik die Forderung formuliert, 20- oder 30-stufige Skalen statt der von Osgood verwendeten sieben-stufigen Skala zu verwenden. Er begründet diese mit den von Osgood et al. (1957) berichteten hohen Übereinstimmungen im Urteilsverhalten bei Retests, die zeigten, daß die verlangte Diskriminationsleistung nicht zu einer Verteilung der Meßwerte führte, die die Bestimmung des Standardmeßfehlers erlaubt.

Die Zahl der tatsächlich verwendeten Skalenkategorien in SD-Untersuchungen schwankt nicht unerheblich um die 7, wobei kaum Skalen mit weniger als 5 und mehr als 10 Kategorien benutzt werden. Schönplflug (1972) hat die Auswirkungen dieser Variation für die Anzahlen von 3 bis 10 Kategorien systematisch untersucht. Für 15 Merkmale der Ertel-Liste (in adjektivischer Form), die zur Beurteilung von 48 Konzepten verwendet wurden, ergaben sich gleiche, dreidimensionale Bedeutungsräume mit gleicher Einlagerung der Konzepte, unabhängig von der kategorialen Differenziertheit der Skala. Die durchschnittlichen Einstufungen waren von der Kategorienzahl nicht abhängig, Urteile von Skalen verschiedener Kategorienzahl korrelierten alle nahezu perfekt.

Nach McKelvies (1978) Befunden aus einer nicht-SD-spezifischen Vergleichsstudie ist eine relativ geringe Zahl von Stufen empfehlenswert: Die Versuchspersonen operierten bei kontinuierlichen Skalen mit 5 oder 6 Stufen; 5-stufige Skalen wiesen die höchsten Reliabilitäten auf; bei stärkerer Abstufung (9 - 12 Stufen) ließen sich keine psychometrischen Vorteile belegen; bei weniger als 5 Stufen zeigte sich ein Verlust an Diskriminationsfähigkeit und Validität. ‚Die magische Zahl 7, plus oder minus 2‘ (Miller 1956), ob sie nun in der Kapazität menschlicher Informationsverarbeitung eine Grundlage hat oder nicht, dürfte als Maßgabe für die Differenzierung von SD-Skalen eine relevante Größe darstellen; dabei dürfte die Frage des Vorzugs einer verbalen Kodierung, wie von Osgood und Mitarbeitern verwendet, gegenüber einer numerischen nicht von Wichtigkeit sein.

Dagegen ist die Verwendung und Definition der Mittelkategorie problematisch. Im Unterschied zu Kaplans (1972) Kritik der mehrfachen Bedeutung dieser Kategorie, erscheint die Verwendung der ‚Neutral‘-Kategorie für den Fall ambivalenter Urteilstendenzen begründet (vgl. Osgood 1977, 16). Forthman (1973) kritisiert allerdings zu Recht, daß diese Kategorie in der Definition als Neutral-, Ambivalenz- und als Irrelevanz-Kategorie eine Art Müllschluckerfunktion zu erfüllen hat. Er isoliert die Ambivalenzfunktion der Mittelkategorie und findet bei Nichtberücksichtigung von Urteilen i. S. der Irrelevanz- und (leider auch gleichzeitig) der Neutral-Instruktion Abweichun-

gen vom Ladungsmuster gemäß der EPA-Struktur, deren Replikation bei Verwendung der Original-Instruktion gelingt.

Die Notwendigkeit einer Irrelevanz-Kategorie wird um so geringer sein, je konzept- und beurteilerspezifischer SD-Indikatoren ausgewählt worden sind. Wenn aufgrund sehr heterogener Beurteiler- und/oder Konzeptstichproben Konzept-Skalenkombinationen zu vermuten sind, die als irrelevant angesehen werden, dürfte eine durch Instruktion angebotene Möglichkeit der Streichung allerdings erwägenswert sein.

Es sollte schließlich noch erwähnt werden, daß Anwender der SD-Technik der Instruktion häufig eine nur geringe Bedeutung beimessen. Abgesehen von der eben genannten Einschränkung ist das von Osgood et al. (1957, 82ff) gegebene Beispiel nachahmenswert.

4.5 Varianten der Technik

Neben den Anwendungen der SD-Technik in verschiedenen Bereichen der Semantik des (verbalen) Urteilsverhaltens (z.B. ‚Persönlichkeits-Differential‘ Kuusinen 1969, Warr & Haycock 1970, Revenstorff 1971, Tzeng 1975, 1977; ‚Angst-Differential‘ Alexander & Husek 1962; ‚Verhaltens-Differential‘ Triandis 1971; ‚Stereotyp-Differential‘ Gardner et al. 1972; ‚Befindlichkeits-Differential‘ Baumann & Dittrich 1972; ‚face differential‘ Hurwitz et al. 1975) ist die Adaptation der SD-Technik zur Bedeutungsanalyse auf der Grundlage sprachfreier Bedeutungsreaktionen, die Entwicklung von ‚Grafischen Differentialen‘ erwähnenswert. Die mit Befunden aus der Analyse von Synästhesien begründete Auffassung, daß EPA die gemeinsamen affektiven Bedeutungsanteile bei Zeichen verschiedener Modalität repräsentiert (Osgood et al. 1957, Osgood 1959, Elliott & Tannenbaum 1963, Osgood et al. 1975), läßt die Wahl nicht-sprachlicher Indikatoren zur Bedeutungsdifferenzierung möglich und gelegentlich wünschenswert erscheinen (Jakobovits 1969). Kontrastive Paare visueller Muster ließen insbes. E, weniger deutlich P und A identifizieren. Die Replikation der EPA-Struktur grafischer Zeichen neben anderen Dimensionen des Bedeutungsraumes (vgl. Bentler & LaVoie 1972a) gelang Bentler & LaVoie (1972 b).

Ein wichtiger Anwendungsbereich der SD-Technik liegt in der Einstellungsforschung. Grundlegend sind dafür die von Osgood et al. (1957) gegebene Begründung, daß die von traditionellen Einstellungsskalen erfaßten Urteilstkontinua sowohl begrifflich wie empirisch mit der E-Dimension der EPA-Struktur übereinstimmen, sowie die Überzeugung, die typischen E-Skalen könnten als ein Standardsatz von Indikatoren für beliebige Einstellungsobjekte verwendet werden. Die Konsequenzen aus der zuletzt genannten Auffassung

sind zwar ‚benutzerfreundlich‘; sie ist jedoch nicht konsistent zu der Auffassung, die hier und von den Urhebern der Technik an anderer Stelle zur Verwendung von SDs vertreten wird (s.o.). Ein Überblick über die Verwendung der SD-Technik in der Einstellungsforschung geben Heise (1970) und Schäfer (1975c).

Kriterien der Konstruktion von Semantischen Differentialen können dazu beitragen, die SD-Technik als ein Verfahren zur Analyse der Bedeutung von Zeichen zu begründen. Die Fortentwicklung der SD-Technik wird allerdings erheblich vom Erfolg abhängen, mit dem die Nützlichkeit von SD-Bedeutungsmaßen für die Erklärung von Verhalten - über Urteilsverhalten hinaus - belegt werden kann.

Literatur

- Adams, F. M. & Osgood, C. E. 1973. A cross-cultural study of the affective meanings of color. *Journal of Cross-Cultural Psychology*, 4, 135-156.
- Alexander, S. & Husek, T. R. 1962. The anxiety differential. Initial steps in the development of a measure of situational anxiety. *Educational and Psychological Measurement*, 22, 325-348.
- Aiken, E. G. 1965. Alternate forms of a semantic differential for measurement of changes in self-description. *Psychological Reports*, 16, 177-178.
- Allison, R. B. 1963a. A two-dimensional semantic differential. *Journal of Consulting Psychology*, 27, 18-23.
- Allison, R. B. 1963 b. Using adverbs as multipliers in semantic differentials. *Journal of Psychology*, 56, 115-117.
- Amsfeld, Elizabeth & Lambert, W. E. 1964. Evaluational reactions of bilingual and monolingual children to spoken languages. *Journal of Abnormal and Social Psychology*, 69, 89-97.
- Amster, Harriet. 1964. Evaluative judgment and recall in incidental learning. *Journal of Verbal Learning and Verbal Behavior*, 3, 466-473.
- Anderson, A. B. 1970. Structure of semantic space. In: Borgatta, E. & Bohrnstedt, G. (eds) *Sociological methodology*. San Francisco: Jossey-Bass, 308-325.
- Anisfeld, M. & Lambert, W. E. 1966. When are pleasant words learned faster than unpleasant words? *Journal of Verbal Learning and Verbal Behavior*, 5, 132-141.
- Anisfeld, M., Bogo, N. & Lambert, W. E. 1962. Evaluation reactions to accented English speech. *Journal of Abnormal and Social Psychology*, 65, 223-231.
- Anisfeld, M., Munoz, S. R. & Lambert, W. E. 1963. The structure and dynamics of the ethnic attitudes of Jewish adolescents. *Journal of Abnormal and Social Psychology*, 66, 31-36.

- Arnold, J. B. 1971. A multidimensional scaling study of semantic distance. *Journal of Experimental Psychology (Monograph)*, 90, 349-372.
- Arthur, A. Z. 1965. Clinical use of the semantic differential. *Journal of Clinical Psychology*, 21, 337-338.
- Arthur, A. Z. 1966. Response bias in the semantic differential. *British Journal of Social and Clinical Psychology*, 5, 103-107.
- Atwood, J. T. & Falkenberg, S. D. 1971. The bipolarity of semantic space as related to concept complexity. *Journal of Psychology*, 79, 119-122.
- Barclay, A. & Thumin, F. J. 1963. A modified semantic differential approach to attitudinal assessment. *Journal of Clinical Psychology*, 19, 376.
- Barnard, J. W. 1966. The effects of anxiety on connotative meaning. *Child Development*, 37, 461-472.
- Barrett, G. V. & Otis, J. L. 1967. The semantic differential as a measure of changes in meaning in education and vocational counseling. *Psychological Reports*, 20, 335-338.
- Baumann, U. & Dittrich, A. 1972. Überprüfung der deutschen Version eines Polaritätenprofils zur Erfassung der Befindlichkeit. *Zeitschrift für klinische Psychologie*, 4, 335-350.
- Baxter, J. C. 1962. Mediated generalization as a function of semantic differential performance. *American Journal of Psychology*, 75, 66.
- Benel, Denise, C. R. & Benel, R. A. 1976. A further note on sex differences on the semantic differential. *British Journal of Social and Clinical Psychology*, 15, 437-439.
- Bentler, P. M. 1969. Semantic space ist (approximately) bipolar. *Journal of Psychology*, 71, 33-40.
- Bentler, P. M. & LaVoie, A. L. 1972a. An extension of semantic space. *Journal of Verbal Learning and Verbal Behavior*, 11, 38-47.
- Bentler, P. M. & LaVoie, A. L. 1972b. A nonverbal semantic differential. *Journal of Verbal Learning and Verbal Behavior*, 11, 491-496.
- Bergler, R. (ed) 1975. *Das Eindrucksdifferential. Theorie und Technik*. Bern: Huber.
- Bettinghaus, E. P. 1963. Cognitive balance and the development of meaning. *Journal of Communication*, 8, 94-105.
- Birch, D. & Erickson, M. 1958. Phonetic symbolism with respect to three dimensions from the semantic differential. *Journal of General Psychology*, 58, 291-297.
- Berlyne, D. E. & Peckham, S. 1966. The semantic differential and other measures of reaction to visual complexity. *Canadian Journal of Psychology*, 20, 125-135.
- Black, H. K. 1975. Semantic differential ratings of impoverished Stimuli: A replication. *Bulletin of the Psychonomic Society*, 5, 81-83.
- Block, J. 1958. An unprofitable application of the semantic differential. *Journal of Consulting Psychology*, 22, 235-236.

- Bobbitt, R. G. & Beck, R. C. 1971. Semantic differential judgments of single and multiple conditioned Stimuli with an aversive delay conditioning paradigm. *Journal of Experimental Psychology*, 89, 398-402.
- Bokander, I. 1966. Semantic description of complex and meaningful stimulus material. *Perceptual and Motor Skills*, 22, 201-202.
- Bortz, J. 1972. Beiträge zur Anwendung der Psychologie auf den Städtebau. II. Erkundungsexperiment zur Beziehung zwischen Fassadengestaltung und ihrer Wirkung auf den Betrachter. *Zeitschrift für experimentelle und angewandte Psychologie*, 19, 226-281.
- Bousfield, W. A. 1961. The problem of meaning in verbal learning. In: Cofer, C. N. (ed) *Verbal learning and verbal behavior*. New York: McGraw-Hill, 81-91.
- Brandt, L. W. 1972. Questions concerning some assumptions underlying the semantic differential. *Psychologische Beiträge*, 14, 61-67.
- Brandt, L. W. 1978. Messung eines Maßstabs: Empirische Untersuchung des Semantischen Differentials (SD). *Probleme und Ergebnisse der Psychologie*, 66, 71-74.
- Brewer, W. F. & Lichtenstein, E. H. 1974. Memory for marked semantic features versus memory for meaning. *Journal of Verbal Learning and Verbal Behavior*, 13, 172-180.
- Brinton, J. E. 1961. Deriving an attitude scale from semantic differential data. *Public Opinion Quarterly*, 25, 289-295.
- Brod, Diane, Kernoff, Phyllis & Terwilliger, R. F. 1964. Anxiety and semantic differential responses. *Journal of Abnormal and Social Psychology*, 68, 570-574.
- Brown, R. 1958. Is a boulder sweet or sour? *Contemporary Psychology*, 3, 113-115.
- Burger, G. K. & Pickett, L. 1976. The California psychological inventory and the semantic differential dimensions. *Journal of General Psychology*, 94, 129-134.
- Burns, R. 1976. The concept-scale interaction problem: A trap for the unwary on the semantic differential. *Educational Studies*, 2, 121-127.
- Bynner, J. & Romney, D. 1972. A method for overcoming the problem of concept-scale interaction in semantic differential research. *British Journal of Psychology*, 63, 229-234.
- Carroll, J. B. 1959. Review of „The measurement of meaning“. *Language*, 35, 58-77.
- Carroll, R. M. & Field, J. 1974. A comparison of the classification accuracy of profile similarity measures. *Multivariate Behavioral Research*, 9, 373-380.
- Carter, R. F., Ruggels, W. L. & Chaffee, S.H. 1969. The semantic differential in opinion measurement. *Public Opinion Quarterly*, 32, 666-674.
- Cassel, R. N. 1970. Development of a semantic differential to assess the attitude of secondary school and College students. *Journal of Experimental Education*, 39, 10-14.
- Chen, Kathleen, C. 1977. Extremity of semantic differential ratings in deaf and hearing subjects. *Journal of General Psychology*, 96, 231-236.

- Clark, H. H. 1969. Linguistic processes in deductive reasoning. *Psychological Review*, **76**, 387-404.
- Clark, Virginia A. & Kerrick, J. S. 1967. A method of obtaining summary scores from semantic differential data. *Journal of Psychology*, **66**, 77-85.
- Clevenger, T., Lazier, G. A. & Clark, Margaret L. 1969. The influence of certain factors on response to the semantic differential. *Public Opinion Quarterly*, **32**, 675-679.
- Cliff, N. 1959. Adverbs as multipliers. *Psychological Review*, **66**, 27-44.
- Cowen, L. 1972. Anxiety, self-concept, and the semantic differential. *Journal of Psychology*, **80**, 65-68.
- Coyne, L. & Holzman, P. S. 1966. Three equivalent forms of a semantic differential inventory. *Educational and Psychological Measurement*, **26**, 665-674.
- Creelman, Marjorie B. 1966. *The experimental investigation of meaning: A review of the literature*. New York: Springer.
- Crockett, W. H. & Nidorf, L. J. 1967. Individual differences in responses to the semantic differential. *Journal of Social Psychology*, **73**, 211-218.
- Cronbach, L. J. & Gleser, Goldine C. 1953. Assessing similarity between profiles. *Psychological Bulletin*, **50**, 456-473.
- Darnell, D. K. 1966. Concept-scale interaction in the semantic differential. *Journal of Communication*, **16**, 106-115.
- Darnell, D. K. 1970. Semantic differentiation. In: Emmert, P. & Brooks, W. D. (eds) *Methods of research in communication*. Boston: Houghton Mifflin Company, 181-196.
- Dawes, R. M. 1977. Suppose we measured height with rating scales instead of rulers. *Applied Psychological Measurement*, **1**, 267-273.
- De Burger, R. A. & Donahoe, J. W. 1965. Relationship between the meanings of verbal Stimuli and their associative responses. *Journal of Verbal Learning and Verbal Behavior*, **4**, 25-31.
- Deese, J. 1964. The associative structure of some common English adjectives. *Journal of Verbal Learning and Verbal Behavior*, **3**, 347-357.
- Deese, J. 1965. *The structure of associations in language and thought*. Baltimore: John Hopkins Press.
- Denmark, Florence L., Shirk, Ethel J. & Riley, R. T. 1972. The effect of ethnic and social class variables on semantic differential Performance. *Journal of Social Psychology*, **86**, 3-9.
- Deutschmann, P. J. 1959. The semantic differential and public opinion research. *Public Opinion Quarterly*, **23**, 435.
- Diab, L. N. 1965. Studies in social attitudes. III: Attitude assessment through SD. *Journal of Social Psychology*, **67**, 303-314.
- Dicken, C. F. 1961. Connotative meaning as a determinant of stimulus generalization. *Psychological Monographs*, **75** (Whole No. 505).

- Diehl, Brigitte & Schäfer, B. 1975. Techniken der Datenanalyse beim Eindrucksdifferential. In: Bergler, R. (ed.) Das Eindrucksdifferential. Bern: Huber, 157-211.
- Dilts, Martha & Taylor, R. E. 1964. The semantic differential of color pyramid test instructions. *Perceptual and Motor Skills*, 19, 968-970.
- DiVesta, F. 1965. Developmental patterns in the use of modifiers as modes of conceptualization. *Child Development*, 36, 186-213.
- DiVesta, F. 1966a. A developmental study of the semantic structures of children. *Journal of Verbal Learning and Verbal Behavior*, 5, 249-259.
- DiVesta, F. 1966 b. A normative study of 220 concepts rated on the semantic differential by children in grades 2 through 7. *Journal of Genetic Psychology*, 109, 205-229.
- DiVesta, F. 1966 c. Norms for modifiers used by children in a restricted word-association task: Grades 2 through 6. *Psychological Reports*, 18, 65-66.
- DiVesta, F. & Dick, W. 1966. The test-retest reliability of children's ratings on the semantic differential. *Educational and Psychological Measurement*, 26, 605-616.
- DiVesta, F. & Stover, D. O. 1962. The semantic mediation of evaluative meaning. *Journal of Experimental Psychology*, 64, 467-475.
- Eisenman, R., Bernard, J. L. & Hannon, J. E. 1966. Benevolence, potency, and God: A semantic differential study of the Rorschach. *Perceptual and Motor Skills*, 22, 75-78.
- Eisenman, R. & Rappaport, Joan. 1967. Complexity preference and semantic differential ratings of complexity-simplicity and symmetry-asymmetry. *Psychonomic Science*, 7, 147-148.
- Eiser, J. R. 1971a. Enhancement of contrast in the absolute judgment of attitude Statements. *Journal of Personality and Social Psychology*, 17, 1-10.
- Eiser, J. R. 1971 b. Categorization, cognitive consistency and the concept of dimensional salience. *European Journal of Social Psychology*, 1, 435-454.
- Eiser, J. R. & Mower White, C. J. 1974. Evaluative consistency and social judgment. *Journal of Personality and Social Psychology*, 30, 349-359.
- Eiser, J. R. & Stroebe, W. 1972. *Categorization and social judgment*. London: Academic Press.
- Elliott, L. L. & Tannenbaum, P. H. 1963. Factor-structure of semantic differential responses to visual forms and prediction of factor-scores from structural characteristics of the Stimulus shapes. *American Journal of Psychology*, 76, 589-597.
- Endler, N. S. 1961. Changes in meaning during psychotherapy as measured by the semantic differential. *Journal of Counseling Psychology*, 8, 105-111.
- Ertel, S. 1964. Die emotionale Natur des „semantischen“ Raumes. *Psychologische Forschung*, 28, 1-32.
- Ertel, S. 1965 a. Standardisierung eines Eindrucksdifferentials. *Zeitschrift für experimentelle und angewandte Psychologie*, 12, 22-58.

- Ertel, S. 1965 b. Weitere Untersuchungen zur Standardisierung eines Eindrucksdifferentials. *Zeitschrift für experimentelle und angewandte Psychologie*, 12, 177-208.
- Ertel, S. 1969. *Psychophonetik: Untersuchungen über Lautsymbolik und Motivation*. Göttingen: Hogrefe.
- Ervin-Tripp, Susan M. & Slobin, D.I. 1966. Psycholinguistics. *Annual Review of Psychology*, 17, 435-474.
- Etkind, A. M. 1979. (Versuch einer theoretischen Interpretation des semantischen Differentials), *Voprosy Psichologii* 1, 17-27 (Orig. russ., engl. Übersetzung: *Soviet Psychology* 1980, Vol. 18, 3-20).
- Everett, A. V. 1973. Personality assessment at the individual level using the semantic differential. *Educational and Psychological Measurement*, 33, 837-844.
- Feger, H. 1974. Die Erfassung individueller Einstellungsstrukturen. *Zeitschrift für Sozialpsychologie*, 5, 242-254.
- Feger, H. 1975. Längsschnittliche Erfassung intraindividuelle Unterschiede bei Einstellungsstrukturen. In: Lehr, U. & Weinert, F. (eds) *Entwicklung und Persönlichkeit*. Stuttgart, 38-49.
- Feger, H. & Wiczorek, R. 1980. Multidimensionale Skalierung in der Einstellungsmessung. In: Petermann, F. (ed.) *Einstellungsmessung - Einstellungsforschung*. Göttingen, 153-174.
- Finley, J. R. & Staats, A. W. 1967. Evaluative meaning words as reinforcing stimuli. *Journal of Verbal Learning and Verbal Behavior*, 6, 193-197.
- Flade, Antje. 1978. Die Beurteilung umweltpsychologischer Konzepte mit einem konzeptspezifischen und einem universellen Semantischen Differential. *Zeitschrift für experimentelle und angewandte Psychologie*, 15, 367-378.
- Flavell, J. H. 1961a. Meaning and meaning similarity: I. A theoretical reassessment. *Journal of General Psychology*, 64, 307-319.
- Flavell, J. H. 1961 b. Meaning and meaning similarity: II. The semantic differential and co-occurrence as predictors of judged similarity in meaning. *Journal of General Psychology*, 64, 321-335.
- Flavell, J. H. & Johnson, Ann. 1961. Meaning and meaning similarity: III. Latency and number of similarities as predictors of judged similarity in meaning. *Journal of General Psychology*, 64, 337-348.
- Fodor, J. A. 1965. Could meaning be an r_m ? *Journal of Verbal Learning and Verbal Behavior*, 4, 73-81.
- Fodor, J. A. 1966. More about mediators: A reply to Berlyne and Osgood. *Journal of Verbal Learning and Verbal Behavior*, 5, 412-415.
- Ford, L. H. & Meisels, M. 1965. Social desirability and the semantic differential. *Educational and Psychological Measurement*, 24, 465-475.
- Forthman, J. H. 1973. The effect of a zero interval on semantic differential rotated factor loadings. *Journal of Psychology*, 84, 23-32.

- Franke, J. 1976. Die Erlebniswirkung von Wohnumgebungen. In: Kaminski, G. (ed.) *Umweltpsychologie*. Stuttgart, 134-143.
- Franke, J. & Bortz, J. 1972. Beiträge zur Anwendung der Psychologie auf den Städtebau. I. Vorüberlegungen und erste Erkundungsuntersuchungen zur Beziehung zwischen Siedlungsgestaltung und Erleben der Wohnumgebung. *Zeitschrift für experimentelle und angewandte Psychologie*, 19, 76-108.
- Friedman, C. J. & Gladden, J. W. 1964. Objective measurement of social role concepts via the semantic differential. *Psychological Reports*, 14, 239-247.
- Friedman, C. J., Johnson, C. A. & Fode, K. 1964. Subjects descriptions of selected TAT cards via the semantic differential. *Journal of Consulting Psychology*, 28, 317.
- Fuchs, A. 1973. Emotionale Bedeutung religiöser Konzepte. Methoden-kritische Untersuchungen mit einem semantischen Differential. Phil. Diss., Universität Bonn (Fotodruck).
- Fuchs, A. 1974. Untersuchungen zu metrischen Problemen der Technik der Bedeutungs-differenzierung. *Archiv für Psychologie*, 126, 114-124.
- Fuchs, A. 1975a. Grundzüge einer Verhaltenstheorie der Bedeutung. In: Bergler, R. (ed.) *Das Eindrucksdifferential*. Bern: Huber, 33-68.
- Fuchs, A. 1975 b. Das Eindrucksdifferential als Instrument zur Erfassung emotionaler Bedeutungsprozesse. In: Bergler, R. (ed.) *Das Eindrucksdifferential*. Bern: Huber, 69-100.
- Fuchs, A. & Schäfer, B. 1972. Kriterien und Techniken der Merkmalsselektion bei der Konstruktion eines Eindrucksdifferentials. *Archiv für Psychologie*, 124, 282-302.
- Gärling, T. 1976. A multidimensional scaling and semantic differential technique study of the perception of environmental settings. *Scandinavian Journal of Psychology*, 17, 323-332.
- Gardner, R. C., Taylor, D. M. & Feenstra, H. J. 1970. Ethnic stereotypes: Attitudes or beliefs? *Canadian Journal of Psychology*, 24, 321-334.
- Gardner, R. C., Wonnacott, E. J. & Taylor, D. M. 1968. Ethnic stereotypes: A factor analytic investigation. *Canadian Journal of Psychology*, 22, 35-44.
- Gardner, R. C., Kirby, D. M., Gorospe, F. H. & Villamin, A. C. 1972. Ethnic stereotypes: An alternative assessment technique, the stereotype differential. *Journal of Social Psychology*, 87, 259-267.
- Gleser, Goldine C., Cronbach, L.J. & Rajaratnam, N. 1965. Generalizability of scores influenced by multiple sources of variance. *Psychometrika*, 30, 395-418.
- Gilpin, A. R. 1973. Lexical marking effects in the semantic differential. *Journal of Psychology*, 85, 277-285.
- Godfrey, R. R. & Natalicio, L. F. S. 1970. Evaluation on semantic differential equals abstraction plus error. *Psychological Reports*, 27, 467-473.
- Goldfried, M. R. 1962. On differences in meaning between normals and neurotics. *Psychological Reports*, 11, 183.

- Goldfried, M. R. 1963. The connotative meaning of some animal symbols for College students. *Journal of Projective Technique*, 27, 60-67.
- Green, P. E., Aheshwari, A. & Rao, V. R. 1969. Dimensional interpretation and configuration invariance in multidimensional scaling: An empirical study. *Multivariate Behavioral Research*, 4, 159-180.
- Grigg, A. E. 1959. Validity study of the semantic differential technique. *Journal of Clinical Psychology*, 15, 179-181.
- Green, R. F. & Goldfried, M. R. 1965. On the bipolarity of semantic space. *Psychological Monographs*, 79 (whole No. 599).
- Grimm, G., Lück, H. E. & Timaeus, E. 1973. Zur internen Validität von Polaritätsprofilen: Untersuchungen zum Standard-Polaritätsprofil von Ertel. *Zeitschrift für experimentelle und angewandte Psychologie*, 20, 547-571.
- Gulliksen, H. 1958. How to make meaning more meaningful. *Contemporary Psychology*, 3, 115-118.
- Hamilton, D. L. 1968. Personality attributes associated with extreme response style. *Psychological Bulletin*, 69, 192-203.
- Hamilton, H. W. & Deese, J. 1971. Does linguistic marking have a psychological correlate? *Journal of Verbal Learning and Verbal Behavior*, 10, 707-714.
- Harman, H. H. 1970. *Modern factor analysis*. Chicago: University of Chicago Press.
- Hastorf, A. H., Osgood, C. E. & Ono, H. 1966. The semantics of facial expressions and the prediction of the meanings of stereoscopically fused facial expressions. *Scandinavian Journal of Psychology*, 7, 179-188.
- Haygood, R. C. 1966. Use of semantic differential dimensions in concept learning. *Psychonomic Science*, 5, 305-306.
- Heaps, R. A. 1972. Use of the semantic differential technique in research: Some precautions. *Journal of Psychology*, 80, 121-125.
- Heise, D. R. 1965. Semantic differential profiles for 1,000 most-frequent words. *Psychological Monographs*, 79, No. 8.
- Heise, D. R. 1969. Some methodological issues in semantic differential research. *Psychological Bulletin*, 72, 406-422.
- Heise, D. R. 1970. The semantic differential and attitude research. In: Summers, G. F. (ed.) *Attitude measurement*. Chicago: Rand McNally, 235-253.
- Herrmann, T. 1962. Urteilsnuanciertheit und Fraktionierung des Bezugssystems. Eine Zweifaktorenhypothese des semantischen Differenzierens. *Psychologische Beiträge*, 7, 539-557.
- Heskin, K. J., Bolton, N. & Smith, F. V. 1973. Measuring the attitudes of prisoners by the semantic differential. *British Journal of Social and Clinical Psychology*, 12, 73-77.
- Hoar, J. R. & Meek, E. E. 1965. The semantic differential as a measure of subliminal message effects. *Journal of Psychology*, 60, 165-169.
- Hörmann, H. 1976. *Meinen und Verstehen*. Frankfurt/M.

- Hoffmann, Kristine. 1976. Die Messung des Erlebens von Wohnumgebungen - Zur Problematik der Reliabilitätsbestimmung. In: Kaminski G. (ed.) *Umweltpsychologie*. Stuttgart, 144-155.
- Hofstätter, P. R. 1955. über Ähnlichkeit. *Psyche*, 9, 54-80.
- Hofstätter, P. R. 1959. *Einführung in die Sozialpsychologie*. Stuttgart.
- Hofstätter, P. R. 1971. *Differentielle Psychologie*. Stuttgart.
- Hogg, J. 1969. A principal components analysis of semantic differential judgments of single colors and color pairs. *Journal of General Psychology*, 80, 129-140.
- Homzie, M. J. & Weimer, J. 1967. Connotative similarity and paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 6, 136-138.
- Hopper, D. & Padden, D. 1965. Psychiatric roles and their meaning. *British Journal of Social and Clinical Psychology*, 4, 35-38.
- Horman, M. 1960. Implicit personality theories of clinicians as defined by semantic structures. *Journal of Consulting Psychology*, 24, 180-186.
- Horn, J. L. 1965. An empirical comparison of methods for estimating factor scores. *Educational and Psychological Measurement*, 25, 313-322.
- Hovland, C. I. & Sherif, M. 1952. Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *Journal of Abnormal and Social Psychology*, 47, 822-832.
- Howe, E. S. 1962. Probabilistic adverbial qualifications of adjectives. *Journal of Verbal Learning and Verbal Behavior*, 1, 225-242.
- Howe, E. S. 1964. Three-dimensional structure of ratings of exploratory responses shown by a semantic differential. *Psychological Reports*, 14, 187-196.
- Howe, E. S. 1965a. Further data concerning the dimensionality of ratings of the therapists' verbal exploratory behavior. *Journal of Consulting Psychology*, 29, 73-76.
- Howe, E. S. 1965b. Uncertainty and other correlates of Osgood's D₄. *Journal of Verbal Learning and Verbal Behavior*, 4, 498-509.
- Howe, E. S. & Pope, B. 1960. Multiple scaling of therapists' responses with a semantic differential. *American Psychologist*, 15, 415.
- Howe, E. S. 1966a. Associative structure of quantifiers. *Journal of Verbal Learning and Verbal Behavior*, 5, 156-162.
- Howe, E. S. 1966b. Verb tense, negatives, and other determinants of the intensity of evaluative meaning. *Journal of Verbal Learning and Verbal Behavior*, 5, 147-155.
- Hurwitz, D., Wiggins, Nancy Hirschberg & Jones, L. E. 1975. A semantic differential for facial attribution: The face differential. *Bulletin of the Psychonomic Society*, 6, 370-372.
- Husek, T. R. & Wittrock, M. C. 1962. The dimensions of attitudes toward teachers as measured by the semantic differential. *Journal of Educational Psychology*, 53, 209.
- Irle, M. 1975. *Lehrbuch der Sozialpsychologie*. Göttingen.

- Jakobovits, L. A. 1966a. Comparative psycholinguistics in the study of cultures. *International Journal of Psychology*, 1, 15-37.
- Jakobovits, L. A. 1966b. Mediation theory and the „single-stage“ S-R model: Different? *Psychological Review*, 73, 376-381.
- Jakobovits, L. A. 1969. The affect of Symbols: Towards the development of a cross-cultural graphic differential. *International Journal of Symbolology*, 1, 28-52.
- Jenkins, J. J. 1960. Degree of polarization and scores on the principal factors for concepts in the semantic atlas study. *American Journal of Psychology*, 73, 274-279.
- Jenkins, J. J., Russell, W. A. & Suci, G. J. 1958. An atlas of semantic profiles for 360 words. *American Journal of Psychology*, 71, 688-699.
- Jenkins, J. J., Russell, W. A. & Suci, G. J. 1959. A table of distances for the semantic atlas. *American Journal of Psychology*, 72, 623-625.
- Johnson, R. L. & Wall, D. D. 1969. Cluster analysis of semantic differential data. *Educational and Psychological Measurement*, 29, 769-780.
- Jones, J. M. 1969. Meanings of ‚most acceptable‘ and ‚most objectionable‘. *Psychological Reports*, 24, 915-921.
- Jones, J. M. 1970. Dimensions of meaning and attitude change. *Psychological Reports*, 26, 955-962.
- Jones, J. M. 1971. Attitudinal valence and semantic differential potency scales. *Psychological Reports*, 28, 991-994.
- Katz, M. 1965. Agreement on connotative meaning in marriage. *Family Process*, 4, 64-75.
- Kahneman, D. 1963. The semantic differential and the structure of inferences among attributes. *American Journal of Psychology*, 76, 554-567.
- Kane, R. B. 1969. Computer generation of semantic differential (SD) questionnaires. *Educational and Psychological Measurement*, 29, 191-192.
- Kane, R. B. 1971. Minimizing order effects in the semantic differential. *Educational and Psychological Measurement*, 31, 137-144.
- Kanungo, R. N. & Lambert, W. E. 1963. Semantic satiation and meaningfulness. *American Journal of Psychology*, 76, 421-428.
- Kaplan, K. J. 1972. On the ambivalence-indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin*, 77, 361-372.
- Kashiwagi, S. 1965. Geometric vector orthogonal solution for the semantic differential scales of Sagara et al. *Psychological Reports*, 16, 914.
- Kaufman, Helen J. 1959. The semantic differential: A critical appraisal. *Public Opinion Quarterly*, 23, 437-438.
- Keil, C. & Keil, Angeliki. 1966. Musical meaning: A preliminary report. *Ethnomusicology*, 10, 153-173.

- Kelly, Jane A. & Levy, L. H. **1961**. The discriminability of concepts differentiated by means of the semantic differential. *Educational and Psychological Measurement*, **21**, 53-58.
- Kentler, H. 1959. Zur Problematik der Profilmethode. *Diagnostica*, 5, 5-18.
- Kilty, K. M. 1972. Attitudinal affect and behavioral intentions. *Journal of Social Psychology*, **86**, 251-256.
- King-Fun Li, Anita. 1966. The Cantonese semantic differential scales. *Journal of Education*, 23.
- Kirby, D. M. & Gardner, R. C. 1973. Ethnic stereotypes: Determinants in children and their parents. *Canadian Journal of Psychology*, **27**, 127-143.
- Kjeldergaard, P. M. 1961. Attitudes toward newscasters as measured by the semantic differential: A descriptive case. *Journal of Applied Psychology*, **45**, 35-40.
- Kjeldergaard, P. M. & Higa, M. 1962. Degree of polarization and the recognition value of words selected from the semantic atlas. *Psychological Reports*, **11**, 629-630.
- Clapper, J. T. 1959. The semantic differential: Its use and abuse. *Public Opinion Quarterly*, **23**, 435-438.
- Klemmack, D. L. & Ballweg, J. A. 1973. Concept-scale interaction with the semantic differential technique. *Journal of Psychology*, **84**, 345-352.
- Koen, F. 1962. Polarization, m, and emotionality in words. *Journal of Verbal Learning and Verbal Behavior*, **1**, 183-187.
- Komorita, S. S. & Bass, A. R. 1967. Attitude differentiation and evaluative scales of the semantic differential. *Journal of Personality and Social Psychology*, **6**, 241-244.
- Korman, M. 1960. Implicit personality theories of clinicians as defined by semantic structures. *Journal of Consulting Psychology*, **24**, 180-186.
- Kostić, D. & Das, Rhea S. 1971. Aspects of meaning revealed by the semantic differential technique. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, **24**, 55-75.
- Krause, Ingrid & Solle, R. 1973. über die faktorielle Struktur des Werte-Raumes bei französischen Studenten. *Zeitschrift für experimentelle und angewandte Psychologie*, **20**, 210-239.
- Krieger, Margery H. 1963. A control for social desirability in a semantic differential. *British Journal of Social and Clinical Psychology*, **2**, 94-103.
- Kroes, W. H. & Libby, W. L. 1971. Relative power of taxonomic, semantic differential, and sense impression categories for the organization of free recall. *Journal of General Psychology*, **85**, 165-169.
- Kubinić, Cathleen M. & Farr, S. D. 1971. Concept-scale and concept-component interaction in the semantic differential. *Psychological Reports*, **28**, 531-541.
- Kumata, Hideya. 1957. A factor analytic investigation of the generality of semantic structures across two selected cultures. Ph. D. Diss., University of Illinois.
- Kumata, Hideya & Schramm, W. 1956. A pilot study of cross-cultural methodology. *Public Opinion Quarterly*, **20**, 229-238.

- Kuusinen, J. 1969. Affective and denotative structures of personality ratings. *Journal of Personality and Social Psychology*, 12, 181-188.
- Lambert, W. E. & Jakobovits, L. A. 1960. Verbal satiation and changes in the intensity of meaning. *Journal of Experimental Psychology*, 60, 376-383.
- Lana, R. E. & Pauling, F. J. 1965. Opinion change when the semantic differential is a pretest. *Psychological Reports*, 17, 730.
- Lane, Silvia T. M. 1973. Semantic differential scales for Portuguese speakers in Brazil. *International Journal of Psychology*, 8, 147-152.
- Lawson, E. D. 1971. Semantic differential analysis of men's first names. *Journal of Psychology*, 78, 229-240.
- Lawson, E. D. & Giles, H. 1973. British semantic differential responses on world powers. *European Journal of Social Psychology*, 3, 233-240.
- Lawson, E. D., Golden, G. H. & Chmura, Kathy J. 1972. Computer programs for the semantic differential. *Educational and Psychological Measurement*, 32, 779-784.
- Levin, J. 1965. Three-mode factor analysis. *Psychological Bulletin*, 64, 442-452.
- Levy, P. 1972. Concept-scale interaction in semantic differential research: Solutions in search of a problem. *British Journal of Psychology*, 63, 235-236.
- Light, C. S., Zax, M. & Gardiner, D. H. 1965. The relationship of age, sex, and intelligence level to extreme response style. *Journal of Personality and Social Psychology*, 2, 907-909.
- Lilli, W. 1975. *Soziale Akzentuierung*. Stuttgart.
- Litt, E. N. 1966. A factorial study of responses to abstract paintings. Unpubl. Master's Thesis. Univ. of Illinois (zit. nach Osgood et al. 1975).
- Lohr, J. M. 1976. Concurrent conditioning of evaluative meaning and imagery. *British Journal of Psychology*, 67, 353-358.
- Long, Barbara, H., Henderson, E. H. & Ziller, R. C. 1968. Self-ratings on the semantic differential: Content versus response set. *Child Development*, 39, 647-656.
- Lyle, J. 1960. Semantic differential scales for newspaper research. *Journalism Quarterly*, 37, 559-562, 646.
- Maclay, H. & Ware, E. E. 1961. Cross-cultural use of the semantic differential. *Behavioral Science*, 6, 185-190.
- Magnusson, D. & Ekman, G. 1970. A psychophysical approach to the study of personality traits. *Multivariate Behavioral Research*, 5, 255-274.
- Maguire, T. O. 1973. Semantic differential methodology for the structuring of attitudes. *American Educational Research Journal*, 10, 295-306.
- Malmstrom, E. J. & French, G. M. 1963. Scale-symmetry and the semantic differential. *American Journal of Psychology*, 76, 446-451.
- Madden, J. E. 1961. Semantic differential rating of self and of self-reported personal characteristics. *Journal of Consulting Psychology*, 25, 183.

- Maltz, H. E. 1963. Ontogenetic change in the meaning of concepts as measured by the semantic differential. *Child Development*, 34, 667-674.
- Manis, M. 1959. Assessing communication with the semantic differential. *American Journal of Psychology*, 72, 111-113.
- Markel, N. N. 1966. The validity of the semantic differential for psycholinguistic analysis. *Journal of Verbal Learning and Verbal Behavior*, 5, 348-350.
- Markel, N. N. & Meisels, M. 1964. Judging personality from voice quality. *Journal of Abnormal and Social Psychology*, 69, 458-463.
- Markel, N. N., Hunt, R. G. & Crapsi, L. A. 1966. The validity of the semantic differential for psycholinguistic analysis. *Journal of Verbal Learning and Verbal Behavior*, 5, 348-350.
- Marks, I. 1965. Patterns of meaning in psychiatric patients: Semantic differential responses in obsessives and psychopaths. London: Oxford University Press.
- Martinez, J. L. jr., Martinez, S. R., Olmedo, E. L. & Goldman, R. D. 1976. The semantic differential technique. A comparison of Chicano and Anglo high school students. *Journal of Cross-Cultural Psychology*, 7, 325-333.
- Maruyama, K. 1971. On the reliability of the data obtained by the semantic differential method. *Japanese Psychological Research*, 13, 51-59.
- Mayerberg, Cathleen K. & Bean, A. G. 1974. The structure of attitude toward quantitative concepts. *Multivariate Behavioral Research*, 9, 311-324.
- McCallon, E. L. & Brown, J. D. 1971. A semantic differential instrument for measuring attitude towards mathematics. *Journal of Experimental Education*, 39, 69-72.
- McKelvie, S. J. 1978. Graphic rating scales - How many categories? *British Journal of Psychology*, 69, 185-202.
- Meisels, M. & Ford, L. H. 1969. Social desirability response set and semantic differential evaluative judgments. *Journal of Social Psychology*, 78, 45-54.
- Messer, S., Jakobovits, L. A., Kanungo, R. & Lambert, W. E. 1964. Semantic satiation of words and numbers. *British Journal of Psychology*, 55, 155-163.
- Messick, S. J. 1957. Metric properties of the semantic differential. *Educational and Psychological Measurement*, 17, 200-206.
- Micko, H. C. 1962. Die Bestimmung subjektiver Ähnlichkeiten mit dem semantischen Differential. *Zeitschrift für experimentelle und angewandte Psychologie*, 9, 242-280.
- Mikula, G. & Schulter, G. 1970. Polaritätenauswahl, verbale Begabung und Einstufung im Polaritätenprofil. *Zeitschrift für experimentelle und angewandte Psychologie*, 17, 371-385.
- Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Miller, P. McC. 1974. A note on sex differences on the semantic differential. *British Journal of Social and Clinical Psychology*, 13, 33-36.
- Miller, S., James, Carol, Rytten, B., Tansill, R. & Thompson, C. 1971. Use of the

- semantic differential in the study of motivation. *Psychological Reports*, 29, **1279-1282**.
- Mills, D. H. 1970. Adjectives pertinent to psychotherapy for use with the semantic differential: An heuristic note. *Psychological Reports*, 26, 211-213.
- Mindak, W. A. 1961. Fitting the semantic differential to the marketing problem. *Journal of Marketing*, 25, 28-33.
- Miron, M. S. 1961. The influence of instruction modification upon test-retest reliabilities of the semantic differential. *Educational and Psychological Measurement*, 21, 883-893.
- Miron, M. S. 1969. What is it that is being differentiated by the semantic differential? *Journal of Personality and Social Psychology*, 12, 189-193.
- Miron, M. S. 1972. Universal semantic differential shell game. *Journal of Personality and Social Psychology*, 24, 313-320.
- Miron, M. S. & Osgood, C. E. 1966. Language behavior: The multivariate structure of qualification. In: Cattell, R. B. (ed.): *Handbook of multivariate experimental psychology*. Chicago: Rand McNally, 790-819.
- Mitsos, S. B. 1961. Personal constructs and the semantic differential. *Journal of Abnormal and Social Psychology*, 62, 433-434.
- Mogar, R. E. 1960. Three Versions of the F scale and Performance on the semantic differential. *Journal of Abnormal and Social Psychology*, 60, 262-265.
- Mordkoff, A. M. 1963. An empirical test of the functional antonymy of semantic differential scales. *Journal of Verbal Learning and Verbal Behavior*, 2, 504-508.
- Mordkoff, A. M. 1965. Functional vs nominal antonymy in semantic differential scales. *Psychological Reports*, 16, 691-692.
- Morimoto, H. 1957. A study of semantic relation by association method and the semantic differential. *Japanese Journal of Educational Psychology*, **4**, **131-137**.
- Moss, C. S. 1960. Current and projected status of semantic differential research. *Psychological Record*, 10, 47-54.
- Mueller, W. S. 1966. Anxiety level, inferred identification **and** response tendencies on a semantic differential. *Journal of Consulting Psychology*, 13, 149-152.
- Muthen, B., Pettersson, T., Olsson, U. & Stahlberg, G. 1977. Measuring religious attitudes using the semantic differential technique: An application of three-mode factor analysis. *Journal of the Scientific Study of Religion*, 16, 275-288.
- Neuringer, C. 1963. Effect of intellectual level and neuropsychiatric status on the diversity of intensity of semantic differential ratings. *Journal of Consulting Psychology*, 27, 280.
- Nickels, S. A. & Shaw, M. E. 1964. Saliency and two measures of attitude. *Psychological Reports*, 14, 273-274.
- Nordenstreng, K. 1968. A comparison between the semantic differential and similarity analysis in the measurement of musical experience. *Scandinavian Journal of Psychology*, 9, 89-96.

- Nordenstreng, K. 1969. Toward quantification of meaning. An evaluation of the semantic differential technique. *Annales Academiae Scientiarum Fennicae*, B 161, 2, Helsinki.
- Nordenstreng, K. 1970. Changes in the meaning of semantic differential scales: Measurement of subject-scale interaction effects. *Journal of Cross-Cultural Psychology*, 1, 217-327.
- Norman, W. T. 1959. Stability-characteristics of the semantic differential. *American Journal of Psychology*, 72, 581-584.
- O'Donovan, D. 1965. Rating extremity: Pathology or meaningfulness? *Psychological Review*, 72, 358-372.
- Oetting, E. R. 1964. Cross-cultural communication and the semantic differential: Research note. *Journal of Counseling Psychology*, 3, 292-293.
- Oetting, E. R. 1967. The effect of forcing response on the semantic differential. *Educational and Psychological Measurement*, 27, 699-702.
- Oles, H. J. 1973. Semantic differential for third through fifth grade students. *Psychological Reports*, 33, 24-26.
- Orlik, P. 1965. Eine Modellstudie zur Psychophysik des Polaritätsprofils. *Zeitschrift für experimentelle und angewandte Psychologie*, 12, 614-647.
- Orlik, P. 1967. Eine Technik zur erwartungstreuem Skalierung psychologischer Merkmalsräume aufgrund von Polaritätsprofilen. *Zeitschrift für experimentelle und angewandte Psychologie*, 14, 616-650.
- Osgood, C. E. 1952. The nature and measurement of meaning. *Psychological Bulletin*, **49**, 197-237.
- Osgood, C. E. 1959a. Semantic space revisited. *Word*, 15, 192-200.
- Osgood, C. E. 1959b. The cross-cultural generality of visual-verbal synesthetic tendencies. *Behavioral Science*, 5, 146-169.
- Osgood, C. E. 1962. Studies on the generality of affective meaning systems. *American Psychologist*, 17, 10-28.
- Osgood, C. E. 1964. Semantic differential technique in the comparative study of cultures. *American Anthropologist*, 66, 171-200.
- Osgood, C. E. 1965. Cross-cultural comparability in attitude measurement via multilingual semantic differentials. In: Steiner, I. & Fishbein, M. (eds) *Current studies in Social Psychology*. New York: Holt, Rinehart and Winston, 95-107.
- Osgood, C. E. 1966. Meaning cannot be r_m ? *Journal of Verbal Learning and Verbal Behavior*, 5, 402-407.
- Osgood, C. E. 1969. On the whys and wherefores of E, P, and A. *Journal of Personality and Social Psychology*, 12, 194-199.
- Osgood, C. E. 1971. Exploration in semantic space: A personal diary. *Journal of Social Issues*, 27, (4), 5-64.
- Osgood, C. E. & Luria, Z. 1954. A blind analysis of a case of triple personality using

- the semantic differential. *Journal of Abnormal and Social Psychology*, 49, **579-591**.
- Osgood, C. E. & Suci, G. 1952. A measure of relation determined by both mean difference and profile information. *Psychological Bulletin*, 49, **251-262**.
- Osgood, C. E. & Suci, G. 1955. Factor analysis of meaning. *Journal of Experimental Psychology*, 50, 325-338.
- Osgood, C. E., Suci, G. & Tannenbaum, P. 1957. *The measurement of meaning*. Urbana: University of Illinois Press.
- Osgood, C. E., May, W. H. & Miron, M. S. 1975. *Cross-cultural universals of affective meaning*. Urbana: University of Illinois Press.
- Osgood, C. E., Ware, E. E. & Morris, C. 1961. Analysis of the connotative meanings of a variety of human values as expressed by American College students. *Journal of Abnormal and Social Psychology*, 62, 62-73.
- Osipow, S. & Grooms, R. R. **1962**. On semantic differential resistance to response bias based on stimulus word position. *Psychological Reports*, **10**, **634**.
- Oyama, T., Tanaka, Y. & Chiba, Y. 1962. Affective dimensions of colors: A cross-cultural study. *Japanese Psychological Research*, **4**, **78-91**.
- Paivio, A. 1969. Mental imagery in associative learning and memory. *Psychological Review*, 76, 241-263.
- Paivio, A. **1971**. *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.
- Peabody, D. 1962. Two components in bipolar scales: Direction and extremeness. *Psychological Review*, 69, 65-73.
- Piaggio, L. **1968**. Einstufungswahrscheinlichkeiten im semantischen Differential. *Zeitschrift für experimentelle und angewandte Psychologie*, **15**, **272-290**.
- Piaggio, L. 1969. Bestimmung der Skalenzahl im semantischen Differential. *Probleme und Ergebnisse der Psychologie*, 31, 5-13.
- Plutchik, R. 1967. The affective differential: Emotion profiles implied by diagnostic concepts. *Psychological Reports*, 20, 19-25.
- Prothro, E. T. & Keehn, J. D. 1957. Stereotypes and semantic space. *Journal of Social Psychology*, 45, 197-209.
- Presly, A. S. 1969. Concept-scale interaction in the semantic differential and its implications for factor scores. *British Journal of Psychology*, **60**, **109-113**.
- Priest, P. N. 1971. The influence of psychiatric status and sex on the semantic differential response style. *Personality: An International Journal*, 2, 9-14.
- Reece, M. W. 1964. Masculinity and femininity: A factor analytic study. *Psychological Reports*, 14, 123-139.
- Reed, T. R. 1972. Connotative meaning of social interaction concepts: An investigation of factor structure and the effects of imagined contexts. *Journal of Personality and Social Psychology*, 24, 306-312.

- Revenstorff, D. 1971. Persönlichkeitsbeurteilung auf dem Polaritätsprofil. *Archiv für Psychologie*, 123, 195-216.
- Revenstorff, D. 1973a. Zur Analyse des konnotativen Raumes. *Zeitschrift für experimentelle und angewandte Psychologie*, 20, 117-152.
- Revenstorff, D. 1973 b. über Profilähnlichkeit. *Archiv für Psychologie*, 125, 203-232.
- Roll, S. & Verinis, J. S. 1971. Stereotypes of scalp and facial hair as measured by the semantic differential. *Psychological Reports*, 28, 975-980.
- Rosenbaum, L. L. & McGinnies, E. 1969. A semantic differential analysis of concepts associated with the 1964 presidential election. *Journal of Social Psychology*, 78, 227-235.
- Rosenbaum, L. L., Rosenbaum, W. B. & McGinnies, E. 1971. Semantic differential factor structure stability across subject, concept, and time differences. *Multivariate Behavioral Research*, 6, 451-469.
- Ross, B. M. & Levy, N. 1960. A comparison of adjectival antonyms by simple card-pattern formation. *Journal of Psychology*, 49, 133-137.
- Ross, J. 1965. Change in the use of semantic differential with a change in context. *Journal of Verbal Learning and Verbal Behavior*, 4, 148-151.
- Rothman, A. I. 1968. The factor analysis of a science-related semantic differential instrument. *Journal of Educational Measurement*, 5, 145-149.
- Rydell, S. T. 1966. Tolerance of ambiguity and semantic differential ratings. *Psychological Reports*, 19, 1303-1312.
- Sagara, M., Yamamoto, K., Nishimura, H. & Akuto, H. 1961. A study on the semantic structure of Japanese language by the semantic differential method. *Japanese Psychological Research*, 3, 146-156.
- Schäfer, B. 1973. Die Messung der ‚Beurteilung von Völkern‘ mit Hilfe eines Eindrucksdifferentials. *Archiv für Psychologie*, 125, 29-38.
- Schäfer, B. 1975a. Konstruktion des Umfrage-Instrumentariums. In: Schweitzer, C. C. & Feger, H. (eds) *Das deutsch-polnische Konfliktverhältnis seit dem Zweiten Weltkrieg*. Boppard, 187-223.
- Schäfer, B. 1975 b. Konstruktion eines Eindrucksdifferentials zur Erfassung der ideologiespezifischen Bewertung politischer Schlüsselwörter. In: Bergler, R. (ed) *Das Eindrucksdifferential*. Bern: Huber, 139-155.
- Schäfer, B. 1975c. Das Eindrucksdifferential als Instrument zur Einstellungsmessung. In: Bergler, R. (ed) *Das Eindrucksdifferential*. Bern: Huber, 101-118.
- Schick, A. 1968. Der Einfluß systematischer Skalenauswahl auf die Verlaufsähnlichkeit von Polaritätsprofilen. *Zeitschrift für experimentelle und angewandte Psychologie*, 15, 146-160.
- Schlosberg, H. 1954. Three dimensions of emotion. *Psychological Review*, 61, 81-88.
- Schludermann, S. & Schludermann, E. 1969. Scale checking style as a function of age and sex in Indian and Hutterite children. *Journal of Psychology*, 72, 253-261.
- Schönpluf, W. 1972. Ein Problem bei der Arbeit mit Kategorienskalen: Welchen

- Einfluß hat die Zahl der Skalenkategorien? Zeitschrift für experimentelle und angewandte Psychologie, 19, 141-171.
- Schuh, A. J. 1966. Use of the semantic differential in a test of Super's vocational adjustment theory. *Journal of Applied Psychology*, 50, 516-522.
- Shell, S. A., O'Mally, J. M. & Johnsgard, K. W. 1964. The semantic differential and inferred identification. *Psychological Reports*, 14, 547-558.
- Shikiar, R., Fishbein, M. & Wiggins, Nancy. 1974. Individual differences in semantic space: A replication and extension. *Multivariate Behavioral Research*, 9, 201-209.
- Sines, J. O. 1962. An indication of specificity of denotative meaning based on the semantic differential. *Journal of General Psychology*, 67, 113-115.
- Singet-, R. D. 1961. A note on the use of the semantic differential as a predictive device in milieu therapy. *Journal of Clinical Psychology*, 17, 376-378.
- Smith, R. G. 1959. Development of a semantic differential for use with speech-related concepts. *Speech Monographs*, 26, 263-272.
- Smith, R. G. 1961. A semantic differential for theatre concepts. *Speech Monographs*, 28, 1-8.
- Smith, R. G. 1962. A semantic differential for speech correction concepts. *Speech Monographs*, 29, 32.
- Smith, R. G. 1963. Validation of a semantic differential. *Speech Monographs*, 30, 50-55.
- Smith, R. G. 1966. Semantic differential dimensions and form. *Speech Monographs*, 33, 17-22.
- Smith, R. G. & Nichols, H. J. 1973. Semantic differential stability as a function of meaning domain. *Journal of Communication*, 23, 64-73.
- Snider, J. G. 1962. Profiles of some stereotypes held by ninth-grade pupils. *Alberta Journal of Educational Research*, 8, 147-156.
- Snider, J. G. & Osgood, C. E. (eds) 1969. *Semantic differential technique: A sourcebook*. Chicago: Aldine.
- Snyder, F. 1967. An investigation of the invariance of the semantic differential across the subject mode. Unpubl. Master's Thesis, Univ. of Illinois (zit. nach Osgood et al. 1975).
- Snyder, F. W. & Wiggins, Nancy 1970. Affective meaning systems: A multivariate approach. *Multivariate Behavioral Research*, 5, 453-468.
- Solar-z, A. K. 1963. Perceived activity in semantic atlas words as indicated by a tapping response. *Perceptual and Motor Skills*, 16, 91-94.
- Sommer, R. 1965. Anchor-effects and the semantic differential. *American Journal of Psychology*, 78, 317-318.
- Springbett, B. M. 1960. The semantic differential and meaning in non-objective art. *Perceptual and Motor Skills*, 10, 231-240.

- Staats, A. W. 1968. Learning, language and cognition. New York: Holt Rinehart and Winston.
- Staats, A. W. & Staats, Carolyn K. 1957. Meaning established by classical conditioning. *Journal of Experimental Psychology*, 54, 74-80.
- Staats, A. W. & Staats, Carolyn K. 1959. Meaning and m: Separate but correlated. *Psychological Review*, 66, 136-144.
- Staats, A. W., Staats, Carolyn K. & Heard, W. G. 1961. Denotative meaning established by classical conditioning. *Journal of Experimental Psychology*, 61, 300-303.
- Stricker, G. 1963. The use of the semantic differential to predict voting behavior. *Journal of Social Psychology*, 59, 159-167.
- Stricker, G. & Zax, M. 1966. Intelligence and semantic differential discriminability. *Psychological Reports*, 18, 775-778.
- Stricker, G., Takahashi, S. & Zax, M. 1967. Semantic differential discriminability: A comparison of Japanese and American students. *Journal of Social Psychology*, 71, 23-25.
- Suci, G. J. 1960. A comparison of semantic structures in American Southwest culture groups. *Journal of Abnormal and Social Psychology*, 61, 25-30.
- Tajfel, H. 1959. Quantitative judgment in social perception. *British Journal of Psychology*, 50, 16-29.
- Tajfel, H. 1975. Soziales Kategorisieren. In: Moscovici, S. (ed.) *Forschungsgebiete der Sozialpsychologie 1*. Frankfurt/M., 345-380.
- Tanaka, Y. 1962. A Cross-cultural study of national stereotypes held by American and Japanese College graduate subjects. *Japanese Psychological Research*, 4, 65-78.
- Tanaka, Y. 1964. Studies on the measurement of meaning and the generality of affective meaning Systems: A review. *Japanese Psychological Research*, 8, 27-69.
- Tanaka, Y. & Osgood, C. E. 1965. Cross-culture, cross-concept, and cross-subject generality of affective meaning Systems. *Journal of Personality and Social Psychology*, 2, 143-153.
- Tanaka, Y., Oyama, T. & Osgood, C. E. 1963. A cross-culture and cross-concept study of the generality of semantic spaces. *Journal of Verbal Learning and Verbal Behavior*, 2, 392-405.
- Tannenbaum, P. H. 1959. Selected applications of the semantic differential. *Public Opinion Quarterly*, 23, 435-439.
- Taylor, C. L. & Haygood, R. C. 1968. Effects of degree of category separation on semantic concept identification. *Journal of Experimental Psychology*, 76, 356-359.
- Taylor, D. M. & Gardner, R. C. 1969. Ethnic stereotypes: Their effects on the perception of communicators of varying credibility. *Canadian Journal of Psychology*, 23, 161-173.

- Taylor, H. F. 1971. Semantic differential factor scores as measures of attitude and perceived attitude. *Journal of Social Psychology*, 83, 229-234.
- Terwilliger, R. F. 1962. Free association patterns as a factor relating to semantic differential responses. *Journal of Abnormal and Social Psychology*, 65, 87-94.
- Triandis, H. C. 1959a. Categories of thought of managers, clerks, and workers about jobs and people in industry. *Journal of Applied Psychology*, 43, 338-344.
- Triandis, H. C. 1959b. Differential perception of certain jobs and people by managers, clerks and workers in industry. *Journal of Applied Psychology*, 43, 221-225.
- Triandis, H. C. 1960. A comparative factorial analysis of job semantic structures of managers and workers. *Journal of Applied Psychology*, 44, 297-302.
- Triandis, H. C. 1971. *Attitude and attitude change*. New York: Wiley.
- Triandis, H. C. & Osgood, C. E. 1958. A comparative factorial analysis of semantic structures in monolingual Greek and American College students. *Journal of Abnormal and Social Psychology*, 57, 187-196.
- Turvey, M. T. & Fertig, Joanne 1970. Polarity on the Semantic differential and release from proactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 9, 439-443.
- Turvey, M. T., Fertig, Joanne & Kravetz, S. 1969. Connotative classification and proactive interference in short-term memory. *Psychonomic Science*, 16, 223-224.
- Tzeng, O. C. S. 1975a. Reliability and validity of semantic differential E-P-A markers for an American English representative sample. *Psychological Reports*, 37, 292.
- Tzeng, O. C. S. 1975b. Differentiation of affective and denotative meaning systems and their influence in personality ratings. *Journal of Personality and Social Psychology*, 32, 978-988.
- Tzeng, O. C. S. 1977. A quantitative method for separation of semantic subspaces. *Applied Psychological Measurement*, 1, 171-184.
- Tzeng, O. C. S. & May, W. H. 1975. More than E, P and A in semantic differential scales: An answer to questions raised by S.T.M. Lane. *International Journal of Psychology*, 10, 101-117.
- Tzeng, O. C. S. & Landis, D. 1978. Three-mode multidimensional scaling with points of view solutions. *Multivariate Behavioral Research*, 13, 181-213.
- Vidali, J. J. 1973. Single-anchor stapel scales versus double-anchor semantic differential scales. *Psychological Reports*, 33, 373-374.
- Vidali, J. J. 1976. Reliability of the semantic differential under practical conditions. *Psychological Reports*, 39, 583-586.
- Vidali, J. J. & Holeway, R. E. 1975. Stapel scales versus semantic differential scales: Further evidence. *Psychological Reports*, 36, 165-166.
- Voyce, C. D. & Jackson, D. N. 1977. An evaluation of a threshold theory for personality assessment. *Educational and Psychological Measurement*, 37, 383-408.
- Walker, Betty A. & Robinson, R. 1977. Using a semantic differential in predicting

- counselor-trainee' success in practicum. *Educational and Psychological Measurement*, 37, 971-975.
- Ware, E. E. 1958. Relationships of intelligence and sex to diversity of individual semantic meaning spaces. Unpubl. Doc. Diss., Univ. of Illinois (zit. nach Osgood et al. 1957; 1975).
- Warr, P. B. & Haycock, Valerie 1970. Scales for a British personality differential. *British Journal of Social and Clinical Psychology*, 9, 328-337.
- Washington, W. N. 1975. A methodology for response style analysis of the semantic differential index. *Journal of General Psychology*, 93, 289-294.
- Weigel, R. G., Weigel, Virginia M., Thornton, G. C. & Magnusson, F. 1975. Assessment of preferences among Company names by semantic differential and free association technique. *Psychological Reports*, 37, 1163-1166.
- Weinreich, U. 1958. Travels through semantic space. *Word*, 14, 346-366.
- Weinreich, U. 1959. A rejoinder (Osgood, C. E., Semantic space revisited.) *Word*, 15, 200-201.
- Weksel, W. & Hennes, J. D. 1965. Attitude intensity and the semantic differential. *Journal of Personality and Social Psychology*, 2, 91-94.
- Wells, W. D. & Smith, Georgianna. 1960. Four semantic rating scales compared. *Journal of Applied Psychology*, 44, 393-397.
- Wickens, D. D. 1970. Encoding categories of words: An empirical approach to meaning. *Psychological Review*, 77, 1-15.
- Wickens, D. D. & Clark, Sandra 1968. Osgood dimensions as an encoding class in short-term memory. *Journal of Experimental Psychology*, 78, 580-584.
- Wiggins, Nancy & Fishbein, M. 1969. Dimensions of semantic space: A problem of individual differences. In: Snider, J. G. & Osgood, C. E. (eds) *Semantic differential technique*. Chicago: Aldine, 183-193.
- Wilcox, R. C. 1966. Effects of context on semantic differential ratings. *Psychological Reports*, 18, 873-874.
- Wildman, R. W. II & Wildman, R. W. 1976. Note on application of the semantic differential to the electoral process. *Psychological Reports*, 38, 1185-1186.
- Williams, J. E. 1964. Connotations of color names among Negroes and Caucasians. *Perceptual and Motor Skills*, 18, 721-731.
- Williams, J. E. 1966. Connotations of racial concepts and color names. *Journal of Personality and Social Psychology*, 3, 531-540.
- Williams, J. E. & Carter, Dorothy J. 1967. Connotations of racial concepts and color names in Germany. *Journal of Social Psychology*, 72, 19-26.
- Williams, J. E., Morland, J. K. & Underwood, W. L. 1970. Connotations of color names in the United States, Europe, and Asia. *Journal of Social Psychology*, 82, 3-14.
- Williams, W. S. 1971. A semantic differential study of the meaning of personality test

- items to children from different socioeconomic groups. *Journal of Psychology*, **79**, 179-188.
- Williams, W. S. 1972. A study of the use of the semantic differential by fifth grade children from different socioeconomic groups. *Journal of Psychology*, **81**, 343-350.
- Winograd, E. 1966. Recognition memory and recall as a function of degree of polarization on the semantic differential. *Journal of Verbal Learning and Verbal Behavior*, **5**, 566-571.
- Zavalloni, Marisa & Cook, S. W. 1965. Influence of judges' attitudes on ratings of the favorableness of Statements about a social group. *Journal of Personality and Social Psychology*, **1**, 43-54.
- Zax, M. & Louiselle, R. M. 1960. Stimulus values of Rorschach inkblots as measured by the semantic differential. *Journal of Clinical Psychology*, **16**, 160-163.
- Zax, M., Gardiner, D. H. & Lowy, D. G. 1964. Extreme response tendency as a function of emotional adjustment. *Journal of Abnormal and Social Psychology*, **69**, 654-657.
- Zippel, B. 1967. Semantic differential measures of meaningfulness and agreement of meaning. *Journal of Verbal Learning and Verbal Behavior*, **6**, 222-225.
- Young, D. D. 1974. The semantic differential: Application as an affective measure. *Journal of Experimental Education*, **42**, 86-91.