

# UNIVERSITÄT TRIER

## FORSCHUNGSBERICHTE ZUM MARKETING

Nr. 9  
Trier 2009

**Daniel Mühlhaus – Robert Hörstrup**

### **Konzeption und Evaluation reduzierter Abfragedesigns für MDS-Anwendungen**

Herausgeber: Univ.-Prof. Dr. Rolf Weiber

## **Forschungsbericht Nr. 9:**

Mühlhaus, D./Hörstrup, R. (2009):

Konzeption und Evaluation reduzierter Abfragedesigns für MDS-Anwendungen,  
Forschungsbericht Nr. 9, hrsg. von Rolf Weiber, Trier.

## **Über die Autoren:**

Dipl.-Volksw. Dipl.-Kfm. Daniel Mühlhaus und Dipl.-Kfm. Robert Hörstrup sind wissenschaftliche Mitarbeiter an der Professur von Univ.-Prof. Dr. Rolf Weiber für Betriebswirtschaftslehre, insb. Marketing und Innovation ([www.innovation.uni-trier.de](http://www.innovation.uni-trier.de)) an der Universität Trier.

## **Kontaktadresse**

Univ.-Prof. Dr. Rolf Weiber  
Universität Trier  
Lehrstuhl für Marketing, Innovation und E-Business  
Fachbereich IV: BWL-AMK  
Universitätsring 15  
D-54286 Trier  
Tel.: 0049-201-2619  
Fax: 0049-201-3910  
E-Mail: [marketing@uni-trier.de](mailto:marketing@uni-trier.de)  
Internet: [www.innovation.uni-trier.de](http://www.innovation.uni-trier.de)

Copyright: Eigenverlag der Professur für Marketing, Innovation und E-Business  
an der Universität Trier  
Univ.-Prof. Dr. Rolf Weiber  
Trier 2009

ISBN 3-930230-28-3

## Vorwort des Herausgebers

Obwohl die Multidimensionale Skalierung (MDS) als das Standardinstrument zur Analyse von Wahrnehmungen gilt und dementsprechend einen festen Platz insbesondere bei der Methodenausbildung von Studierenden der Sozial- und Wirtschaftswissenschaften besitzt, ist sie in der Marktforschungspraxis jedoch deutlich weniger etabliert. Die Gründe hierfür liegen neben Schwierigkeiten bei der inhaltlichen Interpretation von MDS-Konfigurationen oder der Wahl geeigneter Auswertungsmethoden vor allem im Bereich der Gestaltung von MDS-Untersuchungsdesigns. In einem ersten Forschungsbericht konnten die Autoren bereits wertvolle Hinweise zur Verbesserung der Validität von MDS-Konfigurationen im Bereich der Skalierungsmethodik und der Gestaltung des Untersuchungsrahmens geben (vgl. Forschungsbericht Nr. 8: „*Der Einfluss von Untersuchungsrahmen und Messmethodik auf die Güte von MDS-Konfigurationen*“).

Mit dem vorliegenden Forschungsbericht wenden sich die Autoren nun dem in der Anwendungspraxis bedeutsamen **Problem des oftmals hohen Erhebungsaufwandes** bei der Anwendung der MDS zu. Bekanntlich steigt aufgrund der üblicherweise paarweise zu erhebenden Ähnlichkeitsdaten der Erhebungsaufwand exponentiell mit der zu beurteilenden Objektzahl an. Die wissenschaftliche Forschung befasst sich bereits seit Anfang der 70er Jahre mit diesem "Problem" und korrespondierend dazu mit der **Konstruktion von reduzierten Abfragedesigns**. Die abgeleiteten wissenschaftlichen Lösungskonzepte der Literatur (zufällige Designs und zyklische Designs) sind jedoch aus Sicht der Praxis nur bedingt hilfreich, was sich letztendlich auch in der mangelnden praktischen Akzeptanz reduzierter Designs niederschlägt. Vor diesem Hintergrund entwickeln die Autoren drei weitere **Vorschläge zur Ausgestaltung von reduzierten Abfragedesigns**:

- (1) IDR: Individuell determiniertes reduziertes Design (ohne Vorinformation)
- (2) AED: Aggregiert extremobjektorientiertes Design (mit Vorinformation)
- (3) AVD: Aggregiert varianzorientiertes Design (mit Vorinformation)

Gemeinsam mit den Literaturkonzepten (zufällig und zyklisch) werden die von den Autoren entwickelten reduzierten Designs bei drei Objektbereichen (Automobilmarken, TV-Sender, Studiengänge) einer empirischen Güteprüfung unterzogen.

Die gewonnenen Ergebnisse zeigen, dass die in Abhängigkeit der Verfügbarkeit von Vorinformationen entwickelten Konzepte AED und AVD deutlich bessere Ergebnisse erzielen als die in der Literatur vorgeschlagenen zufälligen und zyklischen Designs. Selbst bei sehr hohen Reduktionsgraden von 60-70% ermöglichen AED und AVD die Aufrechterhaltung der individuellen Wahrnehmungsunterschiede. Darüber hinaus entwickeln die Autoren auch **Indikatormodelle zur Abschätzung der tatsächlichen Reproduktionsgüte** bei metrischen MDS-Konfigurationen auf der Basis reduzierter Designs.

Ein besonderer Dank gilt an dieser Stelle den Studierenden der Universität Trier, die über ihre engagierte Mitarbeit und das "zeitaufwändige" Bearbeiten der iteriert vorgelegten Fragebögen eine empirische Prüfung der dargestellten Ansätze und Konzepte überhaupt erst ermöglichten. Darüber hinaus gilt ein Dank aber auch den wissenschaftlichen Hilfskräften der Professur für Marketing, Innovation und E-Business, Frau cand. rer. pol. Katharina Ferreira und Herrn cand. rer. pol. Philip Wegmann, die wiederum einzelne Textfassungen durchgearbeitet, konstruktive Verbesserungsvorschläge unterbreitet und bei der Erstellung einer druckfähigen Fassung dieses Forschungsberichtes mitgewirkt haben.

Auch diesem Forschungsbericht wünsche ich nicht nur in der Scientific Community, sondern auch in der Marktforschungspraxis eine weite Verbreitung, da insbesondere durch die Anwendung der **Design-Vorschläge IDR und AVD** die Anwendungsbreite von metrischen MDS-Studien deutlich erhöht werden kann.

Trier, im März 2009

Rolf Weiber

## Inhaltsverzeichnis:

1	Die Problematik des hohen Erhebungsaufwandes bei der Multidimensionalen Skalierung .....	3
2	Verfahrensvarianten zur Ableitung reduzierter Erhebungsdesigns.....	7
2.1	Verfahren zur Bestimmung reduzierter Designs ohne Vorinformationen.....	7
2.2	Verfahren zur Bestimmung reduzierter Designs bei Vorliegen von Vorinformationen .....	9
3	Empirische Prüfung der Reproduktionsgüte alternativer reduzierter Designs ...	12
3.1	Abbildungsgüte zweidimensionaler aggregierter Konfigurationen als Referenzpunkt.....	15
3.2	Abbildungsgüte der reduzierten Designvarianten anhand des Korrelationskoeffizienten von Pearson .....	16
3.3	Würdigung der empirischen Abbildungsergebnisse und Auswahl der „besten Designs“ .....	21
3.4	Prüfung der Abbildungsgüte der reduzierten Designvarianten unter Verwendung von Probandenbewertungen .....	22
4	Entwicklung von Indikatormodellen zur Abschätzung der tatsächlichen Reproduktionsgüte bei metrischen MDS-Konfigurationen auf der Basis reduzierter Designs .....	26
5	Zusammenfassung und kritische Würdigung der erzielten Ergebnisse .....	29

## Abbildungsverzeichnis:

Abb. 1: Zyklische und zufällig reduzierte Designs in Ring- und Tableaudarstellung...	8
Abb. 2: Reproduktionswerte der Designvarianten bei unterschiedlichen Objektgruppen .....	21

## Tabellenverzeichnis:

Tab. 1: Übersicht der betrachteten Varianten reduzierter Designs .....	12
Tab. 2: Erläuterung der verwendeten Symbole .....	14
Tab. 3: Korrelationswerte der vollständigen Datensätze .....	15
Tab. 4: Reproduktionswerte $r(\Delta^I, D^R)$ für den Automobildatensatz ( $n=25$ ) .....	17
Tab. 5: Reproduktionswerte $r(\Delta^I, D^R)$ für den TV- Datensatz ( $n=27$ ) .....	19
Tab. 6: Reproduktionswerte $r(\Delta^I, D^R)$ für den Studienfachdatensatz ( $n=19$ ) .....	20
Tab. 7: Probandenbezogene Konfigurationsbewertung .....	24
Tab. 8: Prognosemodell bei Anwendung von IDRD II .....	28
Tab. 9: Prognosemodell bei Anwendung von AVD .....	29
Tab. 10: Eignung der unterschiedlichen Erhebungsdesigns .....	31

## 1 Die Problematik des hohen Erhebungsaufwandes bei der Multidimensionalen Skalierung

Obwohl die MDS als das Standardinstrumentarium zur Analyse von Wahrnehmungen gilt und dementsprechend einen festen Platz im universitären Lehrplan, insbesondere bei der Methodenausbildung von Studierenden der Sozial- und Wirtschaftswissenschaften besitzt, ist sie in der Praxis jedoch deutlich weniger etabliert (vgl. Hussey/Hooley 1995, S. 22ff.). Neben verschiedenen Begründungen, die auf Schwierigkeiten bei der inhaltlichen Interpretation von MDS-Konfigurationen (vgl. Carroll/Green 1997, S. 194), dem sehr speziellen Verwendungszweck oder der Wahl geeigneter Auswertungsmethoden beruhen, ist das Kernproblem bei der Anwendung der MDS im Bereich der Datenerhebung zu sehen (vgl. Whipple 1976, S. 99). Hier sind grundsätzlich die nachfolgenden drei zentralen Gestaltungsbereiche zu identifizieren:

- (1) *Festlegung des Untersuchungsrahmens*
- (2) *Auswahl einer Skalierungsmethodik zur Messung der Ähnlichkeitsurteile*
- (3) *Bestimmung eines Abfragedesigns*

Während die Gestaltungsbereiche (1) und (2) bereits im Forschungsbericht Nr. 8 analysiert wurden (vgl. Hörstrup/Mühlhaus 2009), konzentriert sich der vorliegende Beitrag auf die Konzeption und Evaluation reduzierter Abfragedesigns. Aufgrund des exponentiell mit der Zahl der zu beurteilenden Objekte ansteigenden Erhebungsaufwandes – bei 10 Objekten sind gemäß der Formel  $(n \cdot (n-1) / 2)$  insgesamt 45 Paarvergleiche durchzuführen, doch bereits bei 20 Objekten steigt diese Zahl auf 190 an (vgl. Schiffman/Reynolds/Young 1981) – sind Ansätze zur Beherrschung dieses Problems von zentraler Bedeutung im Rahmen der Datenerhebung und für einen Großteil der praktischen Anwendungen relevant. Diese Problematik des bisweilen prohibitiv hohen Erhebungsaufwandes paarweiser Ähnlichkeitsdaten beschäftigt die Forschung bereits seit Anfang der 70er Jahre. Seither ist eine ganze Reihe an Publikationen zur MDS entstanden, die unterschiedliche Ansätze zur Reduktion des erforderlichen Erhebungsaufwandes thematisiert. Grundsätzlich sind dabei zwei Stoßrichtungen zu identifizieren, die beide eine möglichst exakte Reproduktion der Struktur der Ähnlichkeitsdaten bei Minimierung des Abfrageaufwandes zu erreichen versuchen. Hierzu zählen:

### **(a) Ansätze der interaktiven Skalierung**

Bei den Ansätzen der interaktiven Skalierung (vgl. DeSarbo/Young/Rangaswamy 1997; Green/Bentler 1979; Young et al. 1982; Cliff et al. 1977; Young/Cliff 1972; Girard/Cliff 1976) werden simultan zur Datenerhebung auf individueller Ebene einerseits eine Auswahl der erforderlichen Anzahl an Paarvergleichen bestimmt und darüber hinaus die relevanten Paare ausgewählt. Wenngleich derartige Verfahren recht gute Reproduktionsergebnisse liefern (vgl. Young/Cliff 1972), so setzen sie doch die Verwendung spezifischer Software voraus, womit sie für Paper-Pencil-Befragungen nicht geeignet sind und die Verfügbarkeit entsprechender Applikationen nicht uneingeschränkt gegeben ist. Aus diesem Grund konzentrieren sich die nachfolgenden Betrachtungen auf die Anwendung und Ausgestaltung reduzierter Designs, die auch für Paper-Pencil-Befragungen gut verwendbar sind.

### **(b) Verwendung reduzierter Abfragedesigns**

Bei der Verwendung reduzierter Abfragedesigns (vgl. Graef/Spence 1979; Malhotra/Jain/Pinson 1988; Burton 2003; MacCallum 1978; Spence/Domoney 1974; Tsogo/Masson 1998; Spector/Rivizzigno 1982) werden den Probanden nicht alle  $(n \cdot (n-1) / 2)$  Paarvergleiche, sondern lediglich eine Fraktion hiervon zur Beurteilung vorgelegt. Eine solche Vorgehensweise ist nicht nur zur Reduktion des Erhebungsaufwandes sinnvoll, sondern nutzt auch die Existenz von Informationsineffizienzen aus (vgl. Young/Hamer 1987), da nicht alle Paarvergleiche auch relevante Informationen zur Ermittlung einer MDS-Konfiguration liefern (vgl. Spence 1983). Ist eine exakte Darstellung von  $n$  Objekten in einem  $r$ -dimensionalen metrischen Raum möglich, so sind  $z_1 = n \cdot (r+1) - (r+1) \cdot (r+2) / 2$  Ähnlichkeitsdaten zur eindeutigen Positionierung ausreichend (vgl. Young/Cliff 1972; Ross/Cliff 1964). Anhand verschiedener Simulationsstudien kommen Tsogo/Masson/Bardot (2001) zu dem Schluss, dass die notwendige Zahl an Ähnlichkeitsdaten sogar noch geringer ist und  $z_2 = r \cdot (n-2) + 1$  beträgt. Liegen bspw. 15 Objekte vor, die anhand einer 2-dimensionalen Darstellung positioniert werden können, so sind  $z_1 = 39$  bzw.  $z_2 = 27$  der insgesamt 105 Paarvergleiche ausreichend, die restlichen Daten liefern lediglich redundante Informationen. Eine Reduktion des Erhebungsaufwandes um mehr als 60% bzw. fast 75% wäre hier ohne Informationsverlust möglich.

Obwohl die meisten Studien die Verwendung reduzierter Designs befürworten und selbst bei hohen Reduktionsgraden von bis zu 80% gute Reproduktionsergebnisse verzeichnet werden können (vgl. Spence/Domoney 1974; MacCallum 1978; Burton



2003), hat sich dieses Vorgehen in der Praxis jedoch nicht durchsetzen können. Dies spiegelt sich auch in der minimalen Zahl wissenschaftlicher Beiträge wider, in denen reduzierte Designs zur Abfrage paarweiser Ähnlichkeiten verwendet werden. Hier ist lediglich eine einzige Studie von Behrens (1986) zu verzeichnen. Im deutschsprachigen Raum ist zudem kein einziger Hinweis auf die Verwendung oder Eignung reduzierter Abfragedesigns im Rahmen der MDS zu finden. Die Gründe für die geringe Verbreitung reduzierter Abfragedesigns sind vor allem in den wenig praxisorientierten und eher theoretisch gestützten Forschungsbemühungen und Veröffentlichungen zu finden und lassen sich in folgenden Aspekten zusammenfassen:

- Der Großteil der Studien zu reduzierten MDS-Abfragedesigns betrachtet lediglich artifizielle Daten im Rahmen von Monte Carlo Simulationsstudien auf einer aggregierten Ebene. Trotz der überwiegend positiven Ergebnisse (vgl. McCallum 1978; Spence/Domoney 1974) können diese nicht uneingeschränkt auf reale Befragungssituationen und Datensätze übertragen werden, da die Studien von wenig realistischen Annahmen ausgehen, die sich z. B. in der Bekanntheit von Fehlergrößen oder in sehr großen Objektzahlen niederschlagen.
- Abgesehen von der Studie von Malhotra/Jain/Pinson (1988) wurden bisher keine Untersuchungen durchgeführt, in denen der Fokus auf der Rekonstruktion von Individualdaten lag. Größtenteils werden nur aggregierte Daten analysiert, was jedoch nur bei homogenen Antwortstrukturen zweckmäßig ist. Die Betrachtung von lediglich aggregierten Daten ist aber für die meisten Anwendungsfragen wenig zielführend, wenn einerseits Wahrnehmungsunterschiede von Interesse sind oder Segmentierungen vorgenommen werden sollen. Gleiches gilt, wenn die Heterogenität der Daten sehr groß ist und somit wesentliche Charakteristika ohne eine explizite Individualbetrachtung verloren gehen.
- Die Übertragbarkeit der Ergebnisse von Simulationsstudien mit vorab spezifizierten Parametern auf andere Anwendungsfälle ist nur bedingt möglich, wie die Studie von Malhotra/Jain/Pinson (1988) aus dem Bereich der Marketingwissenschaft dokumentiert. Die Autoren kommen in ihrer Untersuchung von realen Datensätzen mit 15, 20 und 25 zu beurteilenden Marken zu dem Schluss, dass die Anwendung von reduzierten Designs mit mehr als 40% fehlenden Angaben bei praktischen Anwendungen, die eine Individualbetrachtung zum Ziel haben, nicht geeignet ist.
- Insbesondere für praktische Anwendungen sind Objektzahlen von über 30 zu meist unzweckmäßig, da anzuzweifeln ist, dass diese Vielzahl an Objekten von den Probanden verlässlich beurteilt werden kann. Darüber hinaus ist selbst unter Verwendung von stark reduzierten Designs eine Zahl von 100 Paarvergleichen unter Berücksichtigung von Monotonie und Anspruchsniveau, die eine derartige Beurteilungsaufgabe mit sich bringt, seriös kaum durchführbar (vgl. Dong 1983; Burton 2003).

Neben den o. g. Gründen besteht bei der Anwendung reduzierter Designs weiterhin das Problem, dass hierfür bisher kein geeignetes Gütekriterium existiert, das eine verlässliche Evaluation entsprechend ermittelter Konfigurationen erlaubt. So sind die klassischen Gütemaße, wie das Stress-Maß (vgl. Kruskal/Carmone 1973) oder der Korrelationskoeffizient  $r$  aus Ähnlichkeitsurteilen und Konfigurationsdistanzen (vgl. Tschudi 1972) ungeeignet, da diese lediglich die erhobenen Ähnlichkeiten berücksichtigen und somit die Gesamtgüte der Darstellung überschätzen. Bei Vorliegen einer vollständigen Ähnlichkeitsmatrix können die empirischen Daten mit den korrespondierenden Modelldistanzen anhand der Stress- oder Korrelationswerte verglichen werden und liefern hier auch gute Aussagen zur „Angemessenheit“ der erzielten Lösung. Eine solche Vorgehensweise ist aber bei einer unvollständigen Datenstruktur nur bedingt geeignet, da die MDS-Konfiguration die Relation der abgefragten Paarvergleiche besser widerspiegelt als die der nicht abgefragten Vergleiche, wobei letztere aber nicht minder wichtig sind.

Vor dem Hintergrund der oben aufgezeigten Probleme verfolgt der vorliegende Beitrag das Ziel, diverse Vorschläge für die Entwicklung von reduzierten Abfragedesigns abzuleiten, wobei auf der Kritik der bisher in der Literatur verwendeten reduzierten Designvarianten aufgesetzt wird (Abschnitt 2). Dabei wird – wie bereits im Forschungsbericht Nr. 8 – eine Einschränkung auf die in praktischen Anwendungen bewährte metrische MDS vorgenommen (vgl. Berekoven/Eckert/Ellenrieder 2004, S. 225), da diese eine geringere Beanspruchung der Probanden aufweist (vgl. Backhaus et al. 2006, S. 630) und darüber hinaus deutlich weniger Zeitaufwand bei der Datenerhebung erfordert (vgl. Taylor/Kinnear 1971, S. 10). Anschließend werden die vorgestellten Designvarianten anhand einer empirischen Untersuchung zu drei Objektbereichen (Automobilmarken, TV-Sender, Studiengänge) mit 10, 12 und 15 jeweils zu beurteilenden Objekten auf ihre Eignung für praktische Anwendungen geprüft (Abschnitt 3). Darüber hinaus liegt ein weiteres Ziel des vorliegenden Beitrags in der Entwicklung von Gütekriterien zur Beurteilung der durch reduzierte Designs ermittelten Konfigurationsgüte, die ebenfalls mit Hilfe der durchgeführten empirischen Untersuchung auf Zuverlässigkeit überprüft werden (Abschnitt 4). Der Beitrag schließt mit einer kurzen Zusammenfassung und kritischen Reflexion der erzielten Ergebnisse.

## 2 Verfahrensvarianten zur Ableitung reduzierter Erhebungsdesigns

In der Literatur finden sich verschiedene Vorschläge zur Bestimmung reduzierter Erhebungsdesigns (vgl. Spector/Rivizzigno 1982; Spence 1983), von denen die „zufälligen“ und „zyklischen Designs“ die prominentesten Vertreter sind. Beiden Varianten ist gemeinsam, dass bei ihrer Ableitung *keine* Vorinformationen hinsichtlich der zu erwartenden Struktur der Ähnlichkeitsurteile berücksichtigt werden, die aber häufig etwa aufgrund von Vorstudien verfügbar sind oder aber auf der Basis von z. B. Experteneinschätzungen leicht generiert werden können. Im Folgenden werden deshalb zunächst die beiden in der Literatur primär diskutierten Designvarianten kurz vorgestellt und einer kritischen Würdigung unterzogen. Aufbauend auf diesen Erkenntnissen werden sodann zwei weitere Varianten zur Ableitung reduzierter Erhebungsdesigns ohne Rückgriff auf Vorinformationen vorgestellt, die hier als IDRD I und IDRD II bezeichnet werden, und drei Verfahren zur Bestimmung von reduzierten Designs bei Verfügbarkeit von Vorinformationen vorgeschlagen.

### 2.1 Verfahren zur Bestimmung reduzierter Designs ohne Vorinformationen

In der Literatur wird implizit unterstellt, dass vor der Durchführung der Erhebung von Ähnlichkeitsdaten *keine* Information oder Vermutung hinsichtlich der zu erwartenden Struktur der Ähnlichkeitsurteile vorliegt. Zur Bestimmung von reduzierten Erhebungsdesigns werden die zu berücksichtigenden Paarvergleiche daher üblicherweise zufallsbasiert ausgewählt (sog. randomisierte reduzierte Designs). Zur Ausgestaltung von randomisierten Designs bei der Erhebung paarweiser Ähnlichkeiten werden unterschiedliche Ansätze diskutiert (vgl. Burton/Nerlove 1976; David 1963; 1982; Gollledge/Raynor 1982; Clatworthy 1955; 1973), deren bekannteste Varianten das zufällige und das zyklische Design darstellen:

#### **(a) Zufällige und zyklische Designs als weit verbreitete Ansätze**

Beim *zufälligen Design* (RAND) werden, ausgehend von einer festgelegten Zahl zu erhebender Ähnlichkeitsdaten, die Paarvergleiche anhand eines auf der Wahrscheinlichkeitstheorie basierenden Auswahlverfahrens ausgewählt. Hierbei kann es vorkommen, dass insbesondere bei hohen Reduktionsgraden und geringen Objektzahlen, also dem bei praktischen Anwendungen angestrebten bzw. oftmals gegebenen Fall, starke Unterschiede in der Häufigkeit  $h_i$  auftreten, mit der die verschiedenen Objekte  $i$  berücksichtigt werden. Das führt zu dem Problem, dass den Objekten mit ho-

hen Nennungsanteilen bei der Ermittlung der MDS-Konfiguration ein stärkeres Gewicht beigemessen wird und Objekte, die nur selten berücksichtigt werden, nicht eindeutig und bei  $h_i = 1$  beliebig radial um das eine, als Referenzpunkt dienende Vergleichsobjekt positioniert sind (vgl. Handl 2002, S. 142ff.).

Dieses Problem wird bei der Anwendung sog. *zyklischer Designs* (CYCL) vermieden (vgl. Spence/Domoney 1974), bei denen die zu nutzenden Paarvergleiche ebenfalls zufällig ausgewählt, aber alle Objekte gleich häufig berücksichtigt werden (vgl. Burton 2003). Die Konstruktion derartiger Designs kann dabei relativ einfach anhand der Ringdarstellung (vgl. McCormick/Bachus 1952; McCormick/Roberts 1952; John 1980) vorgenommen werden. Hierbei ist zum einen die Zahl  $r$  zu spezifizieren, die besagt, wie oft jedes Objekt berücksichtigt werden soll, was den Anteil an Ähnlichkeitsdaten

$\sigma = 2 \cdot r / (n-1)$  bzw. den Reduktionsgrad  $\pi = 1 - \sigma$  bestimmt. Zum anderen ist festzulegen, welche Nachbarbeziehung (d. h. welcher Designtyp) gewählt werden soll.<sup>1</sup>

Abb. 1 veranschaulicht graphisch und in Tableauform die Unterschiede zwischen randomisierten und zyklischen Designs.

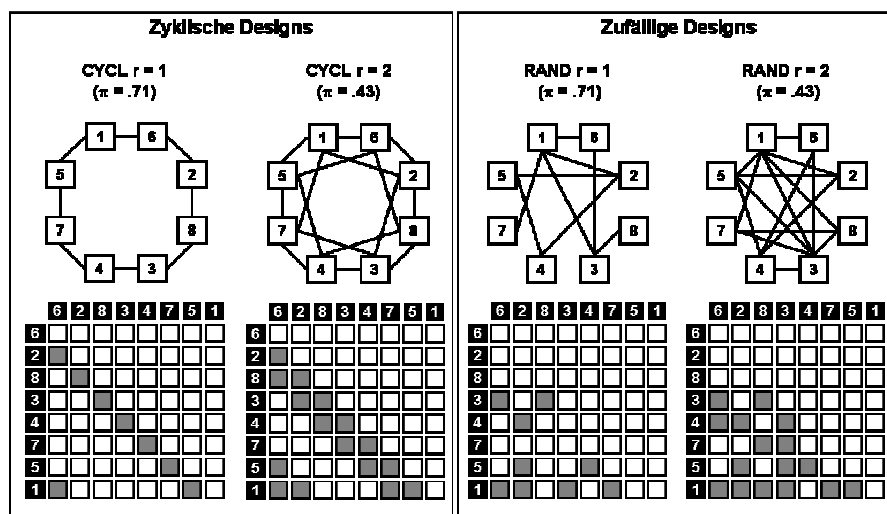


Abb. 1: Zyklische und zufällig reduzierte Designs in Ring- und Tableauform

Im oberen Teil der Abb. 1 ist die sog. Ringdarstellung ausgewiesen, die anhand der Verbindungen aufzeigt, welche Objekte jeweils miteinander verglichen werden sollen. Der untere Teil der Abb. 1 gibt die korrespondierende Darstellung in Tableauform als

<sup>1</sup> Aufbauend auf den Ergebnissen von Spence/Domoney (1974, S. 478ff.) wird hier der Designtyp II gewählt, der den Autoren folgend die besten Ergebnisse liefert. Dieser Typ bezeichnet nach Spence/ Domoney (1974, S. 476) dabei ein zyklisches Design, in dem die Zahl an Dreiecksbeziehungen (Triangles) zwischen jeweils 3 beurteilten Objekten sehr gering ist. Eine eingängige Veranschaulichung dieses Sachverhaltes findet sich bspw. bei Spence (1983, S. 411).

Distanzmatrix an, in der die zu erhebenden Paarvergleiche zwischen den zeilen- und spaltenweise dargestellten Objekten grau hinterlegt sind.

### **(b) Die Konstruktion individuell determinierter reduzierter Designs (IDRD)**

Alternativ zu den in der Literatur diskutierten vollständig randomisierten Designs ist es aber auch denkbar, dass die Auswahl der relevanten Paarvergleiche durch die Probanden – unter Einhaltung bestimmter Restriktionen – selbst gesteuert wird. Wir sprechen in diesem Fall von *individuell determinierten reduzierten Designs* (IDRD) und gehen von folgenden Überlegungen aus: Insbesondere die sehr unähnlichen Paare weisen einen großen Einfluss auf die Gestalt der MDS-Konfiguration auf und sollten zur Sicherstellung einer adäquaten Reproduktionsgüte und zur Steigerung der Informationseffizienz auf jeden Fall abgefragt werden (vgl. Graef/Spence 1979; Borg/Groenen 1997; Isaac 1982). Da dies jedoch ohne Vorkenntnisse und bei Anwendung vollständig randomisierter Designs nicht sichergestellt ist, kann folgende zweistufige Vorgehensweise gewählt werden, die auch unter Gesichtspunkten der praktischen Anwendbarkeit für Paper-Pencil-Befragungen gut geeignet erscheint:

Die Auskunftspersonen werden in einem ersten Schritt gebeten, aus der Menge der zu beurteilenden Objekte, denjenigen Anteil  $t$  an Paaren, die ihnen am unähnlichsten erscheint auszuwählen und hinsichtlich ihrer Ähnlichkeit zu bewerten. In einem zweiten Schritt werden daraufhin weitere Paare vorgelegt, wobei diese entweder anhand eines zyklischen oder eines randomisierten Designs vom Studienleiter bestimmt werden können (Variante: *IDRD I vom Grad  $t$* ). Da weiterhin nicht auszuschließen ist, dass auch sehr ähnliche Paare einen übergeordneten Einfluss auf die MDS-Konfiguration ausüben, können zusätzlich zu den sehr unähnlichen auch die sehr ähnlichen Paare vorab ausgewählt und beurteilt werden (Variante: *IDRD II vom Grad  $t$* ).

## **2.2 Verfahren zur Bestimmung reduzierter Designs bei Vorliegen von Vorinformationen**

In vielen Fällen liegen Vorinformationen bzgl. der mittels MDS zu untersuchenden Objektgruppen vor oder können im Rahmen einer geringzahligen Vorstudie relativ leicht erhoben werden. Die Berücksichtigung von Vorinformationen bei der Ausgestaltung von reduzierten Designs erscheint auch sinnvoll, da wie Graef/Spence (1979) zeigen, die Reproduktionsgüte (also das Ausmaß inwieweit gegebene Ähnlichkeitsstrukturen anhand der MDS-Konfiguration rekonstruiert werden können) stark davon

abhängt, ob kritische Paarvergleiche, etwa zu sehr unähnlichen Paaren, auch berücksichtigt werden. In der Literatur wurde die Berücksichtigung von Vorinformationen allerdings bisher noch nicht diskutiert. Im Folgenden werden deshalb drei Vorschläge unterbreitet, die bei der Konstruktion von reduzierten Designs auf Vorinformationen zurückgreifen. Als Vorinformationen werden dabei solche Informationen bezeichnet, die Rückschlüsse auf die zu erwartende Struktur der Ähnlichkeitsurteile erlauben und bspw. anhand einer repräsentativen Vorstudie oder auf der Basis von Experteneinschätzungen erhoben werden können. In Abhängigkeit der Homogenität der zu erwartenden probandenübergreifenden Ähnlichkeitsurteile sind dabei zwei grundsätzlich verschiedene Ansätze denkbar, die als aggregiert extremobjektorientiertes Design (AED) und als aggregiert varianzorientiertes Design (AVD) bezeichnet werden.

#### **(a) Aggregiert extremobjektorientiertes Design (AED)**

Ist davon auszugehen, dass sich die Beurteilung der Objekte über alle Probanden als relativ homogen darstellt, so sollten, zur Sicherstellung einer adäquaten Reproduktionsgüte bei der Ausgestaltung eines reduzierten Designs, primär diejenigen Paarvergleiche berücksichtigt werden, die sich als sehr unähnlich darstellen. Sollen zur Untersuchung von 15 Objekten insgesamt 40 Paarvergleiche erhoben werden, was einer Reduktionsrate  $\pi$  von  $1-40/105 = .619$  entspricht, so gilt es folglich diejenigen 40 Paarvergleiche abzufragen, bei denen die zu erwartenden Unähnlichkeiten am größten sind. Diese Vorgehensweise wird im Folgenden als *AED I* bezeichnet. Neben dieser Vorgehensweise erscheint es u. U. aber auch zweckmäßig, nicht nur die sehr unähnlichen Paare, sondern zusätzlich auch die sehr ähnlichen Objektpaare zu berücksichtigen. Übertragen auf das oben beschriebene Beispiel wären hierbei die 20 Paare mit den erwarteten höchsten und diejenigen 20 mit den geringsten Ähnlichkeitsurteilen in das reduzierte Design aufzunehmen, was im Folgenden als Methode *AED II* bezeichnet wird. Die Auswahl der zu nutzenden Paare kann bei Vorliegen einer Vorstudie und sich abzeichnender hoher Homogenität der Probandeneinschätzungen anhand der aggregierten Urteile erfolgen. Bei Anwendung von kategorialen Ratingskalen zeigt sich, dass nicht alle Probanden die gesamte Spannweite der Skalen in gleichem Maße ausnutzen. So verwenden einige Probanden bevorzugt die Skalenendpunkte, andere differenzieren im Bereich der mittleren Werte und wieder andere nutzen das gesamte Antwortspektrum zur Widergabe ihrer Einschätzung. Um diesem Aspekt Rechnung zu tragen werden zur Ermittlung der aggregierten Urteile

die individuellen Einschätzungen unter Berücksichtigung des mittleren Angabewertes und der Standardabweichung normiert (vgl. Handl 2002, S. 28).

**(b) Aggregiert varianzorientiertes Design (AVD)**

Sind zusätzlich zu einer Einschätzung der aggregierten oder globalen Ähnlichkeitsurteile auch Informationen verfügbar, hinsichtlich welcher Paarvergleiche eine große Beurteilungsheterogenität vorliegt, so erscheint deren Verwendung bei der Ausgestaltung reduzierter Designs ebenfalls zweckmäßig, um individuelle Unterschiede in der Wahrnehmung zu berücksichtigen. Hierbei sollten gezielt die Ähnlichkeitsurteile der trennschärfsten Paare erhoben werden, bei denen probandenübergreifend die Streuung der standardisierten Urteile am größten ist. Die in diesem Rahmen nicht erhobenen Ähnlichkeitsurteile können dann durch die entsprechenden aggregierten Urteile ersetzt werden. Dies ist erforderlich, da ansonsten bedeutsame Strukturen, die bei allen Probanden vorliegen, nicht berücksichtigt würden. Zusätzlich dazu bleibt der individuelle Charakter der Daten aber erhalten, da all diejenigen Paare, bei denen divergierende Wahrnehmungen vorliegen, explizit und individuell berücksichtigt werden. Diese Vorgehensweise wird im Folgenden als aggregiert varianzorientiertes Design (AVD) bezeichnet und erscheint vor allem bei einem sehr heterogenen Antwortverhalten angemessen. Sie erfordert allerdings sehr differenzierte Vorinformationen, die im Prinzip nur anhand einer Vorstudie (bspw. im Rahmen einer Expertenbefragung) ermittelt werden können.

Während die beiden AED-Varianten versuchen die Berücksichtigung sehr unähnlicher, sowie sehr ähnlicher Paarvergleiche sicherzustellen, werden bei der AVD-Variante diejenigen Paare abgefragt, die die unterschiedlichen Einschätzungen der Probanden möglichst gut widerspiegeln. Die nachfolgende Tabelle gibt einen Überblick zu den in diesem Abschnitt vorgestellten Varianten zur Bildung reduzierter Designs, deren Eignung im Folgenden im Rahmen einer empirischen Untersuchung überprüft werden soll.

Methode	Ausgestaltung	Vor- informationen
<b>Ansätze zur Bestimmung reduzierter Designs aus der Literatur</b>		
RAND	Abfrage eines Anteils $\sigma$ zufällig ausgewählter Paare	Nein
CYCL	Abfrage eines Anteils $\sigma$ zufällig anhand eines zyklischen Designs ausgewählter Paare	Nein
<b>Erweiterte Vorschläge zur Bestimmung reduzierter Designs</b>		
IDRD I (t)	1) Probandenauswahl und Beurteilung des Anteils t unähnlichster Paare 2) Beurteilung des verbleibenden Anteils $\sigma-t$ zufällig ausgewählter Paare	Nein
IDRD II (t)	1) Probandenauswahl und Beurteilung der jeweils $t/2$ Anteile an unähnlichsten und ähnlichsten Paaren 2) Beurteilung des verbleibenden Anteils $\sigma-t$ zufällig ausgewählter Paare	Nein
AED I	Abfrage des Anteils $\sigma$ , der anhand einer Vorstudie als am unähnlichsten identifizierten Paare	Ja
AED II	Abfrage jeweils der Anteile $\sigma/2$ , der anhand einer Vorstudie als am unähnlichsten bzw. am ähnlichsten identifizierten Paare	Ja
AVD	1) Abfrage des Anteils $\sigma$ , der anhand einer Vorstudie als trennschärfsten identifizierte Paare, bei denen die Angaben am heterogensten sind 2) Berücksichtigung des Anteils $1-\sigma=\pi$ der verbleibenden Paare in Form der durchschnittlichen Ähnlichkeitsurteile aus der Vorstudie	Ja

Tab. 1: Übersicht der betrachteten Varianten reduzierter Designs

### 3 Empirische Prüfung der Reproduktionsgüte alternativer reduzierter Designs

Im Rahmen einer empirischen Untersuchung wurden die in Tab. 1 zusammengefassten Varianten zur Ausgestaltung reduzierter Designs auf ihre Eignung bei praktischen Anwendungen geprüft. Zu diesem Zweck wurden eigens reale Daten erhoben, da nur unter realen Bedingungen verlässliche und praxisrelevante Aussagen abgeleitet werden können. Dies ist insbesondere auf die Schwierigkeiten der Modellierung des individuellen Antwortverhaltens und entsprechend korrespondierender Störgrößen zurückzuführen. Darüber hinaus kann die Eignung der auf Vorinformationen basierenden Designvarianten, bei denen die individuellen Einschätzungen in Relation zu den durchschnittlichen Angaben aller Probanden bedeutsam sind, nicht anhand von Simulationsstudien untersucht werden.



Die zur Eignungsprüfung der verschiedenen Designvarianten herangezogenen Daten entstammen zwei Studien, die Anfang 2007 an der Universität Trier zum einen im Rahmen einer wirtschaftswissenschaftlichen Vorlesung und zum anderen mittels Online-Befragungen durchgeführt wurden. Jede der beiden Studien, an der sich insgesamt 71 bzw. 41 Studierende im Alter von 21 bis 31 Jahren beteiligten, basiert auf zwei aufeinander aufbauenden Befragungen: Die jeweils *erste Befragung* umfasste die Erhebung von metrischen Ähnlichkeitsurteilen bezüglich drei Objektgruppen (10 TV-Sender<sup>2</sup>, 12 Automobilmarken<sup>3</sup> und in Studie 2 zusätzlich 15 Studiengänge<sup>4</sup>), wobei sichergestellt wurde, dass alle Objekte den Probanden bekannt waren. Darüber hinaus kann die gewählte Objektanzahl für praktische Anwendungen in verschiedenen Bereichen als „typisch“ angesehen werden. Zur Abfrage wurde eine 7-stufige Ratingskala verwendet, die den Ergebnissen des von Hörstrup/Mühlhaus (2009, S. 13ff.), folgend, zur Anwendung der metrischen MDS sehr gut geeignet ist. Basierend auf den erhobenen Ähnlichkeitsdaten wurden zunächst unter Verwendung der in SPSS 16 implementierten ALSCAL Prozedur verschiedene MDS-Konfigurationen ermittelt. Die Verwendung von ALSCAL wird als zweckmäßig angesehen, da der ALSCAL-Algorithmus nicht nur forschungstechnisch stark durchdrungen (vgl. MacCallum 1978; Takane/Young/De Leeuw 1977), sondern gleichzeitig auch für individuelle metrische Daten mit fehlenden Werten geeignet ist (vgl. Schiffman/Reynolds/Young 1981, S. 251).

In der jeweils *zweiten Befragung* wurden schließlich die erstellten MDS-Karten den Probanden zur Beurteilung vorgelegt, was eine adäquate Möglichkeit zur Ergebnisevaluation darstellt und damit als Gütekriterium durchaus verwendet werden kann (vgl. Hörstrup/Mühlhaus 2009, S. 8ff.). Dabei wurden ausschließlich zweidimensionale Darstellungen gewählt, da diese nahezu für alle Probanden mittels des Elbow-Kriteriums der Stress-Werte (vgl. Wagenaar/Padmos 1971; Green 1975), sowie der Pearson r-Werte anhand der Kategorisierung von Tschudi (1972) als angemessen einzustufen waren und sich darüber hinaus auch von ungeübten Personen visuell gut beurteilen lassen. Die Durchführung einer zweiten Befragung zur Ergebnisevaluation war erforderlich, da bei der Verwendung von reduzier-

---

<sup>2</sup> Zur Untersuchung der TV-Sender wurden ARD, Arte, N 24, Phoenix, Pro 7, RTL, Sat 1, Super RTL, Vox und ZDF zur Beurteilung vorgelegt.

<sup>3</sup> Die zu beurteilenden Automobilmarken waren Alfa Romeo, Audi, BMW, Ferrari, Fiat, Honda, Mazda, Mercedes Benz, Opel, Porsche, Renault und VW.

<sup>4</sup> Als Studienfächer wurden Biologie, BWL, Chemie, Elektrotechnik, Fremdsprachen, Geographie, Geschichte, Informatik, Maschinenbau, Mathematik, Medizin, Philosophie, Physik, Soziologie und VWL ausgewählt.

ten Designs bisher *keine* geeigneten Gütekriterien existieren, anhand derer die entsprechend ermittelten Konfigurationen beurteilt werden können.

Um die Eignung der verschiedenen Varianten zur Erstellung reduzierter Abfragedesigns zu untersuchen und Ansatzpunkte zur Bewertung der entsprechenden Konfigurationsgüte ableiten zu können, wird im Folgenden zunächst die Abbildungsgüte der *vollständigen Datensätze* durch zweidimensionale (aggregierte) Konfigurationen als *Referenzpunkt* aufgezeigt. Anschließend wird die Abbildungsgüte der *reduzierten Designvarianten* in zwei Stufen geprüft:

- (1) Prüfung mit Hilfe des Korrelationskoeffizienten von Pearson der vollständigen Ähnlichkeitsdaten ( $\Delta^I$ ) und den euklidischen Modelldistanzen der anhand der Datenfraktion ermittelten Konfiguration ( $D^R$ ).
- (2) Prüfung unter Verwendung der Probandenbewertungen aus den jeweils zweiten Befragungen.

Die nachfolgende Tabelle fasst die im Weiteren verwendeten Notationen bzgl. Ähnlichkeitsdaten, Distanzen der MDS-Konfigurationen und Korrelationen zusammen.

<b>Symbol</b>	<b>Beschreibung</b>
$\Delta$	Matrix der abgefragten Ähnlichkeitsurteile
D	Distanzmatrix einer MDS-Konfiguration
$\Delta^I$	Vollständige individuelle Ähnlichkeitsurteile
$D^I$	Distanzen der MDS-Konfiguration basierend auf den vollständigen individuellen Ähnlichkeitsdaten
$D^A$	Distanzen der MDS-Konfiguration basierend auf den aggregierten Ähnlichkeitsurteilen über alle Probanden
$\Delta^R$	gemäß Designvariante [R] reduzierte Ähnlichkeitsdaten
$D^R$	Distanzen der MDS-Konfiguration basierend auf den gemäß Designvariante [R] reduzierten Ähnlichkeitsdaten
r	Korrelationskoeffizient nach Pearson

Tab. 2: Erläuterung der verwendeten Symbole

### 3.1 Abbildungsgüte zweidimensionaler aggregierter Konfigurationen als Referenzpunkt

Werden zunächst nur die erhobenen *vollständigen* Datensätze betrachtet, so zeigt sich, dass sowohl die Ähnlichkeitsurteile bzgl. der Automobilmarken als auch bei den TV-Sendern und Studienfächern unter Berücksichtigung der durchschnittlichen  $r(\Delta^I, D^I)$ -Werte von .868, .908 bzw. .844 sehr gut anhand von zweidimensionalen Konfigurationen abgebildet werden können (vgl. Tab. 3).

	$r(\Delta^I, D^I)$	$r(\Delta^I, D^A)$	$r(\Delta^I, D^I)$	$r(\Delta^I, D^A)$	$r(\Delta^I, D^I)$	$r(\Delta^I, D^A)$
Objekt	Automobilmarken (n=25)		TV-Sender (n=27)		Studienfächer (n=19)	
<b>Mittelwert</b>	<b>.868</b>	<b>.681</b>	<b>.908</b>	<b>.730</b>	<b>.844</b>	<b>.723</b>
Oberes Quartil	.914	.778	.949	.820	.849	.760
Median	.881	.690	.928	.743	.842	.745
Unteres Quartil	.826	.617	.872	.671	.828	.732
Anzahl Personen mit $r < .700$	1	13	0	9	0	8

Tab. 3: Korrelationswerte der vollständigen Datensätze

Nach der Kategorisierung von Tschudi (1972), der hierfür einen Richtwert von .700 empfiehlt, erscheint nur für einen der 71 untersuchten Ähnlichkeitsdatensätze die Verwendung einer zweidimensionalen Konfiguration *nicht* angemessen. Weiterhin stellen sich die Angaben über alle Probanden als sehr homogen dar. Die Korrelationen  $r(\Delta^I, D^A)$  der individuellen Ähnlichkeitsdaten mit den Distanzen der aggregierten Konfigurationen, die aus den standardisierten Durchschnittsangaben errechnet wurden, sind sowohl bei den Automobilmarken (.681), den TV-Sendern (.730) als auch den Studienfächern (.723) sehr hoch. Trotz dieser hohen Homogenität zeigt sich aber, dass die Berücksichtigung lediglich der aggregierten Daten nicht empfehlenswert ist. Würde diese Vorgehensweise gewählt, so können abhängig von der Objektgruppe die Angaben von 13, 9 bzw. 8 Personen, also etwa 40% der Auskunftspersonen mit Korrelationswerten unter .700 nicht hinreichend adäquat anhand der aggregierten Konfiguration abgebildet werden.

### 3.2 Abbildungsgüte der reduzierten Designvarianten anhand des Korrelationskoeffizienten von Pearson

Im Folgenden wird geprüft, inwieweit die in Abschnitt 2 vorgestellten Varianten reduzierter Designs geeignet sind um die individuellen Einschätzungen der Probanden zu reproduzieren. Zu diesem Zweck wurden für die verschiedenen Designvarianten zunächst reduzierte Ähnlichkeitsfraktionen  $\Delta^R$  selektiert. Auf dieser Basis wird dann geprüft, inwieweit die wahre Datenstruktur anhand des reduzierten Datensatzes auf Individualebene rekonstruiert werden kann. Als Gütekriterium wird der Korrelationskoeffizient nach Pearson  $r(\Delta^I, D^R)$  herangezogen, der aus den vollständigen Ähnlichkeitsurteilen und den euklidischen Modelldistanzen der anhand der Datenfraktion ermittelten Konfiguration gebildet wurde. Nach MacCallum (1978, S. 70) bildet der Korrelationskoeffizient bei der Anwendung reduzierter Abfragen die verlässlichste und insgesamt eine bessere Einschätzung als das Stress-Gütemaß (vgl. Schiffman/Reynolds/Young 1981, S. 175). Weiterhin wurden je Designvariante, in Abhängigkeit vom Reduktionsgrad  $\pi$ , die individuell reduzierten Ähnlichkeitsfraktionen aus den vollständigen Datensätzen ermittelt und mittels ALSCAL individuelle MDS-Konfigurationen erstellt. Die durchschnittlich erzielten Korrelationen  $r(\Delta^I, D^R)$  der vollständigen Ausgangsdaten mit den Distanzen der so erzielten Konfigurationen werden im Folgenden je Objektgruppe besprochen. Da die Ergebnisgüte bei den zufälligen Designs (RAND) und (CYCL) stark davon abhängt, welche Paare berücksichtigt werden, wurden je Person und Reduktionsgrad 50 zufällig komponierte Designs erstellt. In den folgenden Ergebnistabellen 4 bis 6 sind die jeweiligen mittleren Reproduktionsmaße zur Bereinigung von zufälligen Effekten angegeben.<sup>5</sup> Die letzten drei informationsgestützten Verfahren (AED I, AED II und AVD) basieren auf aggregierten Daten, wobei diese, analog zu einer Vorgehensweise in der Praxis, *nicht* aus den gesamten 25, 27 bzw. 19 individuellen Datensätzen, sondern lediglich aus einer Stichprobe derselben (im vorliegenden Fall vom Umfang 5) ermittelt wurde.<sup>6</sup> Diese Vorgehensweise ist vergleichbar mit Methoden des Splithalf oder Bootstrapping, bei denen ein vorhande-

---

<sup>5</sup> Dabei wurden je Person jeweils 50 verschiedene und zufällig ausgewählte Fraktionen der Ähnlichkeitsdaten ermittelt und entsprechende MDS-Konfigurationen erstellt. Anhand der 50 korrespondierenden Reproduktionsmaße wurde daraufhin ein personenabhängiger durchschnittlicher Reproduktionswert in Form des arithmetischen Mittelwertes berechnet.

<sup>6</sup> Die Stichprobengröße wurde dabei mit einem Anteil von unter 30% bewusst gering gewählt, da auch in realen Anwendungen mitunter starke Abweichungen der Stichprobenwerte von denen der Grundgesamtheit bestehen (sog. Stichprobeneffekt).

ner Datensatz in zwei Gruppen unterteilt wird, wobei die eine Gruppe zum Schätzen und die zweite Gruppe zum Validieren entsprechender Ergebnisse verwendet wird (vgl. Efron/Tibshirani 1986; 1993; Shao/Tu 1995). Um auch hier zufällige Effekte auszuschließen, wurden abermals die Mittelwerte aus jeweils 50 Replikationen gebildet und in den Ergebnistabellen ausgewiesen.

Werden nun für jede Person die Reproduktionswerte betrachtet, die mit den unterschiedlich reduzierten Designs in Abhängigkeit des Reduktionsgrades erzielt werden, so sollten zusätzlich zur absoluten Höhe der Reproduktionsmaße  $r(\Delta^I, D^R)$  auch die Korrelationen  $r(\Delta^I, D^A)$  mit den aggregierten Daten zur Beurteilung herangezogen werden. Liegen diese über den individuellen Werten, so erscheint die Anwendung eines reduzierten Designs, selbst bei absolut betrachtet hohen Korrelationswerten wenig zweckmäßig, da dann die individuellen Charakteristika nicht in ausreichendem Maße aufrechterhalten werden und die aggregierten Konfigurationen besser zur Abbildung der Individualwerte geeignet sind.

**(a) Reproduktionsgüte für die Objektgruppe „Automobilmarken“**

Methode zur Ermittlung der $D^R$	Vorinformation	Art	Reduktionsgrad $\pi$				
			.73	.64	.55	.45	.36
RAND	Nein	Zufällig	.442	.548	.577	.648	<b>.712</b>
CYCL	Nein	Zufällig	.427	.539	.566	.635	<b>.734</b>
IDRD I	Nein	Individuell	.461	.558	.639	<b>.700</b>	<b>.738</b>
IDRD II	Nein	Individuell	.636	.654	<b>.692</b>	<b>.741</b>	<b>.763</b>
AED I	Ja	Aggregiert	.633	.663	<b>.696</b>	<b>.711</b>	<b>.715</b>
AED II	Ja	Aggregiert	.640	.663	<b>.695</b>	<b>.710</b>	<b>.790</b>
AVD	Ja	Aggregiert	<b>.721</b>	<b>.728</b>	<b>.749</b>	<b>.760</b>	<b>.777</b>
Durchschnittswerte:			$r(\Delta^I, D^I) = .868$		$r(\Delta^I, D^A) = .681$		

Tab. 4: Reproduktionswerte  $r(\Delta^I, D^R)$  für den Automobildatensatz ( $n=25$ )

Bezogen auf die Ähnlichkeiten der 12 Automobilmarken zeigt sich, dass die Anwendung von zufälligen und zyklisch reduzierten Designs die schlechtesten, aber sehr ähnliche Reproduktionsergebnisse liefert, was den Erkenntnissen von Malhotra/Jain/Pinson (1988) entspricht, die ebenfalls anhand realer Daten zu diesem Schluss kommen. Aufgrund der überschaubaren Objektzahl ist es im vorliegenden

Fall jedoch überraschend, dass das randomisierte Design leicht bessere Ergebnisse liefert. Insgesamt jedoch sprechen die mittleren Korrelationen gegen die Anwendung hoher Reduktionsgrade bei Verwendung zufälliger Designs, sodass erst bei einem hohen Anteil an Ähnlichkeitsurteilen  $\sigma$  von 64%, also einem geringen Reduktionsgrad  $\pi$  von 36%, akzeptable r-Werte über .700 erzielt werden. Zusätzlich sind die mittleren Korrelationen der individuellen Daten bei stärkerer Datenreduktion geringer als mit den aggregierten Daten ( $r(\Delta^I, D^A) = .681$ ). Das bedeutet, dass der individuelle Charakter der Betrachtung verloren geht.

Deutlich bessere Ergebnisse liefern demgegenüber die beiden individuell determinierten Designs, bei denen zusätzlich zu einer zufälligen Auswahl an Paarvergleichen die 10 unähnlichsten (IDRD I) bzw. die 5 unähnlichsten und 5 ähnlichsten Paare (IDRD II) zur Ermittlung der Konfiguration herangezogen wurden. Um beide Varianten miteinander vergleichen zu können, wurden jeweils 10 individuelle determinierte Paare genutzt. Hier zeigt sich, dass entgegen der Empfehlung von Graef/Spence (1979) nicht nur den sehr unähnlichen Paaren eine große Bedeutung beizumessen ist; denn die Nutzung sowohl der ähnlichsten, als auch der unähnlichsten Paare führt hier zu besseren Ergebnissen. Insgesamt scheint die Methode IDRD II mit einem mittleren r-Wert von .692, der nur leicht unter dem Richtwert von .700, aber über dem bei der aggregierten Konfiguration erzielbaren Wert von .681 liegt, bereits bei einem Reduktionsgrad  $\pi$  von 55% relativ gut geeignet. Die kognitiv höhere Beanspruchung der Probanden, die eine Auswahl der 5 ähnlichsten und unähnlichsten Paare darstellt, scheint dabei durch die Option höhere Reduktionsgrade verwenden zu können kompensiert. Verglichen mit den erforderlichen 42 Paarvergleichen bei Anwendung eines zufallsbasierten Designs ( $\pi = 36\%$ ) wären hier, zusätzlich zu den 10 individuell auszuwählenden Paaren, lediglich noch 20 Beurteilungen ( $\pi = 55\%$ ) zur Erzielung einer annähernd gleichwertigen Reproduktionsgüte ausreichend.

Insgesamt zeigen die drei Verfahren zur Ausgestaltung reduzierter Designs, die auf Vorinformationen zurückgreifen, die höchsten Reproduktionswerte, wobei die beiden extremobjektorientierten Designs ähnliche Werte wie IDRD II liefern und ebenfalls bei einem Reduktionsgrad  $\pi$  von 55% geeignet erscheinen. Weit bessere Ergebnisse liefert die varianzorientierte Variante, die bereits bei einem sehr hohen Reduktionsgrad von  $\pi = 73\%$  außerordentlich gute Ergebnisse liefert, was allerdings nicht überrascht. Dies lässt sich darauf zurückführen, dass für dieses

Verfahren die umfangreichsten Informationen, nicht nur hinsichtlich der Struktur der Ähnlichkeitsurteile, sondern auch bezüglich der Heterogenität der zu erwartenden individuellen Angaben, erforderlich sind. Die großen Unterschiede im Vergleich zu den anderen Verfahren sind aber dennoch bemerkenswert.

**(b) Reproduktionsgüte für die Objektgruppe „TV-Sender“**

Bezogen auf die beurteilten Ähnlichkeiten der 10 deutschen TV-Sender bestätigen sich die vorangegangenen Ergebnisse größtenteils. Auch hier zeigt sich, dass die Nutzung von Vorinformationen zur Ausgestaltung von aggregiert varianzbasierten Designs zu den klar besten Ergebnissen führt. Weiterhin wird deutlich, dass die Verwendung von zufälligen Designs lediglich für mittlere Reduktionsgrade im Bereich von  $\pi \approx 33-44\%$  geeignet erscheint, wobei die Ergebnisse hier stark streuen. Sind keine Informationen verfügbar, so sollten die Probanden gebeten werden die Extrempaare auszuwählen und vorab zu bewerten; denn für den vorliegenden Datensatz zeigt die Designvariante IDRD II auch unter Verwendung größerer Reduktionsgrade von  $\pi = 67\%$  bereits akzeptable und bei  $\pi = 44\%$  recht gute Reproduktionseigenschaften, die gegenüber den beiden AED-Varianten deutliche Vorteile zeigen.

Methode zur Ermittlung der $D^R$	Vorinformation	Art	Reduktionsgrad $\pi$			
			.67	.56	.44	.33
RAND	Nein	Zufällig	.516	.626	.676	<b>.756</b>
CYCL	Nein	Zufällig	.435	.554	.724	.726
IDRD I	Nein	Individuell	.629	.723	<b>.766</b>	<b>.820</b>
IDRD II	Nein	Individuell	.702	<b>.758</b>	<b>.794</b>	<b>.811</b>
AED I	Ja	Aggregiert	.506	.586	.698	<b>.749</b>
AED II	Ja	Aggregiert	.652	.672	<b>.760</b>	<b>.789</b>
AVD	Ja	Aggregiert	<b>.816</b>	<b>.831</b>	<b>.840</b>	<b>.864</b>
Durchschnittswerte:			$r(\Delta^I, D^I) = .908$		$r(\Delta^I, D^A) = .730$	

Tab. 5: Reproduktionswerte  $r(\Delta^I, D^R)$  für den TV- Datensatz ( $n=27$ )<sup>7</sup>

<sup>7</sup> Bei diesem Datensatz wurden lediglich 4 Reduktionsniveaus betrachtet, da die Untergrenze des maximalen Reduktionsgrades für die Anwendung des zyklischen Designs zur Erstellung einer 2-dimensionalen MDS-Konfiguration und 10 Objekten bei 67% liegt. Je Objekt müssen dabei mindestens 3 Paarvergleiche zur eindeutigen Objektpositionierung erhoben werden, was zu einer Minimal-

**(c) Reproduktionsgüte für die Objektgruppe „Studiengänge“**

Methode zur Ermittlung der $D^R$	Vorinformation	Art	Reduktionsgrad $\pi$				
			.71	.64	.57	.42	.35
RAND	Nein	Zufällig	.389	.482	.623	.669	.689
CYCL	Nein	Zufällig	.370	.477	.542	.638	.674
IDRD I	Nein	Individuell	.511	.589	.648	.704	.715
IDRD II	Nein	Individuell	.552	.602	.650	<b>.723</b>	<b>.742</b>
AED I	Ja	Aggregiert	.534	.597	.633	.705	<b>.764</b>
AED II	Ja	Aggregiert	.555	.593	.673	.713	<b>.743</b>
AVD	Ja	Aggregiert	<b>.755</b>	<b>.777</b>	<b>.780</b>	<b>.799</b>	<b>.801</b>
Durchschnittswerte:			$r(\Delta^I, D^I) = .844$		$r(\Delta^I, D^A) = .723$		

Tab. 6: Reproduktionswerte  $r(\Delta^I, D^R)$  für den Studienfachdatensatz ( $n=19$ )

Auch beim dritten Datensatz der beurteilten Ähnlichkeiten von 15 Studiengängen zeigen sich Ergebnisse, die konsistent zu den beiden vorangegangenen sind. Wiederum sind die Rekonstruktionswerte beim aggregiert varianzorientierten Designs deutlich besser, als die alternativer Varianten, und dessen Anwendung scheint auch bei hohen Reduktionsgraden geeignet. Die beiden Arten zufällig reduzierter Designs liefern abermals vergleichbare Ergebnisse, die jedoch selbst bei Rückgriff auf 65% ( $\pi = 35\%$ ) der Ähnlichkeitsurteile mit .689 bzw. .674 nur mäßige Reproduktionswerte aufweisen. Da diese zudem geringer als die mittleren Korrelationen von .723 mit den aggregierten Daten sind, spricht dies gegen deren Verwendung. Deutlich bessere Werte verzeichnet das IDRD II, das bei einem Reduktionsgrad von 42% die Struktur der Individualdaten adäquat aufrechterhält. Alle anderen Designvarianten liefern im vorliegenden Fall nur mäßige bzw. schlechte Ergebnisse und sollten, wie bei den vorab untersuchten Datensätzen in Verbindung mit hohen Reduktionsgraden, nicht genutzt werden.

---

zahl von 15 Gesamtvergleichen und somit zu einem maximalen Reduktionsniveau von  $\pi^{\max} = 1 - ((3 \cdot 10 / 2) / 45) = .67$  führt.



### 3.3 Würdigung der empirischen Abbildungsergebnisse und Auswahl der „besten Designs“

Über alle drei Objektgruppen hinweg zeigt sich, dass in Abhängigkeit davon, ob und welche Informationen a priori verfügbar sind, unterschiedliche Ansätze zur Ausgestaltung reduzierter Designs gewählt werden sollten. Die in der Literatur zur Anwendung der MDS größtenteils diskutierten zufälligen und zyklischen Designs (vgl. MacCallum 1978; Malhotra/Jain/Pinson 1988; Burton 2003) stellen sich dabei jeweils als die schlechtesten Optionen heraus, die nur bei moderaten Reduktionsgraden zwischen 30 und 40% akzeptable Ergebnisse liefern. Als überlegene Alternative hierzu erweist sich das individuell determinierte Design (IDRD II) unter Verwendung der ähnlichsten, sowie der unähnlichsten Paarvergleiche, das auch bei deutlich geringeren Anteilen an Ähnlichkeitsurteilen eine adäquate Rekonstruktion der vollständigen Daten gewährleistet. Einzig die aggregiert varianzorientierten Designs (AVD), bei denen basierend auf einer durchschnittlichen Ähnlichkeitsmatrix nur die trennschärfsten Paarvergleiche explizit erhoben werden, stellen sich als noch besser heraus.

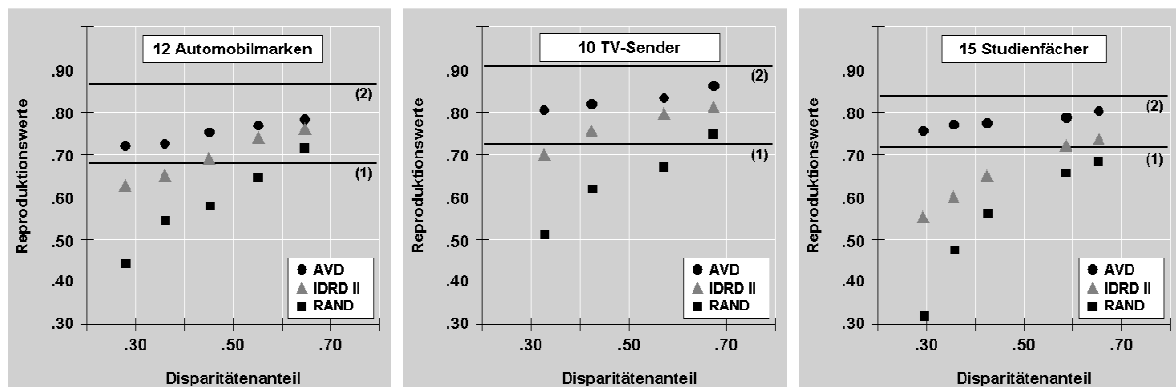


Abb. 2: Reproduktionswerte der Designvarianten bei unterschiedlichen Objektgruppen

Diese Variante liefert selbst dann sehr gute Ergebnisse, wenn nur ein sehr geringer Teil an individuellen Daten genutzt wird, der sogar noch unterhalb der von Tsogo/Masson/Bardot (2001) mit  $(r \cdot (n-2) + 1)$  angegebenen Mindestzahl an Paarvergleichen zur Anwendung der MDS liegt. Abb. 2 veranschaulicht abschließend die Zusammenhänge nochmals graphisch, wobei zusätzlich zur Entwicklung der Reproduktionsmaße, in Abhängigkeit des Reduktionsgrades, auch (1) die mittleren Korrelationen  $r(\Delta^I, D^A)$  der individuellen Angaben mit den anhand der aggregierten Daten er-

stellten sowie (2) mit den vollständigen individuell ermittelten Konfigurationen  $r(\Delta^I, D^I)$  als Bezugswerte ausgewiesen sind.

### **3.4 Prüfung der Abbildungsgüte der reduzierten Designvarianten unter Verwendung von Probandenbewertungen**

Zur Validierung der bisherigen Ergebnisse, die lediglich auf statistischen Kennzahlen beruhen, werden im Folgenden zusätzlich die Probandenbeurteilung der vorgelegten Konfigurationen herangezogen. Obwohl die Eignung einer solchen Vorgehensweise kontrovers diskutiert wird (vgl. Summers/MacKay 1976; Wilkes/Wilcox 1977), erscheint es plausibel, dass diejenigen mittels MDS erstellten Wahrnehmungskarten, die aus Sicht der Auskunftspersonen *nicht* ihrer subjektiven Einschätzung entsprechen, einen nur geringen Wert aufweisen. Ein Untersuchungsdesign, das Daten von zwei aufeinander aufbauenden Erhebungen erfordert, ist sehr aufwendig und komplex, scheint aber insbesondere zur Eignungsprüfung reduzierter Abfragedesigns sinnvoll, da hier nur bedingt Rückschlüsse über die Güte der ermittelten Konfigurationen getroffen werden können. So zeigt sich, insbesondere bei Verwendung von stark reduzierten Designs, dass die wenigen berücksichtigten Ähnlichkeitsurteile, gemessen an den Korrelationswerten besonders gut mittels MDS-Konfiguration abgebildet werden können. Inwieweit jedoch die gesamte Objektmenge angemessen positioniert ist, kann nicht verlässlich abgeschätzt werden, da den vollständigen Distanzwerten, die aus der entsprechenden Konfiguration resultieren, keine empirischen Referenzähnlichkeitsurteile gegenübergestellt werden können. Um jedoch einen Anhaltspunkt über die Eignung der verschieden ausgestalteten, reduzierten Abfragedesigns und einen Vergleich zwischen diesen zu erhalten, erscheint die Verwendung einer probandenbezogenen Konfigurationsbeurteilung zweckmäßig.

Im Folgenden werden von den bisher betrachteten Varianten reduzierter Designs nur jeweils die besten Methoden bei Verfügbarkeit von Vorinformationen (AVD) bzw. ohne Vorinformationen (IDRD II) sowie zu Vergleichszwecken das zufällig reduzierte Design (RAND) herangezogen. Zu diesem Zweck wurde eine weitere Studie via Online-Befragung durchgeführt, an der sich insgesamt 41 Studierende der Universität Trier aus verschiedenen Fachbereichen im Alter von 24 bis 30 Jahren beteiligten. Zunächst wurden bei jedem Befragten die Ähnlichkeitsdaten zu den drei Objektgruppen (TV-Sender, Automobilmarken und Studienfächer) erhoben. Dabei sollten die

Probanden je Objektgruppe in einem ersten Schritt die fünf unähnlichsten und unähnlichsten Paare identifizieren und anhand einer 7-stufigen Ratingskala die Ähnlichkeit beurteilen, was dem ersten Schritt der Vorgehensweise beim IDRDR II entspricht. Hiernach wurden alle verbleibenden Paarvergleiche erhoben, sodass für jede Person und Objektgruppe jeweils eine vollständige Ähnlichkeitsmatrix vorliegt. Basierend auf diesen vollständigen Angaben wurden dann zunächst individuelle MDS-Konfigurationen (RAND mit  $\pi=0$ ) und die aggregierten Konfigurationen je Objektgruppe (AGGR) ermittelt. Zusätzlich dazu wurden für jede Person gemäß den drei Varianten (RAND, AVD und IDRDR II) je Objektgruppe eine MDS-Konfiguration bei unterschiedlichem Reduktionsgrad ermittelt. Basierend auf den Erkenntnissen der ersten Untersuchung, bei der ein überwiegender Teil der Ähnlichkeitsdaten anhand von 2-dimensionalen Konfigurationen adäquat abgebildet werden konnte wurden diesmal auch nur 2-dimensionale Konfigurationen verwendet. Die Auswahl der trennscharfen Paare beim AVD-Design erfolgte dabei anhand der Daten der ersten Untersuchung. Die mittleren Angaben der verbleibenden (wenig trennscharfen) Paare wurden auch der ersten Untersuchung entnommen, was dann dem Vorgehen in der praktischen Umsetzung des AVD entspricht, bei dem eben gerade diese Informationen einer Vorstudie zu entnehmen sind. In der zweiten Erhebungswelle wurden die entsprechend erstellten 18 MDS-Konfigurationen zur Beurteilung präsentiert.<sup>8</sup> Die Bewertung erfolgte dabei anhand einer 6-stufigen Ratingskala (1=spiegelt meine Einschätzung sehr gut wider bis 6=spiegelt meine Einschätzung überhaupt nicht wider), die sich in der sozialwissenschaftlichen Forschung bewährt hat und aufgrund ihrer Ähnlichkeit zum Schulnotensystem durch eine Vielzahl von Untersuchungen befürwortet wird (vgl. Green/Rao 1970, S. 33ff.; Trommsdorff 1975, S. 93ff.; Weiber/Jacob 2000, S. 557ff.). Um bei der Evaluation der Konfigurationen Lern- und Gewöhnungseffekte auszuschließen, die in einem derartigen Kontext auftreten können (vgl. Summers/MacKay, 1976), erfolgte die Reihenfolge der Präsentation der verschiedenen Karten in randomisierter Form.

Werden zunächst die durchschnittlichen Bewertungen der anhand vollständiger Abfragen erstellten Konfigurationen bei einem Anteil an Ähnlichkeitsurteilen von 100% aus der ersten Befragung betrachtet, so zeigt sich über alle Objektgruppen, dass

---

<sup>8</sup> Die 18 Konfigurationen setzen sich dabei folgendermaßen zusammen: Je Objektgruppe wurden die aggregierte und eine zufallsbasierte erstellte Konfiguration, die individuelle Konfiguration unter Rückgriff auf die vollständige Ähnlichkeitsmatrix und jeweils eine Konfiguration für die drei Varianten reduzierter Designs vorgelegt.

diese am besten beurteilt werden. Diese Beurteilungen, die über alle drei Objektgruppen betrachtet zwischen 1.24 und 2.15 liegen, sind verglichen mit den aggregierten (AGGR) und manipulierten Karten (MAN) signifikant besser, was mittels t-Test zu einem Signifikanzniveau von 0.05 bestätigt werden konnte. Diese signifikant bessere Beurteilung kann als Beleg dafür gesehen werden, dass die Auskunftspersonen unterschiedliche Konfigurationen differenziert beurteilen können und darüber hinaus ihre eigenen Karten wieder erkennen, was letztlich dafür spricht, dass eine derartige Vorgehensweise als angemessen bezeichnet werden kann. Dies bestätigt sich auch anhand der objektübergreifend, mit einem steigenden Reduktionsgrad erwartungsgemäß geringer beurteilten Konfigurationsgüte.

	Fallzahl	Reduktionsgrad $\pi$	RAND	AVD	IDRD II	AGGR	MAN
10 TV-Sender	41	.00	1.41	-	-	2.22	5.12
	15	.22	1.73	1.33	1.73		
	13	.44	2.15	1.69	2.08		
	13	.67	2.69	1.69	2.54		
12 Automobilmarken	41	.00	<b>2.15</b>	-	-	<b>2.34</b>	4.54
	12	.27	<b>3.08</b>	2.25	2.17		
	14	.45	3.21	2.43	2.43		
	15	.64	3.73	2.33	2.73		
15 Studienfächer	41	.00	<b>1.24</b>	-	-	2.12	4.78
	14	.21	1.71	1.36	1.50		
	10	.43	2.60	1.40	1.90		
	15	.64	3.27	1.67	2.47		

Tab. 7: Probandenbezogene Konfigurationsbewertung

Bezogen auf die beurteilte Angemessenheit der Konfigurationen, die anhand verschieden reduzierter Designs ermittelt wurden, bestätigen sich die Ergebnisse der vorangegangenen Untersuchung größtenteils: So zeigt sich, dass die Verwendung von zufälligen Designs die schlechtesten Ergebnisse liefert, wobei die Unterschiede zu den anderen Varianten mit zunehmendem Reduktionsgrad stärker ausfallen. Weiterhin ist hier im Vergleich zu den anderen beiden Designvarianten die Konsistenz der Ergebnisse geringer. So zeigt sich, dass bei den Automobilmarken selbst unter Verwendung eines geringen Reduktionsgrades von  $\pi = 27\%$  die anhand des RAND-Designs erstellten Konfigurationen schlechter beurteilt werden

(3.08) als die aggregierten (2.34), was letzten Endes gegen dessen Verwendung spricht, da hier den probandenbezogenen Unterschieden nicht mehr ausreichend Rechnung getragen wird. Bei den Studienfächern ist dies bei einem Reduktionsgrad von 43% und bei den TV-Sendern gar erst bei 67% der Fall. Sehr viel homogener zeigen sich die Bewertungen bei Anwendung des AVD. Hier können insgesamt die besten Ergebnisse erzielt werden, wobei die Konfigurationen abgesehen vom Automobildatensatz auch bei hoher Reduktion von mehr als 60% verglichen mit den aggregierten Karten besser beurteilt werden. Die individuell determinierten Designs (IDRD II), die ohne Vorinformationen auskommen, können analog zu den Korrelationsergebnissen, die in Abschnitt 3.2 dargestellt wurden, die Struktur der Ähnlichkeiten bei mittleren Reduktionsgraden recht gut erhalten. So werden hier bei den TV-Sendern und den Studienfächern bei Reduktionsgraden von 44 bzw. 43% die entsprechenden Konfigurationen deutlich besser als die aggregierten beurteilt, so dass ein Reduktionsgrad von 50% durchaus noch angemessen erscheint.

Insgesamt ist festzuhalten, dass auch die probandenbezogene Validierung gegen die Eignung zufällig reduzierter Designs spricht, sofern weniger als 70-80% der Ähnlichkeitsurteile abgefragt werden, wobei dies auch in der Inkonstanz der Ergebnisse begründet liegt und somit im Vorfeld einer Erhebung nicht abzuschätzen ist, welcher Reduktionsgrad maximal gewählt werden sollte. Sind keine Vorinformationen verfügbar, so sollten in jedem Fall das IDRD II- Design verwendet werden, bei dem unter Verwendung der Hälfte der Ähnlichkeitsurteile bereits akzeptable Ergebnisse erzielt werden können. Liegen hingegen Informationen vor, welche die trennschärfsten Paarvergleiche sind, so sind auch bei hohen Reduktionsgraden von 60-70% mit der Designgestaltung AVD hervorragende Ergebnisse zu erzielen. Daher sollte bei praktischen Anwendungen dem Aufwand für eine notwendige Vorstudie die entsprechende Zeit- und damit Kostenersparnis bei Verwendung von AVD gegenübergestellt werden. Zusätzlich dazu bietet diese Vorgehensweise nicht nur den großen Vorteil der Ermittlung von kritischen oder relevanten Paarvergleichen, sondern erlaubt auch eine Abschätzung des zweckmäßigen Reduktionsgrades und der geeigneten Dimensionalität der MDS-Konfiguration anhand der Daten der Vorstudie.

#### **4 Entwicklung von Indikatormodellen zur Abschätzung der tatsächlichen Reproduktionsgüte bei metrischen MDS-Konfigurationen auf der Basis reduzierter Designs**

Die in Abschnitt 3.4 vorgenommene probandenbezogene Evaluation der MDS-Konfigurationen ist für die Bestimmung der Reproduktionsgüte reduzierter Designs aus wissenschaftlicher Sicht zwar sehr gut geeignet, in der praktischen Anwendung aus Kosten- und Zeitgründen jedoch kaum praktikabel. Im Folgenden werden deshalb für das IDRD II und das AVD, die im vorangegangenen Abschnitt als „beste Varianten“ reduzierter Designs identifiziert wurden, zwei Indikatormodelle entwickelt, mit deren Hilfe auch ohne probandenbezogene Evaluationen eine Abschätzung der Reproduktionsgüte von reduzierten Designs erreicht werden kann. Damit soll eine weitere Forschungslücke im Bereich der Gütebestimmung von MDS-Konfigurationen mit reduzierten Designs geschlossen werden. Die Aussagekraft der beiden Vorschläge wird mit Hilfe der in Abschnitt 3.2 gewonnenen Ergebnisse unter Verwendung der vollständigen Ähnlichkeitsurteile und der entsprechend reduzierten Fraktionen, sowie der korrespondierenden MDS-Konfigurationen geprüft.

##### ***(a) Konfigurationsbeurteilung beim individuell determinierten Design***

Zur Beurteilung der Konfigurationsgüte des individuell determinierten Designs IDRD II, spezifiziert durch den Korrelationskoeffizienten  $r(\Delta^I, D^{IDRD})$ , kann lediglich auf die Korrelationen der erhobenen Ähnlichkeiten mit den korrespondierenden Distanzen der entsprechend ermittelten Konfigurationen  $r(\Delta^{IR^*}, D^{IDRD^*})$  zurückgegriffen werden. Zusätzlich dazu erscheint eine Berücksichtigung des Reduktionsgrades  $\pi$  zweckmäßig, da die Abweichung von der realen Konfigurationsgüte stärker ausfällt, je kleiner die berücksichtigte Datenfraktion ist. Auch sollte die Objektzahl  $N$  in ein Indikatormodell zur Gütebestimmung integriert werden, da zu vermuten ist, dass auch die absolute Zahl der zur Schätzung der MDS-Konfiguration genutzten Paarvergleiche einen Einfluss auf die Reproduktionsgüte ausübt. Berücksichtigt man diese drei Inputvariablen als Indikatoren, so kann folgender linearer Modellansatz zur Prognose des "wahren"  $r$ -Wertes und damit der tatsächlichen Reproduktionsgüte herangezogen werden:

$$r^{prog.}(\Delta^I, D^{IDRD}) = \beta_0 + \beta_1 \cdot r(\Delta^{IR^*}, D^{IDRD^*}) + \beta_2 \cdot \pi + \beta_3 \cdot N$$

Dieser lineare Modellansatz ist dabei dem der logistischen Regression (vgl. Backhaus et al. 2006; Hosmer/Lemeshow 2000) gegenüber vorzuziehen, da alle partiellen Beziehungen zwischen den drei abhängigen und der unabhängigen Größe beim vorliegenden Datensatz für einen linearen Zusammenhang sprechen. Die Linearitätsannahme wurde dabei anhand der entsprechenden partiellen Streudiagramme überprüft. Der Effekt, dass bei dem gewählten linearen Modellansatz Prognosegrößen außerhalb des zulässigen Bereichs von  $-1$  bis  $+1$  des Korrelationskoeffizienten liegen, kann als vernachlässigbar angenommen werden, da derartige Effekte lediglich bei Extremkonstellationen von Reduktionsgraden über 90% oder unter 10%, sowie bei sehr hohen reduzierten Korrelationswerten im Bereich von 1 auftreten, die für praktische Anwendungen jedoch als irrelevant angesehen werden können. Im Fall der vorliegenden Daten trat dieser Fall erwartungsgemäß nicht auf und alle Prognosewerte lagen hier zwischen  $.153$  und  $.972$ .

Unter Rückgriff auf die vollständig erhobenen Ähnlichkeitsdaten aus Abschnitt 3.2 zeigt sich, dass dieser Modellansatz mit einem Bestimmtheitsmaß von  $R^2 = .858$  die Struktur der Korrelationswerte der vollständigen Ähnlichkeitsurteile und Konfigurationsdistanzen sehr gut abbilden kann, was für dessen Eignung zur Prognose der wahren Werte spricht. Bis auf die Objektzahl weisen alle unabhängigen Variablen einen hoch signifikanten Erklärungsgehalt auf, wobei die entsprechenden p-Werte im vorliegenden Fall aussagekräftig sind, da weder Autokorrelation (vgl. Durbin/Watson 1950; 1951) noch ernsthafte Multikollinearität mit Werten des Variance Inflation Factors deutlich unter 10 vorliegen (vgl. Hair et al. 1998, S. 220f.). Werden weiterhin separate Regressionsschätzungen für die drei Objektgruppen durchgeführt, so zeigt sich, dass die Parameterschätzer der insgesamt vier Regressionen nur sehr geringe Unterschiede aufweisen und sich durchweg  $R^2$ -Werte von über  $.821$  ergeben, sodass die ausgewiesenen Schätzer zumindest für die vorliegenden Daten globale Gültigkeit besitzen.

	Konstante ( $\beta_0$ )	$r(\Delta^{IR^*}, D^{IDRD^*})$ ( $\beta_1$ )	$\pi$ ( $\beta_2$ )	N ( $\beta_3$ )	$R^2$
$\beta$ -Schätzer	-.159	1.297	-.590	.001	.858
p-value	.003	<.001	<.001	n.s.	
Variance Inflation Factor	-	2.663	1.942	1.332	
Durbin-Watson Test	2.213 (p<.001)				

Tab. 8: Prognosemodell bei Anwendung von IDRDR II<sup>9</sup>

### **(b) Konfigurationsbeurteilung beim aggregiert varianzorientierten Design**

Auch bei Verwendung des varianzorientierten Designs AVD kann zur Beurteilung entsprechender Konfigurationen ein Regressionsansatz verwendet werden. Da bei dieser Designvariante zur Ermittlung der Konfiguration eine Mischung aus individuellen und aggregierten Ähnlichkeitsdaten ( $\Delta^{AVD}$ ) verwandt wird, können hier zusätzlich zu den Korrelationen der abgefragten Ähnlichkeiten und den korrespondierenden Modelldistanzen noch die r-Werte der aggregiert-individuell komponierten Ähnlichkeitsdaten mit den ermittelten Konfigurationsdistanzen  $r(\Delta^{AVD}, D^{AVD})$  berücksichtigt werden. Da auch hier ein linearer Ansatz am besten geeignet erscheint, wird folgendes Modell unter Rückgriff auf die in Abschnitt 3.2 dargestellten Korrelationen spezifiziert, wobei zusätzlich abermals der Reduktionsgrad  $\pi$  und die Objektzahl N berücksichtigt werden.

$$r^{prog.}(\Delta^I, D^{AVD}) = \beta_0 + \beta_1 \cdot r(\Delta^{IR^*}, D^{AVD^*}) + \beta_2 \cdot r(\Delta^{AVD}, D^{AVD}) + \beta_3 \cdot \pi + \beta_4 \cdot N$$

Mit einem Bestimmtheitsmaß von .747 erweist sich dieses Indikatormodell zwar nicht ganz so gut wie das beim IDRDR II-Design, kann aber ohne nennenswerte Multikollinearität bzw. ernsthafte Autokorrelation immer noch als gut eingestuft werden. Alle genutzten Parameter weisen dabei einen signifikanten Einfluss auf und sollten zur Prognose der unbekanntenen Rekonstruktionsgüte verwendet werden.

<sup>9</sup> Bei diesem Regressionsansatz lag die Fallzahl insgesamt bei 328, da je Person abhängig von der Objektklasse 4 (TV-Datensatz) bzw. 5 Reduktionsniveaus (Automobildaten und Studienfächer) berücksichtigt wurden.



	Konstante ( $\beta_0$ )	$r(\Delta^{IR^*}, D^{AVD^*})$ ( $\beta_1$ )	$r(\Delta^{AVD}, D^{AVD})$ ( $\beta_2$ )	$\pi$ ( $\beta_3$ )	N ( $\beta_4$ )	$R^2$
$\beta$ -Schätzer	-.536	.235	1.270	-.209	.012	.747
p-value	<.001	<.001	<.001	<.001	<.001	
Variance Inflation Factor	-	2.561	2.678	1.332	1.555	
Durbin-Watson Test	2.043 (p<.001)					

Tab. 9: Prognosemodell bei Anwendung von AVD<sup>10</sup>

Da jedoch in diesem Fall die Parameterschätzer für die drei unterschiedlichen Objektgruppen stark variieren, können die angegebenen Werte nicht auf andere Studien übertragen werden und somit auch nicht als Richt- oder Vergleichsmaße dienen. Dies stellt jedoch kein Problem dar, da bei Anwendung der AVD-Methode ohnehin zumeist spezifische Daten aus einer Vorstudie vorliegen, anhand derer ein objekt- und untersuchungsspezifisches Prognosemodell aufgestellt und die entsprechenden Parameter vorab ermittelt werden können.

## 5 Zusammenfassung und kritische Würdigung der erzielten Ergebnisse

Die vorliegenden, auf empirische Daten gestützten Ergebnisse zeigen, dass die Anwendung reduzierter Abfragedesigns zur Durchführung einer metrischen MDS auch bei Objektmengen von geringem Umfang zwischen 10 und 15 möglich und u. U. auch zweckmäßig ist. Dabei konnte klar herausgestellt werden, dass – sofern individuelle Urteile und Unterschiede im Fokus des Interesses stehen – sowohl die Wahl des Reduktionsgrades, d. h. die Festlegung des Ausmaßes zu berücksichtigender Ähnlichkeitsurteile, als auch die konkrete Ausgestaltung (welche Paarvergleiche erhoben werden sollen), einen großen Einfluss auf die erzielbare Konfigurationsgüte aufweisen. Insbesondere die Anwendung zufälliger oder zyklischer Designs, die in diesem Kontext in der Wissenschaft bisher die größte Aufmerksamkeit erfahren haben, ist als kritisch zu bezeichnen. So zeigt sich, dass abhängig von der Struktur der Ähnlichkeiten bereits ein geringer Anteil fehlender Werte zu einer deutlichen Ergebnisverschlechterung führen kann, was insbesondere dann der Fall ist, wenn relevante

<sup>10</sup> Auch in diesem Fall wurden insgesamt 328 abhängige Daten durch die entsprechend unabhängigen Größen erklärt.

Urteile etwa zu Extrempaaren, die sehr ähnlich oder unähnlich sind, nicht abgefragt werden. Da ohne das Vorliegen von ex ante Informationen diese jedoch nicht bekannt sind, wurde eine Methode zur Erstellung von individuell determinierten Designs (Variante: IDRDR II) entwickelt, welche diesem Problem Rechnung trägt. Hier zeigt sich, dass eine deutliche Ergebnisverbesserung verzeichnet werden kann, sofern die Auskunftspersonen diejenigen Paarvergleiche vorab nennen können, die aus ihrer subjektiven Sicht Extrempaare darstellen und diese anschließend auch bewerten. Selbst bei einer überschaubaren Zahl dieser selbstgewählten unähnlichsten und ähnlichsten Paare von insgesamt 10, die vom Probanden als „beherrschbar“ angesehen werden können, ist dieser Effekt deutlich ausgeprägt. Die Vorteilhaftigkeit dieser Vorgehensweise liegt zum einen in der Möglichkeit, höhere Reduktionsgrade zu verwenden. Zum anderen stellen sich die Ergebnisse als stabiler dar und sind damit weniger vom konkreten Datensatz abhängig, was dafür spricht, dass diese Variante grundsätzlich eine gute Eignung aufweist. Die empirische Untersuchung auf Basis der drei Datensätze zu den Automobilmarken, TV-Sendern und den Studiengängen hat weiterhin gezeigt, dass unter Verwendung von IDRDR II bereits 50% der vollständigen Ähnlichkeitsurteile ausreichend, um die ursprüngliche Struktur der Ähnlichkeitsdaten ausgezeichnet reproduzieren zu können. Diese Aussage stützt sich dabei sowohl auf die Korrelationsmaße als auch auf die probandenbezogene Konfigurationsbewertung und ist damit im Gegensatz zu alternativen Studien sowohl theoretisch begründet als auch praktisch geprüft.

Sind weiterhin Informationen aus einer Vorstudie oder Experteneinschätzungen verfügbar, so sollten diese bei der konkreten Ausgestaltung von reduzierten Designs auf jeden Fall berücksichtigt werden. Von den verschiedenen Möglichkeiten, die hierbei untersucht wurden, zeigte jedoch lediglich die Anwendung der in diesem Beitrag entwickelten Methode des aggregiert varianzorientierten Designs (AVD) eine signifikante Ergebnisverbesserung. Mit Hilfe des AVD-Ansatzes konnten insgesamt die besten Ergebnisse erzielt werden. So zeigt sich, dass selbst bei sehr hohen Reduktionsgraden von 60-70% die individuellen Wahrnehmungsunterschiede aufrechterhalten werden. Sowohl gemessen an den Beurteilungen der entsprechend erzielten zweidimensionalen MDS-Karten, die signifikant besser bewertet wurden als die aggregierten Karten, basierend auf den durchschnittlichen Angaben aller Personen, als auch anhand der  $r$ -Werte, die nur geringfügig schlechter als bei Verwendung der vollständigen Ähnlichkeitsurteile ausfielen, konnten mit der Anwendung des AVD-

Ansatzes außerordentlich gute Resultate erzielt werden. Obwohl das AVD aufgrund der Erfordernis von Informationen den aufwendigsten Ansatz darstellt, ist er insbesondere aufgrund der Eignung auch bei sehr hohen Reduktionsgraden und der Stabilität der Ergebnisse für praktische Anwendungen als empfehlenswert zu bezeichnen. Ob der einer Vorstudie geschuldete höhere Aufwand letztendlich durch den verringerten Erhebungsaufwand kompensiert werden kann, ist dabei abhängig von den Charakteristika der Erhebung und sollte fallweise geprüft werden.

Insgesamt kann somit festgestellt werden, dass abhängig von der Verfügbarkeit und dem Ausmaß an Vorinformationen unterschiedliche Optionen bei der Ausgestaltung eines reduzierten Designs gewählt werden sollten. Sind keine Informationen verfügbar, so sollten individuell determinierte Designs genutzt werden. Liegen dagegen Informationen hinsichtlich der trennscharfen Paarvergleiche vor, so sollten diese explizit abgefragt werden. Die nachfolgende Tabelle fasst dies noch einmal zusammen und gibt eine Handlungsempfehlung bzgl. der Wahl der Designvariante in Abhängigkeit vom Untersuchungsfokus, dem erwünschten Reduktionsniveau, sowie der Datenstruktur

Methode	Analysefokus		Reduktionsniveau			Datenstruktur	
	aggregiert	individuell	hoch	Mittel	gering	homogen	heterogen
RAND	*	○	○	○	*	*	○
CYCL	*	○	○	○	*	*	○
IDRD I	**	**	○	*	**	*	**
IDRD II	**	**	*	**	**	*	**
AED I	**	○	○	*	**	**	○
AED II	**	○	○	*	**	**	○
AVD	**	**	**	**	**	*	**
Eignung: ○ wenig geeignet * mittelmäßig geeignet ** sehr gut geeignet							

Tab. 10: Eignung der unterschiedlichen Erhebungsdesigns

Ein bisher in den Studien, die sich mit der Anwendung und Ausgestaltung reduzierter Abfragedesigns im Rahmen der MDS befassen, nicht thematisiertes Problem ist die mangelnde Beurteilbarkeit der Ergebnishüte.

Ohne Kenntnis der Struktur der vollständigen Ähnlichkeitsurteile und ohne eine, für praktische Anwendungen jedoch wenig zweckmäßige, probandenbezogene Ergebnisevaluation kann nicht festgestellt werden, inwieweit anhand reduzierter Abfragen erstellte MDS-Konfigurationen den Wahrnehmungen der Auskunftspersonen entsprechen. Es muss deshalb auf Indikatoren zur Gütebestimmung zurückgegriffen

werden. Bezogen auf die beiden in diesem Beitrag als „besonders geeignet“ identifizierten Varianten IDR2 II und AVD wurde gezeigt, wie unter Rückgriff auf die verfügbaren Informationen bzgl. Korrelationen, Reduktionsgrad oder Anzahl der Objekte gute Indikatormodelle aufgestellt werden können, die Aussagen hinsichtlich der Adäquanz von reduzierten MDS-Konfigurationen erlauben.

Die in diesem Beitrag gewonnenen Ergebnisse sprechen insgesamt für die Anwendung von reduzierten Abfragedesigns, wobei die gewonnenen Erkenntnisse insbesondere aufgrund des Rückgriffs auf empirische Daten aussagekräftiger als die vergleichbarer Simulationsstudien angesehen werden können. Die konkrete Ausgestaltung entsprechender Designs sollte deshalb in Abhängigkeit der Verfügbarkeit von verlässlichen ex ante Informationen erfolgen, wobei die beiden entwickelten Ansätze des IDR2 II und des AVD überaus erfolgversprechend erscheinen. Da die hier verwendete Datenbasis sowohl im Hinblick auf die Fallzahlen als auch im Hinblick auf die unterschiedlichen Objektgruppen beschränkt ist, sollte die Übertragbarkeit der Ergebnisse aber anhand weiterer empirischer Studien geprüft werden. Hierbei sollten der Heterogenität der Ähnlichkeitsurteile bei AVD und die praktische Durchführbarkeit zur Identifizierung von Extrempaaren besondere Aufmerksamkeit gewidmet werden. Auch gilt es zu prüfen, inwieweit die hier für zweidimensionale Konfigurationen erzielten Ergebnisse auch bei mehrdimensionalen Darstellungen bestätigt werden können bzw. wie bei der Anwendung von reduzierten Designs eine geeignete Dimensionalität ermittelt werden kann. Darüber hinaus sollten die dargestellten Ergebnisse, die unter Verwendung der ALSCAL Prozedur erzielt worden, auch anhand von "moderneren" MDS-Varianten (siehe hierzu Borg/Groenen 1997) evaluiert werden. Auch wenn der ALSCAL Algorithmus schon recht alt ist erhielt er für den vorliegenden Beitrag den Vorzug, da er in praktischen Anwendungen am häufigsten eingesetzt und immer noch die am weitesten verbreitete Variante darstellt, was primär darauf zurückzuführen ist, dass er in den gängigen Softwareanwendungen wie z. B. SPSS oder SAS implementiert ist (vgl. Carroll/Green 1997, S. 198).

## Literaturverzeichnis

- Backhaus, K./Erichson, B./Plinke, W./Weiber, R. (2006): *Multivariate Analysemethoden*. 11. Aufl., Berlin: Springer.
- Behrens, C.A. (1986): Shipibo food categorization and preference: Relationships between indigenous and western dietary concepts. *American Anthropologist*, 88, S. 647-658.
- Berekoven, L./Eckert, W./Ellenrieder, P. (2004): *Marktforschung*, 10. Aufl., Wiesbaden: Gabler.
- Borg, I./Groenen, P. (1997): *Modern Multidimensionale Scaling*. Berlin: Springer.
- Burton, M.L. (2003): Too Many Questions? The Uses of Incomplete Cyclic Designs for Paired Comparisons. *Field Methods*, 15, S. 115-130.
- Burton, M.L./Nerlove, S.B. (1976): Balanced designs for triad tests. *Social Science Research*, 5, S. 247-267.
- Carroll, J.D./Green, P.E. (1997): Psychometric Methods in Marketing Research: Part II, Multidimensional Scaling. *Journal of Marketing Research*, 34, S. 193-204.
- Clatworthy, W.H. (1955): Partially balanced incomplete block designs with two associate classes and two treatments per block. *Journal of Research*, National Bureau of Standards, 54, S. 177-190.
- Clatworthy, W.H. (1973): Tables of two-associate-class partially balanced designs. *National Bureau of Standards Applied Mathematics Series*, 63.
- Cliff, N./Girard, R./Green, R.S./Kehoe, J.F./Doherty, L.M. (1977): INTERSCAL: A TSO FORTRAN IV program for subject computer interactive multidimensional scaling. *Educational and Psychological Measurement*, 37, S. 69-90.
- David, H.A. (1963): *The method of paired comparisons*, New York: Hafner.
- David, H.A. (1982): Cyclic designs. In Kotz, S./Johnson, N.L. (Hrsg.), *Encyclopedia of statistical sciences* (S. 256-260). 2. Aufl., New York.
- DeSarbo, W.S./Young, M.R./Rangaswamy, A. (1997): A Parametric Multidimensional Unfolding Procedure for Incomplete Nonmetric Preference & Choice Set Data in Marketing Research. *Journal of Marketing Research*, 34, S. 499-516.
- Dong, H.K. (1983): Method of Complete Triads: An Investigation of Unreliability in Multidimensional Perception of Nations. *Multivariate Behavioral Research*, 18, S. 85-96.

- Durbin, J./Watson, G. S. (1950): Testing for Serial Correlation in Least Squares Regression, I. *Biometrika*, 37, S. 409-428.
- Durbin, J./Watson, G. S. (1951): Testing for Serial Correlation in Least Squares Regression, II. *Biometrika*, 38, S. 159-179.
- Efron, B./Tibshirani, R. (1986): Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1, S. 54-75.
- Efron, B./Tibshirani, R. (1993): *An Introduction to the Bootstrap*, London: Chapman and Hall.
- Girard, R./Cliff, N.A. (1976): A Monte Carlo evaluation of interactive multidimensional scaling. *Psychometrika*, 41, S. 43-64.
- Golledge, R.G./Raynor, J.N. (1982): *Proximity and Preference: Problems in the multidimensional analysis of large data sets*, Minneapolis: University of Minnesota Press.
- Graef, J./Spence, I. (1979): Using distance information in the design of large multidimensional scaling experiments. *Psychological Bulletin*, 86, S. 60-66.
- Green, P.E. (1975): On the Robustness of Multidimensional Scaling Techniques. *Journal of Marketing Research*, 12, S. 71-83.
- Green, P.E./Rao, V.R. (1970): Rating Scales and Information Recovery – How many Scales and Response Categories to use?. *Journal of Marketing*, 34, S. 33-39.
- Green, R.S./Bentler, P.M. (1979): Improving the efficiency and effectiveness of interactively selected MDS data designs. *Psychometrika*, 44, S. 115-119.
- Hair, J.F./Anderson, R.E./Tatham, R.L./Black, W.C. (1998): *Multivariate Data Analysis*, 5. Aufl., New Jersey: Prentice-Hall.
- Handl, A. (2002): *Multivariate Analysemethoden*, Berlin: Springer.
- Hörstrup, R./Mühlhaus, D. (2009): Der Einfluss von Untersuchungsrahmen und Messmethodik auf die Güte von MDS-Konfigurationen, Forschungsbericht Nr. 8, hrsg. von Rolf Weiber, Trier.
- Hosmer, D.W./Lemeshow, S. (2000): *Applied Logistic Regression*, 2. Aufl., New York.
- Hussey, M./Hooley, G. (1995): The diffusion of quantitative methods into marketing management. *Journal of Marketing Practice. Applied Marketing Science*, 1, S. 13-31.
- Isaac, P.D. (1982): Considerations in the selection of stimulus pair for data collection in multidimensional scaling. In Golledge, R.G./Raynor, J.N. (Hrsg.), *Proximity*

- and Preference: Problems in the multidimensional analysis of large data sets* (S. 80-89). Minneapolis: University of Minnesota Press.
- John, P.W.M. (1980): *Incomplete block designs*, New York: Marcel Dekker.
- Kruskal, J.B./Carmona, F.J. (1973): *How to Use MDSCAL, A Program to Multidimensional Scaling and Multidimensional Unfolding* (Version 5M), Bell Laboratories, New York.
- MacCallum, R.C. (1978): Recovery of structure in incomplete data by ALSCAL. *Psychometrika*, 42, S. 69-74.
- Malhotra, N.K./Jain A.K./Pinson, C. (1988): The Robustness of MDS Configurations in the Case of Incomplete Data. *Journal of Marketing Research*, 25, S. 95-102.
- McCormick, E.J./Bachus, J.A. (1952): Paired comparison ratings I: The effect on ratings of reductions of pairs. *Journal of Applied Psychology*, 36, S. 123-127.
- McCormick, E.J./Roberts, W.K. (1952): Paired comparison ratings II: The reliability of ratings based on partial pairings. *Journal of Applied Psychology*, 36, S. 188-192.
- Ross, J./Cliff, N.A. (1964): A generalization of the interpoint distance model. *Psychometrika*, 29, S. 167-176.
- Schiffman, S.S./Reynolds, M.L./Young, F.W. (1981): *Introduction to Multidimensional Scaling*, London: Academic Press.
- Shao, J./Tu, D. (1995): *The Jackknife and Bootstrap*. New York: Springer.
- Spector, A.N./Rivizzigno, V.L. (1982): Sampling designs and recovering cognitive representation of an urban area. In Golledge, R.G./Raynor, J.N. (Hrsg.), *Proximity and Preference: Problems in the multidimensional analysis of large data sets* (S. 47-79). Minneapolis: University of Minnesota Press.
- Spence, I. (1983): Monte Carlo simulation studies. *Applied Psychological Measurement*, 7, S. 405-425.
- Spence, I./Domoney, D.W. (1974): Single subject incomplete designs for nonmetric multidimensional scaling. *Psychometrika*, 39, S. 469-490.
- Summers, J.O./MacKay, D.B. (1976): On the Validity and Reliability of Direct Similarity Judgments. *Journal of Marketing Research*, 13, S. 289-295.
- Takane, Y./Young, F.W./De Leeuw, J. (1977): Nonmetric Individual Differences Multidimensional Scaling: An Alternating Least Squares Method with Optimal Scaling Features. *Psychometrika*, 28, S. 7-67.

- Taylor, J.R./Kinnear, T.C. (1971): Empirical Comparison of Alternative Methods for Collecting Proximity Judgments, *Working Paper No. 46*, University of Michigan.
- Trommsdorff, V. (1975): *Die Messung von Produktimages für das Marketing: Grundlagen und Operationalisierung*. Köln: Carl Heymanns Verlag.
- Tschudi, F. (1972): The latent, the manifest, and the reconstructed in multivariate data reduction models, *unpublished Ph.D. thesis*, University of Oslo.
- Tsogo, L./Masson, M.H. (1998): Multidimensional Scaling Methods for Many-Object Sets: a review, *Workingpaper Université de Technologie Compiègne*, Centre de Recherches de Rayallieu UMR CNRS 6599, Compiègne.
- Tsogo, L./Masson, M.H./Bardot, A. (2001): Recovery of the metric structure of a pattern of points using minimal information, *IEEE Trans. Systems, Man and Cybernetics*, 31, S. 30-42.
- Wagenaar, W.A./Padmos, P. (1972): Quantitative interpretation of stress in Kruskal's multidimensional scaling technique. *British Journal of Mathematical and Statistical Psychology*, 24, S. 101-110.
- Weiber, R./Jacob, F. (2000): Kundenbezogene Informationsgewinnung. In Kleinaltenkamp, M./Plinke, W. (Hrsg.), *Technischer Vertrieb - Grundlagen*, (S. 523-612), 2. Aufl., Berlin: Springer.
- Whipple, T.W. (1976): Variation Among Multidimensional Scaling Solutions: An Examination of the Effect of Data Collection Differences. *Journal of Marketing Research*, 13, S. 98-103.
- Wilkes, R.E./Wilcox, J.B. (1977): On the Validity and Reliability of Direct Similarity Judgments: A Comment. *Journal of Marketing Research*, 14, S. 261-262.
- Young, F.W./Cliff, N. (1972): Interactive scaling with individual subjects. *Psychometrika*, 37, S. 385-415.
- Young, F.W./Hamer, R.M. (1987): *Multidimensional Scaling*, New Jersey: Eribaum Associates.
- Young, F.W./Null, C.H./Sarle, W.S./Hoffman, D.L. (1982): Interactively ordering the similarities among a large set of stimuli. In Golledge, R.G./Raynor, J.N. (Hrsg.), *Proximity and Preference: Problems in the multidimensional analysis of large data sets* (S. 10-28). Minneapolis: University of Minnesota Press.



## **Forschungsberichte zum Marketing**

- Nr. 1: Weiber, Rolf/Jacob, Frank:  
Informationsgewinnung im Business-to-Business-Marketing: Grundlagen der Marktforschung, 2. Aufl., Berlin/Trier 2000.  
89 Seiten.
- Nr. 2: Weiber, Rolf/Kollmann, Tobias/Pohl, Alexander:  
Das Management technologischer Innovationen, 2. Aufl., Berlin/Trier 1999.  
128 Seiten.
- Nr. 3: Weiber, Rolf/Kollmann, Tobias:  
Die Vermarktung von Multimedia-Diensten - Akzeptanzprobleme bei interaktivem Fernsehen, Trier 1995. 67 Seiten. (ISBN 3-930230-08-9)
- Nr. 4: Weiber, Rolf/Kollmann, Tobias:  
Akzeptanz von interaktiven Multimedia-Programmen im universitären Einsatz, Trier 1997. 67 Seiten. (ISBN 3-930230-14-3)
- Nr. 5: Weiber, Rolf/Adler, Jost:  
Ausgewählte Probleme des internationalen Marketing, 2. Aufl., Berlin/Trier 1999. 109 Seiten.
- Nr. 6: Kirschner, Uwe C. M.:  
Zur Notwendigkeit der Neupositionierung schienenengebundener Mobilitätsanbieter: Eine strategische Analyse am Beispiel der Deutschen Bahn Reise&Touristik AG, Trier 2000. 180 Seiten. (ISBN 3-930230-17-8)
- Nr. 7: Pohl, Alexander/Mühlhaus, Daniel:  
Modernes Innovationsmarketing im Kontext von Open Innovation, Trier 2008.  
167 Seiten. (ISBN 3-930230-26-7)
- Nr. 8: Hörstrup, R./Mühlhaus, D. (2009):  
Der Einfluss von Untersuchungsrahmen und Messmethodik auf die Güte von MDS-Konfigurationen, Forschungsbericht Nr. 8, hrsg. von Rolf Weiber, Trier.  
40 Seiten. (ISBN 3-930230-27-5)
- Nr. 9: Mühlhaus, D./Hörstrup, R. (2009):  
Konzeption und Evaluation reduzierter Abfragedesigns für MDS-Anwendungen, Forschungsbericht Nr. 9, hrsg. von Rolf Weiber, Trier.  
43 Seiten. (ISBN 3-930230-28-3)

## **Arbeitspapiere zur Marketingtheorie**

- Nr. 1: Weiber, Rolf:  
Was ist Marketing - Ein informationsökonomischer Erklärungsansatz, 2. Aufl., Trier 1996. 120 Seiten. (ISBN 3-930230-06-2)
- Nr. 2: Weiber, Rolf/Pohl, Alexander:  
Leapfrogging bei der Adoption neuer Technologien - Theoretische Fundierung und empirische Prüfung, 2. Aufl., Trier 1994. 60 Seiten. (ISBN 3-930230-02-X)
- Nr. 3: Adler, Jost:  
Informationsökonomische Fundierung von Austauschprozessen im Marketing, Trier 1994. 110 Seiten. (ISBN 3-930230-03-8)
- Nr. 4: Pohl, Alexander:  
Ausgewählte Theorieansätze zur Erklärung des Nachfragerverhaltens bei technologischen Innovationen, Trier 1994. 163 Seiten. (ISBN 3-930230-04-6)
- Nr. 5: Beinlich, Georg:  
Geschäftsbeziehungen - Ein integrativer Überblick auf Basis der politischen Ökonomie, Trier 1995. 69 Seiten. (ISBN 3-930230-05-4)
- Nr. 7: Kollmann, Tobias:  
Die Akzeptanz technologischer Innovationen - eine absatztheoretische Fundierung am Beispiel von Multimedia-Systemen, Trier 1996. 146 Seiten. (ISBN 3-930230-12-7)
- Nr. 8: Raff, Tilmann:  
Informationsökonomische Fundierung nachfragerseitiger Unsicherheitspositionen im Systemgeschäft, Trier 1998. 105 Seiten. (ISBN 3-930230-15-1)
- Nr. 9: McLachlan, Christopher:  
Die Gestaltung von Informationsasymmetrien durch das Marketing, Trier 1999. 114 Seiten. (ISBN 3-930230-16-X)