

Einige Anmerkungen zur Praxis demoskopischer Gutachten

I. Einleitung

Dr. Robert Schweizer kam 1971 zur damaligen Infratest GmbH & Co., Institut für Markt- und Sozialforschung in München. Wie immer sein Vertrag ausgesehen haben mag, für die Mitarbeiter war er der „Verwaltungschef“ und - wie in vielen Unternehmen mittlerer Größe - gleichzeitig Leiter des Rechnungswesens und letztlich „Mann für alles“, was nicht mit der unmittelbaren Forschungsarbeit zusammenhing. Ein Arbeitsgebiet also, von dem man meinen sollte, daß es selbst einen 10-Stunden-Arbeitstag mehr als ausfüllt. Aber nicht so bei Dr. Robert Schweizer.

Neben all seinen Aufgaben in der Geschäftsleitung entdeckte und konzipierte er für Infratest das Feld der „Rechtstatsachenforschung“, zunächst - weil damals naheliegend - auf dem Gebiet des Kartellrechts. Stichwort war der „relevante Markt“, den er nicht von Juristen und Wirtschafts-Verbänden definiert wissen sondern empirisch untersuchen lassen wollte, durch Befragung der im Markt agierenden Unternehmen selbst. Weiteres Ziel - erst später verwirklicht - war die Anwendung der Rechtstatsachenbefragung als Instrument zur Informationsbeschaffung bei Gesetzesvorhaben, als Begleitforschung während der Umsetzungsphase in der Verwaltung und letztlich zur Erfolgskontrolle hinsichtlich der Erreichung des beabsichtigten Zweckes.

Schon damals aber gab es bei Infratest bereits empirische Forschung auf dem Gebiet des Wettbewerbs- und Warenzeichenrechts, später subsumiert unter dem Begriff „Empirische Rechtsforschung“; eine Bezeichnung, die von den (wenigen) Instituten, die in Deutschland auf diesem Arbeitsgebiet tätig sind, übernommen worden ist.

Und aus dieser Zeit datiert der persönliche Kontakt des Verfassers dieser Zeilen mit Dr. Robert Schweizer, ein Kontakt, der sich über nunmehr Jahrzehnte hinweg zu einer beruflichen Freundschaft entwickelt hat, von der jeder (jeweils

zum Lobe des anderen) behauptet, er habe den größeren Nutzen davon. Die Themen und die Arbeitsweisen der Markt- und Sozialforschung, die Doktor, heute Professor, Robert Schweizer während seiner Tätigkeit für Infratest (heute Infratest Burke) kennenlernte, haben ihn nie mehr losgelassen und sind wohl auch der Grund für sein späteres Engagement in den Vereinigungen der Markt- und Sozialforschungs-Institute und der in Instituten tätigen Markt- und Sozialforscher. Professor Schweizer ist aber nicht nur Freund der Umfrageforschung, sondern auch ihr scharfer und kritischer Beobachter, nicht nur in seiner Funktion als Auftraggeber für alle von ihm beratenen Verlage, sondern auch aus allgemeinem sozialwissenschaftlichen Interesse heraus, einem Interesse, das zunächst zu einem Lehrauftrag und schließlich zu einer Honorarprofessur an der Ludwig-Maximilians-Universität in München geführt hat.

Professor Schweizer blieb juristischer Berater von Infratest Burke, insbesondere für Fragen des Standes- und des Datenschutzrechts, und gibt - wann immer darum gebeten - Hilfestellung zum Verständnis der juristischen Hintergründe bei schwierigen Aufgabenstellungen der Empirischen Rechtsforschung.

Daß das Denken in juristischen Kategorien nicht allen Markt- und Sozialforschern liegt, geschweige denn besondere Freude macht, sei nicht verschwiegen. Die Unbekümmertheit der frühen Marktforschungsjahre, in denen man Umfragen durchführen und Ergebnisberichte verfassen konnte, ohne sich Gedanken über etwaige wettbewerbsrechtliche Konsequenzen machen zu müssen, wo Datenschutz noch keine gesetzliche Vorschrift gleichwohl aber ethischer Grundsatz und unantastbare Praxis war, ist manchen Marktforschern noch in so guter Erinnerung (oder Inhalt von Erzählungen über die vermeintlich, „gute alte Zeiten“), daß es ihnen schwerfällt, in den zahlreichen Regelungen und Beschränkungen, denen Forschungsarbeit heute unterworfen ist, einen Fortschritt zu sehen. So beklagt man denn - in Einklang mit manchen anderen Berufen - die „Verrechtlichung unserer Arbeit“ und weiß doch, daß sich diese Entwicklung nicht zurückschrauben läßt.

Aber wenn es jemandem gelingt, auf rechtliche Gegebenheiten und Erfordernisse nicht nur warnend hinzuweisen, sondern sie auch zu erläutern und obendrein noch praktikable und erfolgversprechende Lösungen aufzuzeigen, dann ist dies Professor Dr. Robert Schweizer. Wobei man manchmal nicht weiß, was man mehr bewundern soll; diese seine Vorzüge oder die Zeit und vor allem Geduld, die er immer wieder aufbringt, wenn jemand zwar gehalten ist, seinen Rat zu suchen, es letztlich aber lieber hätte, ihn nicht befolgen zu müssen, weil es „ohne Juristerei“ doch viel schöner, sprich: einfacher sei.

So bleibt zum Schluß, ihm für all das zu danken, was er der Marktforschung an Rat und Tat gewährt; Dank im eigenen Namen und auch im Namen derer, die vielleicht nicht immer wissen, wieviel sie Robert Schweizer verdanken.

II. Die Problematik

Daß die Ergebnisse demoskopischer Umfragen überall dort ein geeignetes Beweismittel darstellen, wo es auf Wertvorstellungen der Allgemeinheit oder die Auffassung bestimmter beteiligter Verkehrskreise rechtlich ankommt, wird spätestens seit den achtziger Jahren kaum mehr bezweifelt. Die Hersteller von Waren und die Anbieter von Dienstleistungen lassen ihre Chancen in wettbewerbs- und markenrechtlichen Streitfällen durch Umfragen untersuchen, Gerichte erlassen Beweisbeschlüsse, nach denen ein demoskopischer Sachverständiger mittels Umfrage ein Gutachten über ein fest umrissenes Beweisthema vorzulegen hat.

Die generelle Methodik solcher Umfragen, nämlich die Auswahl von Befragungspersonen, die einen repräsentativen Querschnitt der jeweiligen Gesamtheit darstellen und die Befragung dieser Personen nach einem einheitlichen Frageschema durch geschulte Interviewer, stehen - als anerkannte Methoden der Empirischen Sozialforschung - ebenfalls außer Zweifel.

Dennoch werden demoskopische Gutachten des öfteren kritisiert, im allgemeinen unter drei Aspekten:

- Zu wenig Befragungsfälle, deshalb „nicht repräsentativ“
- „Zu enge / falsche Definition“ des Verkehrskreises
- „Suggestive“ Fragestellung

III. Zum Umfang der Stichprobe

Die Kritik an der vermeintlich „zu geringen Fallzahl“ geht meist auf das Mißverständnis zurück, die Repräsentativität einer Stichprobe sei unterhalb einer bestimmten Fallzahl a priori nicht gegeben; wobei oft noch darauf verwiesen wird, daß es „üblich“ sei, mindestens 2.000 Personen zu befragen. Kaum jemand weiß, daß diese „üblichen 2.000“ nur „gegriffen“ sind, nämlich als eine rein unter Nützlichkeitsaspekten festgelegte Zahl von Befragungsfällen bei Umfragen, die sich zwar an einen Querschnitt der Gesamtbevölkerung wenden, deren eigentlicher Wert aber darin besteht, die Ergebnisse daraus gebildeter

Teilgruppen zu betrachten. Solche Teilgruppen werden in der Markt- und Sozialforschung entweder nach demographischen Merkmalen gebildet (Geschlecht, Alter, Beruf, Schulbildung usw.) und / oder nach Einstellungs- und Verhaltensmerkmalen. Derartige Teilgruppen interessieren aber in der Empirischen Rechtsforschung nicht; deshalb können sie auch nicht als Begründung dafür erhalten, mindestens 2.000 Personen zu befragen.

Größere Gesamtstichproben sind nur dann erforderlich, wenn daraus *Verkehrskreise* gebildet werden sollen, die - meistens durch Besitz- oder durch Konsum-Merkmale definiert - nur einen Teil der Gesamtbevölkerung umfassen und wenn es keine Möglichkeit gibt oder es nicht sinnvoll erscheint, diese Populationen durch eine gezielte Stichprobe zu ermitteln, also ohne Fehlkontakte mit nicht betroffenen Personen. So erhält man z. B. (Stand 1998) aus einer Gesamtbevölkerungs-Stichprobe von 2.000 bzw. (seit der Wiedervereinigung) 2.500 Personen ab 14 Jahren (14 oder 16 Jahre sind die übliche untere Altersbegrenzung einer „Erwachsenen“-Stichprobe) ca. 35 % Nutzer von Personal-Computern (PCs), also etwa 875 Personen aus insgesamt 2.500.

Alle Stichproben-Ergebnisse sind zwangsläufig mit einer statistischen Fehlertoleranz behaftet. Die daraus resultierenden Schwankungsbreiten lassen sich berechnen. Demoskopische Gutachten enthalten dazu in der Regel eine sogenannte Fehlertoleranz-Tabelle. Aus dieser ist ablesbar, wie die Fehlertoleranz mit sinkender Fallzahl unterproportional zunimmt (also auch bei steigender Fallzahl nur unterproportional sinkt) und - was vielfach übersehen wird - wie sie zusätzlich von der jeweiligen Merkmalsverteilung abhängt, also vom jeweils gemessenen Ergebnis. Die Fehlertoleranz ist - bei jedwedem Stichprobenumfang - um so größer, je näher die Merkmalsverteilung bei 50:50 liegt (also z. B. bei der 50 % igen Verkehrsdurchsetzung einer Bezeichnung) und sie ist um so geringer, je mehr sich das durch die Umfrage gemessene Ergebnis der Relation 100:0 annähert.

Mit anderen Worten: Die Aussage, daß eine Marke einen Bekanntheitsgrad an 95 % erreicht hat, ist nicht nur vom Ergebnis her besser sondern auch statistisch sicherer als ein Bekanntheitsgrad von 50 %, wenn beide Aussagen auf der jeweils gleichen Zahl von Befragungsfällen beruhen. Handelt es sich in beiden Fällen z. B. um 300 befragte Personen (300 Fälle stellen eine auch unter Kostengesichtspunkten sinnvolle Untergrenze dar, die nur in Ausnahmefällen unterschritten werden sollte), dann beträgt die statistische Schwankungsbreite bei einem 95 % - Ergebnis $\pm 3,6 \%$ und bei einem 50 %- Ergebnis $\pm 8,2 \%$.

Bei gemessenen 50 % liegt der „wahre“ Bekanntheitsgrad also zwischen 41,8 und 58,2 %. Dabei spielt es keine Rolle, ob *insgesamt* nur 300 Interviews stattgefunden haben oder ob die 300 Befragungsfälle als *Teilgruppe* (Verkehrskreis) aus einer größeren Stichprobe (z. B. 2.500) gebildet worden sind. Maßgebend für die Fehlertoleranz ist immer die Fallzahl, die = 100 gesetzt wird, um die Ergebnisse in Prozent darzustellen.

Die Ansicht, eine Umfrage sei „nicht repräsentativ, weil nur 300 Personen befragt wurden“ (oder „nicht die üblichen 2.000“) ist also bezüglich der dabei unterstellten Kausalität falsch. Richtig ist, daß Ergebnisse kleinerer Stichproben mit größeren Fehlertoleranzen behaftet sind als solche, die auf höheren Fallzahlen beruhen. Ob dies die Verwendbarkeit einer Umfrage, die z. B. 300 Fälle umfaßt, beeinträchtigt, hängt einerseits von den Ergebnissen ab und andererseits von rechtlichen Anforderungen. Wenn eine statistisch gesicherte Verkehrsgeltung von mindestens 50 % benötigt wird, dann kann dafür ein *gemessenes* 50 % - Ergebnis niemals ausreichen, denn selbst bei 10.000 Befragungsfällen beträgt die Fehlertoleranz noch $\pm 1,4 \%$; statistisch „sicher“ sind demnach nur $50 \% - 1,4 \% = 48,6 \%$.

Andererseits wäre ein gemessenes Ergebnis von 60 % bereits bei 300 Fällen durchaus beweiskräftig, denn dabei beträgt die Fehlertoleranz $\pm 5,8 \%$, so daß $60 \% - 5,8 \% = 54,2 \%$ als gesicherte Verkehrsgeltung zur Verfügung stünden.

VI. Zur Definition des Verkehrskreises

Richterliche Beweisbeschlüsse, die die Erstellung eines demoskopischen Gutachtens beinhalten, beziehen sich in der Regel auf „die angesprochenen Verkehrskreise“, ohne das Angesprochensein im einzelnen zu definieren. Welche Definition der durchzuführenden Befragung dann zugrundegelegt wird, hängt manchmal von den Vorschlägen des demoskopischen Sachverständigen ab (sofern sie die Billigung des Gerichts finden), wird aber oft auch dahingehend entschieden, daß „jedermann“ angesprochen sei. Dieses „jedermann“ beschränkt sich keineswegs auf Fälle, in denen es um Waren und Dienstleistungen geht, die zum täglichen Bedarf gehören, sondern wird manchmal von Gerichten auch dort unterstellt, wo nachweislich nur ein kleinerer Teil der Bevölkerung zum jeweiligen Zeitpunkt als Konsument oder Nutzer verkehrsbeteiligt ist.

Der demoskopische Sachverständige kann dann allenfalls erreichen, daß zusätzlich zum geforderten Gesamtbevölkerungs-Ergebnis auch ein solches für

den (aus seiner Sicht tatsächlich) beteiligten Verkehrskreis erhoben und im Gutachten ausgewiesen wird. Welches der beiden Ergebnisse das Gericht dann zu seiner Entscheidungsfindung heranzieht, entzieht sich - selbstverständlich - dem Einfluß des Sachverständigen.

In der Praxis läuft das darauf hinaus, daß viele marken- und wettbewerbsrechtliche Streitfälle, in denen demoskopische Umfragen als Beweismittel dienen, schon über die Definition des Verkehrskreises entschieden werden. Je enger nämlich der Verkehrskreis, desto höher die Bekanntheit und die Verkehrsgeltung, desto besser die Information und desto stärker der von der Information ausgehende Einfluß auf Irreführungs- und Verwechslungsgefahr. Je weiter der Verkehrskreis - bis hin zur Gesamtbevölkerung - desto mehr tendieren alle Ergebnisse in die entgegengesetzte Richtung. Die Einbeziehung de facto nicht verkehrsbeteiligter Personen „drückt“ nicht nur die wissensabhängigen Ergebnisse sondern - durch einen erhöhten „Keine Angabe“- Anteil auch die Ergebnisse von Meinungsfragen, auf die es bei der Irreführungs-Problematik vielfach ankommt. (Nicht zuletzt deshalb plädiert in solchen Fällen die beklagte Partei meist für einen möglichst weiten, die Klägerin jedoch für einen möglichst engen Verkehrskreis.)

Bei der Diskussion über die im Beweisbeschluß pauschal erwähnten „angesprochenen Verkehrskreise“ begegnet man immer wieder der Auffassung, eine Werbung wende sich schon deshalb an die gesamte Bevölkerung, weil sie in auflagenstarken Publikumszeitschriften und / oder im Werbe-Fernsehen zu sehen ist, so daß sie von „jedermann“ wahrgenommen werden könne. Daß der tatsächliche Abnehmerkreis der Ware viel kleiner ist (z. B. für PCs, Skiausrüstungen, Cigaretten) tritt dabei völlig in den Hintergrund. Weist der Sachverständige darauf hin, dann wird entgegnet, daß gleichwohl „jedermann“ in die Lage kommen könne, Käufer oder Nutzer zu werden; dies obwohl offenkundig niemals jedermann Cigaretten rauchen oder skifahren wird und selbst bei PCs eine 100 % ige Verbreitung noch viele Jahre auf sich warten lassen dürfte.

Oft erscheint es auch ausreichend, ein Produkt einfach zum Gegenstand des „täglichen Bedarfs“ zu erklären, um dann - unter Berufung auf diesbezügliche BGH-Entscheidungen - die Gesamtbevölkerung als „verkehrsbeteiligt“ zu erklären. Wie wenig diese Unterstellung gerechtfertigt ist, erweist sich am Beispiel der Cigaretten-Raucher, die nachweislich nur ein knappes Drittel der Bevölkerung umfassen, während die große Mehrheit dem Rauchen mehr und mehr ablehnend gegenübersteht und den Gedanken weit von sich weist, irgendwann selbst Raucher zu werden. Ungeachtet dessen bestanden z. B. Hamburger Gerichte darauf, Werbeaussagen der Cigaretten-Industrie, die auf Irre-

führungsfahrer zu prüfen sind, am Verständnis der Gesamtbevölkerung zu untersuchen, wobei manchmal sogar ausdrücklich abgelehnt wird, die Meinung der Cigaretten-Raucher selbst wenigstens als Teil-Gruppe zusätzlich darstellen zu dürfen.

Weitere Beispiele lassen sich leicht finden. Publikumszeitschriften und Fernsehen werben bekanntlich nicht nur für solche Automarken und -modelle, die für die Masse der Pkw-Besitzer in Frage kommen, sondern auch für ausgesprochene Luxusmodelle mit Preisen, die DM 100.000,- weit überschreiten. Dafür „jedermann“ oder jeden Autofahrer als verkehrsbeteiligt anzusehen, widerspricht zwar der allgemeinen Lebenserfahrung, kommt aber gleichwohl vor.

Ähnlich bei der Werbung für Bausparkassen. Folgte man der Argumentation, „jedermann“ könne irgendwann in die Lage kommen, sich für ein Eigenheim oder für eine Eigentumswohnung zu interessieren und deshalb Bausparer zu werden, dann müßte - weil es Werbung für Bausparkassen schon seit vielen Jahrzehnten gibt - inzwischen nahezu jeder Deutsche Eigentümer eines Hauses oder einer Wohnung oder zumindest schon Bausparer sein. Dem ist bekanntlich nicht so: Der Anteil der Bausparer und der Interessenten dafür liegt seit langem nahezu unverändert zwischen 30 % und 40 % der Bevölkerung; wobei es sich zu einem nicht unerheblichen Teil um Personen handelt, die bereits Immobilien-Besitzer oder Bausparer aus früher datierten Verträgen sind, so daß sie den Kreis der Verkehrsbeteiligten nicht vergrößern. Die gegenwärtige Werbung von Bausparkassen hat also für 60 % bis 70 % der Bevölkerung keinerlei markt- oder rechtsrelevante Bedeutung, weil kein Interesse daran besteht. Dennoch fällt es schwer, die *derzeitigen* Nutzer und Interessenten als Verkehrskreis durchzusetzen, also die Nichtbeteiligten außer acht zu lassen.

Um die Einbeziehung von Nicht-Interessenten in den Kreis der (angeblich) Verkehrsbeteiligten zu rechtfertigen, werden manchmal Überlegungen angestellt, die zwar auf Einzelfälle zutreffen mögen, bezüglich der Gesamtbevölkerung aber irrelevant sind. So wurde z. B. beim Deutschen Patentamt darüber diskutiert, bei einer Verkehrsgeltungs-Untersuchung zum Thema Personal-Computer die Gesamtbevölkerung als Verkehrskreis anzusehen. Auf den Einwand, daß (zum damaligen Zeitpunkt) nur knapp ein Viertel der Bevölkerung entweder privat einen PC besitzt oder beruflich damit zu tun hat, wurde entgegnet, es sei nicht auszuschließen, daß „die Oma“ ihrem Enkel einen PC schenken wolle, so daß man um die Gesamtbevölkerung als Verkehrskreis nicht herumkäme, weil „die Oma“ sonst nicht vertreten sei. Daß es sich dabei wirklich um zu vernachlässigende Einzelfälle handeln dürfte und daß selbst dann „die Oma“ kaum diejenige ist, die den Kauf selber tätigt oder hinsichtlich

der Markenwahl beeinflusst, konnte vermittelt werden. Wäre es bei der Gesamtbevölkerung geblieben, dann hätte kaum ein Anmelder auf diesem Gebiet Aussicht auf ein Bekanntheits- und Verkehrsgeltungs-Ergebnis, das die Eintragung qua Verkehrsdurchsetzung ermöglicht; jedenfalls solange nicht, bis nicht tatsächlich in fast jedem Haushalt ein PC zu finden ist und dieser auch von „jedermann“ genutzt wird.

Erstaunlicherweise wird von der Qualitätsseifen-Entscheidung des BGH¹ sowohl in Beweisbeschlüssen als auch bei der Erörterung demoskopischer Untersuchungsvorschläge wenig Gebrauch gemacht. Dieses schon 1982 ergangene Urteil läßt zu, daß für ein Produkt, das - als solches - zweifellos zum täglichen Bedarf gehört, ein engerer Verkehrskreis gebildet werden darf, wenn es sich im speziellen Fall um ein Erzeugnis handelt, das - wegen seiner Hochpreisigkeit - nicht von „jedermann“ gekauft wird.

Soweit demoskopische Gutachten eingeholt werden, gewinnt man den Eindruck, daß sich Gerichte und Parteien - mit Blick auf die mögliche nächste Instanz - auf der sichereren Seite wähnen, wenn sie im Zweifel eher von der Gesamtbevölkerung ausgehen als von einem an sich naheliegenden engeren Verkehrskreis. Wenn der demoskopische Sachverständige in diesen Fällen erreichen kann, daß zumindest *auch* Ergebnisse für einen engeren Verkehrskreis erhoben und dargestellt werden, dann sieht das zwar nach einem befriedigenden Kompromiß aus, führt aber dazu, daß im späteren Gutachten jeweils *zwei* Ergebnisse „zur Auswahl“ stehen, von denen das eine i. d. R. niedriger und das andere höher ausfällt. Die Entscheidung über den „richtigen“ Verkehrskreis ist also nur bis zur Urteilsfindung vertagt. Daß das Gericht dann dem Resultat zuneigt, das sich mit seiner schon vorher gebildeten Auffassung deckt, ist zumindest als Gedanke nicht von der Hand zu weisen. Das demoskopische Gutachten wird in diesen Fällen für die Prozeßbeteiligten zu einer *zweifach* unbekanntem Größe, einmal hinsichtlich dessen, was die Umfrage als solche ergibt (dieses Risiko liegt in der Natur der Sache), zum anderen aber auch im Hinblick darauf, auf welchen Verkehrskreis sich das Gericht später stützen will, wenn ihm das Gutachten - weil weisungsgemäß so angelegt - dafür die Wahl läßt.

Eine dritte Variante besteht im „Abwägen“ zwischen den Ergebnissen *beider* Gruppen, also zwischen der Gesamtbevölkerung und den tatsächlich Verkehrsbeteiligten. Damit erhält die ohnehin bestehende (aus demoskopischer Sicht zweifelhafte) Tendenz, aus unterschiedlichen Untersuchungsergebnissen Mit-

¹ „Aufmachung von Qualitätsseifen“, BGH-Urteil vom 06.05.1982, GRUR 1982/11, S. 672 ff.

telwerte zu errechnen oder zu schätzen und diese dann zur Entscheidungsfindung heranzuziehen, einen weiteren Impuls.

Anzustreben ist vielmehr, in allen Fällen, in denen das tatsächliche Beteiligtsein der Gesamtbevölkerung nicht wirklich offenkundig ist, zur Entscheidung über den (engeren) Verkehrskreis ein Votum des demoskopischen Sachverständigen einzuholen, der dafür die eigenen Marktkenntnisse (Käufer- und Verbraucheranteile) heranziehen muß. Zusätzlich sollte die Überlegung einbezogen werden, daß Verkehrskreise auch je nach *Beweisthema* unterschiedlich zu definieren sind: Wenn es um die Feststellung von *Bekanntheit* und *Verkehrsgeltung* geht, dann ist auf die aktive Beteiligung am Marktgeschehen abzustellen, also auf Personen, die selbst Käufer sind oder den Kauf durch andere nachhaltig beeinflussen. Eine passive Beteiligung am Marktgeschehen, als bloßer Nutzer dessen, was ein anderer gekauft oder veranlaßt hat, erscheint nicht ausreichend, um dem Begriff der „Verkehrsbeteiligung“ gerecht zu werden.

Wenn diese Definition zu eng erscheint, dann sollten allenfalls noch die *Interessenten* an der jeweiligen Waren- oder Dienstleistungsgruppe hinzugenommen werden; wobei dieses Interesse nicht für bestimmte Personengruppen einfach unterstellt werden kann, sondern durch die demoskopische Untersuchung selbst konkret zu ermitteln ist. So sollte z. B. als „Interessent“ für einen teuren Sportwagen nicht einfach derjenige gelten, der bloßes „Interesse“ am Gegenstand bejaht (wem gefiele ein solches Auto nicht?), sondern es muß sich um Personen handeln, die auf eine entsprechende Fragestellung hin erklären, daß der Kauf eines solchen Fahrzeugs für sie in absehbarer Zeit (oder in einem exakt begrenzten Zeitraum) in Frage kommt. Umgekehrt kann man auch „negatives“ Interesse erfragen. So z. B., ob es für den Betroffenen in Frage kommt, in absehbarer Zeit Raucher oder Skifahrer zu werden, also sich am Verkehr mit der betroffenen Ware zu beteiligen. Personen, die das jeweils Erfragte verneinen, sollten nicht „zwangsweise“ zu Verkehrsbeteiligten erklärt werden.

Trotz der aufgezeigten Alternativen geht die demoskopische Empfehlung aber dahin, bei der Messung von *Bekanntheit* und *Verkehrsgeltung* auf diejenigen abzustellen, die sich *bisher schon* am Verkehr mit der Ware beteiligt haben. Nur wenn dies tatsächlich nahezu bei der gesamten Bevölkerung gegeben ist, sollte sich eine engere Verkehrskreis-Definition erübrigen und die Umfrage bei einem repräsentativen Querschnitt der Gesamtbevölkerung durchgeführt werden.

Anders bei Untersuchungen zur *Irreführungs-* und *Verwechslungsgefahr*. Diese zu prüfen erscheint auch und gerade bei den Personen sinnvoll, die zwar noch

nicht aktiv am Verkehr mit der Ware oder Dienstleistung beteiligt sind, sich aber gedanklich bereits damit beschäftigen und in absehbarer Zeit Kaufentscheidungen treffen wollen. Diese Verbraucher sind für anlehrende (und deshalb verwechselbare) Bezeichnungen, Produktgestaltungen usw. und für mißverständliche (und deshalb irreführende) Werbeaussagen eher anfällig, weil sie noch nicht über eigene Kauf- und Produkterfahrung verfügen. Wer aber - laut eigener Aussage - weder bisher noch in absehbarer Zukunft als Käufer oder Interessent der zu untersuchenden Ware oder Dienstleistung in Frage kommt, sollte auch hier nicht „zwangsweise“ zum Verkehrsbeteiligten gemacht werden und durch seine Angaben mit darüber entscheiden, ob eine *rechtsrelevante* Verwechslungs- oder Irreführungsgefahr besteht. Rechtsrelevant deshalb, weil bekanntlich nicht jedweder Irrtum und jedwede Fehlvorstellung, die irgendwo in der Bevölkerung anzutreffen ist, den Wettbewerb beeinflusst, sondern eben nur eine solche, die im Kreis der wirklich Verkehrsbeteiligten besteht und dort die tolerierbare Mindestgrenze übersteigt.

V. Zur Fragestellung

Frageformulierungen, deren Ergebnisse zu demoskopischen Gutachten führen, geraten schon allein deshalb in die Kritik, weil sie bestenfalls *eine* der streitenden Parteien zufriedenstellen können, kaum jemals beide. Leider erschweren stereotyp und unbegründet vorgebrachte Pauschalurteile („hinlenkend“, „suggestiv“) die ernsthafte kritische Auseinandersetzung mit solchen Frage-
texten. Deshalb zunächst zu einigen Mißverständnissen:

Die Mehrzahl aller demoskopischen Fragestellungen - auch in der Empirischen Rechtsforschung - ist per se „hinlenkend“, nämlich hinlenkend auf das *Thema* und - bei gestützter Fragestellung - auf die Antwortmöglichkeiten, die zur Wahl gestellt werden. Selbstverständlich gibt es auch Fragestellungen, in denen eine solche Hinlenkung nicht erfolgen darf, weil das Ziel der Frage gerade darin besteht festzustellen, wohin die Befragten ihre Gedanken selber „lenken“. Letzteres ist z. B. bei den sogenannten offenen Fragen der Fall, aber auch nur bei den *völlig freien* offenen Fragen und nicht bei den *zielgerichteten*. „Völlig frei“ heißt, daß es dem Befragten überlassen bleibt, wohin er seine Überlegungen lenkt und wie er sie formuliert, „zielgerichtet“ heißt, daß ihm eine *Denkrichtung* vorgegeben wird, ihm aber trotzdem keine *Interpretationsinhalte* zur Auswahl vorgegeben werden. Die völlig freie offene Frage lautet etwa: „Woran denken Sie, wenn Sie ... sehen oder hören?“, zielgerichtet würde die Frage

lauten: „Woran denken Sie, wenn Sie ... im Zusammenhang mit ... hören oder sehen?“.

Eine solche Zielrichtung ist immer dann sinnvoll, wenn der Untersuchungsgegenstand als solcher ein Spektrum eröffnet, das vom eigentlichen Ziel der Untersuchung wegführt und keine verwertbaren Angaben mehr erwarten läßt. So hat es z. B. keinen Sinn, ohne jede Zielrichtungs-Vorgabe nach den Vorstellungen von einer bestimmten Farbe zu fragen, wenn es darum geht, zu erfahren, woran der Verkehr denkt, wenn ihm diese Farbe in Werbeanzeigen begegnet. Selbstverständlich stellt die Einengung auf Werbeanzeigen eine „Hinlenkung“ dar, aber eben eine solche, die der Untersuchungsaufgabe entspricht und ohne die die Untersuchungsaufgabe gar nicht zu lösen ist. Der Vorwurf, eine Frage sei „hinlenkend“, kann deshalb (mit dem Anspruch der Ernsthaftigkeit) nicht pauschal erhoben werden (etwa mit dem Satz: „Es ist bekannt, daß immer offen gefragt werden muß“), sondern nur mit einer konkreten Begründung, warum bei dieser Zielsetzung nicht so hätte gefragt werden dürfen.

Damit erübrigt sich auch die Auseinandersetzung mit dem Vorwurf, *jedwede* Frageformulierung, die vorformulierte Antwortmöglichkeiten zur Auswahl stellt (die sogenannte „gestützte“ oder „geschlossene“ Frage) sei „hinlenkend“ und damit „suggestiv“. Im Gegensatz zur offenen Fragestellung, bei der die Befragten - mit oder ohne Zielrichtung - eigene Vorstellungen entwickeln und diese selbst formulieren, besteht der Sinn und der Zweck der gestützten Fragestellung gerade darin, den Befragten den Formulierungsvorgang abzunehmen und sie statt dessen aufzufordern, zu vorgelegten - i. d. R. alternativen - Interpretationsmöglichkeiten Stellung zu nehmen. Die offene Frage soll also beim Befragten eine *Aktion* auslösen, die geschlossene dagegen eine *Reaktion*, nämlich auf das, was im Fragetext (oder auch auf einer Liste oder Karte) zur Wahl gestellt wird. Es macht also keinen Sinn, dem Instrument „gestützte Frage“ das vorzuwerfen, wozu es dient und was es bewirken soll, nämlich das abwägende Nachdenken über bereits Formuliertes.

Damit ist die gestützte Fragestellung keineswegs der Kritik enthoben. Sie bietet - im Gegenteil - dafür viel mehr Anhaltspunkte als jede Form der offenen Fragestellung. Zu überprüfen sind die zur Wahl gestellten Inhalte als solche und die Texte, mit denen diese Inhalte den Befragten vermittelt werden. Neben den allgemeinen demoskopischen Regeln, daß eine Frage (zumindest für die meisten Befragten) klar und verständlich sein muß (wozu auch gehört, daß *alternative* Antwortmöglichkeiten sich verbal tatsächlich gegenseitig ausschließen), gibt es eine Reihe weiterer Grundsätze, die zu beachten sind:

- Die vorgegebenen Antwortmöglichkeiten müssen in sich *neutral* und *wertfrei* formuliert sein.
- Die zur Wahl gestellten Alternativen müssen - im Hinblick auf die Mehrzahl der denkbaren Interpretationsmöglichkeiten - *vollständig* und - im Vergleich untereinander - *ausgewogen* formuliert sein, so daß keine rational oder emotional begünstigt erscheint.
- Die Möglichkeit, sich für *keine* der angebotenen Antwortvorgaben zu entscheiden, darf nicht nur für den Interviewer im Fragebogen enthalten sein, sondern muß auch dem Befragten ausdrücklich angeboten werden.

Ob es sinnvoll ist, die Befragten zusätzlich zu den oder statt der angebotenen Antwortmöglichkeiten auch noch um andere („sonstige“) Interpretationen zu bitten, hängt vom Einzelfall ab, insbesondere davon, ob zuvor schon die Möglichkeit bestand, sich bei offenen Fragen mit eigenen Formulierungen zu äußern. Wenn letzteres der Fall war, dann erscheint „Sonstiges“ überflüssig, außer es bestand eine (nicht behebbare) Unsicherheit darüber, ob die vorgegebenen Antwortmöglichkeiten tatsächlich die wichtigen Aspekte abdecken.

Mit der Antwortmöglichkeit, daß eine zu untersuchende Aussage inhaltlich „gar nichts bedeutet“, sondern „nur Werbung“ ist, sollte vorsichtig umgegangen werden. Sie ist zwar in bestimmten Fällen unerlässlich, um die Befragten nicht zu Deutungen zu veranlassen, die ihnen sonst ferngelegen hätten, sie darf aber auch nicht dazu führen, daß nur die Klischeevorstellung „alles nur Werbung“ bedient wird, obwohl durchaus auch Inhaltliches zu entnehmen ist und entnommen wird.

Angaben, aus denen auf Verwechslungs- oder Irreführungsgefahr geschlossen wird, sollten im Interview unbedingt *begründet* werden, um zu erkennen, ob die Fehlvorstellung tatsächlich auf das streitige Merkmal zurückgeht oder andere Ursachen hat. Zu den „anderen“ Ursachen zählt auch der bei gestützter Fragestellung oft vermutete „Rate- oder Ja-Sage-Effekt“. Er wird zwar hinsichtlich seiner Auswirkung überschätzt, doch wenn es darum geht, ob die Grenze der unbeachtlichen Minderheit erreicht oder überschritten ist, kommt es auf jeden Prozentpunkt an, so daß ein z. B. 2 % iger Rate-Effekt u. U. prozeßentscheidend sein kann. Wenn also Befragte ihre Fehlvorstellung (die beim Interview selbstverständlich nicht so genannt wird) damit begründen, geraten zu haben („Weil Sie mich so gefragt haben, wird es wohl so sein“), muß das zur Korrektur des ursprünglich gemessenen Wertes herangezogen werden. Inwieweit das auch bei anderen Antworten geschehen sollte - etwa bei der allge-

meinen Äußerung „Das denke ich mir so“ - bleibt der rechtlichen Wertung vorbehalten. Begründungen, die explizit auf andere, nicht zum Beweisthema gehörende Ursachen zurückgehen (z. B. „Weil es die bekannteste Marke ist“), sollte dagegen schon der demoskopische Sachverständige bewerten; wobei es sich von selbst versteht, daß jedwede Korrektur an zunächst gegebenen Antworten nachvollziehbar dokumentiert und begründet sein muß.

Ob auch die *Relevanz* einer beim Interview gemessenen Irreführung zu erfragen ist, hängt oft von rechtlichen Überlegungen ab. Aus demoskopischer Sicht ist die Relevanzfrage immer erwünscht, weil die in jeder gestützten Fragestellung enthaltene Aufforderung, sich zwischen verschiedenen - wenn auch neutral und ausgewogen dargebotenen - Alternativen zu entscheiden und nur wenn dies nicht möglich erscheint, mit „Weiß nicht“ zu antworten, unvermeidbar zu Entscheidungen führt, die der Befragte sonst - d. h. ohne Bewußtmachung der Alternativen - möglicherweise nicht getroffen hätte. Je spezifischer ein Beweisthema (insbesondere in richterlichen Beweisbeschlüssen) formuliert ist, desto mehr kann man davon ausgehen, daß etliche Befragte sich beim Interview *erstmal*s mit diesem Gedanken beschäftigt und auseinandergesetzt haben. (so z. B. mit dem Aspekt, daß die Bezeichnung eines Bieres darauf hinweise, daß „die Brauerei vor 1919 dem bayerischen Königshaus gehört habe“).

Wie die Relevanz zu definieren, also zu erfragen ist, wird unterschiedlich beurteilt. Einerseits heißt es, daß sie nicht nur in einem unmittelbaren Einfluß auf die Kaufentscheidung bestünde, sondern auch schon in der durch die Fehlvorstellung hervorgerufenen intensiveren Beschäftigung mit der angebotenen Ware oder Dienstleistung, andererseits wurden aber Relevanzfragen bei Gericht schon verworfen, weil sie nicht explizit auf die Kaufentscheidung abgestellt waren. Wenn auf die Kaufentscheidung abgehoben wird, dann müßte allerdings zusätzlich ermittelt werden, ob die jeweilige Ware / Dienstleistung von Befragten *überhaupt* genutzt wird, denn nur dann macht seine Angabe einen Sinn, die erfragte Vorstellung spiele für ihn beim Kauf eine Rolle bzw. keine Rolle. Dadurch wird aber die Feststellung der tatsächlichen Relevanz noch komplizierter, denn das Kaufinteresse kann ja auch durch die Fehlvorstellung *erstmal*s geweckt werden, so daß man auch *zukünftiges* Verhalten erfragen müßte und dadurch die Ergebnisse immer unsicherer würden.

Aus demoskopischer Sicht geht die Empfehlung deshalb dahin, die Relevanzfrage vom Aspekt „Kauf“ vollständig zu befreien und sie darauf zu beschränken, ob der erfragte Aspekt für den Befragten „wichtig oder unwichtig“ ist, wobei „unwichtig“ dann auch diejenigen mit erfaßt, die an der Ware / Dienstleistung generell kein Interesse haben. Dadurch würden Fehlvorstellungen, die

inhaltlich sehr fern liegen und die nur deshalb untersucht werden, weil sie in einem Beweisantrag so formuliert und so in den Beweisbeschluß übernommen wurden, hinsichtlich ihrer Relevanz auf das richtige Maß zurückgeführt.

Ein oft übersehener Aspekt bei der Untersuchung der Irreführungsgefahr ist die *Verknüpfung* der auf die offenen Fragen gemachten Angaben mit denen, die - gestützt - auf die vorformulierten Antwortmöglichkeiten entfielen. Das Verfahren, zunächst offen und dann gestützt zu fragen, wird sonst in der Markt- und Sozialforschung aus der Erfahrung heraus angewendet, daß die offene Frage nur das zutage fördert, was akut und vordergründig vorhanden ist, während die anschließende gestützte Fragestellung den Wissens- bzw. Vorstellungsrahmen vollständig ausschöpft, so daß die Kombination aus offener und gestützter Fragestellung (um gleichlautende Doppelnennungen bereinigt) das richtige vollständige Ergebnis liefert.

Dies „funktioniert“ allerdings nur bei einfachen Wissensfragen (z. B. bei der Bekanntheit von Markennamen) problemlos. Wenn Vorstellungen und Meinungen erfragt werden, dann lassen sich die Ergebnisse der beiden Frageformen nicht mehr ohne weiteres kumulieren, denn die Befragten antworten „offen“ in aller Regel nicht in den Kategorien und nicht mit denselben Begriffen, die dann in der gestützten Fragestellung verwendet werden. Nur wenn gleichlautend geantwortet wird (z. B. bei Vorstellungen über die geographische Herkunft), kann man eine direkte Verbindung herstellen zwischen dem Anteil der spontanen und der gestützten Antworten. Wobei aber zu berücksichtigen ist, ob der bei der gestützten Frage vorgegebene Antwortenkatalog Alternativen enthält, die zur „Korrektur“ spontan geäußerter Vorstellungen anregen können. So spaltet sich z. B. die spontan geäußerte pauschale Annahme „aus Italien“ bei der gestützten Fragestellung - wenn dort hypothesengemäß differenziert wurde - auf in „nach einem aus Italien stammenden *Rezept*, aber in Deutschland hergestellt“ und „in Italien selbst *produziert*“.

Auch die Umkehr spontaner Antworten ins Gegenteil ist möglich, nämlich dann, wenn bei der gestützten Fragestellung auch die Gegen-Hypothese (des Prozeß-Gegners) mit zur Wahl gestellt wird und - nach erfolgter Abwägung seitens der Befragten- ebenfalls Stimmen erhält. Tauchte allerdings die Gegen-Hypothese vorher in den spontanen Antworten überhaupt nicht auf, dann ist das ein Indiz dafür, daß sich diese Meinung erst während und nur durch das Interview gebildet hat und entsprechend vorsichtiger bewertet werden muß.

Keineswegs darf daraus aber der Schluß gezogen werden, daß *jede* erst während und durch das Interview entstandene Meinung ein Artefakt und rechtlich

nicht relevant sei. Das Interview ist ein Meßinstrument und Meßinstrumente beeinflussen - nicht nur in den Naturwissenschaften - die Ergebnisse immer. Die Vorstellung, eine Umfrage messe nur Meinungen, die „vorher schon“ vorhanden waren, ist eine Illusion; Forderungen, daß dies gewährleistet sein müsse, sind demzufolge unrealistisch. Abgesehen von Alltags- und akuten Sensationsthemen, die ohnehin in aller Munde sind, muß man vielmehr davon ausgehen, daß alles, was bei einem Interview an Meinungen erfragt und zutage gefördert wird, eine Mischung aus vorhandenem „Basisdenken“ und ad hoc angestellten Überlegungen darstellt. Wäre es anders, bedürfte es also statt wohlüberlegter Frageformulierungen nur bestimmter Stichworte als Schlüsselreize, um „fertig“ vorhandene Meinungen gleichsam abzurufen, dann käme der richtigen Frageformulierung nicht die entscheidende Bedeutung zu, die sie nach demoskopischer und juristischer Erfahrung hat und um die - in Schriftsätzen und bei Erörterungsterminen - so intensiv gerungen wird.