

Perfekt repräsentativ?

Haken und Ösen von Wanderstatistiken am Beispiel zweier Doppelstudien

Rainer Brämer

Wandern als Natur- und Selbsterfahrung

Daten und Fakten
zum sanften Natursport

wanderforschung.de
8/2010, akt. 9/2010, erg. 7/2014

Doppelstudien2010

Grenzen der Profilstudien Wandern	1
Grundsätzliche Repräsentativitätsprobleme	2
Deutscher Wanderverband: Zwei parallele Studien	4
Trendscope: Zwei Studien in einer	5

Grenzen der Profilstudien Wandern

Seit Ende der 90er Jahre stammt das generalisierte Bild vom modernen Wanderer, seiner Befindlichkeit, seinen Gewohnheiten und seinen Erwartungen größtenteils aus den "Profilstudien Wandern". An ihnen hat sich die überfällige Modernisierung touristischer Wanderangebote weitgehend orientiert. Die daraus resultierenden Erfolge stellen letztlich eine Art Praxisbeleg für die grundsätzliche Richtigkeit der empirischen Befunde dar.

Das darf nicht darüber hinwegtäuschen, dass jede Erhebung, die anhand einer wie auch immer ausgewählten Stichprobe verallgemeinerte Aussagen über die entsprechende Grundgesamtheit, in diesem Fall also über alle Wanderer, machen will, mit mehr oder weniger großen Repräsentativitätsproblemen zu kämpfen hat. Es geht dabei vor allem um die Doppelfrage: Welche ist die Zielgruppe der Untersuchung und wie kann man sie in einer Auswahlstichprobe optimal abbilden?

Die Profilstudien Wandern hatten ihre Premiere im Rahmen eines wandertouristischen Pionierprojektes (Rothaarsteig) und sind von daher prinzipiell auf touristische Bedürfnisse zugeschnitten, was nicht heißt, dass ihre Ergebnisse nicht auch außerhalb dieses Bereichs von Bedeutung sind. Konsequenterweise handelt es sich bei ihnen also um Zielort- bzw. Standortstudien: Angesprochen werden aktive Wanderer in der Ausübung ihres Hobbys, also während des Wanderns. Das geschieht an touristisch relevanten Wandwegen in einem solchen Abstand von Ortschaften oder Parkplätzen, der zufällig Vorbeikommende möglichst weitgehend von der Befragung ausschließt. Die Selektion der Stichprobe erfolgt also über die erkennbar ausgeübte Aktivität.

Dabei ist bewusst einkalkuliert, dass Vielwanderer statistisch gesehen eine größere Chance haben, einem Befrager zu begegnen als Wenigwanderer, die Zielgruppe in der Stichprobe mithin entsprechend ihrem Aktivitätsgrad gewichtet wird. Das entspricht der touristischen

Notwendigkeit, auf die Wünsche der Gäste nach Maßgabe ihrer Angebotsnutzung einzugehen, also nicht nur um jeden Wanderer, sondern um jeden ihrer Wandertage bemüht zu sein - so wie sich touristische Erfolg generell eher in Übernachtungszahlen als in Gästezahlen messen. Schließlich soll jede Wanderung das optimale Erlebnis bieten, und die Bedürfnisse derer, die häufiger unterwegs sind, müssen entsprechend stärker berücksichtigt werden als die der Wenigwanderer. Auf diese Weise spielt in die Profilstudien eine Statistik der Wanderungen und nicht nur der Wanderer, oder verallgemeinert der Nutzungen und nicht nur der Nutzer hinein.¹

Für ihre besonders tourismusadäquate Aussagekraft handelt sich eine derartige Zielortstudie nicht unerhebliche Probleme bei der Auswahl der Befragungsorte ein: Wie und mit welcher Gewichtung sind sie über die deutschen Landschaftsformationen und Tourismusregionen zu verteilen, an welchen Wanderwegen findet sich das für das Befragungsziel relevante Publikum, zu welchen Zeitpunkten ist es typischerweise anzutreffen? Diese und ähnliche Fragen können nur auf Grund bestenfalls plausibler Vorannahmen beantwortet werden, über die sich trefflich streiten lässt.

Die Profilstudien mussten überdies bislang mit vergleichsweise geringen Mitteln auskommen, was die Differenzierungsmöglichkeiten beim Zuschnitt Stichprobe weiter begrenzte. Die Auswahl der Befragungsstandorte beschränkte sich zumeist auf wenige wandertouristische Kerndestinationen einschlägiger deutscher Mittelgebirge in der Mitte sowie im Norden, Süden, Osten und Westen Deutschlands. In der Regel wurden die Fragebögen während der Wandersaison an mit Bänken hinreichend ausgestatteten Rastpunkten im Verlauf bekannter Wanderwege ausgegeben.

Insofern geben die Ergebnisse nur eine grobe Orientierung über die Situation in stark frequentierten Wanderzentren, haben damit aber auch eine gewisse touristische Leitfunktion, die durch eine vergleichsweise hohe Konsistenz der Daten im Zeitvergleich untermauert wird. Statistisch gesehen positiv schlägt die geringe Quote an Teilnahmeverweigerungen vor Ort (in der Größenordnungen von 10 bis 15%) sowie die Möglichkeit zu Buche, viele Fragebögen parallel ausgeben und so punktuell hohe Wandererdichten und komplette Wandergruppen erfassen zu können.

Grundsätzliche Repräsentativitätsprobleme

Nicht selten wird den Profilstudien Wandern der Vorwurf mangelnder Repräsentativität gemacht. Dabei wird allerdings nicht nur die Schwierigkeit der Erfüllung eines solchen Anspruchs, sondern auch der Umstand übersehen, dass es je nach Zielsetzung unterschiedliche Arten von Repräsentativität gibt. In aller Regel wird mit diesem Begriff allein die Vorstellung von Bevölkerungsrepräsentativität verbunden, wie sie uns aus Meinungsumfragen geläufig ist. Diese ist allerdings nicht - wie bei den touristisch orientierten Profilstudien - zielort-, sondern wohnortbezogen und wird meist über die zufallsgesteuerte Ziehung von nach soziode-

¹ Schon Busch u.a. (1992, S.30) weisen im Abschlussbericht ihres Forschungspraktikums in Angewandter Geographie, der eine Befragung von 475 Wanderern in dreie Wanderdestinationen zugrunde liegt, auf diesen Umstand hin: „Es stellt sich bei der Auswertung der Befragung/Zählung also die Problematik, ob wir Wanderer (als Personen) oder Wanderungen (als Tätigkeit) erfasst hatten. Es besteht die Gefahr, dass diejenigen, die häufiger im Gebiet wanderten, bei der Befragung/Zählung überrepräsentiert waren.“ Grundsätzlicher wird auf den damit angesprochenen Unterschied von Nutzer- und Nutzungsstatistiken unter dem Titel "Was heißt repräsentativ?" in der Rubrik "Profilstudien Wandern kurz" eingegangen.

mografischen Kriterien geschichteten Stichproben durch Haustür- oder Telefonbefragung realisiert.

Tatsächlich gab es bislang so gut wie keine hinreichend breit angelegte bevölkerungsrepräsentative Wanderstudie in diesem Sinne. Den regelmäßigen Markt- und Freizeitanalysen der großen Meinungsforschungsinstitute sind lediglich wenige Daten zum Umfang der Wanderneigung bzw. -aktivität zu entnehmen. Selbst wenn in der Regel unerwähnt, sehen auch sie sich mit erheblichen methodischen Schwierigkeiten konfrontiert. So liegt etwa die Teilnehmerverweigerung bei Telefonbefragungen nicht selten unter 50%, was dann zum Problem wird, wenn sie erst nach der Nennung des Umfragethemas "Wandern" erfolgt: Scheiden dabei Nichtwanderer eher als gelegentliche und diese eher als Vielwanderer aus? Einen nicht geringen Einfluss auf die Ergebnisse haben auch die Vorgaben für die Schichtung der Stichprobe. So kommt es, dass sich die ermittelten Wanderquoten der verschiedenen Studien selbst angesichts von jeweils 20.000 in die Befragung einbezogenen Bundesbürgern nicht unwesentlich unterscheiden.²

Ein weiterer Nachteil von Haustür- oder Telefonerhebungen: Wanderer stellen nur einen Teil der Stichprobe. Ob sie sich dazuzählen, muss erst erfragt werden. Die Antwort aber hängt vom individuellen Wanderverständnis ab, welches daher eigentlich zusätzlich abgefragt werden müsste - es sei denn, im Interview wird eine Wanderdefinition vorgegeben. Die aber kann nur willkürlich sein und prägt überdies die Haltung der Befragten zum Befragungsgegenstand vor.

Demgegenüber basiert die Profilstudie Wandern auf einer in der Literatur so genannten „Klumpenstichprobe“. Sie setzt voraus, dass möglichst ausnahmslos alle, die die Voraussetzungen für die Befragung erfüllen (in diesem Falle vor Ort wandernd angetroffen werden), in die Stichprobe einbezogen werden. Angesichts von geringen Verweigerungsquoten von um die 10% (etwa wegen fehlender Brille oder anderer Kommunikationsprobleme) kann diese Voraussetzung für die Profilstudien als erfüllt unterstellt werden. Damit kann ohne Umstände eine Nutzungsstatistik erhoben werden, deren Repräsentativität dann hauptsächlich von der Auswahl der Befragungsstandorte abhängt.

Im Ergebnis rekrutieren sich die Stichproben der Wohnort- und Zielortbefragung also sehr unterschiedlich: Hier die sich selbst zu solchen erklärenden Wanderer als Repräsentanten der Nutzer, dort die manifest aktiven Wandertouristen als Repräsentanten der Nutzungen. Im Vergleich haben hier die Wenig-, dort die Vielwanderer das relative Übergewicht.

Diese grundsätzlichen Unterschiede der beiden Befragungsalternativen und deren jeweils spezifische Defizite in Hinblick auf die Verallgemeinerbarkeit ihrer Ergebnisse bleiben nicht ohne Folgen für die Bewertung der beiden groß angelegten Wanderstudien des Deutschen Wanderverbandes und der Firma Trendscape aus den Jahren 2009/10. In beiden Fällen handelt es sich um Doppelstudien, die jeweils eine "bevölkerungsrepräsentative" Telefon-Erhebung "CATI" (computer assisted telephone interviews) mit einer Vorort-Befragung im Stil der Profilstudien Wandern verbinden. Welchem Zweck diese Verdoppelung dient, bleibt ebenso unerörtert wie die Frage nach dem unterschiedlichen statistischen Design und Aussagewert der Erhebungen.

In beiden Fällen wird nachdrücklich darauf hingewiesen, dass damit erstmals ausgiebige bevölkerungsrepräsentative Daten zum Wandertema erhoben worden sind. Der Hauptimpuls

² Ein studienübergreifender Vergleich dieser Zahlen für das Stichjahr 2008 findet sich in der Rubrik Wandermarkt>Wanderdemoskopie.

scheint demnach darin zu bestehen, eine vermeintliche Schwäche bzw. Lücke der Profilstudien zu füllen, ohne dass darüber Rechenschaft abgegeben wird, in welche Richtung sich das Erhebungsdesign verändert und welche neuen Probleme sich damit auftun. Warum dann gleichwohl jeweils der größere Befragungsaufwand einer ergänzenden Vorort-Studie gewidmet wird, bleibt ebenfalls unerfindlich. Keine Aussage findet sich in den vorliegenden Materialien auch über die Vereinbarkeit oder besser Nichtvereinbarkeit beider Erhebungsvarianten.

Deutscher Wanderverband: Zwei parallele Studien³

Für die Doppelstudie des Wanderverbandes wurde zum einen eine bevölkerungsrepräsentative Stichprobe von 3.032 Bundesbürgern ab 16 Jahren mittels CATI befragt, wovon sich 1.698 zum Wandern bekannten und als "aktive Wanderer"⁴ in Erscheinung treten. Zum anderen wurden 4.022 Wanderer ganzjährig an ausgewählten Wanderwegen in allen landschaftlichen Hauptformen von der Küste bis ins Hochgebirge "vor Ort" angesprochen. Hinzu kamen "weitere 500 Personen im Modul 'Winterwandern'". Die Ergebnisse beider Studien werden erkennbar getrennt ausgewiesen, ihr unterschiedlicher Stellenwert aber nicht erläutert.

Immerhin konnte anhand einer entsprechenden Frage geklärt werden, was seitens der Bevölkerung (also nicht nur der befragten Wanderer) "unter dem Begriff Wandern zusammengefasst" wird. Es sind dies zu den folgenden Stichproben-Anteilen:

Spazieren 51%, Trekking 44%, Nordic Walking 43%, Walking 42%, Pilgern 42%, Klettern 17%, Geocaching 15%, Sonstiges 3%

Der Begriff wird also teilweise sehr weit gefasst, woran die Wandergewohnten einen wesentlichen Anteil haben dürften. Gleichwohl liegt es nahe, dass auch unter denjenigen, die sich als Wanderer geoutet haben, viele den Wanderbegriff ähnlich auslegen. Mithin werden sich darunter nicht wenige Spaziergänger, Walker und Kletterer befinden. Tatsächlich spricht die Studie in diesem Zusammenhang - offenbar mit Blick auf die Kurzstreckler - von "wohnnahem Wandern". Treffender ließen sich die CATI-Befragten im Gegensatz zu den im ursprünglichen Sinne touristischen Wanderern bzw. Wandertouristen der Vorort-Befragung auch als "Allround-Wanderer" bezeichnen.

In welchem Maße sich durch diese Selbstdefinition das Wanderspektrum in Richtung von Kurzstrecken erweitert hat, macht ein besonders herausgestellter Befund der Verbandsstudie deutlich. So errechnete sich für die durchschnittliche Tourendauer ein Wert von 2:45 h. "Pro Wanderungen werden dabei etwas weniger als 10 km zurückgelegt".⁵ Diese geringe Wanderdauer lässt sich mit so gut wie keiner der bisherigen wandertouristischen Vorort-Erhebungen (auch jenseits der Profilstudien Wandern) zur Deckung bringen, deren mittlere Tourenlängen überwiegend bei ca. 4 h bzw. 12-15 km kumulieren (mit wenigen Ausreißern nach oben und unten). Die sich hierauf berufende Aussage "Die starke Ausdehnung der Nachfrage hat in den letzten Jahren somit zu einer deutlichen Verkürzung der Gehzeiten geführt" ist auf dieser Datenbasis nicht haltbar – ganz abgesehen davon, dass hierzu kaum Vergleichsdaten aus früheren bevölkerungsrepräsentativen Umfragen vorliegen.

³ Dokumentation Zukunftsmarkt Wandern – Erste Ergebnisse der Grundlagenuntersuchung Freizeit- und Urlaubsmarkt Wandern. Kassel 2010

⁴ Dieser Begriff lässt offen, was man sich abgrenzend unter "passiven Wanderern" vorzustellen hätte.

⁵ Dabei bleibt offen, ob die zugrunde liegenden Zahlen nur über alle Wanderer gemittelt oder zusätzlich auch mit der Zahl der von ihnen durchgeführten Wanderungen pro Jahr gewichtet wurden. Nur im letzteren Fall könnte man von der "durchschnittlichen Dauer einer Wanderung" sprechen.

Auf dieser Allround-Basis dann auch noch – wie geschehen - den per se zweifelhaften Versuch zu unternehmen, Wandern neu (verbindlich?) zu definieren und durch "Beschluss" zu untermauern, macht wenig Sinn. Vielmehr hätte spätestens an dieser Stelle auf die grundlegend andere statistische Herkunft dieses Ergebnisses verwiesen werden müssen. Das mit großer Wahrscheinlichkeit höhere Paralleldatum aus der Vorort-Befragung des Wanderverbandes ist der vorliegenden Erstpublikation nicht zu entnehmen.

Trendscope: Zwei Studien in einer⁶

Auch Trendscope kombiniert eine CATI mit 1.000 deutschen Haushalten und die "Hauptbefragung" von 2.343 Wandertouristen in Ausübung ihres Hobbys, deren prinzipielle Unterschiede nicht benannt werden. Stattdessen werden in der Hauptbefragung Wanderer während einer Wanderreise und solche "im Quellgebiet", die dort nur eine Tageswanderung unternehmen, aber in der Vergangenheit bereits auch eine Wanderreise unternommen haben", unterschieden. Der Fokus der Studie stellt also Wanderreisen in den Mittelpunkt und ist insofern ein eindeutig touristischer.

Der als "scheinbar zufällig" klassifizierten Befragung an Wanderwegen wird allerdings per se keine wie auch immer geartete Repräsentativität zugeschrieben, so dass man sich vor dem Dilemma sieht, der gleichwohl zur Hauptsache erklärten Erhebung vor Ort nachträglich einen repräsentativen Zuschnitt geben zu müssen. Allein diesem Zweck soll die Telefonbefragung dienen, indem sie eine Reihe von für maßgeblich gehaltenen Schlüsseldaten liefert, anhand derer die Vorort-Stichprobe nachträglich in eine bevölkerungsrepräsentative umgerechnet werden kann.

Die grundlegenden Unterschiede in der Rekrutierung der beiden Stichproben berühren indes nicht nur einige wenige, sondern mehr oder weniger alle abgefragten Indikatoren. Selbst wenn man es könnte, lassen sich die beiden Stichproben daher nicht einfach miteinander vermischen. Das gilt umso mehr, als offen bleibt, wie seinerseits nach der Akzeptanz des Anrufes die Auswahl der CATI-Befragten erfolgte. Die dazu gestellt Selektionsfrage bestimmt maßgeblich über den Zuschnitt der CATI-Stichprobe (s.o.). Hierzu wird ebenso wenig gesagt wie zur demgegenüber relativen Überrepräsentanz der Vielwanderer vor Ort. Von daher stellt sich auch nicht die Frage, welche der beiden Grundgesamtheiten für den touristischen Gebrauch tatsächlich die relevantere ist. Trendscope geht (nach dem obigen schwer nachvollziehbar) davon aus, dass die CATI-Erhebung die eigentliche "Potenzialanalyse des deutschen Wanderreisemarktes" darstellt.

Folgerichtig werden die Ergebnisse der Hauptstichprobe anhand einiger weniger, bezüglich ihrer Auswahl nicht näher begründeten CATI-Schlüsselindikatoren angepasst. Trendscope spricht in diesem Zusammenhang von Quotierungsvorgaben, mit deren Hilfe ein "möglichst genaues Abbild der Grundgesamtheit" konstruiert werden soll. Von der Quotierung betroffen und damit auch in der Vorort-Stichprobe letztlich CATI-Daten sind die Variablen

- Geschlecht und Alter nach drei Altersklassen,
- die Aufteilung in Tageswanderern sowie Urlaubswanderer, die ihrerseits gelegentlich differenziert wurden nach
 - Wanderurlaubern, die das Wandern als Hauptgrund für ihren Urlaub angaben, und

⁶ Trendscope Marktstudie "Wandern in Deutschland"

- Urlaubern mit Wanderaktivität, "bei denen das Wandern nur eine von mehreren Urlaubsaktivitäten darstellte".
- die Art des bevorzugten Wandergebietes (Flachland, Mittelgebirge, Hochgebirge),
- die regionale Herkunft auf Basis der Nielsen-Gebiete⁷.

In der Kombination ergab sich hieraus eine auf der Basis von 144 "Quotierungstöpfen" anzupassende Stichproben-Vorgabe. Wo die Vorort-Stichprobe dem nicht gerecht werden konnte, erfolgte die Anpassung per rechnerischer Gewichtung.

Die 100 Befragungsstandorte waren gleichmäßig über das Bundesgebiet verteilt, also nicht nach den von Land zu Land beträchtlich schwankenden Wandereranteilen an der Bevölkerung quotiert. Ihre konkrete Auswahl erfolgte nach der Art der bevorzugten Landschaftsform, also nicht nach deren wandertouristischer Frequentierung. Die Befragung geschah nach einem nicht genannten bzw. begründeten Schlüssel an ausgewiesenen Wanderwegen unter Einschluss von Premium-, Qualitäts- und unsertifizierten Wanderwegen sowie von Fernwanderwegen, Rundtouren und kürzeren oder unbekannteren Rund- oder Streckenwanderungen. Befragt wurde zwischen Mitte April und Anfang Oktober 2009, vorzugsweise aber in den Sommerferien - also in den Hauptreisemonaten, weit weniger im Frühjahr und Herbst, den dominierenden Zeiten des Freizeitwanderns.

Mit ihrem Hauptfokus auf touristisches Wandern in einem "weiten Verständnis" (offenbar ähnlich dem der Wanderverbandsstudie) tut sich die Trendscape-Studie mit Tageswanderungen schwer. Ihnen will man um des touristischen Erkenntnisinteresses willen nur ein zweitrangiges Gewicht zugestehen. Dies geschieht mit Hilfe von willkürlich erscheinenden Ausgrenzungen: "Obwohl die im März 2009 durchgeführte CATI-Befragung ergeben hatte, dass rund 14% der Deutschen (entspricht etwa 18% der wandernden deutschen Bevölkerung)⁸ angaben, lediglich im Wohnumfeld zu wandern und nicht im Rahmen von Urlauben, wurde sich zu Beginn der Hauptbefragung dennoch dafür entschieden, den Anteil der Tageswanderer an der Hauptstichprobe auf mindestens ein Drittel der Stichprobe festzusetzen. Dies geschah aus dem Grunde, dass insgesamt 75,5% der telefonisch befragten Personen angaben, Tageswanderungen durchzuführen.⁹ Insgesamt wurden auf diese Weise 1.481 Urlauber (542 Wanderurlauber und 939 Urlauber mit Wanderaktivität) und 862 Tageswanderer anhand des Fragebogens befragt". Diese im Detail schwer nachzuvollziehenden Proportionierungen tragen im Ergebnis dem hohen Stellenwert, welche die Münchener Fremdenverkehrswirtschaftler des dwif Tagesausflügen bzw. -wanderungen zuzumessen pflegen, kaum Rechnung.

Die Tageswanderungen wurden durch zwei weitere Vorabfestlegungen beschnitten:

- Berücksichtigt wurden (bei welcher Stichprobe?) nur Tageswanderungen, "die außerhalb des eigentlichen Wohnortes unternommen wurden". Gehören Touren durch den Stadtpark bzw. Stadtwald dazu oder war bis zum Wanderstart einen Mindestabstand einzuhalten?
- Die Mindestdauer für eine zu wertende Tagweswanderung sollte 2h betragen. Galt dieses Selektionskriterium für beide Studien, mussten auch die CATI-Wanderer mindestens einmal oder im Schnitt oder immer 2h pro Tour zurücklegen, um in die Stichprobe aufgenommen zu werden, und ihre Aussagen darauf beziehen?

⁷ Unklar bleibt, inwieweit hierbei die unterschiedlichen Wanderneigungsquoten in den Nielsen-Gebieten Berücksichtigung fanden.

⁸ Wenn 14% der Deutschen 18% der wandernden Deutschen entsprechen, müssten rund 50 Mio Deutsche oder 80% der Bevölkerung Wanderer sein – weit mehr als in allen vergleichbaren Studien.

⁹ Auch diese Zahl lässt sich nicht ansatzweise mit anderen bevölkerungsrepräsentativen Umfragen zur Deckung bringen, womit sich die Frage aufdrängt, ob es bereits bei der CATI Vorab-Auswahlkriterien gab.

Der perfektionistische, aber statistisch unzulässige Wunsch, beide Stichproben gewissermaßen zu verschmelzen und damit zu einer bevölkerungsrepräsentativen Vorort-Stichprobe zu kommen, hat die Dinge nicht einfacher, sondern unübersichtlicher gemacht. Es mussten zahlreiche, bestenfalls plausible Zusatzvorgaben gemacht werden, welche die Vorort-Stichprobe in unkontrollierbarer Weise verzerrt haben. Bei alledem wurden grundsätzliche statistische Rahmenbedingungen übersehen.

Im Effekt sind lediglich die an der CATI orientierten Rahmenwerte von ihrer Herkunft her rekonstruierbar – als eher tourismusfernere Freizeitdaten. Die Befunde der zielbezogenen Erhebung sind angesichts eines Wustes fragwürdiger Zusatzvorgaben weder nachvollziehbar noch in ihrer Aussagekraft bewertbar. Grundsätzlich führt an der getrennten Auswertung beider Originalstichproben kein Weg vorbei. Wechselseitig abweichende Befunde wären zu interpretieren und könnten so aufschlussreiche Zusatzeinsichten liefern. Angesichts ihrer artifizialen Vermischung kann man nur auf die immer wieder wundersamen Kräfte der Statistik setzen, die nicht selten auch gravierende methodische Verstöße auszugleichen in der Lage sind.