

## Fragebogenkonstruktion (I) – Grundlagen, Arten von Fragen, Arten von Skalen

*Rolf Porst – unter Mitarbeit von Ruth Holthof*

Nachdem wir uns im letzten Beitrag zur Serie „Forschen in der Schule“ (Porst 2015) mit Arten von Befragungen und mit Befragungstechniken beschäftigt haben, wenden wir uns nun der Fragebogenkonstruktion zu. Da ein guter Fragebogen die unabdingbare Voraussetzung für gute Daten und damit für gute Ergebnisse ist, und da sich Ihre SchülerInnen in dieser Phase des Forschungsprojekts besonders intensiv einbringen und zielgerichtet arbeiten können, haben wir uns entschlossen, dem Thema „Fragebogenkonstruktion“ zwei Beiträge zu widmen.

In dem ersten, dem heute vorliegenden Beitrag, werden wir etwas über die kognitionspsychologischen Grundlagen der Befragung und über kommunikative Aspekte der Befragungssituation erfahren; darüber hinaus werden wir Arten von Fragen und Arten von Skalen näher kennenlernen. Im zweiten Beitrag zur Fragebogenkonstruktion werden wir uns im nächsten Heft mit der Formulierung von Fragebogen-Fragen beschäftigen und ein paar Hinweise darauf geben, was bei der Konstruktion eines Fragebogens noch von Bedeutung ist.

### Kognitionspsychologische und kommunikative Grundlagen der Befragung

Ausgangspunkt aller Überlegungen zur Qualität von Fragen ist ein Modell<sup>2</sup>, demzufolge Personen beim Beantworten von Fragen mehrere Aufgaben lösen müssen. Befragungspersonen müssen...



**Rolf Porst**  
Markt- und Sozialforscher

1. die gestellte Frage verstehen
2. relevante Informationen zum Beantworten der Frage aus dem Gedächtnis abrufen
3. auf der Basis dieser Informationen ein Urteil bilden
4. dieses Urteil ggfs. in ein Antwortformat einpassen (wir nennen das „formatieren“) und
5. ihr „privates“ Urteil vor Weitergabe an den Interviewer/die Interviewerin bzw. den Fragebogen ggfs. modifizieren (wir nennen das „editieren“).

Ob und wie gut nun Befragungspersonen ihre gerade beschriebenen Aufgaben lösen können, hängt sehr stark von der Qualität der gestellten Frage und damit nicht zuletzt von der Qualität der Frageformulierung ab. Wir werden im Folgenden die fünf Aufgaben näher darstellen, die Befragungspersonen beim Beantworten unserer Fragen lösen müssen (wir müssen dies eher knapp machen, detailliert finden Sie das bei Porst 2014: 20ff).

## 1. Die gestellte Frage verstehen

Wenn wir eine Frage formulieren, stellt sich damit zugleich auch für uns eine Frage: Werden die Befragungspersonen die Frage so verstehen, wie es von uns intendiert ist, also wie wir möchten, dass sie von ihnen verstanden wird?<sup>3</sup> Nur wenn uns dies gelingt, ist eine erste entscheidende Klippe umschifft. Es geht also weniger um „Verstehen“, sondern um „Verständnis“ einer Frage.

Verständnis hat dabei zwei Dimensionen:

- a) *Semantisches Verständnis*: Was soll eine Frage oder ein Begriff in einer Frage „heißen“?
- b) *Pragmatisches Verständnis*: Was will der/die Fragende eigentlich „wissen“?

Eine Frage ist dann gut, wenn sie sowohl von der Semantik als auch von der Pragmatik her für die Befragungspersonen einleuchtend und nachvollziehbar, „verständlich“ ist. Ist dies nicht der Fall, sehen sich die Befragungspersonen einer Vielzahl von Problemen gegenüber.

Probleme mit dem *semantischen Verständnis* treten bei Befragungspersonen u.a. dann auf, wenn...

- Begriffe in Fragen *unbekannt* sind („In welchem Alter beginnt Ihrer Meinung nach bei Männern die ☞ Andropause?“),
- Fragen oder Begriffe (zwar „bekannt“, aber) *unklar formuliert* sind („Leben in Ihrer ☞ Nachbarschaft ☞ eher junge oder ☞ eher alte Leute?“),
- Fragen zu *schwierig formuliert* sind („Hat Sie schon einmal jemand durch Kritisieren Ihres Alkoholtrinkens ärgerlich gemacht?“),
- Begriffe in Fragen *mehrdeutig* sind („Wie viele Minuten brauchen Sie, um zu Fuß zu der ☞ Bank zu kommen, die Ihrer Wohnung am nächsten ist?“),
- Begriffe verwendet werden, die legitimer Weise von jeder Befragungsperson *individuell verstanden und individuell interpretiert* werden können („Wie häufig hatten Sie in der letzten Woche so richtig ☞ Stress an Ihrem Arbeitsplatz?“) oder
- Begriffe verwendet werden, die von unterschiedlichen Befragungsgruppen unterschiedlich verstanden werden („An wie vielen Tagen in der letzten Woche sind Sie einer ☞ bezahlten Arbeit nachgegangen?“).

Mit der Frage, wie man mit Problemen dieser Art bei der Konstruktion von Fragebogen umgeht, werden wir uns im nächsten Heft beschäftigen; hier sollen die Probleme zunächst nur benannt werden. Damit kommen wir zum pragmatischen Verständnis.

Probleme mit dem *pragmatischen Verständnis* treten dann auf, wenn eine Frage nicht erkennen lässt, was der Frager/die Fragerin eigentlich „wissen“ will. Nehmen Sie die Frage „Was haben Sie heute Morgen zwischen 7 und 8 Uhr alles getan?“.

Die Semantik ist eindeutig. Da will jemand wissen, was ich heute Morgen zwischen 7 und 8 Uhr alles gemacht habe. Aber: *Was will der wirklich wissen?* Doch wohl nicht, dass ich mich beim Frühstück an der Nase gekratzt habe, oder doch? Dass ich mich beim Rasieren geschnitten habe? Wird doch wohl kaum interessieren, oder doch? Na ja, so recht weiß ich nicht, was ich auf diese Frage antworten soll.

Sie sehen, mit Fragen von unklarer Pragmatik bringen wir unsere Befragungspersonen dann doch in Schwierigkeiten. Was soll hier angegeben werden? Um das Problem zu lösen, versuchen Personen, die pragmatische Bedeutung einer Frage zu erschließen. Dazu greifen sie auf Regeln der Alltagskonversation zurück.

Wenn Alltagskonversation kooperativ sein soll, liegt ihr modellhaft folgender Ansatz zugrunde (Grice 1975): Ein Sprecher sollte versuchen, informativ und eindeutig zu sein, der Wahrheit zu folgen und gesprächsrelevante Beiträge zu leisten. Der Zuhörer interpretiert die Äußerungen des Sprechers so, als ob und dass dieser diesen Zielen gerecht zu werden versucht.

Was bedeutet das für unsere Befragung? Die Befragungsperson unterstellt bei Fragen und Antwortvorgaben grundsätzlich einen Sinn, und wenn sie ihn nicht unmittelbar erkennt, sucht sie danach. Bei der Suche nach dem Sinn einer Frage lotet die Befragungsperson den Fragenkontext aus bzw. nutzt den Fragenkontext, um Unsicherheit zu reduzieren. Wenn sie also bei einer Frage nicht sofort erkennt, wie ein dort zum Einsatz gebrachter Begriff gemeint sein könnte, versucht sie, diesen Begriff durch Rückgriff auf vorangegangene Fragen(formulierungen) zu „verstehen“.

Die Aufgabe des Fragebogenkonstruktors ist es nun, genau dies zu verhindern. Die Befragungsperson darf nicht nach dem Sinn einer Frage suchen müssen. Es ist vielmehr unsere Aufgabe, die Fragen so zu formulieren, dass sie von der Semantik und von der Pragmatik her eindeutig sind.

## 2. Informationen aus dem Gedächtnis abrufen

Nachdem die Befragungsperson nun weiß (oder schließt), was der Sinn einer gestellten Frage ist, muss sie diese Frage auch beantworten. Damit sie dies kann, muss sie jetzt Informationen aus dem Gedächtnis abrufen, die zur Beantwortung der Frage von Bedeutung sein könnten.

Der Idealfall, in dem die Befragungsperson die Antwort auf eine Frage sofort und ohne längeres Nachdenken abrufen kann, ist dabei eher die Ausnahme als die Regel und funktioniert am ehesten noch bei Fragen wie „In welchem Monat und welchem Jahr sind Sie geboren?“, bei Fragen nach dem Familienstand oder nach der Anzahl der eigenen Kinder unter 18 Jahren. Hier gelingt es, die Frage direkt und ohne Nachdenken zu beantworten.

Leider ist dieser Idealfall nicht der Normalfall. Normalerweise müssen Personen erst nach relevanten Informationen suchen, um eine angemessene Antwort auf eine Frage zu generieren. Freundlicherweise übernimmt unser Gehirn diese Aufgabe für uns.

Bei *Einstellungsfragen* sucht unser Gehirn alles zusammen, was auf die Stimuli der Fragen passt: persönliche Erfahrungen mit dem Thema, die Erfahrungen Dritter, mit denen wir uns unterhalten oder von denen wir erfahren haben, die Behandlung des Einstellungsgegenstandes in den Medien und sogar kollektive Vorstellungen über den Gegenstand der Frage.

Bei Verhaltensfragen gilt es, die für die Beantwortung der Fragen relevanten Ereignisse zu erinnern, sie ggfs. zu datieren, die Zahl der Ereignisse zu berechnen oder ggfs. zu schätzen. Beantworten Sie einfach mal die Frage „Wie häufig sind Sie in diesem Schuljahr von Eltern Ihrer SchülerInnen telefonisch kontaktiert worden?“. Und beobachten Sie sich selbst: Wie sind Sie beim Beantworten der Frage vorgegangen?

Nun ist die Suche nach Informationen natürlich nicht grenzenlos. Ihr Gehirn hört auf zu suchen, wenn es das „Gefühl“ hat, dass es genügend Informationen erinnert hat, damit Sie in der Lage sind, eine Frage mit hinreichender subjektiver Sicherheit zu beantworten.

### 3. Ein Urteil bilden

Auf der Basis der erinnerten Informationen bilden sich Befragungspersonen jetzt bei Einstellungsfragen ein Urteil über den fraglichen Sachverhalt; bei Verhaltensfragen konstruieren oder rekonstruieren sie ihr (mutmaßliches) Verhalten und nutzen die so gewonnenen Informationen zur Urteilsbildung.

### 4. Das Urteil in ein Antwortformat einpassen

Das Urteil, das sich die Befragungsperson gebildet hat, muss dann an den Interviewer/die Interviewerin oder den Fragebogen kommuniziert werden. Bei *offenen* Fragen, bei denen es keinerlei Antwortvorgaben gibt, macht die Person das in ihren eigenen Worten. Bei geschlossenen, kategorisierten Fragen muss sie ihr Urteil in ein Antwortformat einpassen, z. B. in eine Skala oder in ein Set vorgegebener Antwortkategorien. Die Befragungsperson gibt also nicht ihr eigentliches Urteil preis, sondern einen Wert, der ihr Urteil ihrer Ansicht nach am besten repräsentiert. Dies kann eine Ziffer aus einer Skala sein oder eine von mehreren vorgegebenen Antwortkategorien.

### 5. Das Urteil editieren

Wenn die Befragungsperson am Ende des gesamten Prozesses ein Urteil gebildet hat und es formatiert, macht sie nichts anderes, als ihr „privates“ Urteil „öffentlich“ zu machen, d.h. es dem Interviewer/der Interviewerin oder dem Fragebogen zu übermitteln. Dabei können vor allem bei Befragungen, die von InterviewerInnen persönlich-mündlich oder telefonisch durchgeführt werden (interviewer-basierte Befragungsmodi) Erwägungen der *sozialen Wünschbarkeit* zu einer „Editierung“ der Antwort führen, d.h. die Befragungsperson sieht „gute Gründe“ dafür, nicht das wahre Resultat ihrer Urteilsbildung zu kommunizieren, sondern vom „wahren Wert“ abzuweichen – und tut dies bei ihrer Antwort auf eine Frage auch.

Bei sozialer Wünschbarkeit, die sich auf das Antwortverhalten einer Befragungsperson im Sinne einer „Editierung“ auswirken kann, unterscheiden wir zwischen befragungssituations-spezifischen Aspekten und allgemeinen Wünschbarkeitsaspekten;

erstere leiten sich aus der konkreten Befragungssituation ab, zu letzteren gehört alles, was eine Befragungsperson als gesellschaftlich akzeptiert oder eben als gesellschaftlich wünschbar interpretiert. In Befragungssituationen werden häufig beide Aspekte aktualisiert und relevant.

Schauen wir uns ein Beispiel an. Die (erfundene) Frage lautet:

„Sollten die Krankenkassen bei Menschen, die älter als 80 Jahre sind, ein künstliches Hüftgelenk bezahlen oder nicht?“

Nun ist es durchaus nachvollziehbar, wenn eine jüngere Befragungsperson, der ein alter Interviewer gegenüber sitzt, sich hier vielleicht etwas „vorsichtiger“ ausdrücken würde, als wenn sie von einem gleich jungen Interviewer befragt würde. Wenn dann noch gilt, dass man dem gesellschaftlichen Anspruch nach dem Alter gegenüber ganz allgemein Achtung haben sollte, müsste das Urteil einer befragten jungen Person einem alten Interviewer gegenüber milder ausfallen, als es ohne entsprechende Editierung geschehen würde.

Nun stellen Sie sich aber – bei gleicher allgemeiner gesellschaftlicher Norm – die Diade junge Befragungsperson und junger Interviewer vor. Warum sollte eine junge Befragungsperson jetzt ein Blatt vor den Mund nehmen, wenn sie ohnehin der Ansicht ist, alte Menschen lebten mehr oder weniger auf Kosten der Jungen?

Damit beenden wir den „theoretischen“ Teil dieses Beitrages und wenden uns konkret dem Fragebogen zu. Erfahren wir zunächst etwas über Arten von Fragen, dann über Arten von Skalen.

## Arten von Fragen

Wir unterscheiden Fragebogen-Fragen nach ihrer Form in *geschlossene*, *halboffene* und *offene* Fragen.

Bei *geschlossenen* Fragen gibt es eine begrenzte und definierte Anzahl von Antwortkategorien, in welche die Befragungsperson ihre Antwort einpassen („formatieren“) muss. Wir unterscheiden dabei zwischen Fragen mit nur einer zulässigen Antwort (*Einfachnennung*), bei denen sich die Befragungsperson für *eine* der vorgegebenen Alternativen entscheiden *muss* und Fragen mit mehr als einer zulässigen Antwort (*Mehrfachnennungen*), bei denen sie *mehrere* der vorgegebenen Alternativen *auswählen* kann.

Beispiel für eine *geschlossene* Frage mit *Einfachnennung*:

Wie stark interessieren Sie sich für Politik.....

- |                        |                          |
|------------------------|--------------------------|
| <i>sehr stark</i>      | <input type="checkbox"/> |
| <i>stark</i>           | <input type="checkbox"/> |
| <i>mittel</i>          | <input type="checkbox"/> |
| <i>wenig</i>           | <input type="checkbox"/> |
| <i>überhaupt nicht</i> | <input type="checkbox"/> |

Beispiel für eine *geschlossene* Frage mit *Mehrfachnennung*:

Welche Bundeskanzler/in gehören bzw. gehörten der CDU an? Sie können hier mehrere Kreuze machen.

- |                             |                       |
|-----------------------------|-----------------------|
| <i>Konrad Adenauer</i>      | <input type="radio"/> |
| <i>Ludwig Erhard</i>        | <input type="radio"/> |
| <i>Kurt Georg Kiesinger</i> | <input type="radio"/> |

<i>Willy Brandt</i>	<input type="radio"/>
<i>Helmut Schmidt</i>	<input type="radio"/>
<i>Helmut Kohl</i>	<input type="radio"/>
<i>Gerhard Schröder</i>	<input type="radio"/>
<i>Angela Merkel</i>	<input type="radio"/>

Geschlossene Fragen sind der Fragentyp par excellence in standardisierten Befragungen. Ihr Vorteil: Sie sind sowohl bei der Befragung selbst als auch bei der späteren Datenauswertung schnell abzuarbeiten. Ihr Nachteil: Befragungspersonen finden sich gelegentlich nicht in den vorgegebenen Antwortkategorien wieder.

Nebenbei: Wenn bei der Frage nach dem politischen Interesse Kästchen zum Ankreuzen vorgegeben waren, bei der Frage nach den Bundeskanzlern dagegen Kreise, steckt dahinter in aller Kürze die generelle Empfehlung: Gestalten Sie beim Layouten Ihres Fragebogens Dinge, die gleich sind, gleich, Dinge, die unterschiedlich sind, unterschiedlich. Hier am Beispiel: Verwenden Sie bei Fragen mit Einfachnennungen immer Kästchen, bei Fragen mit Mehrfachnennungen immer Kreise zum Ankreuzen. Das macht den Befragten die Aufgabe leichter, reduziert also „Kosten“.

Bei *offenen* Fragen gibt es nur den Fragentext, aber keine vorgegebenen Antwortkategorien. Die Befragungsperson antwortet in ihren eigenen Worten.

Beispiel für eine *offene* Frage:

Welche ist Ihrer Ansicht nach die beste Rockband aller Zeiten? *Bitte tragen Sie nur den Namen einer Band ein:* \_\_\_\_\_

Offene Fragen haben den Vorteil, dass die Befragungsperson so antworten kann, wie sie es gewohnt ist, wie ihr „der Schnabel gewachsen“ ist. Als Nachteile gelten gemeinhin, dass die Ergebnisse sehr stark von der Verbalisierungsfähigkeit der Befragungspersonen abhängen und dass Personen häufig Schwierigkeiten haben im Umgang mit offener Sprache oder Schrift. Und vor allem: Der Auswertungsaufwand ist bei offenen Fragen sehr hoch.

Bei *halboffenen* Fragen schließlich wird einer an sich geschlossenen Frage eine zusätzliche Kategorie angehängt, die wie eine offene Frage beantwortet werden kann, wenn sich die Befragungsperson nicht in eine der vorgegebenen Antwortkategorien einordnen kann oder will.

Beispiel für eine *halboffene* Frage:

Welcher deutsche Politiker/welche deutsche Politikerin setzt sich Ihrer Ansicht nach am meisten für den Mindestlohn in Deutschland ein? *Bitte wählen Sie nur eine Person aus.*

<i>Sigmar Gabriel</i>	<input type="checkbox"/>
<i>Bernd Lucke</i>	<input type="checkbox"/>
<i>Angela Merkel</i>	<input type="checkbox"/>
<i>Cem Özdemir</i>	<input type="checkbox"/>
<i>Horst Seehofer</i>	<input type="checkbox"/>
<i>Sahra Wagenknecht</i>	<input type="checkbox"/>
<i>Christian Lindner</i>	<input type="checkbox"/>

Andere/r Politiker/in, bitte eintragen:

\_\_\_\_\_

Halboffene Fragen sind eigentlich eher eine Folge der Unfähigkeit des Fragebogenentwicklers, eine Frage zu schließen, kommen aber in der Praxis der Umfrage sehr häufig vor.

Wenn Sie sich nun fragen, ob Sie besser offen, halboffen oder geschlossen fragen sollen: Die (!) Antwort auf Ihre Frage gibt es nicht. Vielmehr muss in jedem Einzelfalle und bei jeder Fragebogenfrage immer aufs Neue überdacht werden, in welches Format man sie bringen soll. Als Faustregel gebe ich Ihnen mit auf den Weg:

Verwenden Sie *geschlossene* Fragen immer dann, wenn Sie das Universum der Antworten sicher kennen und es aus einer bestimmten und bestimmbar Menge besteht, wenn die Anzahl der Antwortmöglichkeiten nicht allzu groß ist, wenn Ihre Fragen sich mit mehr oder weniger bekannten Sachverhalten beschäftigen und – ganz pragmatisch – wenn es schnell gehen soll.

Verwenden Sie *halboffene* Fragen immer dann, wenn das tatsächliche Universum möglicher Antworten auf eine Frage zwar gut geschätzt, aber nicht definitiv bestimmt werden kann.

Verwenden Sie offene Fragen immer dann, wenn Sie sich mit einem neuen, Ihnen selbst noch nicht sehr vertrauten Thema beschäftigen, wenn das Universum der möglichen Antworten unbekannt oder sehr groß ist und/oder wenn Sie eine Beeinflussung des Antwortverhaltens durch Lenkung vermeiden wollen.

## Arten von Skalen: Nominal-, Ordinal-, Intervall- und Ratio-Skalen

Bevor wir uns mit Skalen beschäftigen, wie sie in sozialwissenschaftlichen und sonstigen Fragebogen zu finden sind, wollen wir uns ein wenig mit den Grundlagen der Skalierung und den unterschiedlichen Skalenniveaus befassen.

Der Beantwortung einer Frage liegt – technisch betrachtet – grundsätzlich der Prozess des *Messens* zugrunde. Unter „Messen“ verstehen wir jegliche regelhafte und kodifizierte Zuordnung von Symbolen oder Ziffern zu Aspekten oder Ausprägungen manifester oder latenter Variablen. Das Bezugssystem, das dem Messen zugrunde liegt, bezeichnen wir als Skala. Skalen können unterschiedliche Skalenniveaus haben. In Abhängigkeit vom Skalenniveau unterscheiden wir Skalen in *Nominal*-Skalen, *Ordinal*-Skalen, *Intervall*-Skalen und *Ratio*-Skalen.

Um eine *Nominal*-„Skala“ zu bilden, ordnet man den Ausprägungen einer Variablen Ziffern oder Symbole derart zu, dass sie die gegenseitige Ausschließlichkeit der unterschiedlichen Ausprägungen zu erkennen geben. Man ist also entweder das eine oder das andere. Nehmen wir als Beispiel die Variable „Geschlecht“: Sie besteht üblicherweise aus zwei Ausprägungen, welche genauso üblicherweise als „weiblich“ und „männlich“ bezeichnet werden. Alleine diese Konvention macht „Geschlecht“ zur nominalskalierten Variablen – entweder man ist weiblich oder man ist männlich.

Bei der Ordinal-Skala unterliegen die Merkmalsausprägungen einer Rangordnung. Nehmen wir noch einmal die Frage nach dem politischen Interesse; die Antwortkategorien „sehr stark – stark – mittel – wenig – überhaupt nicht“ sind ranggeordnet: „Sehr stark“ ist nun mal mehr als „stark“, „wenig“ ist mehr als „überhaupt nicht“. Oder denken Sie an die Zustimmung zu einer Aussage mit den Ausprägungen „stimme voll und ganz zu – stimme eher zu – stimme eher nicht zu – stimme überhaupt nicht zu“. Ein besonderes Merkmal der Ordinalskala ist, dass die Antwortkategorien zwar ranggeordnet, die Abstände zwischen den Kategorien aber nicht gleich sind.

Über die relationalen Eigenschaften der Ordinal-Skala hinausgehend zeichnet sich die *Intervall*-Skala dadurch aus, dass die Abstände zwischen den Skalenpunkten gleich sind. Für die „gleichabständige“ Intervall-Skala gilt: Wenn A, B, C und D aufeinander folgen, dann ist  $B - A = D - C$ . Eine typische Intervall-Skala ist die Temperaturskala

in Grad Celsius (der Abstand zwischen 22 und 23 Grad ist genauso groß wie der Abstand zwischen 27 und 28 Grad). Intervall-Skalen finden wir in Fragebogen häufig als „endpunktbenannte Skala“ zur Beurteilung vorgegebener Aussagen:

trifft überhaupt nicht zu        trifft voll und ganz zu

Ratio-Skalen schließlich sind – einfach ausgedrückt – Intervall-Skalen mit einem echten *Nullpunkt*. Echte Ratio-Skalen finden wir in Fragebogen etwa zur Messung des Netto-Einkommens einer Person oder zur Messung der Anzahl der Kinder unter 18 Jahren im Haushalt oder zur Anzahl der im Jahre 2014 gelesenen Bücher. Wir versuchen dort gar nicht erst, die Skala optisch als Skala anzubieten, sondern fragen offen und ohne Vorgabe von Kategorien oder Skalenpunkten ab.

Im Zusammenhang mit Skalenniveaus ist entscheidend: Je höher das Skalenniveau ist, umso mehr und bessere Auswertungsverfahren können zum Einsatz gebracht werden. Alleine die Berechnung des Mittelwerts einer Skala setzt „eigentlich“ mindestens intervallskalierte Daten voraus.

## Skalen für sozialwissenschaftliche (und andere) Fragebogen

Um Meinungen, Einstellungen, Wertorientierungen oder vergleichbare latente oder manifeste Variablen differenziert zu messen, kommen in Fragebogen zumeist mehrstufige Skalen zum Einsatz, vermittels derer die Befragungsperson ihre Position zu bestimmten vorgegebenen Themen zum Ausdruck bringen kann. Die Mehrstufigkeit der Skalen entspricht zum einen dem Bedürfnis befragter Personen nach differenzierter Darstellung ihrer Position, zum anderen aber auch dem Bedürfnis der Auswerter nach möglichst elaborierten Auswertungsverfahren.

Die Diskussion um die „richtige“ Skala ist so alt wie die Umfrageforschung selbst, und wir können schon jetzt den Schluss ziehen, dass es diese „richtige“ Skala gar nicht gibt. Vielmehr muss bei jeder einzelnen Frage und bei jeder zu bewertenden Aussage von Neuem überlegt werden, welche Skala man wirklich zum Einsatz bringen will, und damit ist nicht nur gemeint, dass man die Skalenpunkte richtig benennt, sondern auch die Art und die Struktur der Skala selbst. Wenn wir im Folgenden einige in sozialwissenschaftlichen Fragebogen gebräuchliche Skalentypen und deren Vor- und Nachteile diskutieren, müssen wir uns mit mindestens vier Fragen auseinandersetzen:

1. Sollen wir alle Skalenpunkte verbalisieren – oder nur die Skalenendpunkte?
2. Sollen wir Skalen mit einer geraden Anzahl von Skalenpunkten verwenden – oder solche mit einer ungeraden Anzahl von Skalenpunkten?
3. Wie viele Skalenpunkte soll eine Skala überhaupt haben?
4. In welcher Richtung sollen wir Skalen formulieren?

### 1. Verbalisierte vs. endpunktbenannte Skalen

Verbalisierte Skalen sind Skalen, bei denen jeder einzelne Skalenpunkt mit einer verbalen Benennung versehen ist, z.B.:

Trifft überhaupt nicht zu <input type="checkbox"/>	Trifft eher nicht zu <input type="checkbox"/>	Trifft eher zu <input type="checkbox"/>	Trifft voll und ganz zu <input type="checkbox"/>
---	--	--	---



Die verbalisierte Skala hat den Vorteil, dass sie der Befragungsperson genau vorgibt, was diese unter den unterschiedlichen „Kästchen“ verstehen soll. Von daher kommt dieser Skalentyp den Befragungspersonen entgegen. Für den Forscher/die Forscherin hat diese Skalenform allerdings den Nachteil des relativ niedrigen weil ordinalen Skalenniveaus, zum anderen besteht das Problem der angemessenen Verbalisierung der Skalenpunkte, vor allem mit deren steigender Anzahl: Je breiter die Skala ist, je mehr Skalenpunkte also zu verbalisieren sind, umso schwieriger ist diese Aufgabe zu lösen.

Genau diese Schwächen der verbalisierten Skala machen den Charme der endpunktbenannten Skala aus. Endpunktbenannte, auch als „numerische“ bezeichnete Skalen verzichten auf die Verbalisierung jedes einzelnen Skalenpunktes und benennen nur die beiden Extrempunkte, z.B.:

Trifft überhaupt nicht zu <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> trifft voll und ganz zu
---

Weil nur die Endpunkte zu verbalisieren sind, beschränkt sich das Problem der Verbalisierung eben nur auf diese beiden Extrempunkte; die endpunktbenannte Skala kann – im Prinzip, aber nur im Prinzip (siehe die Ausführungen zur Breite von Skalen an späterer Stelle) – beliebig breit sein.

Dazu kommt, dass solche Skalen ausnahmslos als intervallskaliert gelten können, was sie für die AuswerterInnen interessanter macht als verbalisierte Skalen. Und: die meisten Befragungspersonen sind durchaus in der Lage, mit diesem Skalentyp zu arbeiten, wenn man ihnen erklärt, wie sie damit umzugehen haben.

Als Nachteil gilt, dass – sieht man von den Endpunkten ab – jeder Befragungsperson selbst überlassen bleibt, wie sie die nicht-verbalisierten Zwischenpunkte interpretieren will. Das ist – ehrlicherweise gesagt – dem Ziel der standardisierten Befragung natürlich abträglich. *Entscheidend ist aber, dass die meisten Befragungspersonen die Skalenpunkte einer Intervallskala als gleichabständig interpretieren, wenn wir die Skalen in möglichst neutraler Darstellungsweise präsentieren. Wenn die Personen verstehen, dass eine Intervallskala eine Rangordnungsskala mit gleichen Abständen zwischen den Skalenpunkten ist, ist das Ziel der Skala erreicht.*

## 2. Gerade vs. ungerade Skalen

Bei geraden Skalen ist die Anzahl der Skalenpunkte geradzahlig, z.B. mit sechs Skalenpunkten:

trifft überhaupt nicht zu <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> trifft voll und ganz zu
---

Ungerade Skalen dagegen verfügen über eine ungerade Anzahl von Skalenpunkten; der eigentliche Unterschied zwischen den beiden Skalenformen besteht darin, dass bei der ungeraden Skala ein „Skalenmittelpunkt“ vorhanden ist:

trifft überhaupt nicht zu <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> trifft voll und ganz zu
--

Dieser „Skalenmittelpunkt“ ist oft gar kein richtiger Mittelpunkt in dem Sinne, dass er die Skala in zwei Teile oder zwei Richtungen trennt. Ungeachtet dessen wird er von Befragungspersonen häufig als Skalenmittelpunkt interpretiert und genutzt. Und nicht nur inhaltlich, um die eigene Position anzuzeigen, sondern oft auch als „Fluchtkategorie“, weil man sich entweder nicht einordnen kann oder nicht einordnen will. Die Entscheidung für das „mittlere“ Kästchen ist oft nichts anderes als der Versuch,

zur nächsten Frage zu gelangen, ohne seine wirkliche Meinung oder auch Meinungslosigkeit zu der Frage kundtun zu müssen.

Bringt man eine geradzahlige Skala zum Einsatz, verschwindet die Fluchtmöglichkeit natürlich, aber man kauft sich ein anderes Problem dafür ein: Personen möchten sich bewusst und wohlüberlegt „in der Mitte“ platzieren, aber sie haben nicht die Chance dazu, weil wir den „mittleren“ Skalenpunkt nicht anbieten.

Wenn Sie mich fragen, mir ist der Fehler bei der ungeraden Skala lieber als der Fehler bei der geraden, weil ich für die ungerade Skala eine gute Lösung anzubieten habe: Ich gebe außerhalb der eigentlichen Skala eine Restkategorie vor, z.B. „weiß nicht“ oder „will ich nicht beantworten“. Dann erkennen die Befragungspersonen, dass sie nicht in die Mitte der Skala fliehen müssen, sondern dass es einen legitimen Punkt außerhalb der Skala gibt, den sie als „Fluchtkategorie“ nutzen können, wenn sie nicht antworten wollen oder können. *Natürlich gäbe es diese Lösung auch für gerade Skalen, aber dort brauchen wir sie nicht, weil es keine Fluchtkategorie gibt, der wir entgegenwirken müssen. Wenn wir bei geraden Skalen eine legitime Restkategorie angeben, dann eher, um den Befragten zu zeigen, dass sie sich nicht innerhalb der Skala einstufen müssen, wenn sie das nicht möchten – aber das wollen wir ja eigentlich nicht.*

### 3. Zur optimalen Skalenbreite

Wie breit eine Skala sein kann, sollten wir von der vermuteten Abstraktionsfähigkeit der potenziellen Zielpersonen abhängig machen. Machen wir die Skala zu breit, überfordern wir die Befragungspersonen rasch; machen wir die Skala zu eng, sind die Befragten nicht mehr in der Lage, eine Aussage hinreichend differenziert zu bewerten.

Mein Vorschlag: Endpunktbenannte Skalen sollten nicht weniger als 5 und nicht mehr als 9 Skalenpunkte haben, wenn Befragungspersonen sinnvoll damit umgehen sollen. Ob Sie nun 5, 7 oder 9 Skalenpunkte einsetzen wollen, hängt, wie gesagt, davon ab, wie „klug“ Ihre Zielgruppe ist. Bei eher „klugen“ Befragten können wir mit 9 Skalenpunkten arbeiten, bei „weniger Klugen“ besser mit 5 Skalenpunkten, und wenn Sie in allen Bildungsgruppen fragen, verwenden Sie am besten 7-er Skalen. Oder, falls Sie SchülerInnen befragen wollen: Setzen Sie in der Orientierungsstufe eher 5er-Skalen ein, in der Oberstufe eher 9er-Skalen.

Bei verbalisierten Skalen sollte die Breite von 4 bis höchstens 6 Skalenpunkten reichen, weil die Verbalisierung von mehr als 6 Skalenpunkten meistens unzulänglich wird, wenn sie überhaupt gelingt.

### 4. Zur Richtung der Skala

Wenn wir von „Richtung“ der Skala sprechen, geht es uns darum, ob wir die Skala „von links nach rechts“ oder „von rechts nach links“ anbieten wollen. „Von links nach rechts“ heißt, der niedrigste Skalenwert steht auf der linken und der höchste Skalenwert auf der rechten Seite der Skala. Entsprechend bedeutet „von rechts nach links“ die Umkehrung dieser Vorgabe.

Bei der Frage nach der Richtung der Skala ist entscheidend, ob wir mit verbalisierten oder mit endpunktbenannten Skalen arbeiten: Endpunktbenannte Skalen sollten grundsätzlich „von links nach rechts“ präsentiert werden, also vom – links – niedrigsten zum – rechts – höchsten Skalenpunkt. Wir denken von links nach rechts; inso-

fern entspricht diese Skalenrichtung unserer allgemeinen Denkweise, und von daher gibt es keinen Grund, eine Skala anders als „von links nach rechts“ zu präsentieren.

Ganz anders stellt sich das bei verbalisierten Skalen dar. Vor allem wenn verbalisierte Skalen vorgelesen werden, nimmt sie die Befragungsperson nicht wie Skalen wahr, sondern wie „normale“ Antwortkategorien (was sie de facto ja auch sind). Hier geht es darum, dass die Vorgabe in den Sprachfluss der Frage passt. Und da ist es meistens besser, sie vom höchsten Wert bis zum niedrigsten Wert laufen zu lassen.

## 5. Die optimale Antwortskala

... gibt es nicht. Um Sie mit dieser wenig tröstlichen Nachricht aber nicht ganz alleine zu lassen, möchte ich meine kurze Darstellung von Vor- und Nachteilen unterschiedlicher Skalen mit einer Empfehlung beenden: Auch wenn die endpunktbenannte Skala den Nachteil hat, dass sie die Interpretation der Skalenpunkte zwischen den verbalisierten Endpunkten den Befragten überlässt und sie gelegentlich auch „verunsichert“, haben sich in der Umfragepraxis numerische, endpunktbenannte Skalen mit 5 bis 9 Skalenpunkten bewährt. Wenn Sie statt ungeraden Skalen gerade Skalen einsetzen wollen, bieten sich Skalen mit 4, 6 oder 8 Skalenpunkten an.

Beachten Sie die Richtung der Skalen von links nach rechts und vermeiden Sie die Verbalisierung des linken Extrempunktes mit negativen Begriffen oder Zeichen, z.B. „-5“; Befragungspersonen neigen dazu, sich positiv darzustellen und vermeiden es deshalb auch, sich in Kategorien mit negativen Benennungen zu platzieren. Benennen Sie den linken Extrempunkt der Skala also besser mit „überhaupt nicht zufrieden“ als mit „sehr unzufrieden“, mit „kein Erfolg“ statt mit „Misserfolg“.

## Zum Schluss

Soweit die kurze Einführung in die kognitionspsychologischen und kommunikativen Grundlagen der Befragung, in Arten von Fragen und Arten von Skalen. Im nächsten Heft werden wir uns, immer noch unter der Überschrift „Fragebogenkonstruktion“, mit Question Wording beschäftigen, also mit der Formulierung guter Fragen und Antwortkategorien für Fragebogen.

## Anmerkungen

- 1 Ruth Holthof ist als Studienrätin im Fach Sozialkunde am Eleonoren-Gymnasium in Worms tätig.
- 2 Das „model of information processing in a survey situation“ von Strack & Martin (1987:124ff); an Beispielen erläutert bei Porst (2014: 20ff).
- 3 Also nicht „wie wir möchten, dass sie sie beantworten“, sondern „wie wir möchten, dass sie sie verstehen“.

## Literatur

- Cole, P. & J. L. Morgan (1975): Syntax and Semantics 3: Speech Acts. New York: Academic Press  
 Grice, H.P. (1975): Logic and Conversation. S. 41 – 58 in Cole/Morgan, Hrsg. (1975)

- Hippler, H.-J., Schwarz, N. & S. Sudman, Hrsg. (1987): *Social Information Processing and Survey Methodology*. New York: Springer
- Porst, R. (2014): *Fragebogen. Ein Arbeitsbuch*. 4. Auflage. Wiesbaden: Springer VS
- Porst, R. (2015): Arten von Befragungen und Befragungstechniken. S. 93 – 103 in *Gesellschaft – Wirtschaft – Politik (GWP) 64, Heft 1*
- Strack, F. & L. L. Martin (1987): Thinking, Judging and Communicating: A Process Account of Context Effects in Attitude Surveys. S. 123 – 148 in Hippler/Schwarz(Sudman, Hrsg. (1987)