

Pretesting – Verfahren zur Evaluation von Fragebogen(entwürfen)

Rolf Porst¹

Durch die Lektüre der bisherigen Beiträge zur Serie „Forschen in der Schule“ haben wir uns – hoffentlich – das Rüstzeug angeeignet, um einen guten Fragebogen für unser schulisches Umfrageprojekt zu konstruieren. Wir haben die *zentralen Begrifflichkeiten* kennengelernt, die uns auf dem Weg von der Fragestellung zum Fragebogen begleiten: Theorien, Konzeptualisierung, Operationalisierung, Variablen, Hypothesen (Porst 2014a). Wir haben *Arten von Befragungen* und *Befragungstechniken* kennengelernt (Porst 2015a); wenn wir eine standardisierte Befragung durchführen wollen, können wir uns jetzt begründet für eine Befragungstechnik – persönlich-mündlich, schriftlich oder telefonisch – entscheiden. Wir haben dann (Porst 2015b, c) die unterschiedlichen *Bestandteile eines Fragebogens* (Arten von Fragen, Arten von Skalen) und *Regeln zur Formulierung von Fragebogen-Fragen* kennengelernt.

Wir sollten nun also in der Lage sein, einen „guten“ Fragebogen für unser schulisches Umfrageprojekt zu konstruieren bzw. von unseren SchülerInnen konstruieren zu lassen.

Aber Vorsicht: Ein Fragebogen ist meist ein recht komplexes und nicht einfach zu erstellendes wissenschaftliches Produkt, und es ist keineswegs Understatement, wenn renommierte Umfrageforscher wie Sudman & Bradburn (1982: 283) schreiben: „*Even after years of experience, no expert can write a perfect questionnaire.*“ Oder noch deutlicher: „*If you don't have the resources to pilot test your questionnaire, don't do the study*“ (ebenda: 283).

Wer also nach einem wirklich guten Fragebogen strebt – und wer wollte oder tut das nicht? – wird nicht umhinkommen, seinen oder ihren Fragebogenentwurf einem



Rolf Porst, Markt- und Sozialforscher
Römerberg

(oder mehreren) systematischen Pretest(s) zu unterziehen. Wir halten einen Pretest für eine unabdingbare Voraussetzung für die Vorbereitung der Hauptstudie. Was ist also ein Pretest, und welche Ziele werden mit dem Pretest verfolgt?

Definition und Zielsetzung von Pretests

Unter Pretest verstehen wir – ganz allgemein gesagt – die Testung und Evaluation eines Fragebogens oder einzelner seiner Teile vor ihrem Einsatz in der Hauptbefragung: „Praktisch handelt es sich bei einem Pretest um einen Testlauf eines ‚Fragebogen-Prototyps‘, also eines mutmaßlich noch nicht ausgereiften Fragebogens.“ (Porst 2000: 65).

Pretests geben uns – im Vorfeld unserer Haupterhebung, also bevor es zu spät ist – Auskunft über...

- die Verständlichkeit von Fragen
- die Reihenfolge der Fragen
- Kontexteffekte
- das Zustandekommen von Antworten
- Probleme der Befragten mit ihrer Aufgabe
- Interesse und Aufmerksamkeit der Befragten bei einzelnen Fragen
- das Wohlbefinden der Befragten
- Probleme der Interviewer mit dem Fragebogen
- technische Probleme bei Fragebogen und Befragungshilfen
- Häufigkeitsverteilungen der Antworten
- die Zeitdauer der Befragung

Diese Liste mit ihren sehr unterschiedlichen Zielrichtungen lässt erahnen, dass es *den* Pretest gar nicht gibt. Und tatsächlich ist es so, dass wir zum Erreichen der einzelnen Ziele recht unterschiedliche Pretest-Wege gehen müssen.

Zunächst einmal unterscheiden wir zwei große Gruppen von Pretests, nämlich *Standardbeobachtungspretests* und *kognitive Pretesttechniken*. Da es sich bei Fragebogen für schulische Projekte zumeist um kurze, wenig umfängliche Instrumente handeln wird, werden Standardbeobachtungspretests, bei denen umfangreichere Fragebogen komplett getestet werden, üblicherweise nicht zum Einsatz kommen. Wir werden uns deshalb im Folgenden ausschließlich mit kognitiven Pretestverfahren beschäftigen.²

Kognitive Pretesttechniken

Kognitive Pretestinterviews zählen zu den „aktiven“ Pretesttechniken; sie werden zum Einsatz gebracht, um einen Einblick in die kognitiven Prozesse zu bekommen, die beim Beantworten von Fragen ablaufen:

- Wie interpretieren Befragungspersonen Fragen oder Begriffe?
- Wie rufen sie Informationen und Ereignisse aus dem Gedächtnis ab?
- Wie kommen sie zu einer Entscheidung darüber, wie sie antworten?
- Wie ordnen sie ihre „intern“ ermittelte Antwort formalen Antwortkategorien zu?

Dazu ein Beispiel zum Frageverständnis. Bei einem Pretest sollte folgendes Item bewertet werden:

„Das Ziel der Nationalsozialisten war, durch den Krieg gegen die Sowjetunion den jüdischen Bolschewismus zu beseitigen“.³

Unser Einwand, dass diese Aussage nicht geeignet sei, bei einer Allgemeinpopulation abgefragt zu werden, wurde erst akzeptiert, nachdem wir das Item zwanzig Testpersonen vorgelegt und sie gebeten hatten, es vermittels einer Skala von 1 = „stimme überhaupt nicht zu“ bis 7 = „stimme voll und ganz zu“ zu bewerten. Das Interessante daran war, dass nur zwei der zwanzig Testpersonen angaben, sie könnten diese Bewertung nicht vornehmen; alle anderen 18 Testpersonen haben spontan einen Skalenwert vergeben, und mit Hilfe kognitiver Pretesttechniken konnten wir zeigen, dass alle diese 18 Testpersonen zwar eine Bewertung abgegeben, aber keine klare Vorstellung vom Inhalt der Aussage hatten.

Wie kann es zu einem solchen Ergebnis kommen?

Zunächst einmal deshalb, weil Befragungspersonen in Umfragen üblicherweise nicht ermuntert werden, Probleme beim Beantworten von Fragen zu benennen. Ihre Aufgabe besteht im Beantworten der Fragen, nicht darin, die Fragen selbst zu hinterfragen.

Dazu kommt, dass Befragungspersonen ein Problem natürlich nur dann benennen können, wenn sie sich selbst dieses Problems bewusst sind. Nehmen sie ein Problem gar nicht als solches wahr, kann es natürlich auch nicht zu entsprechenden Reaktionen (wie etwa Nachfragen nach dem Sinn der Frage) kommen.

Und schließlich tendieren Personen in Befragungen dazu, sich selbst möglichst positiv darzustellen, und dazu gehört manchmal auch, dass sie Fragen beantworten, die sie eigentlich nicht wirklich verstanden haben.

Für die Evaluation von Fragebogen heißt das: Wir müssen beim Pretesten aktiv vorgehen und den Befragten *Fragen zu den Fragen* stellen – und genau das machen wir mit kognitiven Pretesttechniken.

Die wichtigsten kognitiven Pretesttechniken sind das *think aloud* (Technik des lauten Denkens), das *paraphrasing* (Paraphrasieren), das *sorting* (Sortiertechniken) und das *probing* (Nachfragetechniken). Wir wollen uns ganz kurz mit diesen Techniken beschäftigen, um Ihnen zu zeigen, wie Sie selbst mit kognitiven Pretesttechniken zur Verbesserung Ihrer Fragen beitragen können.

1. Think aloud

Beim *think aloud* werden die Befragungspersonen aufgefordert, während des Beantwortens einer Frage „laut zu denken“, also sämtliche Gedankengänge, die zur Antwort führen, zu formulieren. Damit soll der Antwortprozess erkennbar werden und dadurch auch eventuell vorhandene Probleme mit dem Fragenverständnis.

Das laute Denken fällt den meisten Befragungspersonen schwer. Sie sind häufig nicht in der Lage, die Prozesse zu artikulieren, die zu einer Antwort führen oder geführt haben. Deshalb ist eine detaillierte Anleitung für die Befragten wichtig. Wichtig sind auch wiederholte Aufforderungen, laut zu denken. Die Instruktion zum lauten Denken könnte etwa wie in dem folgenden Beispiel aussehen:

„Bitte sagen Sie mir bei der folgenden Frage alles, an was Sie denken oder was Ihnen durch den Kopf geht, bevor Sie die Frage beantworten. Sagen Sie bitte auch Dinge, die Ihnen vielleicht unwichtig erscheinen. Die Frage lautet:

(zu testende Frage)

„Wie oft waren Sie in den letzten 12 Monaten beim Zahnarzt?“

(Antwort einer Testperson)

„Oh, lassen Sie mich nachdenken. Ich glaube, ich war einmal zum Nachsehen und um die Zähne reinigen zu lassen. An ein anderes Mal kann ich mich nicht erinnern. Ich bin mir nicht sicher. Aber irgendwann um diese Zeit habe ich Probleme mit meinem Kiefer gehabt, aber sie konnten nichts finden. Ich bin mir aber nicht sicher, welcher Monat das war. Ich würde also sagen ein Mal.“

Aus Antworten wie dieser müssen Sie dann Ihre Schlüsse ziehen: Hat die Befragungsperson die Frage richtig verstanden? Wie gehen Befragungspersonen vor, wenn sie diese Frage beantworten? Ist Ihre Frage geeignet, diese Vorgehensweise so zu provozieren, dass die Frage in Ihrem Sinne „richtig“ verstanden und beantwortet wird?

2. Paraphrasing

Beim *paraphrasing* soll die Testperson nach Beantworten einer Frage den Fragetext in eigenen Worten wiederholen. Diese Vorgehensweise führt zu Informationen darüber, ob die Person die Frage überhaupt verstanden und sie so verstanden hat, wie von dem Forscher/der Forscherin intendiert. Wichtig dabei ist: Die Befragungsperson soll nicht versuchen, den Text wörtlich zu erinnern; es geht nicht um das Gedächtnis der Person, sondern um das Verständnis einer Frage. Paraphrasing eignet sich von daher nicht für den Test kurzer Faktfragen.

Für *paraphrasing* spricht, dass es leicht anwendbar ist und wenig Vorbereitung bedarf; auch reagieren die Testpersonen zumeist recht amüsiert.

Auch hier ein Beispiel; die zu testende Frage hieß:

„Was ist aus Ihrer Sicht die ideale Kinderzahl für eine Familie in Deutschland?“

Nachdem die Testperson die Frage beantwortet hat, wird sie gebeten, die Frage zu paraphrasieren:

„Bitte wiederholen Sie die Frage, die ich Ihnen eben vorgelesen habe, noch einmal in Ihren eigenen Worten.“

Entscheiden Sie selbst, ob die im Folgenden zitierten Testpersonen die Frage verstanden haben oder nicht:

- „Wie viele Kinder heutzutage zu einer Familie gehören.“
- „Wie viele Kinder die ideale Familie hat in Deutschland.“
- „Was die ideale Anzahl der Kinder in Deutschland für eine Familie ist.“
- „Wenn ich das aus meiner Sicht sehe, wie viele Kinder ich akzeptieren würde.“
- „Was in Deutschland am besten ist von der Anzahl der Kinder, was man sich vielleicht noch erlauben kann und was nicht mehr.“

3. Sorting

Mit Sortiertechniken wird exploriert, wie Testpersonen Begriffe oder Situationen bestimmten Überbegriffen oder Kategorien zuordnen. Wir unterscheiden zwischen freiem und dimensionalem Sortieren: Beim *freien Sortieren* sollen die Testpersonen vorgegebene Items nach eigenen Kriterien gruppieren, beim *dimensionalen Sortieren* sollen vorgegebene Items nach von Ihnen vorab festgelegten Dimensionen sortiert werden.

Beim folgenden Beispiel geht es um die Zuordnung von Alltagssituationen zu übergeordneten Begriffen; es handelt sich um ein dimensionales Sortieren, weil wir den Testpersonen die übergeordneten Begriffe – „Verkehrsunfall“ bzw. „kein Verkehrsunfall“ vorgegeben haben:

„Auf diesen Kärtchen sind verschiedene Situationen beschrieben, die sich auf der Straße abspielen können. Sagen Sie mir bitte für jede Beschreibung, ob Sie diese Situation als „Verkehrsunfall“ bezeichnen würden oder als „kein Verkehrsunfall“.“

Die zu beurteilenden Situationen haben wir jeweils einzeln auf Kärtchen geschrieben, und die Testpersonen sollten die Kärtchen auf ein großes Blatt sortieren, das jeweils ein Feld mit der Inschrift „Verkehrsunfall“, „kein Verkehrsunfall“ und „unentschieden“ enthielt. Vorgegeben waren unter anderem die Kärtchen:

- „Zwei Autos stoßen auf einer Kreuzung zusammen“
- „Ein Radfahrer gerät ohne Fremdverschulden mit seinem Fahrrad in eine Straßenbahnschiene und stürzt“
- „Ein Busfahrgast stürzt beim Aussteigen und bricht sich ein Bein“
- „Ein Fußgänger rutscht auf dem Gehweg auf einer Bananenschale aus und verletzt sich“.

Durch diese Vorgehensweise erfahren wir, was Personen als „Verkehrsunfall“ bezeichnen würden und was nicht. Ergebnis im Übrigen: Ein „Verkehrsunfall“ zeichnet sich im Wesentlichen dadurch aus, dass mindestens zwei Akteure beteiligt sind und das Ereignis sich im öffentlichen Verkehrsraum abspielt.

4. Probing

Mit *probing* bezeichnen wir kognitive Techniken, bei denen Begriffe, Fragetexte oder Antwortkategorien mittels einer oder mehrerer Zusatzfragen hinterfragt werden mit dem Ziel, Informationen zum Verständnis der Fragen und Begriffe zu erhalten. Dabei unterscheiden wir vier zentrale *probing*-Techniken

- Comprehension Probing: Nachfragen zum Verständnis
- Category Selection Probing: Nachfragen zur Wahl der Antwortkategorie
- Information Retrieval Probing: Nachfragen zur Informationsbeschaffung
- Confidence Rating: Bewertung der Verlässlichkeit der Antwort

Beginnen wir mit einem Beispiel für *comprehension probing*, also für Nachfragen zum Verständnis; die Frage, deren Inhalt hinterfragt werden sollte, lautet:

„In der folgenden Frage geht es um Familie und Partnerschaft. Glauben Sie, dass man eine Familie braucht, um wirklich glücklich zu sein, oder glauben Sie, man kann alleine genauso glücklich sein?“

Unmittelbar nachdem die Testpersonen die Frage beantwortet hatten, wurden sie mit der folgenden *probing*-Frage konfrontiert:

„In der Frage ist ja von ‚Familie‘ die Rede. Was verstehen Sie unter ‚Familie‘?“

Ohne die Ergebnisse dieses Tests im Detail darstellen zu wollen, zeigte sich, dass sich die Testpersonen drei Arten von Verständnis von Familie zuordnen ließen; die meisten Testpersonen verstanden unter „Familie“ entweder eine Form der Kernfamilie („Familie ist verheiratet sein, mit Kind oder ohne Kinder“ – „Unter Familie verstehe ich ein Ehepaar oder auch zwei Lebenspartner mit Kind“ – „Frau und Mann mit Kind oder Kindern“) oder eine Art der „modified extended family“ („Eltern, Geschwister, enge Verwandte“ – „Meine Ehefrau, meine Kinder, meine Eltern und Geschwister“); nur vereinzelt aufgetreten, aber dennoch interessant ist die Definition von Familie ohne Rückgriff auf deren Personalinventar: „Familie ist für mich, dass ich sagen kann, ich kann viel einbringen und bekomme auch viel zurück. Es ist ein Geben und Nehmen. Und ein gewisser Halt.“ Oder: „Zusammenhalt, Sicherheit“.

Comprehension probing ist bestens geeignet, um das Verständnis von Begriffen zu ermitteln, aber auch dazu, zu erkennen, wenn Begriffe falsch verstanden werden. Betrachten wir uns dazu das folgende (warnende) Beispiel:

„Volksbegehren und Volksentscheide sind eine notwendige Ergänzung der repräsentativen Demokratie.“

Unmittelbar nach der Bewertung dieses Items vermittels einer Skala wurde nachgefragt:

„Was – würden Sie sagen – ist das, eine „repräsentative Demokratie?“

Dass Sie den Begriff „repräsentative Demokratie“ zumindest in Fragebogen für Allgemeinpopulationen nicht verwenden sollten, zeigen die folgenden Antworten auf diese Frage:

- „Die Demokratie wird hier repräsentiert, sie wird eben repräsentiert, dargestellt.“
- „Eine Demokratie, die nach außen gerne gezeigt wird, weil sie in vielen Dingen gut ist und eine Demokratie mit Menschennähe.“
- „Was sich darstellt. Und Demokratie ist das, was wir haben. Die präsentiert sich halt.“

Sie erkennen selbst das Verständnisproblem, das mit dem Begriff „repräsentative Demokratie“ verbunden ist.

Das nächste Beispiel dient der Veranschaulichung der Vorgehensweise beim *category selection probing*, also dem Nachfragen bei der Wahl der Antwortkategorie. Vermittels einer Skala von 1 = „stimme überhaupt nicht zu“ bis 7 = „stimme voll und ganz zu“ sollten Testpersonen das folgende – Ihnen bereits bekannte – Item bewerten:

„Das Ziel der Nationalsozialisten war, durch den Krieg gegen die Sowjetunion den jüdischen Bolschewismus zu beseitigen“

Wie schon dargestellt, haben 18 von 20 Testpersonen spontan eine Bewertung dieses Items vorgenommen. Unmittelbar danach wurden die Testpersonen gefragt, warum sie sich gerade für den von ihnen gewählten Skalenwert entschieden hätten.

Eine Auswahl der Begründungen für die Entscheidung für einen bestimmten Skalenwert zeigt uns – und das ist eines der zentralen Probleme, mit denen sich die Umfrageforschung herumschlagen hat – dass Personen Fragen beantworten, ohne

überhaupt zu wissen, um was es dabei geht⁴ (in Klammern jeweils der gewählte Skalenwert):

- (4) „Das weiß ich nicht. Ich kann damit nichts anfangen.“
- (5) „Es war immer das Ziel der Nationalsozialisten, gegen die Juden zu sein.“
- (2) „Ich weiß es nicht, aber ich glaube es nicht. Habe die „2“ vergeben auf Grund meines Bauchgefühls“
- (6) „Ich vergebe ungern „1“ und „7“. Mit der „6“ möchte ich relativieren und ich glaube, dass es das Ziel der Nationalsozialisten war.“
- (4) „Bin mir relativ unsicher. Ich dachte, ich nehme einen Mittelweg. Die „4“ deshalb, weil ich es nicht weiß.“
- (1) „Weil ich mir einfach nicht vorstellen kann, dass die Nationalsozialisten ein anderes Ziel hatten. Ich vermute es, habe keinen blassen Schimmer.“

Beim *information retrieval probing* geht es darum zu erfahren, wie sich Personen die Informationen beschaffen, die sie zum Beantworten einer Frage brauchen. Dieses Verfahren wird häufig bei retrospektiven (Verhaltens-)Fragen zum Einsatz gebracht. Auch hier ein Beispiel:

Zu testende Frage:

„Wie oft sind Sie in diesem Jahr bisher zur Behandlung bei einem Arzt gewesen, egal ob im Krankenhaus oder in der Praxis?“

Antwort:

„Fünfmal.“

Nachfrage:

„Wie haben Sie sich daran erinnert, dass Sie in diesem Jahr fünfmal beim Arzt gewesen sind?“

Antwort:

„Ich habe im Januar angefangen und dann Monat für Monat überlegt, wie oft ich beim Arzt war, aber an manche Monate konnte ich mich nicht richtig erinnern.“

Und schließlich die vierte zentrale Probing-Technik, das *confidence rating*. Bei dieser Technik soll die Testperson nach der eigentlichen Antwort den Grad der Verlässlichkeit dieser Antwort selbst bewerten, entweder offen, zumeist aber auch mit Hilfe einer Skala, z.B. so:

„Wie sicher waren Sie sich beim Beantworten der Frage – sehr sicher, eher sicher, eher unsicher oder sehr unsicher?“

Oder:

„Wie genau ist Ihre Antwort – sehr genau, ziemlich genau, eher ungenau oder nur grob geschätzt?“

Und dann, immer wenn eine Antwortkategorie gewählt wird, die Unsicherheit erkennen lässt, wird die Nachfrage gestellt, warum man beim Beantworten der Frage nicht „sehr sicher“ oder „eher sicher“ gewesen ist bzw. warum man die eigene Antwort für „eher ungenau“ hält oder „nur grob geschätzt“ habe.

Wie Sie und Ihre SchülerInnen kognitive Pretestinterviews durchführen könn(t)en

Zunächst einmal sollten Sie wissen, dass es zur Durchführung kognitiver Interviews eigentlich keine „festen Regeln“ gibt. Die Spannweite der Vorgehensweisen reicht hier von völlig unstrukturiert und vollkommen unstandardisiert (quasi wie ein Tiefeninterview) bis hin zu völlig strukturiert und vollkommen standardisiert (quasi wie ein standardisiertes Interview); am wirkungsvollsten sind Vorgehensweisen, die zwischen diesen beiden extremen Polen liegen.

Weiterhin gibt es (fast) auch keine Regeln über die Auswahl und Anzahl der Testpersonen. Bei der Anzahl der Testpersonen sollten Sie aber daran denken, dass Sie mit kognitiven Interviews qualitative Daten(massen) produzieren, die es später interpretativ auszuwerten gilt.

Die Frage, welche Personen Sie für den Test auswählen, beantwortet sich über die Zielgruppe Ihrer Hauptbefragung; die Testpersonen sollten grundsätzlich Personen sein, die später auch in der Hauptbefragung befragt werden könnten. Planen Sie eine Befragung einer relativ homogenen Zielpopulation, ist die Definition der Testpersonen eher einfach, planen Sie eine Befragung einer Allgemeinpopulation, sollten Sie bei der Auswahl Ihrer Testpersonen darauf achten, dass Sie zumindest Personen unterschiedlichen Geschlechts, unterschiedlichen Alters und unterschiedlicher formaler Bildung testen. Was die Bildung angeht, empfehle ich Ihnen, eher auf Personen mit niedriger und Personen mit hoher Bildung zurückzugreifen, weniger auf Personen mit mittlerer Bildung; denn, etwas platt ausgedrückt, sind es die „Dümmeren“ und die „Klügeren“, die uns im kognitiven Pretest die besseren Informationen liefern, nicht die Personen mit „mittlerer“ Bildung.

Unabhängig von dem Befragungsmodus Ihrer geplanten Befragung muss das kognitive Interview persönlich-mündlich erfolgen, oder sagen wir mal vorsichtiger: Den größten Nutzen aus kognitiven Pretestinterviews erzielen Sie, wenn Sie der Testperson gegenüber sitzen und dadurch nicht nur deren verbale Reaktionen erfassen können, sondern auch deren Mimik und Gestik.

Wie sollten oder könnten Sie vorgehen, wenn Sie einen kognitiven Pretest durchführen wollen? Ich würde Ihnen folgende Vorgehensweise empfehlen:

1. Legen Sie fest, welche Ihrer Fragen im kognitiven Pretest getestet werden sollen. Testen Sie Fragen, die für Sie von besonderer Wichtigkeit sind oder/und solche Fragen, bei denen Sie Probleme erwarten, z.B. Probleme mit dem Verständnis der Frage oder einzelner ihrer Bestandteile. Bei kurzen Fragebogen können Sie natürlich alle inhaltlichen Fragen testen.
2. Denken Sie dabei daran, dass die Gesamtdauer Ihres kognitiven Pretests möglichst nicht länger als 60 Minuten sein sollte. Die kognitive Vorgehensweise ist für die Testpersonen sehr anstrengend, und wenn Sie zu lange testen, müssen Sie mit Abstrichen bei der Antwortqualität rechnen.
3. Arbeiten Sie halbstandardisiert. Erstellen Sie zunächst einen „kognitiven Fragebogen“, der die zu testenden Fragen und die zugeordneten kognitiven Techniken enthält. Lassen Sie Platz für Kommentare der Testperson sowie Mitschriften und Kommentare des Testleiters/der Testleiterin (also z.B. für Ihre eigenen Erkenntnisse aus der speziellen Testsituation). Die halbstandardisierte Vorgehensweise erleichtert Ihnen später die ansonsten (und auch so noch immer) sehr aufwändige Auswertung.
4. Zeichnen Sie die kognitiven Interviews mit – wenn das möglich ist – einer Videokamera auf, zumindest aber mit einem Tonbandgerät oder digitalem Aufnahmegerät.

5. Wählen Sie die Testpersonen so aus, dass sie zur Menge Ihrer späteren Zielpersonen gehören könnten. Instruieren Sie die Testpersonen dahingehend, dass sie nicht als „Befragungspersonen“ im üblichen Sinne mit Ihnen arbeiten, sondern als Quasi-„Experten“, die Ihnen helfen sollen, Ihren Fragebogen zu verbessern.
6. Wenn es Ihnen möglich ist, sollten Sie den Testpersonen ein Honorar für ihre Mitarbeit zahlen; angemessen erscheint ein Betrag von 30 Euro für eine Stunde Teilnahme am kognitiven Interview. Stehen Ihnen keine Mittel zur Bezahlung der Testpersonen zur Verfügung, müssen Sie stärker auf Ihren Charme und Ihre Überredungskunst bauen, vielleicht auch auf eine Art „Mitleidseffekt“.
7. Sie brauchen normalerweise mindestens 5, höchstens 30 Testpersonen, wobei Sie bei der Anzahl der Testpersonen immer den Aufwand für die Verarbeitung der im Interview gewonnenen Texte berücksichtigen sollten. Stellen Sie schon nach den ersten Interviews bestimmte Schwächen bei einer Frage fest, müssen Sie sich natürlich diese Schwächen nicht auch noch von weiteren Testpersonen aufzeigen lassen.
8. Zum Ablauf: Lassen Sie die Testperson die zu testende Frage beantworten. Bringen Sie dann die im kognitiven Fragebogen vorbereiteten und geplanten kognitiven Techniken zum Einsatz. Schieben Sie bei Bedarf spontan weitere, nicht vorbereitete kognitive Techniken und Fragen nach. Fragen Sie so lange nach, bis Sie sich sicher sind, die Testperson „verstanden“ zu haben.
9. Zur Auswertung: Erstellen Sie für jede zu testende Frage eine fallspezifische Auflistung, in der Sie die Antworten auf die getestete Fragen festhalten, spontane Äußerungen der Testperson zur Frage, die Antworten und Reaktionen der Testperson auf die kognitiven Nachfragen und die Anmerkungen des Testleiters/der Testleiterin (also z.B. Ihre eigenen Anmerkungen).
10. Ziehen Sie auf der Grundlage dieser fallspezifischen Auflistungen Ihre Schlüsse auf die Qualität Ihrer Frage. Beachten Sie aber: Wir haben es bei der Interpretation der Daten aus kognitiven Interviews mit qualitativen Daten zu tun, deren Auswertung interpretativ und oft nicht ohne Beeinflussung durch subjektive Komponenten erfolgt. Wenn Sie also nicht möchten, dass Sie aus den Daten das herauslesen, was Sie schon vor dem Test erwartet hatten, sollten Sie – das wäre grundsätzlich zu empfehlen – die Auswertung eines kognitiven Pretests nie alleine vornehmen. Lassen Sie zwei oder mehrere Ihrer SchülerInnen die Auswertung parallel vornehmen. Vergleichen Sie dann die Interpretationen. Haken Sie ab, wo es Übereinstimmungen gibt; kommt es nicht zu einer übereinstimmenden Interpretation, und können Sie eine solche im Diskurs auch nicht herbeiführen, sollten Sie weitere Personen zu Rate ziehen, die die Texte zu der entsprechenden Frage ebenfalls bearbeiten und ihre Schlüsse ziehen.

Jetzt können Sie aufgrund der Ergebnisse Ihres kognitiven Pretests Ihren Fragebogenentwurf überarbeiten. Natürlich wäre es wünschenswert, wenn Sie die aufgrund des Pretest veränderten Fragen jetzt noch einmal kognitiv testen würden, aber weil das natürlich den Aufwand an Zeit und Geld erhöhen würde, werden Sie das im Normalfalle nicht tun. Aber wenn Sie die gerade beschriebenen zehn Schritte absolviert haben, haben Sie schon sehr viel für die Verbesserung Ihres Fragebogenentwurfes geleistet.

Wenn Sie jetzt vielleicht resigniert sein sollten, weil Sie den Aufwand eines kognitiven Interviews mit den Ihnen verfügbaren Mitteln nicht schultern können: Es gibt durchaus weitere, weniger aufwändige Verfahren, um die Qualität eines Fragebogenentwurfes zu prüfen und zu verbessern: Lassen Sie auf jeden Fall eine Kollegin oder

einen Kollegen einen Blick auf Ihren Fragebogenentwurf werfen, wenn Sie zwei oder drei KollegInnen haben, die Ihnen hier helfen – umso besser. Sie wissen, dass man seinen eigenen Texten gegenüber sehr schnell betriebsblind wird.

Nehmen Sie Ihren Fragebogen-Entwurf, suchen Sie sich ein paar wenige potenzielle Befragungspersonen und „simulieren“ Sie auf einfache Weise einen kognitiven Pretest: Lassen Sie diese Personen Ihre Fragen beantworten und sprechen Sie dabei oder danach mit ihnen: Wo hat es Probleme gegeben? Welche Fragen oder Begriffe in Fragen waren unklar gewesen und warum? Welche Fragen könnte man anders und besser formulieren?

Damit kommen Sie ohne großen Aufwand schon ein gutes Stück weiter. Lassen Sie sich selbst Strategien einfallen, wie Sie Ihren Fragebogenentwurf verbessern könnten. Alles, was Sie in dieser Phase der Fragebogenentwicklung tun, ist besser als nichts zu tun.

Zum Schluss

Jetzt, da wir unseren Fragebogen konstruiert und durch Pretests evaluiert haben, können wir uns auf die Durchführung unserer Befragung konzentrieren. Bevor wir aber unsere Fragebögen „in’s Feld“ schicken, sollten wir uns im Klaren darüber sein, wen wir eigentlich befragen wollen. Vom Ablauf des Forschungsprozesses muss dieser Schritt natürlich lange vor dem Fragebogen stehen. Wir müssen uns schon sehr früh darüber im Klaren sein, wer die Zielgruppe unserer Befragung ist und über wen wir Aussagen treffen wollen.

Im nächsten Beitrag zur Serie „Forschen in der Schule“ werden wir uns mit den Begriffen Population, Vollerhebung, Stichprobe und Repräsentativität beschäftigen. Wenn wir im Fernsehen oder in der Zeitung oder sonst wo Ergebnisse einer Befragung zur Kenntnis nehmen, stellen wir gerne mit einer gewissen Erleichterung fest, dass es sich bei der jeweiligen Studie um eine „repräsentative“ Befragung handelt. Warum erleichtert uns das? Was ist das überhaupt: „Repräsentativität“? Dazu müssen wir Begriffe wie „Population“, „Vollerhebung“ und „Stichprobe“ näher erläutern. Insbesondere sollten wir wissen, welche Arten von Stichproben es gibt und welche Arten von Aussagen welche Arten von Stichproben zulassen. Und natürlich die Frage: Können SchülerInnen „repräsentative“ Befragungen durchführen (Antwort: Ja!) und sollten SchülerInnen „repräsentative“ Befragungen durchführen (Antwort: Warum sollten sie?).

Anmerkungen

- 1 Für die kritische Durchsicht des Manuskripts und Vorschläge zu seiner Verbesserung danke ich Ruth Holthof, Studienrätin im Fach Sozialkunde am Eleonoren-Gymnasium in Worms.
- 2 Der Verzicht auf Standardbeobachtungspretests bringt es mit sich, dass wir einige der Informationen auf der gerade dargestellten Liste nicht erhalten können, vor allem Häufigkeitsverteilungen, technische Probleme bei Fragebogen und Befragungshilfen oder Zeitdauer der Befragung. Wenn Sie Näheres zu Standardbeobachtungspretests erfahren möchten, empfehlen wir Ihnen die kurze Darstellung bei Porst (2014b: 191f).
- 3 Wenn Sie beim Lesen dieser Aussage jetzt den Kopf schütteln und der Ansicht sind, niemand würde ernsthaft ein solches Item in einer Befragung vorlegen, muss ich Sie leider enttäuschen: Dieses Item wurde nicht von mir zur besseren Illustration erfunden, sondern

war in einem von mir beratenen Fragebogenentwurf zur Befragung einer Allgemeinpopulation enthalten; durch die Ergebnisse unseres Pretests konnten wir seine Verwendung in der Hauptbefragung allerdings verhindern.

- 4 Der Nachweis zu dieser Behauptung lässt sich mit kognitiven Pretests eigentlich einfach leisten; kognitive Pretests helfen aber auch, Lösungen für dieses Problem zu finden, indem sie Verständnisprobleme überhaupt erst aufzeigen, in Zusammenarbeit mit den Testpersonen zu alternativen Formulierungen oder – wenn es gar nicht anders geht – zu der Erkenntnis führen können, dass man auf eine Frage oder ein Item besser gleich ganz verzichten sollte, bevor man mit unverständlichen Fragen „Rauschen“ produziert.

Literatur

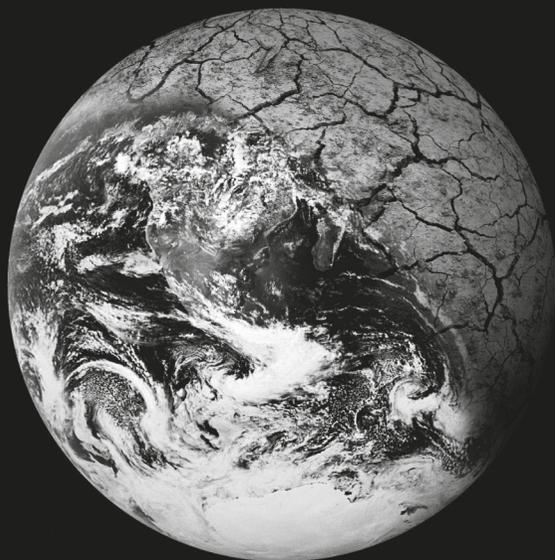
- Porst, R. (2000): Praxis der Umfrageforschung. 2. Auflage. Wiesbaden: Teubner
- Porst, R. (2014a): Von der Fragestellung zum Fragebogen – Zentrale Begrifflichkeiten. S. 221 – 232 in Gesellschaft. Wirtschaft. Politik 63, Heft 2
- Porst, R. (2014b): Fragebogen. Ein Arbeitsbuch. 4. Auflage. Wiesbaden: Springer VS
- Porst, R. (2015a): Arten von Befragungen und Befragungstechniken. S. 93 – 103 in Gesellschaft. Wirtschaft. Politik 65, Heft 1
- Porst, R. (2015b): Fragebogenkonstruktion (I) – Grundlagen, Arten von Fragen, Arten von Skalen. S. 239 – 250 in Gesellschaft. Wirtschaft. Politik 64, Heft 2
- Porst, R. (2015c): Fragebogenkonstruktion (II) – Question Wording und was sonst noch wichtig ist. S. 381–392 in Gesellschaft. Wirtschaft. Politik 64, Heft 3
- Sudman, S. & Bradburn, N.M. (1982): Asking Questions. San Francisco: Jossey-Bass

Ohne Wachstum ist alles nichts?

Blätter für deutsche und internationale Politik (Hg.)

MEHR GEHT NICHT!

Der Postwachstums-Reader



editionBlätter

Alberto **Acosta**
Elmar **Altvater**
Maude **Barlow**
Ulrich **Brand**
Jayati **Ghosh**
David **Harvey**
Tim **Jackson**
Naomi **Klein**
Serge **Latouche**
Niko **Paech**
Vandana **Shiva**
Harald **Welzer**
u.v.a.

336 Seiten, 18,00 Euro
ISBN: 978-3980492591

Ab sofort im Buchhandel – und als Dankeschön für
jedes neue »Blätter«-Abo. Mehr auf www.blaetter.de