



QuickStart

«Pretest»

Holen Sie sich die Sicherheit, einen geprüften Fragebogen ins Feld zu schicken!

Einleitung

Die Entwicklung eines Fragebogens ist eine komplexe und nicht ganz einfache Angelegenheit. Haben Sie die Fragen Ihres Fragebogens erst einmal erfasst, ist es allerdings nicht ratsam, gleich mit der Befragung zu starten. Auch wenn Sie bereits Erfahrung im Konstruieren von Fragebögen jeglicher Art haben, bleibt immer ein Restrisiko. Wir empfehlen darum, Ihren Fragebogen vor der eigentlichen Feldphase unbedingt testen zu lassen, um Mängel mit hoher Wahrscheinlichkeit auszuschliessen. Die Frage, die Sie sich dabei stellen sollten, lautet immer: „Werden mir die Daten des Fragebogens gute, sprich valide und reliable Daten liefern?“ Diese Frage können Sie mit an Sicherheit grenzender Wahrscheinlichkeit nicht ohne einen oder sogar mehrere Pretests beantworten.

„If you don't have the resources to pilot test your questionnaire, don't do the study.“
(Sudmann/Bradburn 1982)

Ähnlich wie bei der Konstruktion der Fragen Ihres Fragebogens gibt es auch bei der Durchführung eines Pretests keine empirisch fundierten Konstruktionsrichtlinien. Die Entscheidung darüber, welche Pretest-Variante(n) für Ihren Fragebogen am Geeignetsten ist, müssen Sie von Fall zu Fall neu treffen und die Entscheidung dabei Ihrem Fragebogentyp anpassen. Damit Sie wissen, welche Möglichkeiten Ihnen dabei zur Verfügung stehen, geben wir Ihnen in diesem Dokument einen Überblick über die verschiedenen Varianten zur Prüfung Ihres Fragebogens. Wir lehnen uns dabei an die Ausführungen von Prüfer und Rexroth (1991).

Testerhebungen im Feld

In der Folge stellen wir Ihnen Methoden vor, wie Sie Ihren Fragebogen „im Feld“ testen können. „Im Feld“ bedeutet dabei, dass der Fragebogen anhand einer ausgewählten Stichprobe systematisch getestet wird. Dabei stehen Ihnen folgende Verfahren zur Verfügung:

Standard-Pretest: Der Standard-Pretest wird sehr häufig angewendet und ist ein etabliertes Verfahren in der Umfrageforschung. Ziel dieser Methode ist es, mithilfe eines Interviews aus der Reaktion bzw. Antwort der Befragten Rückschlüsse auf das Frageverständnis zu ziehen. Haben die Befragten Probleme bei der Beantwortung der Fragen, lässt sich dies immer auf Konstruktionsmängel zurückführen und zeigt dem Forscher so, wo er seine Fragen noch optimieren muss. Kennzeichnend für einen solchen Test sind folgende Charakteristiken:

- Einmalige Erhebung eines Fragebogens unter möglichst realistischen Bedingungen
- 20-50 Interviews
- Interviewer haben die Aufgabe, Probleme und Auffälligkeiten zu beobachten und dem Forscher zu berichten.
- Passives Verfahren, bei dem der Interviewer nur beobachtet, ohne aktiv zu hinterfragen.
- Man kann die Befragten entweder über die Test-Strategie informieren (participating pretest) oder aber nicht (undeclared pretest).

Wer das Interview durchführt, ist nicht festgelegt. Manche Autoren empfehlen nur Experten, andere sagen, dass auch diverse Mitglieder der Projektgruppe die Befragung durchführen können. Wie so häufig müssen Sie auch hier für sich persönlich entscheiden, was die beste Strategie ist. Die Stärken dieses Verfahrens sind, dass es relativ schnell und problemlos durchführbar ist, es einen geringen organisatorischen Aufwand braucht, die Kosten niedrig sind und eine annähernd realistische Schätzung der Befragungsdauer möglich ist. Auf der anderen Seite kann trotz „formal“ korrekter Antworten ein falsches Fragenverständnis vorliegen. Zudem ist die Instruktion der Interviewer meist wenig präzise, die Berichte also inhaltlich wie formal wenig systematisch. Insgesamt gesehen ist der Standard-Pretest ein sehr grobes Verfahren.

Behaviour-Coding: Das Prinzip dieses Verfahrens basiert auf der Klassifizierung von Verhalten. Mit einem mehr oder weniger detaillierten Codesystem wird das Verhalten von befragten Personen und Interviewer bewertet und analysiert. Somit lassen sich Rückschlüsse auf die Qualität einer Frage ziehen. Das Behaviour-Coding ist also eine gute Alternative wenn es darum geht, die Fragenqualität zu bestimmen. Zentrales Element bei diesem Pretest ist das Codesystem, das mehr oder weniger differenziert das Verhalten erfassen soll. Hier ein Beispiel eines Codesystems.

Interviewer

<u>Code</u>	<u>Beschreibung</u>	
-------------	---------------------	--

E	Exact	Interviewer liest Frage exakt
---	-------	-------------------------------

S	Slight change	leichte Änderungen
---	---------------	--------------------

M	Major change	starke Änderungen
---	--------------	-------------------

Befragte(r)

❶	Interruption	Befragte(r) antwortet vorzeitig
---	--------------	---------------------------------

	Clarification	Befragte(r) will Wiederholung der Frage oder Klärung der Frage oder macht Bemerkung, die auf Verständnisproblem schließen läßt
--	---------------	--

	Adequate answer	Befragte(r) antwortet adäquat
--	-----------------	-------------------------------

	Qualified answer	Antwort ist adäquat, zusätzliche Bemerkung läßt jedoch auf Unsicherheit schließen
--	------------------	---

	Inadequate answer	Inadäquate Antwort
--	-------------------	--------------------

	Don` t know	Weiß nicht
--	-------------	------------

❷	Refusal to answer	Befragte(r) verweigert Beantwortung der Frage
---	-------------------	---

Vor der Vercodung sollte festgestellt werden, welches Verhalten überhaupt berücksichtigt werden soll. Es gibt nach Oskenberg, Cannell und Kalton (1991) grundsätzlich drei Varianten:

- Vercodet wird nur die erste Reaktion des Befragten
- Vercodet wird das gesamte Befragtenverhalten, es sind also mehrere Codes pro Frage möglich
- Vercodet wird das gesamte Befragtenverhalten, im Unterschied zur zweiten Variante werden aber identische Codes nur einmal vergeben, auch wenn sie mehrmals auftreten sollten

Sie sollten verschiedene Codes einbauen, wovon einer ein adäquates Antwortverhalten suggeriert. Wenn man dieses Item dann bei der Auswertung herausnimmt, gilt folgende Regel: Je mehr Codes vergeben wurden, desto schlechter (unverständlicher) ist der Fragebogen. In diesem Beispiel wären es ohne die Nennungen für adäquates Antwortverhalten (Item 3) 13 Code-Nennungen. Es gäbe bei diesem Fragebogen also noch einiges an Verbesserungspotenzial.

Code-ziffer	ID 1	ID 2	ID 3	ID 4	ID 5	ID 6	ID 7	ID 8	ID 9	ID 10	N
1		x						x			2
2	x				x						2
3			x	x	x			x	x	x	6
4	x			x	x						3
5	x	x					x				3
6		x					x				2
7						x					1
											19

Im Vergleich zu den anderen Pretest-Techniken ist die Behaviour-Coding-Technik die mit Abstand reliabelste Technik, besonders im Vergleich zum Standard-Pretest, bei dem die Pretest-Erkenntnisse oft von der subjektiven Wahrnehmung der Interviewer abhängen. Dagegen ist die Behaviour-Technik viel objektiver und systematischer. Dafür erfasst dieses Verfahren keine Hinweise auf mögliche Ursachen für inadäquates Verhalten. Auch hier kann zudem einer formal korrekten Antwort ein falsches Fragenverständnis zugrunde liegen.

Problem-Coding: Das Problem-Coding ist eine Variante des Behaviour-Coding. Die Bewertung der Verhaltensweisen der Befragten wird dabei nicht vom Coder nach dem Interview, sondern vom Interviewer selbst während des Interviews vorgenommen. Es gibt also kein ausführliches,

detailliertes Codesystem, sondern ein reduziertes Codesystem, welches spontanes Befragtenverhalten nur noch im Hinblick darauf, ob es im Sinne der Fragestellung adäquat oder nicht adäquat ist, bewertet wird. Es handelt sich also um ein binäres, 0-1 codiertes Codesystem. Bei einer inadäquaten Verhaltensweise beschreibt der Interviewer im Erfahrungsbericht dann die Situation möglichst genau. Der Vorteil dieser Variante ist, dass man neben dem quantitativen Häufigkeitswert als Qualitätsindikator auch Hinweise auf die konkreten Ursachen für den Qualitätsmangel einer Frage erhalten kann.

Random-Probe: Jeder Interviewer wählt vor dem Interview nach einem Zufallsverfahren eine bestimmte Anzahl an Fragen aus, bei denen jeweils Zusatzfragen zum Verständnis gestellt werden. So kann überprüft werden, ob die Befragten die Frage auch wirklich richtig verstanden haben.

Intensive Interview: Auch bei formal korrekten Antworten kann ein falsches Verständnis des Frageinhalts zugrunde liegen. Deshalb werden die Probanden im Intensive-Interview nach der Durchführung eines Standard-Pretests zum Verständnis von drei bis vier bereits vorher festgelegten Fragen intensiv befragt. Dabei wird ein zweistufiges Vorgehen angewendet:

1. Der Interviewer liest die Frage und die bereits vorliegende Antwort nochmals vor. Dann soll der Befragte berichten, wie seine Antwort zustande kam. Falls nötig, sollte der Interviewer extensiv nachfragen.
2. In einem zweiten Schritt stellt der Interviewer vorgegebene Fragen, um festzustellen, wie bestimmte Aspekte interpretiert wurden.

Unstrukturierte/ qualitative Interviews/ Tiefeninterviews: Diese Verfahren können vor allem in einer frühen Entwicklungsphase des Fragebogens eingesetzt werden, denn sie besitzen einen explorativen und experimentellen Charakter. Bei dieser Vorgehensweise wird kein standardisiertes Interview durchgeführt, sondern der Interviewer kann unterbrechen, nachfragen oder hinterfragen. Diese Interviews dienen dazu, Hinweise und Informationen zur Fragenkonstruktion zu generieren.

Analyse der Antwortverteilungen: Mithilfe von Häufigkeitsverteilungen von Antwortalternativen können Rückschlüsse auf die Qualität der Frage ziehen. Indikatoren dabei sind nicht-besetzte oder minimal-besetzte Antwortkategorien, extreme Häufigkeitsverteilungen über alle Antwort-Kategorien oder hohe Häufigkeitswerte bei „Ausweichkategorien“ (z.B. „weiss nicht“). Dieses Verfahren ist nur sinnvoll bei einer grossen Fallzahl.

Split-Ballot: Bei einem Split-Ballot-Pretest werden zwei Varianten einer Frage jeweils einer Teilgruppe zur Beantwortung präsentiert. Unterschiede in Antwortverteilungen werden dann auf unterschiedliche Fragevarianten zurückgeführt. Ziel dieser Methode ist es, eine Entscheidung für diejenige Fragevariante zu treffen, die zum Einsatz kommen soll. Diese Entscheidung kann auch anhand von statistischen Analysen durchgeführt, weswegen der Stichprobenumfang mindestens 100 Interviews betragen sollte.

Kognitive Laborverfahren

Waren die Pretest-Methoden früher eher von Feld-Verfahren geprägt, so veränderten vor allem kognitionspsychologische Theorien und Techniken die Benutzung von Pretests. Heute sind kognitive Laborverfahren ein bewährtes und oft angewandtes Mittel zum Prüfen eines Fragebogens. Bei diesen Verfahren werden Hinweise über kognitive Prozesse während des Antwort-Prozesses gesammelt. Diese Informationen geben dann Hinweise darüber, wie die Befragten eine Frage oder ein bestimmtes Element davon verstehen und interpretieren. Kognitive Laborverfahren zeichnen sich dementsprechend durch eine geringe Teilnehmerzahl aus. In der Folge stellen wir Ihnen die wichtigsten kognitiven Pretest-Verfahren vor:

Think-Aloud: Die Think-Aloud-Technik ist die zentrale kognitive Methode überhaupt. Bei dieser Technik soll der Befragte laut denken und sämtliche Gedankengänge, die zur Antwort führen, formulieren. Das Ziel ist, aus den Äusserungen des Befragten Hinweise über die Verständlichkeit der Frage oder einzelner Begriffe zu erhalten (verbal protocols). Dabei gibt es zwei zentrale Vorgehensweisen: zum einen werden die Befragten dazu aufgefordert, während der Antwortformulierung laut zu denken (Concurrent-Think-Aloud-Methode). Zum anderen sollen die Befragten nach der Beantwortung der Frage nachdenken, wie diese zu Stande kam (Retrospektive-Think-Aloud-Technik). Die Think-Aloud-Technik wurde in den unterschiedlichsten Bereichen bereits erfolgreich eingesetzt:

- Bei retrospektiven Fragen macht die Think-Aloud-Technik den Antwortprozess transparent, wodurch sich die Frageformulierung verbessern lässt.
- Bei Meinungsfragen demonstriert die Think-Aloud-Technik häufig eine Unzulänglichkeit der Itemformulierung (z.B. eine zu breite Generalisierung der zu bewertenden Situationen).
- Zur Überprüfung von Hypothesen

Probing: Beim Probing wird die gegebene Antwort vom Interviewer durch eine oder mehrere Zusatzfragen (Probes) hinterfragt, um mehr Informationen zu erhalten. Das Probing-Verfahren ist auch ein wichtiges Element bei der Random-Probe und dem Intensiv-Interview. Es gibt zwei unterschiedliche Probing-Arten:

- Das Follow-Up-Probing, das während des Interviews sofort nach der spontanen Antwort durchgeführt wird.
- Das Post-Interview-Probing, welches nach dem Interview durchgeführt wird.

Unabhängig vom Zeitpunkt kann nach der Aufgabenstellung, auf die sich das Probing bezieht, unterschieden werden. Dadurch entstehen zwei weitere Varianten:

- Comprehension-Probing, das zum Frageverständnis verwendet wird. Die Befragten sollen dabei die Bedeutung eines bestimmten Begriffs oder Aspekte ihrer Antwort erläutern und somit zeigen, wie verständlich ein Begriff war.

- Retrieval-Probing, welches zur Informationsbeschaffung dient. Dieses Verfahren ist besonders bei retrospektiven Fragen sinnvoll (z.B. „Wann waren Sie das letzte Mal beim Zahnarzt?“)

Confidence-Ratings: Bei dieser Technik sollen die Befragten nach der eigentlichen Antwort den Grad der Verlässlichkeit ihrer Antwort bewerten. Zur Hilfe wird dabei meist eine Skala verwendet.

Paraphrasing: Beim Paraphrasing sollen die Befragten nach der Beantwortung die Frage mit eigenen Worten wiederholen, entweder Wort für Wort oder sinngemäss. Diese Technik kann dem Forscher aufschlussreiche Hinweise darüber geben, welche inhaltlichen Aspekte die Befragten mit einer Frage verbinden, und andererseits kann die Technik zeigen, ob der Fragetext in allen Aspekten erinnert werden kann.

Sorting-Verfahren: Diese Verfahren sollen vor allem Hinweise darüber geben, wie der Befragte Begriffe kategorisiert, beziehungsweise als Konzept versteht. Es gibt drei Arten von Sorting-Verfahren:

- Free Sort: Die vorgegebenen Items sollen nach eigenen Kriterien gruppiert werden.
- Dimensional Sort: Gleiches Vorgehen wie bei Free Sort, nur dass hier die vorgegebenen Items nach vorher festgelegten Kriterien sortiert werden sollen.
- Vignette Classifications: Ist eine Variante des Dimensional Sort. Die Befragten sollen z.B. kurze Situationsbeschreibungen (Vignettes) lesen und entscheiden, ob diese in die Überlegungen zum Beantworten einer Frage mit einbezogen werden sollen oder nicht. Diese Klassifizierungen werden vor allem zur Bestimmung des Verständnisses bestimmter Begriffe eingesetzt.

Response Latency: Bei diesem Verfahren wird die Zeit zwischen der Präsentation der Frage und der Antwort des Befragten gemessen. Dabei gilt eine lange Reaktionszeit als Indikator für Fragemängel.

Wir haben Ihnen nun verschiedene Möglichkeiten beschrieben, wie Sie Ihren Fragebogen einem kognitiven Pretest unterziehen können. In der folgenden Tabelle sehen Sie eine Liste der jeweiligen Vor- und Nachteile von kognitiven Pretests.

Vorteile	Nachteile
Schnelle Durchführung	Techniken beschränken sich vorwiegend auf die Evaluation einzelner Fragen und nicht auf den Fragebogen als Ganzes. Sie können also nicht den Test des gesamten Instruments ersetzen!
Niedrige Kosten	
Techniken können innerhalb verschiedener Stadien der Fragebogenkonstruktion angewandt werden	Unsicherheitsrisiko bezüglich der Generalisierbarkeit der Ergebnisse durch geringe Fallzahl

Andere Verfahren

Zum Schluss stellen wir Ihnen noch eine Reihe von anderen Verfahren vor, die Sie ebenfalls für Ihren Pretest anwenden können.

Focus Groups: Bei dieser Technik werden Gruppen gebildet. Jedes der Gruppenmitglieder bearbeitet dabei zuerst für sich den Fragebogen (ohne Rückfragen). Danach können die Mitglieder allgemeine Eindrücke zu Thematik, Dauer, Schwierigkeiten und anderem äussern. Anschliessend wird der Fragebogen Frage für Frage durchgearbeitet. Zusätzlich werden vom Moderator bereits vorbereitete Fragen zu einzelnen Fragen gestellt, insbesondere über das Verständnis von Fragen oder einzelnen Begriffen. Focus Groups können auf zwei Arten sinnvoll eingesetzt werden:

- In der frühen Entwicklungsphase des Fragebogens können sie wertvolle Hinweise zur Akzeptanz oder dem Verständnis des Befragungsthemas, einzelner Themenbereiche, einzelner Fragen oder einzelner Begriffe liefern.
- Sie eignen sich vor allem zum Testen von schriftlichen Fragebögen.

Es empfiehlt sich, die Gespräche der Gruppen auf einem Tonträger festzuhalten, um die Analyse besser durchführen zu können. Der Nachteil dieses Verfahrens ist, dass es für die Evaluation einzelner Fragen nur bedingt geeignet ist. Zudem entsteht in Gruppen immer soziale Interaktion und Gruppendynamiken, welche das Individualverhalten beeinflussen und so die Meinung verzerrt darstellen. Demgegenüber liegt der Vorteil darin, dass mehrere Personen gleichzeitig befragt werden können.

Experten: Ein probates Mittel ist sicherlich auch das Herbeiziehen von Umfrageexperten, idealerweise von mehreren.

Leistungsfähigkeit einzelner Verfahren

Wie bei der Frageformulierung gibt es auch beim Pretest keine Methode, die in allen Problembereichen zufriedenstellend arbeitet. Presser und Blair (1994) beispielsweise sagen, dass der Standard-Pretest im Vergleich zu anderen Verfahren am wenigsten reliabel ist. Im Gegensatz dazu steht das Behaviour-Coding, das allerdings wiederum keine Hinweise auf die Ursachen der Probleme gibt. Kognitive Verfahren liefern die meisten Lösungen für Verständnisprobleme, während Expertenrunden allgemein vergleichsweise die meisten Erkenntnisse liefern und am kostengünstigsten sind, dafür aber Defizite bei Hinweisen auf Interviewer-Probleme liefern. Es empfiehlt sich also zusammenfassend, mehrere Varianten sinnvoll zu kombinieren. Dabei sollten Sie bei jedem Fragebogen wieder aufs Neue schauen, welche Methoden für den jeweiligen Fragebogen am meisten Sinn macht.

Ihre Ansprechpartner

Für Fragen zu unseren Pretest-Produkten und Verfahren kontaktieren Sie uns einfach per E-Mail info@onlineumfragen.com oder per Telefon unter +41 44 500 5054. Wir stehen Ihnen gerne für Auskünfte, Beratung und Unterstützung zur Verfügung.

Wer sind wir?

Onlineumfragen.com ist ein führender Schweizer Partner für Onlineumfragen. Unsere ner Kunden erstellen selbst oder mit unserer Unterstützung eigene Umfragen: schnell und einfach.

Jährlich mehrere Millionen von Teilnehmer-Feedbacks erleichtern unseren Kunden seit 1999 fundierte, mehrperspektivische Entscheidungen, viele davon für grosse Europäische Unternehmen, Banken, Versicherungen, öffentliche Institutionen und Universitäten.

Onlineumfragen.com GmbH
Untere Gründlistrasse 26
6055 Alpnach (Obwalden)
Schweiz

Telefon +41 44 500 5054 Telefax +41 44 500 5105
E-Mail info@onlineumfragen.com
<http://www.onlineumfragen.com>

Literatur:

Oskenberg, L./Cannell, Ch./Kalton, G. (1991): „New Strategies for Pretesting Survey Questions“, in: *Journal of Official Statistice*, 7, S. 349-365.

Presser, S./Blair, J., (1994): Survey Pretesting: „Do different Methods produce different Results?“, in: *Sociological Methodology*, S. 73-104.

Prüfer, Peter und Margrit Rexroth (1996): Verfahren zur Evaluation von Survey-Fragen: ein Überblick. Mannheim: Zentrum für Umfragen, Methoden und Analysen ZUMA.

Sudman, S./Bradburn, N. (1982): Asking Questions. A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass.