

---

# Das klassische Delphi-Verfahren: Konzept und Vorgehensweise

## 2

Das Delphi-Verfahren wurde als dialogorientiertes Verfahren von der RAND Co. in den 50er und 60er Jahren entwickelt (vgl. Dalkey & Helmer 1963). Ursprünglich in Experimenten für militärische Zwecke eingesetzt, wird es mittlerweile vor allem als Prognose- und Zukunftsforschungsinstrument, in der Betriebswirtschaftslehre aber auch in pflege- und gesundheitswissenschaftlichen Kontexten verwendet (vgl. Häder & Häder 1995: 9; Mintroff und Turoff 1975; Benarie 1988; Häder 2014; Meister & Oldenburg 2008, Cuhls 2012, Ab Latif et al. 2016; Robertson et al. 2017; Ammon 2009).

Eine einheitliche Begriffsdefinition gibt es bis heute nicht. Nach Dalkey und Helmer eignet sich das Verfahren “to obtain the most reliable consensus of opinion of a group of experts ... by a series of intensive questionnaires interspersed with controlled feedback” (1963: 458). Linstone und Turoff (1975: 3) charakterisieren ein Delphi „as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem“. Häder und Häder (1995: 12) definieren die Delphi-Methode als einen vergleichsweise stark strukturierten Gruppenkommunikationsprozess, in dessen Verlauf Sachverhalte durch Experten beurteilt werden, über die unsicheres und unvollständiges Wissen existiert.

Ganz allgemein kann ein Delphi als ein Verfahren verstanden werden, bei dem in einem iterativen Prozess Expertenurteile zu einer bestimmten Fragestellung ermittelt werden, mit dem Ziel Konsens und Dissens in den Urteilen zu erfassen und zu begründen. Vor allem sollen zufällige Abweichungen, unterschiedliches Verständnis der Fragestellungen, Missverständnisse in der Repräsentation von Zahlen zu Sachverhalten (etwa bei Beurteilungsskalen) von tatsächlichen Differenzen in der Beurteilung eines Sachverhaltes oder einer Zielsetzung getrennt werden. Das Verfahren dient also zur Kalibrierung von pluralen Urteilen von Experten. Momentan wird das Delphi-Verfahren als das populärste Verfahren zur Konsensfindung unter Experten wahrgenommen (Ab Latif et al. 2016; Gordon

2015). „The major elements of the conventional Delphi are iteration with feedback of responses into the successive round, assessment of group judgment, anonymity of participants, and opportunity for participants to revise their views“ (Webler et al. 1991: 257). Die Validität eines Delphi-Verfahrens wird in der Regel über die „Theorie der Fehler“ belegt. Diese besagt, dass „die aggregierten Gruppenantworten eine Aussage repräsentieren, die der Mehrheit der einzelnen ExpertInnen überlegen ist“ (Aichholzer 2002: 4).

## Der Ablauf eines Delphis

Auch wenn, wie Häder (2014) ausführlich darstellt, bis heute verschiedene Varianten des Delphi-Verfahrens zu finden sind, kann ein allgemeines Ablaufmodell postuliert werden (vgl. hierzu auch Renn & Webler 1998; Cuhls & Blind 1999, Ammon 2009).

### Schritt 1: Entwicklung eines Fragebogens

Ein Forschungsteam entwickelt einen standardisierten Fragenbogen über den relevanten Sachverhalt. Dieser beinhaltet in der Regel die Abfrage zu erwartender Konsequenzen einer Maßnahme oder einer Entscheidungsoption sowie die Abfrage der subjektiven Sicherheit der Experten.



Bei einem Delphi-Verfahren werden auch die Delphi-Ausführenden, also das **Forschungsteam**, zu „Quasi-Experten“ des Themas. Nur so ist eine Kommunikation „auf gleicher Augenhöhe“ möglich (vgl. Pfadenhauer 2005). Die Forscher müssen sich oftmals schon bei der Einladung als Quasi-Experten präsentieren, also als kompetente Gesprächspartner, denen das Fachgebiet der Experten vertraut ist. So kann die Teilnahmekchance der Experten am Verfahren erhöht werden.

Neben standardisierten Fragen, können ergänzend offene Items formuliert werden, bei denen die Befragten selbstständig ihre Antworten eintragen können. Mit diesen Fragen werden meist inhaltliche Begründungen für numerische Abschätzungen erfragt. Sie werden aber auch eingesetzt, wenn das Forschungsteam keine standardisierten Antwortdimensionen formulieren kann. Dies kann unter anderem auf den Innovationsgrad der Frage oder auf die Komplexität einer fachfremden Materie zurückgeführt werden. Derartige Fragen haben in einem Delphi einen eher

explorativen Charakter. Gerade bei sehr komplexen Themen wie Nano- oder Computertechnologien kann die Abfrage offener Fragestellung eine geeignete Strategie zur Reduktion von Komplexität und Unsicherheit auf Seiten des Forschungsteams bzw. des Auftraggebers des Delphi-Verfahrens sein. Die Nachteile sind der hohe kognitive Anspruch an die befragten Experten und der relativ hohe Zeitaufwand für die Auswertung. Aus diesen Gründen werden offene Fragen in der Regel nur ergänzend eingesetzt.

Die konkrete Auswahl der relevanten Items erfolgt in der Regel auf Basis einer vorangegangenen Literaturrecherche oder über qualitative Experteninterviews. Alternativ können auch inter- und transdisziplinäre Expertenworkshops im Vorfeld des Delphi-Verfahrens durchgeführt werden (vgl. Niederberger & Wassermann 2015).

**Tab. 1** Beispielfrage für eine klassische Delphi-Frage

Wie beurteilen Sie die folgenden Thesen hinsichtlich Ihrer Eintrittswahrscheinlichkeit? Zu welchem Zeitpunkt tritt diese ein?	Wahrscheinlichkeit des Auftretens				Jahr des Eintretens		
	Hoch	Mittel	Gering	Gar nicht	Bis 2020	Bis 2040	Bis 2060
<b>Globalisierung geht mit folgenden Aspekten einher:</b>							
Internationale Kooperationen von Unternehmen im Bereich der Ausbildung nehmen zu							
Unternehmen werden ihre Ausbildungsaktivitäten stärker ins Ausland verlagern							
...							

**Schritt 2: Erste Befragung der Experten**

Der Fragebogen wird an eine Gruppe von anerkannten Experten des jeweiligen Fachgebietes verschickt. Die Experten beantworten die Fragen anonym nach bestem Wissen und schätzen die „subjektive Gewissheit“, d. h. die geschätzte Validität ihrer eigenen Antworten. Wer dabei als Experte gilt, ist nicht immer einfach zu beantworten. Angenommen wird, dass Experten für bestimmte Wissensgebiete und Realitätsausschnitte über ein detailliertes und spezialisiertes Sonderwissen verfügen, das andere nicht haben (vgl. Wassermann 2015b). Diese Zuschreibung erfolgt in der Regel auf Basis der beruflichen Position. Bei der Beantwortung der Fragen spielen aber auch persönliche Merkmale und Meinungen der Experten eine Rolle (vgl. Martens & Brüggemann 2006).

! Es ist für Experten bei Befragungen nicht möglich, die eigenen in der persönlichen Biographie und Lebenswelt ausgebildeten Bezüge zu einem Thema zu ignorieren oder zurückzustellen, in dem sie sich sozusagen einen „Astronauten-Anzug“ anziehen und die Analyse in der Schwerelosigkeit durchführen. Dies muss in der Auswertung der Ergebnisse zumindest als möglicher Interpretationsfaktor berücksichtigt werden (vgl. Köck-Hódi & Mayer 2013; Gebauer 2009; Kühn & Koschel 2011)

### Schritt 3: Auswertung durch das Forschungsteam

Das Forschungsteam wertet die ausgefüllten Fragebögen aus und ermittelt die Mittelwerte und Streuungen. Berechnet werden in der Regel die Maße der zentralen Tendenz: der Modus, der Median und das arithmetische Mittel. Bei der Streuung können beispielsweise der Interquartilsabstand und die Varianz berechnet werden (vgl. Benninghaus 2005, Hinweise zur Berechnung Abschnitt 2.1). Weiterführende bi- oder multivariate Analysen, die Assoziationen oder Korrelationen zwischen den Items untersuchen, werden in der Regel nicht durchgeführt.

Welche der erwähnten Maßzahlen konkret verwendet werden, hängt maßgeblich vom Messniveau der Skalen ab. Dies muss bereits bei der Entwicklung des Fragebogens bedacht werden. In der Praxis dominieren bei Delphi-Befragungen ordinale Skalen und dementsprechend wird häufig mit dem Median und dem Quartilsabstand gearbeitet (siehe Abschnitt 2.1).

- ! Der Informationsgehalt von Daten wird durch das Messniveau (oder auch Skalenniveau) beeinflusst. Für die statistischen Auswertungen im Rahmen eines Delphi-Fragebogens ist die Unterteilung in drei Messniveaus entscheidend:

Messniveau	Erläuterung	Beispiele	Relevante statistische Maßzahlen
Nominal	Bei Nominalskalen kann nur die Verschiedenartigkeit zwischen den Antworten (Werte) analysiert werden. Die statistische Auswertung konzentriert sich auf tabellarische oder grafische Häufigkeitsauszählungen. Elaborierte Analysen sind damit nicht möglich.	<ul style="list-style-type: none"> <li>• Geschlecht</li> <li>• Beruf</li> </ul>	<ul style="list-style-type: none"> <li>• Modus</li> </ul>
Ordinal	Bei Ordinalskalen kann zusätzlich eine sinnvolle Rangfolge dargestellt werden. Die Differenzen bzw. Abstände zwischen den einzelnen Antworten sind nicht berechenbar.	<ul style="list-style-type: none"> <li>• Zufriedenheit</li> <li>• Berufliche Position</li> </ul>	<ul style="list-style-type: none"> <li>• Median</li> <li>• Quartilsabstand</li> </ul>
Metrisch (Intervall und Ratio)	Bei metrischen Variablen können neben der Verschiedenartigkeit und der Rangfolge auch die Abstände zwischen den einzelnen Antworten angegeben werden. Diese Abstände müssen gleich groß sein. Damit sind alle mathematischen Prozeduren möglich.	<ul style="list-style-type: none"> <li>• Gehalt</li> <li>• Zeitangaben</li> </ul>	<ul style="list-style-type: none"> <li>• Arithmetische Mittel</li> <li>• Varianz bzw. Standardabweichung</li> </ul>

#### Schritt 4: Zweite Befragung der Experten

Der ursprüngliche Fragebogen wird zusammen mit der Auswertung der ersten Befragung an die Experten zurückgesandt. Dabei werden alle Namen der Experten anonym gehalten, um Beeinflussungen durch Status oder Seniorität auszuschalten. Die Experten werden gebeten, den Fragebogen ein zweites Mal auszufüllen, dieses Mal aber mit der Vorgabe, die Ergebnisse der ersten Befragung als Korrektiv der eigenen Urteile mit in die erneute Urteilsbildung einzubeziehen. Zweck dieser zweiten Befragung ist es, die Streuung der möglichen Antworten zu reduzieren und die

kollektive Urteilssicherheit zu erhöhen. Ein Beispiel, wie solche Ergebnisse präsentiert und gleichzeitig wieder abgefragt werden können, zeigt folgende Darstellung:


Sehr geehrte Damen und Herren,

vielen Dank für Ihre Bereitschaft, an der zweiten Runde der Delphi-Befragung zum Thema ..... teilzunehmen.

Mit diesem Fragebogen werden Ihnen zum einen die Ergebnisse der ersten Runde vorgelegt, zum anderen werden Sie um die erneute Beantwortung der Fragen gebeten. Die (meisten) Fragen sind identisch mit der ersten Fragerunde. Wir bitten Sie neben der inhaltlichen Einschätzung ebenfalls um die Bewertung der Eintrittswahrscheinlichkeit sowie um ihre Urteilsicherheit.

In dem Fragebogen sind die Antworten der ersten Runde direkt eingetragen. Im untenstehenden Beispiel kennzeichnet ein \* den Median (mittlerer Wert) der Antworten, der Balken präsentiert die Quartilsspanne. Im Beispiel war der Median bei „gering“. Die Mehrzahl der Antworten (mind. 50 %) liegen im Balken. Analog dazu sehen Sie die Antworten für das Jahr des Eintritts.

Bitte geben Sie bei den folgenden Fragen nun erneut Ihre Antwort bezüglich der Eintrittswahrscheinlichkeit und des Jahres des Eintritts an. Dabei bitten wir Sie Ihre Einschätzung vor dem Hintergrund der Antworten Ihrer Kollegen aus der ersten Delphi-Runde kritisch zu prüfen. Analog zu dem unten stehenden Beispiel können Sie Ihre Antworten auch in dem grauen Bereich angeben. Wenn Sie eine Antwort außerhalb des grauen Bereiches, also eine Antwort abweichend von der Mehrheit der Antworten abgeben, interessiert uns die Begründung Ihrer Meinung. In diesem Fall können Sie diese in dem Feld darunter eintragen.

Wie beurteilen Sie die folgenden Thesen hinsichtlich Ihrer Eintrittswahrscheinlichkeit Zu welchem Zeitpunkt tritt diese ein?	Wahrscheinlichkeit des Auftretens				Jahr des Eintretens		
	Hoch	Mittel	Gering	Gar nicht	Bis 2020	Bis 2040	Bis 2060
<b>Globalisierung geht mit folgenden Aspekten einher:</b>							
Internationale Kooperationen von Unternehmen im Bereich der Ausbildung nehmen zu.			*				
Begründung (falls Antwort außerhalb des grauen Balkens)							
_____ 							

**Abb. 1** Beispiel für ein Anschreiben zweite Delphi-Runde (dunkles grau = IQR, \* = Median)

### **Schritt 5: Ggfs. Wiederholung der Befragung der Experten**

Die Schritte 2, 3 und 4 werden solange wiederholt, bis die Experten keine Änderungen ihrer Urteile mehr vornehmen. In der Praxis werden zwei oder drei Wiederholungen durchgeführt.

Im Idealfall identifiziert das Delphi-Verfahren die Bewertungen, die innerhalb der Expertengruppe konsensfähig sind oder einen Dissens begründen. Durch die Anonymisierung der Teilnehmer und den iterativen Prozess der Befragung kann der jeweilige Kenntnisstand ohne Rücksicht auf Prestige oder Status eines jeden Teilnehmers am Delphi-Prozess dargestellt werden.

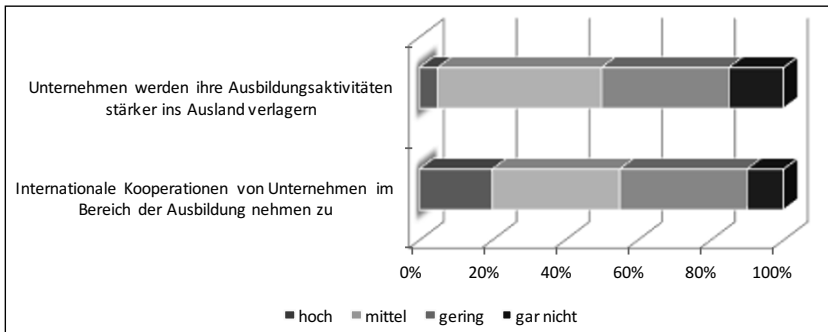
Diese kurze Darstellung verdeutlicht die wesentlichen Kennzeichen eines klassischen Delphi-Verfahrens (vgl. Webler et al. 1991: 257; Häder 2014). Sie können folgendermaßen zusammengefasst werden:

- Verwendung eines standardisierten Fragebogens
- Befragung von Experten
- Anonymität der Einzelantworten
- Ermittlung einer statistischen Gruppenantwort über univariate Befunde
- Rückkopplung der statistischen Ergebnisse an die Teilnehmer mit der Möglichkeit der Revision ihrer Urteile
- (Mehrfache) Wiederholung der Befragung

---

## **2.1 Auswertungsstrategie des klassischen Delphis**

Die Auswertung eines Delphis erfolgt in der Regel mittels einfacher mathematischer Prozeduren. Es geht nicht um anspruchsvolle statistische Analysen, die das Aufdecken neuer Zusammenhänge, Kausalitäten oder Typen erlauben. Stattdessen geht es oftmals um univariate Häufigkeitsauszählungen, sprich um Häufigkeitstabellen und -grafiken in unterschiedlichen Darstellungen (z. B. Balkendiagramme, Kreisdiagramme) (vgl. Cuhls & Kimpeler 2008; Cuhls et al. 2007). Ziel sind Verteilungsaussagen, mit denen Mehrheits- und Minderheitsvoten identifiziert werden können.



**Abb. 2** Beispiel einer grafischen Auswertung: Balkendiagramm (fiktive Zahlen)

In der statistischen Analyse können drei Maße der zentralen Tendenz und verschiedene Streuungsmaße berechnet werden.

### Maße der zentralen Tendenz

Es gibt drei Maße der zentralen Tendenz: den Modus, den Median und das arithmetische Mittel. Welches Maß berechnet wird, hängt vom Messniveau und der inhaltlichen Zielstellung ab.

1. Der **Modus** ist der häufigste Wert einer Verteilung und kann immer berechnet werden. Er eignet sich besonders gut, wenn eine Antwortmöglichkeit deutlich überrepräsentiert ist. Wenn mehrere Antworten gleich häufig angegeben werden, ist der Modus wenig aussagekräftig. Gerade bei kontinuierlichen Variablen, also bei Fragen mit sehr vielen Antwortmöglichkeiten (z. B. Einkommen), kann dies vorkommen.
2. Der **Median** (2. Quartil) entspricht dem mittleren Wert einer Verteilung. Er halbiert die der Größe nach geordneten Fälle der Verteilung. Er erfordert mindestens ordinales Messniveau, d. h. die Antworten müssen in eine sinnvolle Rangfolge, beispielsweise von gering bis hoch, geordnet werden können. Für die Bestimmung des Medians ist die Unterscheidung nach einer geraden oder ungeraden Anzahl an Fällen erforderlich. Bei einer geraden Anzahl von Fällen müssen wie in dem Beispiel der folgenden Tabelle die Werte rechts und links neben der Mitte angegeben werden. Wenn metrische Daten vorliegen, kann auch die Mitte dieser beiden Zahlen angegeben werden.



	<b>Wahrscheinlichkeit des Auftretens:</b> Globalisierung geht mit folgenden Aspekten einher: 1. Internationale Kooperationen von Unternehmen im Bereich der Ausbildung nehmen zu.				
Befragte (geordnet nach ihren Antworten)	1	2	3	4	5
Antwort	Gar nicht	Gering	<b>Gering</b>	Mittel	Hoch

**Abb. 3** Bestimmung des Medians bei ungerader Anzahl von Fällen (n=5), der Median ist fett markiert

	<b>Wahrscheinlichkeit des Auftretens:</b> Globalisierung geht mit folgenden Aspekten einher: 1. Internationale Kooperationen von Unternehmen im Bereich der Ausbildung nehmen zu.					
Befragte (geordnet nach ihren Antworten)	1	2	3	4	5	6
Antwort	Gar nicht	Gering	<b>Gering</b>	<b>Mittel</b>	Hoch	Hoch

**Abb. 4** Bestimmung des Medians bei gerader Anzahl von Fällen (n=6), der Median ist fett markiert

3. Das **arithmetische Mittel** ist umgangssprachlich der Durchschnitt und entspricht dem gewichteten Schwerpunkt der Verteilung. Es wird berechnet als die Summe der einzelnen Antworten, geteilt durch die Anzahl der Antworten. Das arithmetische Mittel kann nur bei metrischen Variablen berechnet werden.<sup>1</sup>

	<b>Wie viele Jahre beschäftigen Sie sich beruflich mit dem Thema Globalisierung?</b>					
Befragte (geordnet nach ihren Antworten)	1	2	3	4	5	6
<b>Antwort in Jahren</b>	2	3	3	4	5	10
<b>Arithmetische Mittel: <math>(2+3+3+4+5+10)/6 = 27/6 = 4,5</math></b>						

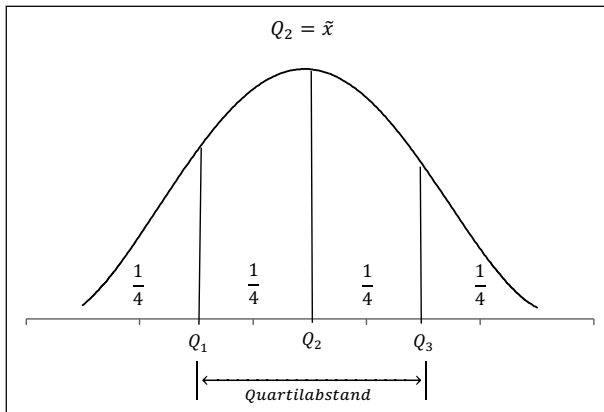
**Abb. 5** Berechnung des arithmetischen Mittels

1 Allerdings wird diese statistische Anforderung in der Forschungspraxis oftmals aufge-  
weicht. So werden ordinale Skalen mit mindestens fünf Ausprägungen als quasi-metrisch  
interpretiert und Berechnungen auf metrischem Niveau durchgeführt (vgl. Urban &  
Mayerl 2008).

## Streuungsmaße

Neben den drei Maßen der zentralen Tendenz werden auch Streuungsmaße berechnet. Sie zeigen an, wie aussagekräftig die Maße der zentralen Tendenz sind und stützen die Interpretation. Allgemein formuliert, geben die Streuungsmaße an, wie sich die einzelnen Antworten der Experten um den Mittelwert verteilen. Für die Berechnung ist ordinales oder metrisches Messniveau notwendig.

1. Eine sehr leicht berechenbare Größe ist der **Range** (Spannbreite). Er zeigt die Differenz aus dem größten und kleinsten Wert an. Bei Delphi-Befragungen ist er eher ungeeignet, weil nicht selten alle Antwortmöglichkeiten angekreuzt werden.
2. Häufiger eingesetzt wird bei Delphi-Befragungen dagegen der **Interquartilsabstand (IQR)**. Er beschreibt den Abstand zwischen dem 1. Quartil (bei 25 %) und dem 3. Quartil (bei 75 %) einer Verteilung und umschließt damit die mittleren 50 %. Auch der mittlere Quartilsabstand wird bisweilen berechnet. Bei ihm wird der Abstand zwischen dem 1. und 3. Quartil durch zwei dividiert. Quartilsabstände zeigen, in welchem Bereich die meisten Befragten geantwortet haben.



**Abb. 6** Darstellung von Quartilen

3. Die **Varianz  $S^2$  und Standardabweichung  $S$**  sind mittelwertbezogene Streuungsmaße. Sie besagen, ob die Antworten nahe beieinanderliegen oder weit auseinander streuen. Je geringer die Streuung ist, desto aussagekräftiger ist das arithmetische Mittel. Die Varianz berechnet sich aus den Abweichungsquadraten aller Werte vom arithmetischen Mittel dividiert durch die Gesamtfallzahl

(großes N). Die Standardabweichung ist die positive Wurzel aus der Varianz. Beide Maße können nur bei metrischen Daten berechnet werden.

$$s^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}$$

**Abb. 7** Formel zur Berechnung der Varianz

	Wie viele Jahre beschäftigen Sie sich beruflich mit dem Thema Globalisierung?					
Befragte (geordnet nach ihren Antworten)	1	2	3	4	5	6
Antwort in Jahren	2	3	3	4	5	10
Arithmetische Mittel: (2+3+3+4+5+10) / 6 = 27/6 = 4,5 d. h. Im Durchschnitt beschäftigen sich die Befragten seit 4,5 Jahren mit dem Thema Globalisierung.						
Varianz: 1/6 / (2-4,5) <sup>2</sup> + (3-4,5) <sup>2</sup> + (3-4,5) <sup>2</sup> + (4-4,5) <sup>2</sup> + (5-4,5) <sup>2</sup> + (10-4,5) <sup>2</sup> = 1/6 / 6,25 + 2,25 + 2,25 + 0,25 + 0,25 + 30,25 = 41,5 / 6 = 6,92 Die Varianz beträgt 6,92.						

**Abb. 8** Berechnung der Varianz und der Standardabweichung

Die vorgestellten Maße können über alle befragten Experten hinweg oder vergleichend für verschiedene Expertengruppen (z. B. Vergleich der Antworten zwischen Experten verschiedener Fachrichtungen) berechnet und dargestellt werden. Da in Delphi-Befragungen häufig ordinale Skalen eingesetzt werden, beruhen die statistischen Berechnungen meist auf dem Median und dem Quartilsabstand.

Neben den quantitativen Auswertungen gilt es auch, eventuelle offene Fragen systematisch und regelgeleitet auszuwerten. Die Auswertung dieser Fragen ist nicht immer einfach. Denn die Experten können selbst entscheiden, wie detailliert und tiefgründig sie antworten. Eventuell müssen deshalb Bilanzaussagen mit spezifischen Angaben verglichen werden. Hierfür gilt es eine nachvollziehbare Methode zu wählen, die es erlaubt die Ergebnisse kurz und knapp, im Idealfall standardisiert für die nächste Befragungswelle, darzustellen. Zu diesem Zweck können beispielsweise qualitative und quantitative Inhaltsanalysen (vgl. Früh 2007; Mayring 2007;

Lamnek 2005) eingesetzt werden. Dabei werden die Antworten kodiert und bei der quantitativen Variante ausgezählt und standardisiert (vgl. Bodensohn et al. 2007).



Die **Inhaltsanalyse** ist ein Oberbegriff für verschiedene Verfahren zur systematischen und regelgeleiteten Analyse von Texten, beispielsweise in Form von verschriftlichten Interviews oder Zeitungsartikeln. Zentrales Ziel ist die Reduktion des Materials und die Identifikation zentraler Aussagen. Dabei wird zwischen qualitativen und quantitativen Inhaltsanalysen unterschieden.

In der *quantitativen Inhaltsanalyse* wird das Material nach einem meist theoretisch hergeleiteten Schema kodiert und ausgewertet. Dabei geht es um Frequenzanalysen, d. h. um die Auszählung von Schlagworten in bestimmten Kontexten. Berücksichtigt wird ausschließlich der manifeste Inhalt. Damit sind deskriptive statistische Analysen möglich.

In der *qualitativen Inhaltsanalyse* geht es nicht um eine zahlenmäßige Analyse, sondern um eine hermeneutische Textinterpretation, bei der der latente Sinn des Textes analysiert wird. Maßgeblich entwickelt hat die Methode Mayring (2007). Anliegen ist es offen und explorativ, aber auch systematisch und regelgeleitet das Material auszuwerten.

---

## 2.2 Varianten des klassischen Delphi-Verfahrens

Bis heute wird das Delphi-Verfahren beständig weiterentwickelt und modifiziert (vgl. Häder 2014). Die verschiedenen Varianten beziehen sich vor allem auf die Anzahl der Befragungswellen, die Auswahl der Experten, die Gestaltung der Feedbackrunden sowie die Ermittlung der Selbsteinschätzung der Experten (vgl. Häder 2014; Cuhls & Blind 1999). Standardmäßig wird das Delphi-Verfahren als postalische Befragung konzipiert, d. h. die Experten erhalten in mehreren Wellen einen standardisierten Fragebogen. Dieses Vorgehen hat den Vorteil, dass eine große Anzahl an Experten eingebunden werden kann. Dennoch werden in der aktuellen Literatur verschiedene Arten von klassischen Delphi-Verfahren diskutiert. Sie werden differenziert auf Basis spezieller Erkenntnisinteressen oder aufgrund methodischer Besonderheiten im Ablauf. Einige wichtige Varianten seien kurz vorgestellt (weitere Varianten in Häder 2014: 24ff):

**Decision-Delphi:** Bei dieser Variante fungieren die Experten auch als Entscheidungsträger, welche die Ergebnisse umsetzen sollen (vgl. Keeney et al. 2011).

**Policy-Delphi:** Hier liegt das Ziel nicht nur in der Konsensfindung, sondern in der Identifikation und Präsentation stark kontrastierender Ansichten (vgl. Keeney et al. 2011; Turoff 1970). Ziel ist neben der Generierung von Expertenmeinungen die Maximierung der Bandbreite an Expertenmeinungen. „The Policy Delphi also rests on the premise that the decision maker is not interested in having a group generate his decision; but rather, have an informed group present all the options and supporting evidence for his consideration. The Policy Delphi is therefore a tool for the analysis of policy issues and not a mechanism for making a decision“ (Turoff 1970: 80).

**Breitband-Delphi-Methode:** Bei einer überschaubaren Anzahl an relevanten Experten, kann auch die Breitband-Delphi-Methode eingesetzt werden. Dieses Verfahren wird oftmals firmenintern bei komplexen und anspruchsvollen betriebswirtschaftlichen Fragestellungen eingesetzt. Bei diesem Verfahren kennen sich die Experten in der Regel. Inhaltlich geht es bei diesem Verfahren oft um Schätzungen, beispielsweise um kalkulatorische Fragestellungen. Das Breitband-Delphi kann dabei im Rahmen einer Schätzklausur durchgeführt werden.



**Schätzklausuren** sind Verfahren zur Einschätzung der Kosten und Beurteilung der Zeitplanung. Sie können Grundlage für die Weiterführung aber auch zur Aussonderung von Projektideen sein. Damit Gruppenmitglieder realistische Einschätzungen vornehmen können, werden die Schätzungen auf einem Workshop vorgenommen.

Auch bei der Breitband-Methode beantworten die Experten den Fragebogen allein. Die Schätzklausur dient vor allem dem Austausch zwischen den Befragungswellen. Dadurch können Rückfragen und Argumente ausgetauscht werden und das Ziel einer realistischen und tragfähigen Schätzung unterstützt werden.

**Real-Time-Delphi:** Auf dem Vormarsch sind Online-Delphi-Verfahren, die sogenannten Real-Time-Delphis (vgl. Gnatzy et al. 2011, Cuhls 2012). Das Besondere ist, dass die Fragebögen online verschickt werden und in Echtzeit von einem entsprechenden Computer- bzw. Statistikprogramm ausgewertet und die Ergebnisse den teilnehmenden Experten angezeigt werden. Das Verfahren ist damit „rundenlos“, d. h. die Teilnehmer können ihr Urteil jederzeit revidieren, und sind nicht gezwungen, sich ggfs. nur ein- oder zweimal festzulegen. Dabei besteht das Risiko, dass einige Experten häufiger teilnehmen als andere: „Rein theoretisch könnten sie sogar extremen Einfluss nehmen, indem sie besonders stark abweichende Antworten geben und

dadurch alle Anderen in die Irre führen. Dies ist allerdings nur eine theoretische Möglichkeit, in der Realität wurde ein solches Verhalten bisher noch nicht beobachtet“ (Cuhls 2012: 142). Der Vorteil ist die enorme Zeitersparnis bei der Durchführung und Auswertung. Die Anonymität der Experten bleibt gewahrt. Der Nachteil ist, dass es unwahrscheinlicher wird, dass sich die Teilnehmer erneut mit ihren Urteilen auseinandersetzen und diese ggf. überdenken und/oder, dass sie zu einem späteren Zeitpunkt den Online-Fragebogen erneut aufrufen und revidieren. Insgesamt zeigt aber ein Vergleich zwischen dem klassischen Delphi und einem Online-Delphi, dass beide Verfahren zu ähnlichen Ergebnissen führen. „...the comparison analyses showed no significant differences between conventional and real-time Delphi survey methods. Therefore, we argue that the result of the real-time Delphi survey tool are as robust as the results generated by a conventional Delphi survey“ (Gnatzy et al. 2011: 1692).

**Verbindung von social media big data Analyse und Delphi Verfahren:** Eines der erstaunlichen Ergebnisse der big data Forschung besteht darin, dass bei einfachen statistischen Schätzverfahren (etwa das Gewicht eines Sparschweins nur auf der Basis eines Bildes und einer Hintergrundinformation (gefüllt mit 10-Cent-Stücken) abzuschätzen) der Mittelwert bei hinreichend großer Teilnehmerzahl dem wahren Wert approximiert ist. Diesen Umstand machen sich neuartige Delphi-Verfahren zu Nutze und spiegeln den Experten nicht nur die Kennwerte innerhalb der beteiligten Expertengruppe, sondern auch die durch social media Analyse berechneten Maße auf der Basis der eingehenden Schätzwerte wider (vgl. Servan-Schreiber 2012; Page 2007). Diese Kombination hat sich vor allem im Bereich wirtschaftlicher Prognostik bewährt.

**Hybrid-Delphi:** Hybrid-Delphis stellen eine Kombination des Delphi-Verfahrens mit anderen Methoden der empirischen Sozialforschung dar. Verwendet und systematisch eingeführt, haben den Begriff Landeta et. al für eine Kombination von drei Verfahren: Fokusgruppe (FG), Nominalgruppentechnik (NGT) und Delphi-Methode (vgl. Landeta et al. 2011).



Eine **Fokusgruppe** (FG) ist ein moderiertes Diskursverfahren mit einer Kleingruppe, die typischerweise sechs bis acht Personen umfasst. Durch eine ungezwungene und offene Atmosphäre werden konstruktive und kreative Gruppenprozesse unterstützt. Strukturiert wird die Diskussion anhand eines Leitfadens, der die zentralen Fragestellungen enthält. Angeregt wird die Gesprächsrunde durch einen Stimulus, beispielsweise in Form eines kurzen Films, eines Bildes oder eines Vortrages. Eine Fokusgruppe dauert insgesamt zwei bis drei Stunden. (vgl. Schulz & Mack 2012).

Durch die **Nominalgruppentechnik (NGT)** kann eine Gruppe strukturiert Entscheidungen treffen oder Lösungen für eine Problemstellung entwickeln. „Within the NGT, ideas are generated in a short period of time and participants can see at first hand the process of reaching consensus“ (Keeney et al. 2011: 15). Die Gruppe besteht dabei nominal – also dem Namen nach. In keiner Phase interagieren die Gruppenmitglieder frei miteinander. So wird die Dominanz einzelner Gruppenmitglieder ebenso unterbunden wie die Fokussierung auf einen einzelnen Aspekt oder Gedanken. Die Methode basiert im Kern auf der Sammlung von Ideen und der anonymen Abstimmung der Teilnehmer (vgl. Keeney et al 2011: 15).

Durch die Kombination von Delphi-Verfahren, Fokusgruppe und Nominalgruppentechnik wird versucht, die Nachteile des jeweiligen Verfahrens durch die Vorteile des anderen aufzufangen. So kann die NGT neue und innovative Ideen hervorbringen, die bei einem klassischen Delphi aufgrund der starken Strukturierung und Standardisierung keinen Platz haben. Teilweise werden auch nur zwei der drei Verfahren kombiniert. Dabei können die FG oder die NGT vor oder nach dem Delphi stattfinden (vgl. Landeta et al. 2011: 1632f). Landeta et al. (2011) präferieren allerdings die Reihenfolge: FG; NGT und Delphi, wobei die Ergebnisse des vorherigen Verfahrens den Input für das Folgende darstellen. „The final output is that of the Delphi, which has the capacity to reorder and improve the outputs of the other two techniques, which it has received in the form of input“ (Landeta et al. 2011: 1637).

**Fazit:** Gemeinsam ist allen genannten Delphi-Verfahren, ob postalisch oder online, dass die Experten den standardisierten Fragebogen allein ausfüllen und durch eine Feedbackschleife zur erneuten Befragung gebeten werden, mit der Bitte, das eigene Urteile noch einmal zu überdecken. Dieses Vorgehen unterscheidet die klassischen Delphi-Varianten von dem Gruppendelphi. Doch ehe dieses vorgestellt wird, sollen einige Praxisbeispiele für klassische Delphi-Verfahren dargestellt werden.

---

## 2.3 Beispiele für klassische Delphi-Verfahren

Es gibt eine Vielzahl an Beispielen für klassische Delphi-Verfahren. Die konkreten Vorgehensweisen unterscheiden sich in Details, orientieren sich aber an den zentralen Merkmalen eines Delphi-Verfahrens: Expertenbefragung, Wiederholung der Befragung und Rückkopplung der Ergebnisse. Die Beispiele aus der Praxis zeigen,

dass das klassische Delphi-Verfahren mittlerweile häufig im Rahmen einer Methodentriangulation eingesetzt werden, d. h. es mit anderen Erhebungsmethoden (z. B. einem Workshop) kombiniert wird. Einen Überblick über einige Beispiele aus sehr unterschiedlichen Fachdisziplinen fasst die folgende Tabelle zusammen (weitere z. B. in Köck-Hódi & Mayer 2013; Keeney et al. 2011; Vidgen & Gallegos 2014).

**Tab. 2** Beispiele für klassische Delphi-Verfahren

**Durchführung eines Delphis zum Thema Wirksamkeit und Machbarkeit von Leistungstests in der Sportwissenschaft (vgl. Robertson et al. 2017)**

Durchgeführt wurde ein Delphi mit zwei Runden. Eingebunden wurden 33 Experten, unter anderem aus der Sportwissenschaft. Am Ende konnte bei allen Punkten Konsens zwischen den Experten hergestellt werden. Auf Basis dieser Informationen wurden Definitionen festgelegt und eine Checkliste entwickelt bzw. veröffentlicht.

**Delphi zur Erstellung einer Concept Map zum Thema Diabetes Mellitus (vgl. Ab Latif et al. 2016)**

Durchführung von drei Delphi-Runden mit einem strukturierten Fragebogen. Die Befragung erfolgte per Email. Befragt wurden 10 Experten, in diesem Fall Studierende aus dem Pflegebereich.

**Durchführung der Studie „Demographischer Impact in städtischen Wohnquartieren“ zur Analyse und Erfassung wesentlicher Akteure, Einflussfaktoren und Zukunftstrends der Quartiersentwicklung (vgl. Schnur & Markus 2010)**

Der Expertenkreis bestand aus Persönlichkeiten aus Kommunen, Ländern und Bundesinstitutionen, aus der Wohnungswirtschaft (Verbände, Unternehmen), aus dem Bereich wissenschaftlicher oder kommunaler Dienstleistungen und aus der Wissenschaft. Durchgeführt wurde eine schriftliche Delphi-Befragung in zwei Runden ( $n_1=47$  und  $n_2=36$ ). Basis waren vorab durchgeführte Einzel-Expertengespräche und Fallstudienuntersuchungen in verschiedenen Wohnquartieren der Städte Berlin, Leipzig, Essen und Brandenburg an der Havel.

**Erarbeitung von regionalpolitischen Maßnahmen zur Förderung der Dienstleistungswirtschaft im Erzgebirge (vgl. Hesse et al. 2009)**

1. *Offene Befragung zur Generierung von Vorschlägen:* Die Ergebnisse dieser ersten Befragungsstufe werden klassifiziert und als Basis zur Erstellung eines standardisierten Fragenprogramms verwendet, das in einen geschlossenen Fragebogen mündete.
2. *Geschlossene Befragung zur Konkretisierung und Verdichtung der Vorschläge.* Die Befragung erfolgt wiederum anonym und ohne Abstimmung innerhalb der Gruppe.
3. *Persönliche Expertenrunde zur gemeinsamen Diskussion der Empfehlungen.* Die persönliche Expertenrunde bringt die Teilnehmer schließlich zu einem Workshop zusammen. Hier werden ausgewählte Sachverhalte detailliert diskutiert sowie branchenspezifische Problemlagen und Lösungsansätze eruiert.



**Studie zu Nanotechnologien in den Bereichen Nahrungsmittel, kosmetische Erzeugnisse und Bedarfsgegenstände (vgl. Zimmer et al. 2009)**

1. Zur Vorbereitung des Delphi-Prozesses wurden *eine Literaturanalyse, Experteninterviews und ein internes Expertengespräch* „Nanotechnologie, ihre Produkte und Risiken für den Verbraucher“ durchgeführt.
2. *Erste Delphi-Runde mit standardisierten Fragebogen.* Die Teilnehmenden haben die Aufgabe, die Beantwortung des Fragebogens organisationsintern zu koordinieren, d. h. sie können sich mit Kollegen austauschen und den Fragebogen gemeinsam ausfüllen. Von den 100 angeschriebenen Experten schickten 71 einen beantworteten Fragebogen zurück.
3. *Zweite Delphi-Runde mit standardisierten Fragebogen und den Ergebnissen der vorherigen Befragung.* Hier werden Vertiefungsfragen gestellt, da einige Bewertungen in der ersten Runde starke Varianzen ausweisen oder Unklarheiten in den Fragestellungen aus Runde 1 sichtbar sind. Von den 71 angeschriebenen Experten, die sich in Runde 1 beteiligt hatten, antworteten 56 Personen. Das entspricht einer Rücklaufquote von 78 % in Runde 2.
4. *Experten-Workshop.* Ziel ist es, die qualitativen Begründungszusammenhänge zu hinterfragen und stakeholderübergreifend das Wissen zu vertiefen.
5. *Interner Experten Workshop* mit dem Projektteam und dem Auftraggeber zur Vorstellung und Diskussion der Befunde.

**Zukunftstechnologien in drei Delphi-Studien (vgl. von Oertzen et al. 2006; Cuhls et al. 2007; Cuhls & Kimpeler 2008)**

**Vorgehen (am Beispiel der dritten Delphi-Studie)**

1. Durchführung von Expertenworkshops zur Entwicklung von Thesen
2. Durchführung der ersten standardisierten Online-Befragung der Experten (n=2.064)
3. Durchführung einer zweiten standardisierten Online-Befragung der Experten (n=681) mit dem gleichen Fragebogen und Auswertung

**Entwicklung eines für die Lehramtsstudiengänge spezifischen Studieneingangstests (vgl. Bodensohn, et al. 2007)**

1. *Versand eines halbstrukturierten Befragungsbogens.* Inhaltsanalytische Auswertung und Übersetzung der Ergebnisse in Itemform
2. *Bewertung der extrahierten Items* in standardisierter Form von den gleichen Experten bezüglich ihrer Relevanz im Hinblick auf den Schulalltag

Diese Daten werden in aufbereiteter statistischer Form zum Gegenstand einer *abschließenden Gesamtkonferenz* der teilnehmenden Experten gemacht. So können etwa Redundanzen, fehlende oder zu schwach repräsentierte Aspekte sowie Widersprüche im Itemset im Plenum diskutiert werden.

**Bundesweite Erhebung zur Ermittlung des zukünftigen Forschungs- und Entwicklungsbedarfs in der beruflichen Aus- und Weiterbildung (vgl. Brosi et al. 2003)**

1. *Bestandsaufnahme von Forschungsideen, -vorschlägen und -erfordernissen über eine offene Frage:* „Welche Forschungs- und Entwicklungsaufgaben halten Sie für erforderlich, damit die Menschen in Zukunft umfassend beruflich qualifiziert und künftigen Herausforderungen gewachsen sind?“ Die Experten beschreiben in eigenen Worten ihre Vorschläge.
2. *Bewertung der Forschungsideen:* In der zweiten Runde werden die Experten gebeten, zu jedem der Vorschläge Stellung zu beziehen.
3. *Rückkoppelung, Vertiefung und Diskussion:* In der dritten Runde werden die Resultate der beiden ersten Delphi-Runden mit einem kleineren Kreis von Experten thematisiert. Besonders werden dabei umstrittene Ergebnisse diskutiert, mit dem Ziel Interpretationshinweise für das Verständnis mehrdeutiger Ergebnisse zu gewinnen. Die Auswahl der Teilnehmer erfolgt als Zufallsstichprobe aus dem Kreis derer, die sich an der zweiten Runde beteiligt haben.

**Zukünftigen Veränderungen im wissenschaftlichen Zeitschriftenwesen (vgl. Keller 2001)**

Drei schriftliche Fragerunden mit 45 internationalen Experten. In der ersten Befragungsrunde werden offene Fragen formuliert und in den folgenden Wellen wird der Fokus stärker auf quantitative Fragen gelegt.

## 2.4 Fazit

Die Strategie der Delphi-Methode kann charakterisiert werden: durch die Konzentration auf zentrale Aussagen, den mehrstufigen, rückgekoppelten Prozess, den Experten als Untersuchungseinheit und die statistische Konzentration auf das ordinale Messniveau. Störende Einflüsse werden meist durch die Anonymisierung der Experten sowie der schriftlichen individuellen Befragung eliminiert.

**Zentrale Kennzeichen eines klassischen Delphi-Verfahrens sind:**

- Expertenbefragung
- Anonyme schriftliche oder Online-Befragung
- Mehrfachbefragung mit Rückspiegelung der Ergebnisse auf aggregiertem Niveau
- Verwendung von Ratingskalen mit ordinalem Messniveau
- Auswertung auf Basis deskriptiver univariater Statistik

Das klassische Delphi-Verfahren wird vor allem bei Schätzfragen, in Zukunftsstudien und anspruchsvollen technologischen Themen eingesetzt. Es gehört zum festen Instrumentarium der Technikvorausschau.

In der Praxis werden meist zwei standardisierte Befragungswellen durchgeführt. Im Vorfeld oder im Nachgang ist die Kombination mit Expertenworkshops (beispielsweise mit Fokusgruppen) möglich. Sie dienen entweder der Identifikation relevanter Statements oder der inhaltlichen Unterfütterung der statistischen Ergebnisse.

Ein häufiges Problem besteht bei einem klassischen Delphi-Verfahren darin, dass die Experten ihre einmal geäußerte Meinung in den folgenden Runden trotz Anonymität nicht ändern, sodass der Zusatznutzen weiterer Runden oft klein ist. Zudem fehlen die Begründungen für die Antworten. Außerdem ist der Ressourcenaufwand sehr hoch. Der gesamte Delphi-Prozess kann Jahre dauern und ist sehr kostenintensiv. Diese Nachteile waren der Auslöser für die Überlegung das klassische Delphi-Verfahren zu modifizieren und bilden die Ausgangslage für die Entwicklung des Gruppendelphis.

## Relevante Literatur

- Ammon, U. (2009). Delphi-Befragung. In S. Kühl, P. Strodtholz & a. Taffertshofer (Hrsg.), *Handbuch Methoden der Organisationsforschung. Quantitative und Qualitative Methoden* (S. 458-476). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Cuhls, K. (2008). *Methoden der Technikvorausschau – eine internationale Übersicht*. Stuttgart: IRB-Verlag.
- Cuhls, K. (2009). Delphi-Befragungen in der Zukunftsforschung. In R. Popp & E. Schüll (Hrsg.), *Zukunftsforschung und Zukunftsgestaltung. Beiträge aus Wissenschaft und Praxis* (S. 207–221). Berlin, Heidelberg: Springer-Verlag.
- Cuhls, K. (2012). Zu den Unterschieden zwischen Delphi-Befragungen und „einfachen“ Zukunftsbefragungen. In R. Popp (Hrsg.), *Zukunft und Wissenschaft. Wege und Irrwege der Zukunftsforschung* (S. 139-157). Berlin, Heidelberg: Springer-Verlag.
- Häder, M. (2014). *Delphi-Befragungen: Ein Arbeitsbuch* (2. Aufl.). Wiesbaden: Springer Fachmedien Wiesbaden.
- Seeger, T. (1979). *Die Delphi-Methode: Expertenbefragung zwischen Prognose und Gruppenmeinungsbildungsprozessen*. Freiburg: Hochschul-Verlag.
- Steinmüller, K. (1997). *Grundlagen und Methoden der Zukunftsforschung: Szenarien, Delphi, Technikvorausschau*. Gelsenkirchen: Sekretariat für Zukunftsforschung.

Das Gruppendelphi-Verfahren

Vom Konzept bis zur Anwendung

Niederberger, M.; Renn, O.

2018, XI, 207 S. 28 Abb., Softcover

ISBN: 978-3-658-18754-5