

Der Einfluss der Expertise auf das Antwortverhalten in Delphi-Studien: ein Hypothesentest

Blind, Knut; Cuhls, Kerstin

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Blind, K., & Cuhls, K. (2001). Der Einfluss der Expertise auf das Antwortverhalten in Delphi-Studien: ein Hypothesentest. *ZUMA Nachrichten*, 25(49), 57-80. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-207945>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

DER EINFLUSS DER EXPERTISE AUF DAS ANTWORTVERHALTEN IN DELPHI-STUDIEN: EIN HYPOTHESENTEST¹

KNUT BLIND UND KERSTIN CUHLS

Basierend auf den Ergebnissen von Delphi'98, einer Delphi-Studie zur globalen Entwicklung von Wissenschaft und Technik im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF), wurde der Einfluss der Expertise der beteiligten Experten auf das Antwortverhalten getestet. Eine erste Hypothese, dass Experten mit geringem Expertenwissen mit geringerer Wahrscheinlichkeit an der zweiten Befragungsrunde teilnehmen, konnte für die Mehrheit der zwölf Technikfelder bestätigt werden. Auf Basis der individuellen Einschätzungen der möglichen Realisierungszeiträume von über 1.000 zur Diskussion gestellten Thesen im Delphi'98 wurde eine Reihe weiterer Hypothesen getestet. Für die Bewertung und Weiterentwicklung der Delphi-Methode besonders interessant ist die weit gehende Bestätigung der Hypothese, dass bei höherer Expertise die Experten weniger bereit sind, sich der Durchschnittsmeinung anzuschließen, und damit den Konvergenzprozess nicht unterstützen. Diesem Zusammenhang muss sowohl bei der Durchführung von Delphi-Untersuchungen als auch bei der Analyse der Ergebnisse Rechnung getragen werden.

The second wave of a Delphi study on the global development of science and technology (Delphi '98) commissioned by the German Federal Ministry of Education, Research, Science and Technology (BMBF) was used to investigate relationships between the level of expertise and response and non-response behaviours. For the majority of the 12 technological fields covered in the study, the level of expertise correlates positively with participation in the second wave of the Delphi panel. Respondents' assessments of the expected realisation time frame for more than 1,000 topics were used to investigate a number of hypotheses. Our findings support the hypothesis that the greater their expertise, the less likely respondents are to agree with the average assessment of a topic. Such experts therefore do not support the process of

1 Die Autoren danken zwei anonymen Gutachtern für wertvolle Hinweise.

convergence. This relationship has to be taken into account in the conduct of Delphi studies and the analysis of the results.

1. Einleitung

Delphi-Studien werden bereits seit den fünfziger Jahren durchgeführt. Anfangs als reines Prognose-Instrument betrachtet, traten später andere Ziele wie Konsensfindung oder Anregung von Kommunikation in den Vordergrund. Erst mit den umfassenderen nationalen Studien der neunziger Jahre ist es möglich geworden, ausreichend große Samples für die Überprüfung von Hypothesen bezüglich des individuellen Antwortverhaltens der Experten zu erhalten. Die Studie Delphi '98, die das Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI) im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) durchgeführt hat, ist eine derartige Studie und liegt der folgenden Hypothesenprüfung zu Grunde.

Der Kern des Delphi-Verfahren besteht aus mindestens zwei so genannten „Runden“. Von Fachkommissionen erarbeitete Thesen werden einer großen Anzahl an Experten zur Bewertung vorgelegt. Deren Antworten werden ausgewertet und denselben Personenkreisen noch einmal zugeschickt. In dieser zweiten Runde sollen die Experten ihre Antworten unter dem Einfluss der Einschätzungen ihrer Fachkollegen noch einmal überdenken und ihre Meinung ändern - oder nicht. Anonymität ist gewährleistet, so dass bei einer Meinungsänderung niemand sein Gesicht verliert oder sich rechtfertigen muss.

In den letzten Jahren wuchsen Zweifel, ob es richtig sei, für eine Delphi-Befragung nur „Experten“ zu Rate zu ziehen und nur diese an den nationalen Foresight-Übungen teilnehmen zu lassen, denn somit wird ihnen eine grundlegende Entscheidungsvorbereitung überlassen (Cuhls 2000). Im Hintergrund steht die Annahme, dass Experten dazu neigen, ihr eigenes Arbeitsgebiet zu überschätzen (Bias). Dies konnte in den vom ISI durchgeführten Delphi-Studien bisher jedoch nur in ausgewählten Themenfeldern bestätigt werden, z. B. für die Energie oder die Biotechnologie (Grupp 1995), für die Mehrzahl der Felder fanden sich dagegen keine grundsätzlichen Auffälligkeiten (Cuhls/Blind/Grupp 2001).

Der vorliegende Artikel greift vier Hypothesen zur Bedeutung des Expertenstatus für das Verhalten in Delphi-Studien auf, die anhand der Delphi '98-Daten differenziert nach Themenfeldern getestet werden. Jedoch wird zunächst kurz beschrieben, wie das Delphi '98 durchgeführt wurde. Der Beitrag schließt mit Folgerungen für die zukünftige Durchführung von Delphi-Studien und die Analyse der Ergebnisse.

2. Hintergrund: Delphi`98

In Deutschland werden seit Anfang der neunziger Jahre auf nationaler Ebene Vorausschau-Studien durchgeführt. Verschärfter Wettbewerb auf den Weltmärkten, Globalisierung der Wirtschaftsunternehmen und ein immer unüberschaubar werdender technischer Fortschritt zwingen dazu, nach Orientierungswissen zu suchen. Die wesentliche Fragestellung lautet: Welchen technischen Möglichkeiten kommt in Zukunft die größte Bedeutung zu, um die ökonomische Wettbewerbsfähigkeit, die Erhaltung oder Wiederherstellung einer lebenswerten Umwelt oder die soziale Wohlfahrt zu gewährleisten?

Mit Delphi-Studien kann ein fundierter Blick in die Zukunft gewagt werden. Wie die Zukunft tatsächlich aussehen wird, weiß niemand. Es ist jedoch möglich, bereits jetzt bestimmte Entwicklungen einzuschätzen, sie gemeinsam zu überprüfen, zu diskutieren und dann Maßnahmen zu überlegen, um sie Wirklichkeit werden zu lassen - oder sie gegebenenfalls zu verhindern. Damit wird Zeit gewonnen, evidente Fehlentwicklungen zu bremsen oder Innovationen, die dringend benötigt werden, anzustoßen oder rascher voranzutreiben. Delphi-Studien liefern also nicht einfach ein Bild von der Zukunft, sondern eine Informationsgrundlage für die Entscheidung, was heute zu tun oder zu lassen ist. Dabei wird nicht mehr davon ausgegangen, dass grundsätzlich ein Konsens erreicht wird oder erreicht werden kann, sondern bereits die Feststellung, dass bezüglich eines bestimmten Zukunftsthemas Konsens herrscht (oder eben nicht), ist eine für die teilnehmenden Akteure wichtige Information.

Die erste deutsche Delphi-Studie zur Entwicklung von Wissenschaft und Technik (BMFT 1993) wurde 1993 im Auftrag des BMFT vom Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI), das mit dem japanischen National Institute of Science and Technology Policy (NISTEP) zusammenwirkt, durchgeführt. Da sich die erste deutsche Delphi-Studie in der Umsetzung und strategischen Nutzung für die Wirtschaft, aber auch für staatliche Entscheidungsträger als sehr hilfreich erwiesen hat, wurde mit den Mini-Delphi-Studien (Cuhls/Breiner/Grupp 1995) von 1995 in acht wissenschaftlich-technischen Problemlösungsbereichen die japanisch-deutsche Kooperation fortgesetzt und die Methode weiterentwickelt. Frankreich, Südkorea und Großbritannien sowie seit kurzem auch andere Länder wie Österreich, Ungarn oder Südafrika sind inzwischen der Initiative, Delphi-Befragungen für eine umfassende Technikvorausschau zu nutzen, gefolgt. In Japan ist inzwischen bereits die 7. Delphi-Studie veröffentlicht worden (NISTEP 2001). Auf europäischer Ebene existiert mittlerweile ein Netzwerk der beteiligten Institutionen, um Informationen zu neuer Technologie, Technologiepolitik und auch Methodentheorie auszutauschen.

1996 startete das Delphi '98. Diese neuerliche Studie sollte eine Aktualisierung der Daten ermöglichen, denn in den letzten Jahren sind in Deutschland und in anderen hoch entwickelten Industrieländern einige Weichen für die Zukunft neu gestellt worden, so dass die zukünftigen Entwicklungsmöglichkeiten aus der heutigen Sicht womöglich anders zu beurteilen sind.

Zuerst wurde ein Lenkungsausschuss am BMBF gegründet, der zukunftsrelevante Themenfelder definiert hat. Die detaillierte inhaltliche Vorbereitung der Studie (Themenfindung usw.) fand in Fachausschüssen statt, die mit mehr als 100 fachkundigen Personen aus Industrie, Hochschulen und anderen Einrichtungen besetzt waren. Koordiniert wurde der gesamte Prozess vom Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI) in Karlsruhe. Von dort wurde auch die weitere Erarbeitung der Fragebögen, der Versand, der Aufbau einer Adress-Datenbank und die analytische Arbeit betreut. Der Originalbericht (Cuhls/Blind/Grupp 1998) steht für Interessierte zur Verfügung (www.isi.fhg.de/ti/projektbeschreibungen).

2.1 Die Themenfelder und ihre Inhalte

Die in Tabelle 1 aufgeführten Themenfelder wurden für so wichtig erachtet, dass ein Blick in ihre Zukunft lohnt. Für jedes Themenfeld wurde ein separater Fragebogen mit bis zu 114 Thesen erstellt. Insgesamt haben die zwölf Fragebögen 1.070 Thesen enthalten. Stärker als beim ersten Mal wurden beim Delphi '98 die deutschen Belange berücksichtigt. Um internationale Vergleiche zu ermöglichen, werden aber auch gemeinsame Themen mit dem sechsten japanischen Delphi, das parallel zu der deutschen Studie durchgeführt wurde, betrachtet. Eine genauere Analyse findet sich in Cuhls/Blind/Grupp (2001). Die jeweilige Anzahl der in Deutschland und in Japan identischen Thesen ist in der Tabelle 1 ebenfalls angegeben. Die Thesen sind in Japan allerdings zum Teil in anderen Themenfeldern zur Diskussion gestellt worden. Um auch beobachten zu können, ob sich bestimmte Einschätzungen in den letzten fünf Jahren geändert haben, wurden einige Thesen aus dem ersten deutschen Delphi-Report unverändert übernommen und noch einmal in die verschiedenen Fragebögen integriert.

Tabelle 1: Anzahl der Thesen in den Delphi'98 - Themenfeldern

Delphi'98 Themenfelder	Anzahl der Thesen im Delphi'98	gleiche Thesen in Japan	gleiche Thesen im Delphi 93
1. Information & Kommunikation	111	68	16
2. Dienstleistung & Konsum	78	21	6
3. Management & Produktion	71	14	7
4. Chemie & Werkstoffe	104	33	24
5. Gesundheit & Lebensprozesse	104	36	7
6. Landwirtschaft & Ernährung	101	20	7
7. Umwelt & Natur	76	15	2
8. Energie & Rohstoffe	114	32	14
9. Bauen & Wohnen	75	32	9
10. Mobilität & Transport	107	18	5
11. Raumfahrt	78	24	9
12. Großexperimente	51	10	7
Summe	1.070	323	113

2.2 Wie sahen die Fragebögen aus?

Die erarbeiteten Fragebögen umfassten Thesen, die durch Ankreuzen nach verschiedenen Kriterien zu beurteilen waren. Die erste Spalte, die im Fragebogen anzukreuzen war, betraf die Selbsteinschätzung der *Fachkenntnis* als „groß“ (bezeichnet mit Fk1), „mittel“ (Fk2), „gering“ (Fk3) oder „fachfremd“ (Fk4). Aufgrund dieser Selbstbewertung wurde der Wissensstand auf Basis der Antwortenden pro These berechnet. Personen, die sich als „fachfremd“ einstufen, wurden bei dieser Berechnung nicht einbezogen, da sie die Fragen nicht weiter zu beantworten hatten.

Die Thesen sollten dann nach den folgenden Kriterien von den Experten beurteilt werden. Zum Ersten wurde nach der Wichtigkeit der Lösungsbeiträge gefragt. Dabei wurde unterschieden nach der Bedeutung für die Erweiterung menschlichen Wissens, für die wirtschaftliche Entwicklung, für die gesellschaftliche Entwicklung, für die Lösung der ökologischen Probleme und für Arbeit und Beschäftigung. Als zweiter Punkt sollte der mögliche Zeitraum der Verwirklichung eingeschätzt werden, um festzulegen, ob es sich um einen langen oder eher kürzeren Zeitraum handelt, auf den sich die weitere Beurteilung bezieht, und um den zeitlichen Rahmen für Handlungsspielräume festzulegen. Der Zeithorizont war in Fünf-Jahres-Intervalle unterteilt, ergänzt um den Zeitraum bis zum Jahre 2000 und ein offenes Intervall nach 2025.

Die nächste Frage bezog sich auf den Stand der Forschung und Entwicklung. Welches Land wird als führend angesehen, die USA, Japan, Deutschland, ein anderes Land der EU oder ein anderes Land außerhalb der EU? Außerdem wurde danach gefragt, wie die Rahmenbedingungen eingeschätzt werden bzw. welche Maßnahmen zu ihrer Verbesserung angebracht sind. Als Optionen standen verbesserte Ausbildung, der Austausch von Personal zwischen Wissenschaft und Wirtschaft, die internationale Kooperation, die Verbesserung der Forschungsinfrastruktur, eine verbesserte staatliche Projektförderung oder Regulationsänderungen zur Auswahl. Schließlich wurde nach möglichen Folgeproblemen bei einer Realisierung der Visionen gefragt. Alternative Lösungsmöglichkeiten konnten ebenfalls genannt werden.

2.3 Wer wurde befragt?

Wer waren nun die in der Studie befragten Personen? Wer ist Experte für die Zukunft? Da niemand genau wissen kann, wie die Zukunft werden wird, sollten möglichst viele Personen beteiligt sein. Denn es ist erwiesen, dass bei einer großen Anzahl von Antworten individuelle Schätzfehler ausgemittelt werden können und damit die Wahrscheinlichkeit einer „treffsicheren Prognose“ größer ist.² Die „Treffsicherheit“ darf daher keinesfalls als Alleinkriterium für den „Erfolg“ einer Studie angesehen werden, sondern die Tatsache, dass Aktivitäten angestoßen werden (Kooperationen, neue Projekte etc.) ist gemessen an den Zielen der Studie als ebenso wichtig anzusehen.

Die Befragten mussten sich mit der These auseinander setzen können, um überhaupt ein Urteil abgeben zu können. Deshalb wurde versucht, eine Adress-Datenbank aufzubauen, die Personenkreise umfasst, welche sich auf den unterschiedlichen Gebieten auskennen. Diese Personen sollten möglichst zu jeweils einem Drittel aus der Industrie, aus Hochschulen sowie aus anderen Forschungseinrichtungen, dem öffentlichen Dienst und Verbänden stammen.

Um pro Themenfeld etwa 100 auswertbare Bögen zu erhalten bzw. sogar möglichst mindestens 100 Antworten pro Fragestellung, wurden zunächst ca. 7.000 Personen angeschrieben. Ihnen wurde angeboten, einen zusätzlichen Fragebogen anzufordern, falls sie sich auch für ein anderes Gebiet kompetent fühlten, oder ihren Bogen gegen einen anderen auszutauschen. Insgesamt haben über 2.000 Personen 2.453 Bögen ausgefüllt (Tabelle 2). Dies ist ein erfreuliches Ergebnis, da ein auswertbarer Rücklauf von mehr als 30 % (in einigen Gebieten des Delphi '98 sind es sogar fast 50 %) die

² Dies wurde bereits früh von Dalkey, Helmer und anderen in den USA getestet (siehe z.B. Dalkey 1969a und b; Dalkey/Brown/Cochran 1969; Dalkey 1968; Dalkey/Helmer 1963).

Erfahrungswerte in Deutschland übertrifft. In einigen Feldern, z.B. „Chemie und Werkstoffe“, „Gesundheit und Lebensprozesse“, „Energie und Rohstoffe“ sowie „Information und Kommunikation“, sind deutlich mehr als 200 Antworten eingegangen. In Feldern wie „Großexperimente“ oder „Raumfahrt“ dagegen konnten so viele Antworten nicht erwartet werden, denn die Zahl der Forscher liegt auf diesen Gebieten in Deutschland niedriger als auf anderen. In der folgenden Tabelle findet sich die Anzahl der in der ersten Runde eingegangenen Bögen und damit in der zweiten Runde beteiligten Experten. Diejenigen Personen, die anonym geantwortet haben, konnten leider an der zweiten Runde nicht mehr beteiligt werden. Der Rücklauf am Ende der zweiten Runde findet sich in der rechten Spalte der Tabelle 2 und beträgt ca. 75 % der in der ersten Runde eingegangenen Bögen.

Tabelle 2: Anzahl der auswertbaren Bögen nach Themenfeldern

Delphi Themenfelder	Anzahl der auswertbaren Fragebögen	
	1. Runde	2. Runde
1. Information & Kommunikation	287	206
2. Dienstleistung & Konsum	215	163
3. Management & Produktion	229	179
4. Chemie & Werkstoffe	260	206
5. Gesundheit & Lebensprozesse	255	182
6. Landwirtschaft & Ernährung	206	140
7. Umwelt & Natur	282	209
8. Energie & Rohstoffe	246	187
9. Bauen & Wohnen	110	94
10. Mobilität & Transport	150	122
11. Raumfahrt	97	77
12. Großexperimente	116	91
Gesamt	2.453	1.856

2.4 Die Expertise der Beteiligten

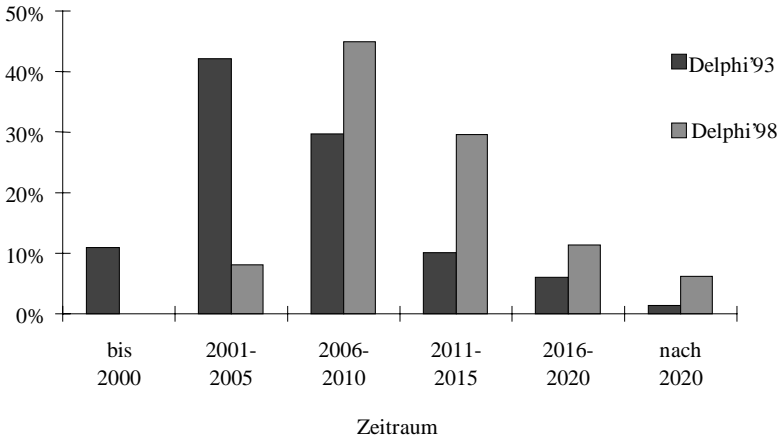
Bei der Auswahl der angeschriebenen Experten war das Ziel, viele Personen mit einer möglichst hohen Expertise in den jeweiligen Bereichen zu identifizieren. Jedoch sollten auch Personen mit so genannter „sekundärer“ Fachkenntnis in den Kreis der Befragten aufgenommen werden, um ein gewisses korrigierendes Gegengewicht zu den Wissenschaftlern mit höchster Expertise zu erhalten.

Im Endergebnis haben sich durchschnittlich 10 % der Antwortenden als Experten mit höchster Fachkenntnis eingestuft, d.h. sie arbeiten auf dem betreffenden Gebiet. Weitere 30 % haben mittlere (d.h. frühere Arbeit auf dem Gebiet und/oder Lesen der Primärliteratur) und 60 % geringe Fachkenntnis (Lesen der Sekundärliteratur bzw. im Gespräch mit Fachleuten stehend) angegeben. Dies mag auf den ersten Blick unbefriedigend erscheinen, jedoch gilt es zu bedenken, dass die zwölf Fragebögen eine so große inhaltliche Spannbreite umfassen, dass der einzelne Experte lediglich bei sehr wenigen Thesen glaubwürdig eine *hohe* Fachkenntnis (arbeitet auf dem Gebiet) bekunden kann. Ferner ist anzumerken, dass der Anteil der Experten mit höchster Fachkenntnis in der ersten Runde noch bei 13,5 % lag. Es haben sich wohl mehrere „Fachkenner“ in der zweiten Runde nicht mehr beteiligt bzw. in ihrer Selbsteinschätzung zurückgestuft. Eine grundsätzliche Tendenz zur Überschätzung der eigenen Fachkenntnis konnte auf der Basis einzelner Fragebögen nicht beobachtet werden und spiegelt sich daher auch im aggregierten Ergebnis nicht wider.

2.5 Erwarteter Realisierungszeitraum der Visionen

Die Verteilung der Realisierungszeiträume der Thesen unterscheidet sich gemäß Abbildung 1 in ihrer generellen Form nicht von der Verteilung bei der ersten Delphi-Studie. Einzige Ausnahme ist, dass sie sich um etwa fünf Jahre in die Zukunft verschoben hat. Nahezu die Hälfte der zur Einschätzung vorgelegten Thesen kann nach Meinung der befragten Experten im Zeitraum zwischen 2006 und 2010 realisiert werden, ein weiteres Drittel zwischen 2011 und 2015. Bis zum Jahr 2000 sollte es nach der durchschnittlichen Einschätzung der Experten noch *keine einzige* Realisierung geben. Aber auch nach 2025 werden nicht einmal 30 Visionen eingestuft. Damit ist der angestrebte Zeithorizont bis 2025 bei der Mehrzahl der Thesen eingehalten worden.

Während zwischen den Experten in Unternehmen und Hochschulen kein Unterschied in der Beurteilung des Realisierungszeitraums auszumachen ist, zeigen sich geringe Differenzen zwischen den verschiedenen „Expertentypen“. So sehen die auf Basis der Einschätzungen allgemeiner Megatrends bestimmten Umweltpessimisten und Bevölkerungsoptimisten die Innovationen zwischen 2007 und 2015 als realisierbar an, während die Fortschrittsskeptiker ihre Verwirklichung im Durchschnitt ein Jahr später erwarten - vor dem Hintergrund eines Zeithorizontes von 30 Jahren kein großer Unterschied (Blind/Cuhls/Grupp 2001). Ferner neigen die Antwortenden mit hoher Expertise dazu, die Realisierung der Thesen früher einzuschätzen.

Abbildung 1: Zeitraumverteilungen Delphi'93 und Delphi'98

Da die durchschnittlichen Realisierungszeiträume von der Zusammensetzung der Thesen hinsichtlich der Innovationsstufen nicht ganz unabhängig sind, soll an dieser Stelle auch hierauf eingegangen werden. Die Thesen in den Fragebögen enthalten stereotype Wörter, die zwischen wissenschaftlichem, technischem oder wirtschaftlichem Fortschritt unterscheiden. Es dominieren mit nahezu 40 % die Thesen, die sich auf bereits vermarktete, aber noch nicht weit verbreitete Produkte und Dienstleistungen beziehen. Weitere 30 % umfassen Visionen zur ersten wirtschaftlichen Anwendung bereits entwickelter und als Prototyp vorliegender Technologien. Weitere 18 % haben die Entwicklung von Prototypen zum Gegenstand, und lediglich etwas mehr als 11 % der Thesen sind der Grundlagenforschung zuzurechnen. Damit ist im Vergleich zur ersten deutschen Delphi-Studie eine höhere Anwendungsorientierung zu bemerken.

3. Test von Hypothesen zur Rolle des Expertenstatus

Bevor die Hypothesen zur Rolle des Expertenstatus getestet werden, muss eine Reihe methodischer Vorbemerkungen gemacht werden. Die befragten Experten im Bereich von Wissenschaft und Technologie wurden auf der Basis von öffentlich zugänglichen Datenbanken und durch Nennungen von anderen Experten ausgewählt. Damit sollte die Durchführung einer Expertenbefragung sichergestellt werden. Jedoch wurden

verschiedene Wissensgebiete in zwölf Fragebögen zusammengefasst, um der Interdisziplinarität nahe stehender Gebiete Rechnung zu tragen, um ausreichende Samples zu gewährleisten und um Kosten zu sparen. Solche, mehrere Themenfelder umfassende Fragebögen, bedeuten für die Experten, dass sie nur in ihren ausgewählten Teilbereichen eine hohe Expertise besitzen, während sie in den angrenzenden Randbereichen eine geringere und in noch entfernteren Feldern unter Umständen gar kein Expertenwissen mehr besitzen. Die Angaben der Experten zu ihrer Expertise bezüglich einer bestimmten Vision beruhen allein auf ihrer Selbsteinschätzung, die mit objektivierbaren Kriterien im Einzelfall nicht überprüfbar ist.³ Deshalb muss generell davon ausgegangen werden, dass diese Selbsteinschätzung korrekt vorgenommen wurde und zwischen den Experten eine gewisse Übereinstimmung besteht.

Methodische Konsequenz aus diesem Sachverhalt ist, dass den befragten Experten kein genereller Expertenstatus zugeordnet und deshalb auch nicht danach segmentiert werden kann. Die Analyse wird bei der Beobachtungseinheit deshalb nicht auf die ungefähr 2.000 Experten zurückgreifen, sondern auf deren knapp 100.000 Antworten aus den beiden Runden der Delphi-Befragung bzw. ca. 65.000 Antworten von Experten, die in der ersten und zweiten Runde teilgenommen haben.

In der Vergangenheit wurde die Panelmortalität bei Delphi-Untersuchungen in einem engen Zusammenhang mit der Expertise der Befragten gesehen. Schon Bardecki (1984) konnte nicht bestätigen, dass Experten mit hoher Expertise bzw. starkem ego-involvement mit einer geringeren Wahrscheinlichkeit aus einem Delphi-Prozess ausscheiden. Häder (2000) konnte in seiner Untersuchung dieser Fragestellung auch keine signifikanten Unterschiede zwischen der Abbruchwahrscheinlichkeit und dem Expertenstatus der Befragten ausmachen. Im Kontext dieser Untersuchungen ist die folgende erste Hypothese zu sehen.

Hypothese 1: Experten mit geringem Expertenwissen werden mit geringerer Wahrscheinlichkeit an der zweiten Befragungsrunde teilnehmen. Hypothese 1 lässt sich dadurch begründen, dass mit zunehmender Expertise das Interesse an einer Einflussnahme und den Endergebnissen selbst ausgeprägter ist, weil sich die Fachleute davon unter Umständen auch neue Erkenntnisse für die eigene zukünftige Forschung versprechen. Ferner dürften Teilnehmer bei zu geringer Expertise befürchten, eine „falsche“ Antwort zu geben, und deshalb eher in der zweiten Runde nicht mehr teilneh-

³ Es wurde jedoch geprüft, ob einzelne Experten in allen von ihnen ausgefüllten Fragen die höchste Expertise angegeben haben. Da dies aus den oben geschilderten Gründen nicht möglich ist, wären solche Antworten aus dem Datensatz entfernt worden. Dieser Fall kam aber nicht vor.

men. Daher wird im Folgenden geprüft, ob Experten mit geringerem Fachwissen tatsächlich in größerer Zahl von einer Teilnahme an der zweiten Runde absehen.

Das Gesamtsample wurde differenziert in diejenigen Antworten, die nur in der ersten Runde, und diejenigen, die auch in der zweiten Runde abgegeben wurden. Abbildung 2 stellt die Verteilung der Expertise nach Themenfeldern basierend auf den Antworten zu einzelnen Thesen in der ersten Runde dar. Als Test-Methode wurde der Chi-Quadrat nach Pearson verwendet. Tabelle 3 stellt das aggregierte Ergebnis über alle zwölf Themenfelder und das Ergebnis des Chi-Quadrat-Tests dar. Es wird deutlich, dass die Durchschnittsexpertise bei den Antworten der Abbrecher unter derjenigen liegt, die sich auch an der zweiten Runde beteiligt haben.

Eine getrennte Durchführung des Tests nach den zwölf Themenfeldern ergibt jedoch differenzierte Ergebnisse. Während in den Themenfeldern „Information & Kommunikation“, „Dienstleistung & Konsum“, „Management & Produktion“, „Gesundheit & Lebensprozesse“, „Umwelt & Natur“, „Bauen & Wohnen“, „Mobilität & Transport“ und „Großexperimente“ ein negativer Zusammenhang zwischen der Expertise und der Abbruchwahrscheinlichkeit besteht, haben in den Themenfeldern „Chemie & Werkstoffe“, „Energie & Rohstoffe“ und „Raumfahrt“ an der zweiten Runde weniger Personen mit hoher Expertise teilgenommen. Hier könnten sich die Experten mit hoher Expertise an den Inhalten, nach methodischen Aspekten oder nach Analyse der Ergebnisse der ersten Runde dazu entschieden haben, sich an der zweiten Runde nicht mehr zu beteiligen.⁴ Im Themenfeld „Landwirtschaft & Ernährung“ konnte keinerlei Zusammenhang zwischen Expertise und Abbruch-Wahrscheinlichkeit festgestellt werden.

⁴ Hier kann auch Zufall nicht ausgeschlossen werden, da viele Experten einfach aufgrund von Zeitmangel/Arbeitsüberlastung den zweiten Bogen nicht mehr ausfüllen konnten.

Abbildung 2: Verteilung der Experten-Anteile in den Themenfeldern differenziert nach Abbrechern (2. Zeile) und Nicht-Abbrechern (1. Zeile)

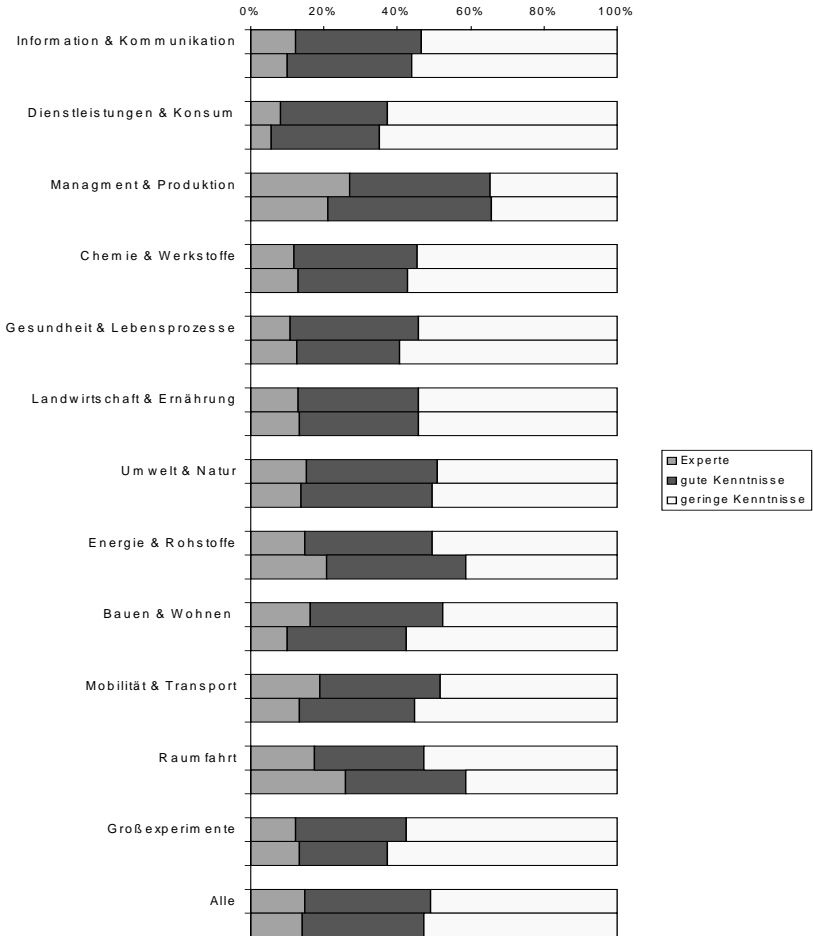


Tabelle 3: Zusammenhang zwischen Abbruch und Grad der Expertise in der ersten Runde⁵

Themenfelder	Ergebnis des Chi-Quadrat-Tests	Anzahl der Beobachtungen
1. Information & Kommunikation	---	11.201
2. Dienstleistung & Konsum	---	8.352
3. Management & Produktion	---	9.013
4. Chemie & Werkstoffe	++	7.901
5. Gesundheit & Lebensprozesse	---	11.809
6. Landwirtschaft & Ernährung	?	8.011
7. Umwelt & Natur	-	9.028
8. Energie & Rohstoffe	+++	12.239
9. Bauen & Wohnen	---	3.660
10. Mobilität & Transport	---	7.970
11. Raumfahrt	+++	3.348
12. Großexperimente	--	1.816
Gesamt	---	94.348

Obwohl in zwei Dritteln der Themenfelder die Abbruchwahrscheinlichkeit mit fallender Expertise steigt, müssen in diesem Zusammenhang noch zwei weitere Hypothesen für einen Abbruch genannt werden (Häder 2000). So besagt die Dissonanz-Hypothese (Festinger 1978), dass Aussteiger die Sachverhalte systematisch anders bewerten und diese kognitiven Dissonanzen ein Abbruchmotiv darstellen. In dieselbe Richtung, jedoch mit umgekehrter Kausalität, argumentieren die Befürworter der Nonkonformitäts-Hypothese, dass Aussteiger aufgrund ihrer Distanz zum Thema extremere Urteile abgeben als der Durchschnitt der Experten. Die Stärke der Dissonanz hängt allerdings von verschiedenen Faktoren ab, ebenso wie die Reaktionen der Einzelpersonen. So haben diese nach Bardecki (1984) die verschiedenen Möglichkeiten, 1. den psychologischen Anker des zurückgespielten Mittelwertes oder Medians zu ignorieren, 2. sich bewusst in Kontrast zum psychologischen Anker zu stellen oder 3. seine Meinung in Richtung des psychologischen Ankers zu verändern (Assimilation). Generell kann nur gesagt werden „agreeing with people decreases dissonance, disagreeing increases it“ (Bardecki 1984, S. 283). Wie die Delphi-Experten im Einzel-

⁵ Abbruch ist einfach dadurch definiert, dass der Experte an der zweiten Runde nicht mehr teilgenommen hat. Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/-- < 0,01; ++/-- < 0,05, +/- < 0,1. Ein Fragezeichen bedeutet kein signifikanter Zusammenhang.

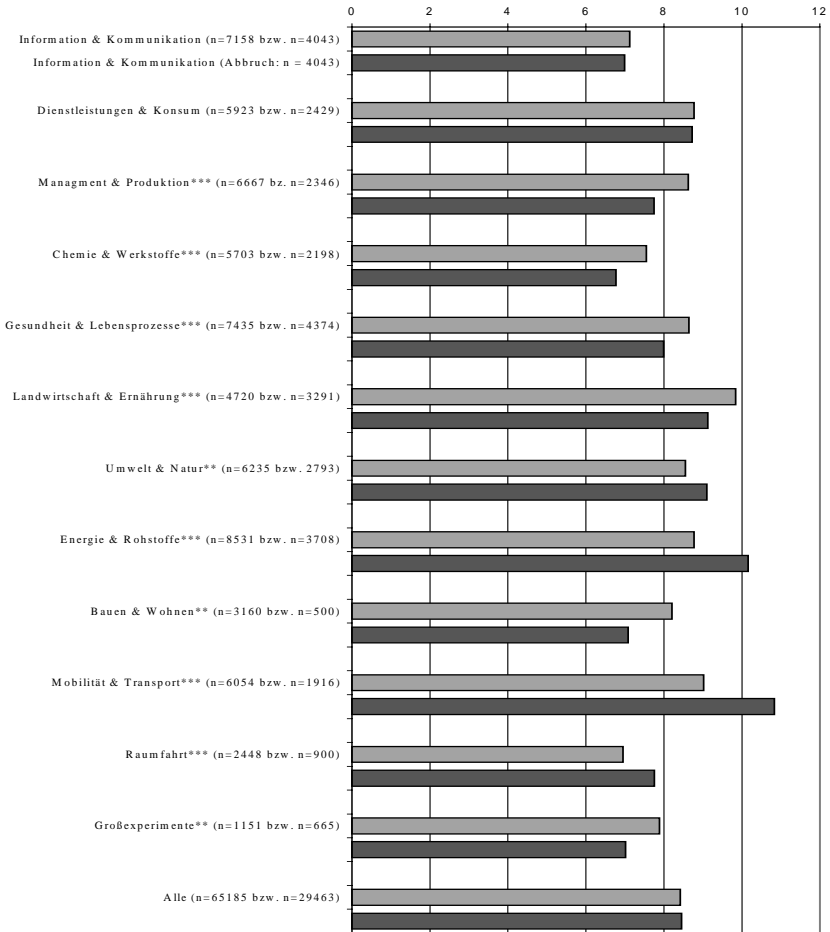
nen reagieren, kann daher nicht pauschalisiert werden, sondern scheint sich nach der Einzelperson zu richten.

Extreme Einschätzungen lassen sich im Delphi'98 am besten an den Antworten zum erwarteten Realisierungszeitraum ablesen, während die Skalen zu den anderen Fragen in der Regel nur Ja/Nein-Optionen erlauben. Die Differenz zwischen der Einzeleinschätzung bezüglich des Realisierungszeitraums und des Medianwertes wurde als Absolutwert berechnet. In Abbildung 3 sind die durchschnittlichen Differenzen der Abbrecher bzw. Nicht-Abbrecher von den jeweiligen Medianen differenziert nach Themenfeldern angegeben. Während im Aggregat aller Antworten kein Unterschied beim Mittelwert der Abweichungen festzustellen ist, sind in zehn der zwölf Themenfelder signifikante Unterschiede auszumachen. Jedoch halten sich die Themenfelder, in denen die Abbrecher extremere Urteile gefällt haben, und die Themenfelder, in denen die Nicht-Abbrecher stärker vom Median abweichen, die Waage. Dieses Ergebnis unterstreicht den Bedarf an weiteren Untersuchungen zur Abbruchwahrscheinlichkeit bei Delphi-Untersuchungen. Unmittelbar an den Zusammenhang zwischen Abbruchwahrscheinlichkeit und Extremurteilen schließt sich der in Hypothese 2 unterstellte Zusammenhang an.

Hypothese 2: Je höher die Expertise, desto größer sind die Unterschiede zwischen der eigenen Einschätzung und der mittleren Einschätzung. Bei dieser Hypothese wird davon ausgegangen, dass Experten vorgefasste Meinungen haben, die gegebenenfalls extrem sein können und von denen sie nicht abweichen wollen.⁶ Dies trifft auf ein allgemein verbreitetes Vorurteil gegenüber Wissenschaftlern (diese seien in ihr Thema „verliebt“ und ließen sich selten von ihrer Meinung abbringen) und bezieht sich bezüglich Delphi'98 auf die zweite, bei anderen Delphi-Untersuchungen eventuell auch auf spätere Runden. Wie oben schon erwähnt, lassen sich extreme Einschätzungen am besten an den Antworten zum erwarteten Realisierungszeitraum ablesen. Für die Überprüfung der Hypothese wurde die Differenz zwischen der Einzeleinschätzung und des Medianwertes als Absolutwert ermittelt und mit der jeweiligen Expertise korreliert.

⁶ Ein unfreiwilliger Test ergab sich diesbezüglich im Mini-Delphi von 1995, bei dem in einem Fragebogen durch einen technischen Fehler falsche Daten zu Thesen zurückgespielt wurden. Die meisten Experten korrigierten ihre Meinung in Richtung der falschen Angabe. Nur einige der Fk1-Experten bestanden vehement auf ihrer vorherigen Meinung und wunderten sich ob der seltsamen Ergebnisse (Kommentare, Cuhls/Breiner/Grupp 1995).

Abbildung 3: Mittelwerte der Abweichungen von den Medianen in den Themenfeldern differenziert nach Abbrechern (2. Zeile) und Nicht-Abbrechern (1. Zeile)⁷



⁷ Abbruch ist einfach dadurch definiert, dass der Experte an der zweiten Runde nicht mehr teilgenommen hat. Die Anzahl der Sternchen repräsentiert das Signifikanzniveau: ***<0,01; **<0,05.

Tabelle 4: Zusammenhang zwischen der Abweichung von der Durchschnittseinschätzung und dem Grad der Expertise⁸

Themenfelder	Signifikanzniveaus des Spearman-Rho-Tests						
	Abweichung in Runde 1 mit Expertise in Runde 1 (mit Abbrecher)	Abweichung in Runde 1 mit Expertise in Runde 1 (ohne Abbrecher)	Anzahl der Beobachtungen	Abweichung in Runde 2 mit Expertise in Runde 2	Abweichung in Runde 1 mit Durchschnitts-Expertise	Abweichung in Runde 2 mit Durchschnitts-Expertise	Anzahl der Beobachtungen
1. Information Kommunikation	---	?	11.201	+++	--	?	7.158
2. Dienstleistung Konsum	---	---	8.352	?	--	?	5.923
3. Management Produktion	---	?	9.013	?	?	+	6.667
4. Chemie Werkstoffe	---	--	7.901	+++	?	+++	5.703
5. Gesundheit Lebensprozesse	---	?	11.809	+++	---	+	7.435
6. Landwirt-Ernährung	---	---	8.011	?	---	?	4.720
7. Umwelt Natur	---	?	9.028	?	--	-	6.235
8. Energie Rohstoffe	---	+++	12.239	+++	+	?	8.531
9. Bauen Wohnen	?	++	3.660	?	?	?	3.160
10. Mobilität & Transport	---	--	7.970	?	?	?	6.054
11. Raumfahrt	---	?	3.348	+++	--	?	2.448
12. Großexp.	?	?	1.816	?	?	?	1.151
Gesamt (Pearson)⁹	+++	?	94.348	+++	---	?	65.185

⁸ Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/-- < 0,01; +/--- <0,05, +/- <0,1. Ein Fragezeichen bedeutet kein signifikanter Zusammenhang.

⁹ Aufgrund der hohem Fallzahl erlaubt SPSS die Berechnung des den Daten angemessenen Spearman-Rho-Tests nicht.

In Tabelle 4 sind die Signifikanzniveaus der Korrelationsanalyse dokumentiert, wobei zum einen zwischen den Unterschieden in der ersten und zweiten Befragungsrunde differenziert wurde. Ferner wurden in den Analysen eine Durchschnittsexpertise basierend aus den Angaben aus der ersten und zweiten Runde berechnet und mit den Unterschieden korreliert, weil – wie oben schon ausgeführt – die Experten dazu tendierten, ihre in der ersten Runde postulierte Expertise nach unten zu revidieren. Wer sich als wahrer Experte (Fk1) betrachtet, hat demnach häufiger das Standing, auch bei einer anderen von der Mehrheitsmeinung abweichenden Zeiteinschätzung zu bleiben.

Die Ergebnisse zeigen auf den ersten Blick ein völlig heterogenes Bild. Während für die Antworten in der ersten Runde Hypothese 2 eindeutig widerlegt werden kann, weil die Abweichungen der individuellen Einschätzungen mit abnehmender Expertise in den meisten Themenfeldern signifikant zunehmen, gilt dies für die Antworten in der zweiten Runde nicht mehr. Hier kann aber auch für kein Technikfeld die Gegenhypothese zu Hypothese 2 bestätigt werden. Dagegen kann in der Hälfte der Themenfelder eine signifikante Bestätigung der Hypothese 2 nachgewiesen werden. Hier nimmt mit zunehmender Expertise der Abstand zur Median-Meinung zu. Zwei Effekte können hier wohl als Gründe für die Bestätigung der Hypothese herangezogen werden. Zum einen haben wir gesehen, dass viele Experten – angesichts des sehr niedrigen Anteils an ausgewiesenen Experten – ihre Selbsteinschätzung bezüglich ihrer Expertise revidiert und in vielen Fällen reduziert haben. Dies bedeutet, dass nur noch die „wahren“, auf dem Gebiet arbeitenden Experten das Selbstbewusstsein haben, diese Kategorie im Fragebogen auch anzukreuzen. Zum Zweiten haben die Experten nun die Durchschnittsmeinung ihrer Kollegen kennen gelernt. Auch hier bedarf es eines ausgeprägten Selbstbewusstseins, um sich weiterhin signifikant von der Durchschnittsmeinung zu unterscheiden. Das Zusammenspiel dieser beiden Effekte kann somit erklären, dass es in vielen Themenfeldern doch noch zu einer Bestätigung von Hypothese 2 kommt. Zur Information sind auch noch die Signifikanzniveaus der Korrelationsanalyse der Differenzen mit der Durchschnittsexpertise aus beiden Runden dokumentiert. Angesichts der oben beschriebenen Effekte ergeben sich deshalb lediglich für die Einschätzungen der zweiten Runde in manchen Themenfeldern die erwarteten Ergebnisse.

Hypothese 3: Je höher die Expertise, desto geringer sind die Unterschiede zwischen den Einschätzungen in der 1. und der 2. Runde. Aufgrund der bereits zu Hypothese 2 gemachten Erläuterungen kann davon ausgegangen werden, dass die Antwortenden mit einer hohen Fachkenntnis nur in einem geringeren Umfang ihre in der ersten Befragungsrunde abgegebenen Einschätzungen revidieren werden. Ähnlich wie bei

Häder (2000) liegt Hypothese 3 die Annahme zu Grunde, dass Personen, die auf dem Gebiet arbeiten, ein gesundes Selbstbewusstsein hinsichtlich ihrer (Eigen-) Einschätzung besitzen.

Dazu wurden die Ergebnisse der 1. und 2. Runde hinsichtlich einer Veränderung bei der Angabe zur „Fachkenntnis“ überprüft. Die in Tabelle 5 dargestellten Ergebnisse machen deutlich, dass Hypothese 3 in knapp der Hälfte der Themenfelder bestätigt werden kann. Es herrscht daher die Tendenz, dass diejenigen Personen mit geringer Expertise ihre Einschätzungen stark revidieren (d. h. der Mehrheitsmeinung anpassen), während die „wahren“ Experten nur in sehr geringem Ausmaß ihre Einschätzung der ersten Runde verändern.

Tabelle 5: Zusammenhang zwischen der Veränderung der eigenen Einschätzung und dem Grad der Expertise¹⁰

Themenfelder	Ergebnisse des Spearman-Rho-Tests mit Expertise der 2. Runde	Ergebnisse des Spearman-Rho-Tests mit Durchschnittsexpertise	Anzahl der Beobachtungen
1. Information & Kommunik.	++	+++	7.158
2. Dienstleistung & Konsum	++	++	5.923
3. Management & Produktion	?	?	6.667
4. Chemie & Werkstoffe	?	?	5.703
5. Gesundheit & Lebensprozesse	+++	+	7.435
6. Landwirtschaft & Ernährung	?	++	4.720
7. Umwelt & Natur	+++	++	6.235
8. Energie & Rohstoffe	?	-	8.531
9. Bauen & Wohnen	+++	+	3.160
10. Mobilität & Transport	?	?	6.054
11. Raumfahrt	?	?	2.448
12. Großexperimente	?	?	1.151
Gesamt	?	?	65.185

¹⁰ Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/--- < 0,01; ++/-- < 0,05, +/- < 0,1. Ein Fragezeichen bedeutet kein signifikanter Zusammenhang.

Hypothese 4: Je höher die Expertise, desto geringer ist die Konvergenz (Differenz zwischen der Abweichung vom Median in der ersten Runde und der Abweichung vom Median in der zweiten Runde). In engem Zusammenhang mit Hypothese 3 steht Hypothese 4, wonach die Konvergenz, d. h. die Annäherung an die Median-Meinung in der zweiten Befragungsrunde, mit zunehmender Fachkenntnis schwächer ausgeprägt ist.

Tabelle 6: Zusammenhang zwischen der Konvergenz und dem Grad der Expertise¹¹

Themenfelder	Ergebnisse des Spearman-Rho-Tests mit Expertise der 2. Runde	Ergebnisse des Spearman-Rho-Tests mit Durchschnittsexpertise	Anzahl der Beobachtungen
1. Information & Kommunik.	---	?	7.158
2. Dienstleistung & Konsum	?	---	5.923
3. Management & Produktion	-	?	6.667
4. Chemie & Werkstoffe	?	?	5.703
5. Gesundheit & Lebensprozesse	---	---	7.435
6. Landwirtschaft & Ernährung	?	?	4.720
7. Umwelt & Natur	---	?	6.235
8. Energie & Rohstoffe	---	+	8.531
9. Bauen & Wohnen	--	?	3.160
10. Mobilität & Transport	?	?	6.054
11. Raumfahrt	---	--	2.448
12. Großexperimente	?	?	1.151
Gesamt	---	---	65.185

Auch hierzu wurden die Ergebnisse der ersten mit denen der zweiten Runde verglichen. Tabelle 6 zeigt eindrücklich, dass Hypothese 4 für zwei Drittel der Themenfelder bestätigt werden kann. Dies bedeutet, dass die Befragten in ihren Feldern mit hoher Expertise sich dem der Delphi-Methode typischen Konvergenz-Prozess nicht in dem Ausmaß unterwerfen, wie es für die Experten mit mittlerer oder geringer Kenntnis der Fall ist. Dieses Ergebnis ruft unmittelbar einen der Methode immanenten Widerspruch hervor. Zum einen will man für Delphi-Befragungen im Bereich von Wissenschaft und Technik möglichst viele Antwortenden gewinnen, die sich durch eine hohe Expertise

¹¹ Die Anzahl der Plus- bzw. Minuszeichen repräsentiert das Signifikanzniveau: +++/-- < 0,01; ++/-- < 0,05, +/- < 0,1.

auszeichnen. Andererseits deuten die Ergebnisse bezüglich des Tests von Hypothese 4 darauf hin, dass sich gerade die „wahren“ Experten in ihren speziellen Fachgebieten durch eine gewisse Resistenz gegenüber dem Konvergenzprozess auszeichnen und dadurch das Ziel einer mehrstufigen Delphi-Befragung unterlaufen.¹² Dies unterstreicht die neueren Ziele, Delphi-Befragungen nicht mehr anzuwenden, um Konsens herbeizuführen, sondern um festzustellen, ob bereits ein Konsens besteht oder nicht (Cuhls/Blind/Grupp 1998). Dies war insbesondere im Delphi'98 eine der wichtigen Informationen für diejenigen Akteure, welche die Studie für sich selbst ausgewertet haben.

Abbildung 4 bestätigt die Ergebnisse des Hypothesentest. Denn mit zunehmender Fachkenntnis¹³ für die einzelnen Themenfelder nimmt die Varianz (Differenz zwischen dem 3. und dem 1. Quartil) bei der Einschätzung der Realisierungszeiträume zu. Ausnahme ist hier das Feld „Management & Produktion“.

Angesichts dieser Ergebnisse ist es noch wichtiger, die Gesamtergebnisse von Delphi-Befragungen differenziert nach der Expertise der Beteiligten auszuweisen. Es wäre natürlich auch überprüfenswert, ob und wie die Experten ihre Meinungen bei Einzelthesen geändert haben, zum Beispiel, inwieweit die Fk1-Experten im Einzelnen bzw. gesamt konvergieren. Dies kann und darf aufgrund der geringen Fallzahlen nicht ausreichend nachgewiesen werden.

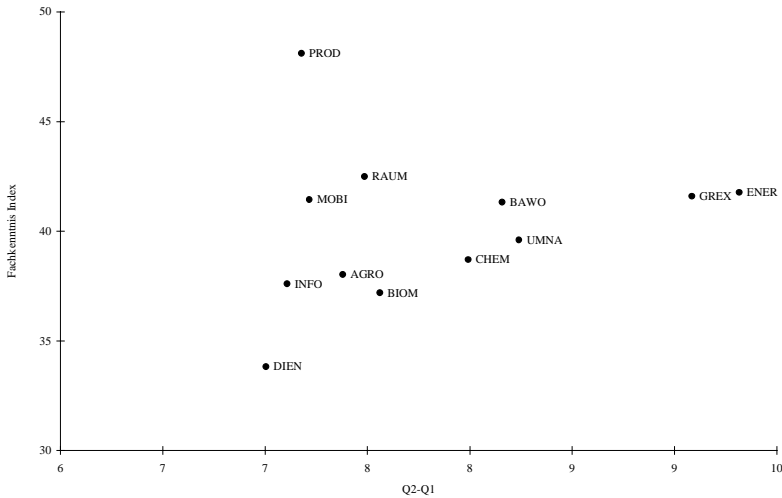
4. Zusammenfassung und Schlussfolgerungen

Delphi '98 hatte nicht zum Ziel, methodisch und statistisch genaue Trennungen der Fachexperten sozialwissenschaftlich zu untersuchen. Da sich aber die einmalige Gelegenheit eines großen Samples bietet (bisher sind nur sehr kleine Samples für Begründungen und Tests herangezogen worden), wurde der Versuch einer detaillierten Prüfung des Expertenverhaltens unternommen.

¹² Theoretisch könnte man auch das Konvergenzverhalten der Experten bezüglich deren mittlerer Einschätzung untersuchen. Jedoch ist die Anzahl der abgegebenen Einschätzungen insbesondere in der 2. Runde bezüglich einzelner Thesen so gering, dass die Berechnung des Medians oder Mittelwertes nicht gerechtfertigt ist.

¹³ Der Fachkenntnisindex berechnet sich aus den Anteilen der Fk1-, Fk2- und Fk3-Experten, wobei die Fk1-Experten mit dem Gewicht 4 und die Fk2-Experten mit dem Gewicht 2 in die Indexberechnung eingehen, so dass Wert 100 bei einem 100%-Anteil der Fk1-Experten und der Wert 25 bei einem 100%-Anteil der Fk3-Experten erreicht wird.

Abbildung 4: Zusammenhang zwischen Fachkenntnis und Varianz (Differenz zwischen dem ersten und dritten Quartil)



Die vier für die Delphi-Methodik grundlegenden Hypothesen konnten nur teilweise bestätigt werden. Das heißt, die der Methodik zu Grunde liegenden Annahmen müssen im Einzelfall sehr sorgfältig überprüft werden. Hypothese 1 konnte nur in einigen Themenfeldern nachgewiesen werden. Es ist also nur bedingt anzunehmen, dass die Experten mit geringem Wissensstand häufiger zu sogenannten Drop-outs werden und sich an der zweiten Runde nicht mehr beteiligen. Es scheinen also diverse Gründe eine Rolle zu spielen, sich weiter zu beteiligen oder nicht. Die Expertise ist nur eine Möglichkeit unter mehreren. Damit unterstreicht dieses Ergebnis generell das Funktionieren der Delphi-Methodik, weil es nicht zu einem systematischen Ausscheiden der Teilnehmer mit geringer Expertise kommt.

Auch die zweite Hypothese kann nur bedingt belegt werden. Während in der ersten Runde kein Beleg erfolgt, kann für die zweite Runde konstatiert werden, dass die Experten mit hohem Kenntnisstand stärker das Selbstbewusstsein haben, zu ihrer Meinung zu stehen, wenn diese vom mittleren Wert abweicht. In diesem Zusammenhang steht auch die Hypothese 3, die in knapp der Hälfte der Themenfelder belegt werden konnte: In diesen Feldern revidierten die Personen mit geringer und mittlerer

Expertise ihre Meinungen stark, während die „wahren“ Experten ihre Einschätzungen kaum veränderten.

Dieses Ergebnis hat für neue Vorausschau-Experimente durchaus Konsequenzen, bedeutet es doch, dass eine getrennte Ausweisung von Ergebnissen sehr wichtig ist, weil ein Bias in den Antworten nicht ausgeschlossen werden kann, dass aber auch die Funktion von „Laien“ oder Personen mit mittlerer Expertise darin bestehen kann, ein Korrektiv darstellen zu können.

Für zwei Drittel der Themenfelder bestätigt sich sogar Hypothese 4: Experten mit besonders hohem Wissensstand scheinen sich dem Delphi-typischen Konvergenzprozess nicht in dem Ausmaß zu unterwerfen wie andere Personenkreise. Dies hat weit reichende Konsequenzen. Wollte man bisher für Delphi-Befragungen immer die Personen mit möglichst hoher Expertise gewinnen, so zeigt dieses Ergebnis, dass dies nicht immer notwendig ist, sondern in Fällen, in denen hochgradige Expertise notwendig ist (z. B. in bestimmten Technologiebereichen) für die Vorausschau besser auf Einfachbefragungen, Interviews oder andere Verfahren zurückgegriffen werden sollte. Wenn das Ziel also Konsensbildung heißt, sollten demnach nicht nur die kenntnisreichen Experten, sondern möglicherweise auch Laien („Betroffene“) oder Personen mit geringerem Kenntnisstand hinzugezogen werden, da diese eher konsensbereit scheinen. Der Inhalt des Themas muss jedoch Berücksichtigung finden, da zu vermuten ist, dass in einem kontroversen Gebiet wie der Energieversorgung auf allen Seiten eine geringe Konsensbereitschaft anzutreffen sein wird.

Lautet das Ziel einer Delphi-Studie jedoch festzustellen, ob es überhaupt einen Konsens im Thema gibt, ist eine der wichtigen Informationen, ob die Expertenwelt im Vergleich zur „Laienwelt“ oder den „Anwendern“ eine andere Meinung vertritt und diese auch in der zweiten Runde nicht ändert. Gleichzeitig bestätigt das Ergebnis das Vorurteil, dass (insbesondere technische) Experten nur schwer von ihrer einmal vorgefassten Meinung abzubringen sind. Dies kann insbesondere für die Forschungspolitik im Bereich der angewandten Forschung eine sehr wichtige Information sein und bestätigt indirekt den Ansatz des Bundesministeriums für Bildung und Forschung, im neuen Futur-Prozess nicht mehr nur Experten zu Rate zu ziehen, sondern in Diskursen auch Personen mit geringerem Vorwissen teilnehmen zu lassen. Auf diese Weise soll eine größere Bedarfsorientierung erreicht werden, damit nicht nur die technisch-naturwissenschaftlichen Angebote bei der Förderung von Einzelprojekten Berücksichtigung finden.

Korrespondenzadresse

Dr. Knut Blind, Dr. Kerstin Cuhls

Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI)

Breslauer Str. 48

76139 Karlsruhe

Tel: 0721-6809-212 bzw. -14

Fax: 0721-6809-260

kb@isi.fhg.de; cu@isi.fhg.de

Literatur

Bardecki, M. J., 1984: Participants' Response to the Delphi Method: An Attitudinal Perspective. *Technical Forecasting and Social Change* 24: 281-292.

Blind, K./Cuhls, K./Grupp, H., 2001: The Influence of Personal Attitudes on the Estimation of the Future Development of Science and Technology: A Factor Analysis Approach, erscheint in: *Technical Forecasting and Social Change* 2001.

Bundesministerium für Forschung und Technologie, BMFT (Hg.), 1993: *Deutscher Delphi-Bericht zur Entwicklung von Wissenschaft und Technik*, Bonn.

Cuhls, K., 2000: Opening up Foresight Processes. *Économies et Sociétés, Série Dynamique technologique et organisation* 5: 21-40.

Cuhls, K./Blind, K./Grupp, H., 2001: *Innovations for our Future*, Heidelberg: Physica Verlag.

Cuhls, K./Blind, K./Grupp, H. (Hg.), 1998: *Delphi '98 Umfrage. Zukunft nachgefragt. Studie zur globalen Entwicklung von Wissenschaft und Technik*, Karlsruhe.

Cuhls, K./Breiner, S./Grupp, H., 1995: *Delphi-Bericht 1995 zur Entwicklung von Wissenschaft und Technik - Mini-Delphi -*, Karlsruhe (auch herausgegeben als BMBF-Broschüre, Bonn 1996).

Dalkey, N.C., 1969a: Analyses From a Group Opinion Study, *Futures*, 2, Jg., 12: 541-551.

Dalkey, N.C., 1969b: *The Delphi Method: An Experimental Study of Group Opinion*, prepared for United States Air Force Project Rand, Santa Monica.

Dalkey, N.C./Brown, B./Cochran, S., 1969: *The Delphi Method, III: Use Of Self Ratings To Improve Group Estimates*, Santa Monica: Rand Corporation Paper RM-6115-PR.

Dalkey, N.C., 1968: *Predicting the Future*, Santa Monica.

Dalkey, N./Helmer, O. 1963: An Experimental Application Of The Delphi-Method To The Use Of Experts. *Journal of the Institute of Management Sciences, Management Science*, 9: 458-467.

Festinger, L., 1978: *Theorie der kognitiven Dissonanz*, Bern/Stuttgart/Wien: Huber.

Grupp, H. (Hg.), 1995: *Der Delphi-Report*, unter Mitarbeit von Breiner, S. und Cuhls, K. Stuttgart: dva-Verlag.

Häder, M., 2000: Subjektiv Sicher und Trotzdem Falsch? Methodische Ergebnisse einer Delphi-Studie zur Zukunft des Mobilfunks. *ZUMA-Nachrichten* 46: 89-116.

National Institute of Science and Technology Policy (NISTEP), 1997: Dai 6 kai Gijutsuyosoku Chôsa, NISTEP Report No. 52, Tokyo.

National Institute of Science and Technology Policy (NISTEP), 2001: Dai 7 kai Gijutsuyosoku Chôsa, NISTEP Report No. 71, Tokyo.