



Vergleichende Risikobewertung

Konzepte, Probleme und Anwendungsmöglichkeiten

H. Schütz, P.M. Wiedemann, W. Hennings, J. Mertens, M. Clauberg



Schriften des Forschungszentrums Jülich
Reihe Umwelt/Environment

Band/Volume 45

Forschungszentrum Jülich GmbH
Programmgruppe Mensch, Umwelt, Technik

Vergleichende Risikobewertung

Konzepte, Probleme und Anwendungsmöglichkeiten

Abschlussbericht zum BfS-Projekt StSch 4217
„Risikobewertung und -management:
Ausarbeitung von Konzepten eines integrierten
und vergleichbaren Risikoansatzes“

Holger Schütz, Peter M. Wiedemann, Wilfried Hennings,
Johannes Mertens, Martin Clauberg

Schriften des Forschungszentrums Jülich
Reihe Umwelt/Environment

Band / Volume 45

ISSN 1433-5530 ISBN 3-89336-350-5

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte Bibliografische Daten sind im Internet
über <<http://dnb.ddb.de>> abrufbar.

Herausgeber
und Vertrieb: Forschungszentrum Jülich GmbH
Zentralbibliothek
D-52425 Jülich
Telefon: 02461 61-5368 · Telefax: 02461 61-6103
e-mail: zb-publikation@fz-juelich.de
Internet: <http://www.fz-juelich.de/zb>

Umschlaggestaltung: Grafische Betriebe, Forschungszentrum Jülich GmbH

Druck: Grafische Betriebe, Forschungszentrum Jülich GmbH

Copyright: Forschungszentrum Jülich 2004

Der Bericht gibt die Meinung des Auftragnehmers wieder und muss nicht mit der Meinung des Auftraggebers (Bundesminister für Umwelt, Naturschutz und Reaktorsicherheit) übereinstimmen.

Schriften des Forschungszentrums Jülich
Reihe Umwelt/Environment Band/Volume 45

ISSN 1433-5530
ISBN 3-89336-350-5

Alle Rechte vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie oder in einem anderen Verfahren) ohne schriftliche Genehmigung des Verlages reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Vergleichende Risikobewertung

Konzepte, Probleme und Anwendungsmöglichkeiten

Abschlussbericht zum BfS-Projekt StSch 4217
„Risikobewertung und –management:
Ausarbeitung von Konzepten eines
integrierten und vergleichenden Risikoansatzes“

Autoren:

H. Schütz, P. Wiedemann, W. Hennings, J. Mertens, M. Clauberg

Forschungszentrum Jülich GmbH
Programmgruppe „Mensch, Umwelt, Technik“ (MUT)

Dezember 2003

Der Bericht gibt die Meinung des Auftragnehmers wieder und muss nicht mit der Meinung
des Auftraggebers (Bundesminister für Umwelt, Naturschutz und Reaktorsicherheit)
übereinstimmen.

Vorwort

Die vorliegende Arbeit ist der Abschlussbericht zum Projekt StSch 4217 „Risikobewertung und –management. Ausarbeitung von Konzepten eines integrierten und vergleichenden Risikoansatzes“, das vom Bundesamt für Strahlenschutz gefördert wurde.

Gegenstand dieses Projekts war die Aufarbeitung, Bewertung und Weiterentwicklung von Konzepten für die vergleichende Risikobewertung und deren Möglichkeiten und Probleme in der praktischen Anwendung.

Die Ergebnisse der Projektarbeiten wurden Ende Oktober 2003 auf einem Workshop mit Expertinnen und Experten aus dem Bereich Risikoabschätzung, Risikobewertung und Risikomanagement diskutiert. An diesem Workshop haben – neben den Autoren des Berichts – teilgenommen:

BDir Andreas Bieber, Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, Referat WA I 4, Bonn; Prof. Dr.-Ing. Ulrich Hauptmanns, Otto-von-Guericke-Universität Magdeburg, Institut für Apparate- und Umwelttechnik, Magdeburg; Prof. Dr. Dietrich Henschler, Würzburg; Prof. Dr. Rolf F. Hertel, Bundesinstitut für Risikobewertung, Berlin; Helmut Jahraus, Bundesamt für Strahlenschutz, Fachbereich Strahlenhygiene, Oberschleissheim (Neuherberg); Prof. Dr. Andreas Kappos, Behörde für Arbeit, Gesundheit und Soziales, HPHC, Hamburg; Dr. Klaus Köberlein, Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) mbH Forschungsgelände, Garching bei München; Dip.-Biol. Rainer Konietzka, Umweltbundesamt, Fachgebiet II 4.7, Toxikologie des Trink- und Badebeckenwassers, Berlin; Prof. Dr. Werner Lilienblum, Niedersächsisches Landesamt für Ökologie, Hildesheim; Dr. Hans-Jochen Luhmann, Wuppertal Institut für Klima, Umwelt, Energie GmbH, Wuppertal; Dr. Odile Mekel, Landesinstitut für den Öffentlichen Gesundheitsdienst NRW, Abteilung Umweltmedizin/ Umweltthygiene, Bielefeld; Dr. Olaf Mosbach-Schulz, Universität Bremen FB 3: Risikoforschung, Umwelt, Gesundheit, Bremen; Dr. Hermann Neus, Behörde für Umwelt und Gesundheit, Abteilung Verbraucherschutz, Hamburg; Prof. Dr. Ortwin Renn, Akademie für Technikfolgenabschätzung in Baden-Württemberg, Stuttgart; Dr. Paul Schön, Otto-von-Guericke-Universität Magdeburg, Institut für Apparate- und Umwelttechnik, Magdeburg; Dr. Michael Schümann, Arbeitsgruppe Epidemiologie AgE der BAGS, Hamburg; PD Dr. Joachim Schüz, Institut für Medizinische Biometrie, Epidemiologie und Informatik, Universität Mainz und Dr. Michael Winter, Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft, Bonn.

Wir danken den Teilnehmerinnen und Teilnehmern für ihre zahlreichen kritischen und konstruktiven Anmerkungen und Verbesserungsvorschlägen, die wir für den vorliegenden Endbericht genutzt haben.

Schließlich danken wir Herrn Helmut Jahraus vom Bundesamt für Strahlenschutz, der dieses Projekt fachlich begleitet hat, für seine Unterstützung.

Die Autoren

Inhalt

Vorwort.....	3
Empfehlungen zu einer integrierten und vergleichenden Risikobewertung (CRA)	7
1. Einleitung.....	15
2. Konzepte und Definitionen.....	21
2.1 Risiko.....	21
2.2 Gefahr / Schädigungspotenzial – Gefährdungspotenzial.....	24
2.3 Ungewissheit und Unsicherheit.....	24
2.4 Risikoabschätzung.....	25
2.5 Risikobewertung	26
2.6 Risikovergleich.....	28
2.7 Risikomanagement	29
2.8 Risikoregulierung	31
2.9 Risikokommunikation	33
3. CRA in der Praxis	35
3.1 Das politische Umfeld der CRA.....	35
3.1.1 CRA in den USA	35
3.1.2 CRA in Europa	37
3.1.3 CRA in Deutschland.....	39
3.2 Projekte als Fallbeispiele	41
3.2.1 US EPA „Unfinished Business“ Projekt	41
3.2.2 Comparative Risk Framework Methodology and Case Study (US-EPA)	48
3.2.3 Projekt „Global Burden of Disease (GBD) Study“	54
3.2.4 ExterneE-Projekt.....	63
3.2.5 Projekt „Ganzheitliche Betrachtung von Energiesystemen (GaBE)“	74
3.2.6 Klassierung krebserzeugender Luftschadstoffe für die TA-Luft-Novelle.....	83
3.2.7 Zusammenfassung.....	84
4. Empirische Grundlagen der CRA.....	87
4.1 Ausgangspunkt für Risikovergleiche	87
4.2 Risikovergleiche als Mittel der Risikokommunikation	88
4.3 Verfahrensprobleme	93
4.4 Verhandlungskonflikte bei der CRA	108

5.	Konzeptioneller Rahmen für eine integrierte und vergleichende Risikobewertung	115
5.1	Methodische Probleme bei einer CRA	115
5.1.1	Problem: Unsicherheit und Ungewissheit	115
5.1.2	Problem: Bewertungskriterien	119
5.1.3	Konsequenzen für die Durchführung einer CRA.....	121
5.2	Methodik vergleichender Risikobewertung	124
5.2.1	Multiattribute Risikobewertung: quantitative und qualitative Ansätze.....	125
5.2.2	Durchführung einer multiattributen Bewertung	127
6.	Umsetzung der CRA in der Praxis	145
6.1	Grenzen der Vergleichbarkeit	145
6.2	Ziele für eine vergleichende Bewertung von Risiken	148
6.2.1	Vorverfahren	148
6.2.2	Risikoabschätzung	152
6.2.3	Risikobewertung	154
6.2.4	Risikomanagement	154
6.3	Beteiligungsmodelle.....	155
6.4	Akteure der CRA unter Beteiligung der Öffentlichkeit.....	156
6.5	Ablauf der CRA.....	159
6.6	Organisation der Kommunikation	160
	Anhang 1: Risikometriken.....	165
	Anhang 2: Multiattribute vergleichende Risikobewertung.....	175
	Anhang 3: Vergleichende Bewertung bei undeutlichen Risiken	189
	Anhang 4: Vergleichende Bewertung von Störfallrisiken.....	197
	Literatur	203
	Glossar.....	225

Empfehlungen zu einer integrierten und vergleichenden Risikobewertung (CRA)

1. Die CRA sollte im Prozess der Risikoregulation fest verankert sein

Die deutsche Risikokommission (2002, S. 22) fordert für den Prozess der Risikoregulation eine systematische Risikobewertung, wenn sie feststellt: „Die Vorgehensweise bei der Prioritätensetzung, bei der Veranlassung einer Risikoabschätzung und auch bei der Auswahl und Terminierung der einzuleitenden Maßnahmen ist uneinheitlich, unsystematisch und wird häufig von einer zufälligen oder interessengeleiteten Problemwahrnehmung durch die allgemeine Öffentlichkeit oder die Medien bestimmt.“

Zwar ist die Kommission beim Einsatz von Risikovergleichen für Kommunikationszwecke vorsichtig; sie hält aber solche Vergleiche für die Risikobewertung für unverzichtbar. Sie äußert sich dazu wie folgt: „Risikovergleiche in der öffentlichen Risikokommunikation sind fragwürdig. Für eine rationale Prioritätensetzung beim Einsatz öffentlicher und privater Mittel zur Risikominderung oder -vermeidung sind sie aber essentiell“ (Risikokommission 2002, S. 58).

Dennoch spielt die vergleichende Risikobewertung (*Comparative Risk Assessment*; CRA) in Deutschland bislang keine besondere Rolle, weder im Bereich der Risikoabschätzung und -bewertung, noch im Bereich des Risikomanagements. Die Gründe dafür sind vielfältig. Zum einen wird auf die methodischen Schwierigkeiten der CRA hingewiesen, und zum anderen gibt es ganz unterschiedliche Einschätzungen des Nutzens von CRA für die Umweltpolitik.

Die Abstinenz bezüglich der CRA hängt zum einen mit der Regeldichte zusammen, zum anderen mit der in Deutschland verfolgten Regulationsphilosophie, Risiken aus verschiedenen Regulationsbereichen (nicht-ionisierende Strahlung, ionisierende Strahlung und Chemikalien, Verkehr etc.) und aus verschiedenen Umweltpolitikbereichen (Naturschutz, Bodenschutz, Gewässerschutz, Klimaschutz und Luftreinerhaltung, Abfallwirtschaft sowie Gefahrenstoffe) nicht gemeinsam zu betrachten. Diese sektorale Betrachtung hat aber auch Nachteile. Nicht zuletzt verunsichert es die Bürger, wenn in einem Bereich Risiken hingenommen werden, die in einem anderen Bereich nicht toleriert werden. Das Verständnis von Risiken leidet auch unter dem in Deutschland verfolgten Grundprinzip, bei technischen Anlagen die Regulierung über Sicherheitsnachweise zu operationalisieren, ohne einen quantitativen Bezug zu Risiken herzustellen.

Insbesondere die separate Behandlung von Einzelsektoren ohne hinreichende Berücksichtigung des übergreifenden Sicherheits- und Risiko-Kontextes gilt es zu ändern. Hierbei kann die CRA, wenn sie in den Prozess der Risikoregulation eingebaut wird, einen Beitrag leisten; die Ergebnisse von Risikoabschätzungen sollten mit Hilfe einer CRA zusammenfassend betrachtet werden. Als Grundlage für ein Rahmenkonzept bei solch übergreifenden Betrachtungen wird häufig das weithin akzeptierte Leitbild der „Nachhaltigen Entwicklung“ vorgeschlagen.

Aus der Zusammenschau von Risiken unterschiedlichen Ursprungs lässt sich ein Risiko-Ranking ableiten, das als ein Input für Entscheidungen über Risikomanagement-Maßnahmen dienen kann. Voraussetzung für ein derartiges Ranking sind harmonisierte Vorgehensweisen bei der Risikoabschätzung. Die Managementmaßnahmen sind das wesentliche Ergebnis eines Risikoregulierungs-Prozesses. Die Akzeptanz dieser Ergebnisse ist umso höher, je umfassender alle Interessengruppen („Stakeholder“) am Prozess beteiligt waren.

2. Bei der Planung von CRAs müssen unterschiedliche Typen von Risiken unterschieden werden

In der öffentlichen Diskussion um Risiken finden sehr unterschiedliche Typen von Risiken Beachtung. Breiten Raum nehmen Risiken ein, die aus der Exposition mit Schadstoffen resultieren. Hier geht es um die Frage, wie wahrscheinlich und wie schwerwiegend mögliche adverse Gesundheitseffekte sind. Ein anderes Risikothema betrifft die Störfallrisiken technischer Anlagen. Entsprechende Analysen konzentrieren sich daher auf die Anlagentechnik und das Zustandekommen von Störfällen. Dabei können auch sehr seltene, außerhalb des direkten Erfahrungsbereichs liegende Störfälle von Bedeutung sein, wenn sie mit besonders gravierenden Folgen verbunden sind. Und schließlich werden auch „Risiken“ diskutiert, für die noch gar nicht geklärt ist, ob sie überhaupt Risiken sind. Die Frage ist, ob bestimmte Substanzen bzw. Expositionssituationen überhaupt zu adversen Gesundheitseffekten führen können. Gerade hier ist die wissenschaftliche und in der Folge auch die öffentliche Diskussion oft besonders kontrovers.

Wegen dieser grundlegenden Unterschiede plädieren wir dafür, vergleichende Risikobewertungen nur innerhalb dieser drei typisierenden Gruppen vorzunehmen. Außerdem können Risikovergleiche grundlegende Unterschiede im Wissen über die Risiken (also das Ausmaß von Unsicherheiten) verdecken; wir empfehlen daher, erhebliche Wissensunterschiede klar zu benennen und gegebenenfalls auf eine vergleichende Bewertung zu verzichten.

3. Die CRA kann in allen Schritten der Risikoregulierung wesentliche Informationen liefern

Die einzelnen Schritte im Prozess der Risikoregulierung haben spezifische Zielsetzungen. In allen Schritten kann die CRA wesentliche Informationen liefern; dabei wird deutlich, dass auch einzelne Bestandteile einer CRA – im Sinne risikobezogener Vergleiche – für die jeweiligen Zielsetzungen nützlich sein können.

Im *Vorverfahren* geht es um eine Abschätzung neuartiger Risikopotenziale, um das öffentliche Mobilisierungspotenzial neuer Technologien sowie um eine erste Einschätzung der Dringlichkeit einer Risikobewertung. Der Nutzen einer CRA liegt hier im Vergleich neuartiger Technologiefelder, im Vergleich der öffentlichen Risikowahrnehmung in unterschiedlichen Fällen sowie in einem Vergleich von Stoffen mit Gefährdungspotenzial hinsichtlich Emissionsdaten, Expositionscharakteristika und Toxizität.

Bei der *Risikoabschätzung* steht die Bewertung von Evidenzen im Fokus. Da hier häufig wissenschaftliche Kontroversen bestehen, kann der Vergleich unterschiedlicher Evidenzbewertungen, etwa unter Nutzung erprobter Leitlinien und Evidenzkategorien, wesentlich zur Problemlösung beitragen. Als Verfahren ist hier die *vergleichende Bewertung undeutlicher Risiken (BUR)* zu nennen, bei der Gemeinsamkeiten und Unterschiede, die sich zwischen verschiedenen Experten bzw. Expertengruppen in der Evidenzeinschätzung für das Vorliegen einer Gefährdung zeigen, verglichen werden.

Im Regulierungsschritt der *Risikobewertung* bestehen vier verschiedene Möglichkeiten für risikobezogene Vergleiche: (1) die Bewertung der Potenz einer Noxe, (2) die Bewertung der Exposition mit einer Noxe, (3) die Bewertung der Vulnerabilität von Populationen und (4) die vergleichende Bewertung verschiedener Risiken.

Die zentrale Variante ist die vergleichende Bewertung verschiedener Risiken. Zu nennen ist hier insbesondere die *multiattribute vergleichende Risikobewertung (MRB)*, bei der bekannte Risiken auf verschiedenen Bewertungsdimensionen (Attributen) von einem oder mehreren Bewertern (Stakeholdern) verglichen werden. Dies entspricht im wesentlichen den klassischen Ansätzen der CRA. Ergebnis einer solchen MRB ist eine Rangfolge von Risiken auf der Basis einer (möglichst quantitativen) Abschätzung der Gesundheitsfolgen und deren Bewertung.

Auch für das *Risikomanagement* bieten sich eine Reihe risikobezogener Vergleiche an, etwa bei der Auswahl unter technischen Optionen oder bei der Suche nach Standorten für Anlagen mit Risikopotenzial. In derartige Vergleiche werden auch Kosten- und Nutzenaspekte einbezogen.

4. Die CRA als Kombination von naturwissenschaftlich basierten Risikoabschätzungen und Werturteilen erfordert die Kooperation von Experten und gesellschaftlichen Stakeholdern

Experten – als Autoren naturwissenschaftlich basierter Risikoabschätzungen – und die Öffentlichkeit verstehen und interpretieren Risikoabschätzungen häufig unterschiedlich. Ein wesentliches Problem aus Sicht der Experten besteht darin, dass sich Endergebnisse von Analysen verselbständigen und prinzipielle Einschränkungen, methodische Unsicherheiten und Gültigkeitsbereiche der Öffentlichkeit nicht bewußt werden.

Grundlegend für das Verständnis und die Rolle von Risikoabschätzungen ist darüber hinaus der Risikobegriff selbst. Es hat sich gezeigt, dass der technische Begriff der Experten aus Sicht der Öffentlichkeit sehr eng ist und nur einen Teil der Aspekte und Werte umfaßt, die die Allgemeinheit – vertreten in ihrer Breite durch gesellschaftliche Stakeholder – bei der Beurteilung von Risiken für wichtig hält. Schon die aus dem Versicherungswesen stammende Gleichgewichtigkeit der Risikokomponenten Häufigkeit und Schaden ist – auch in Expertenkreisen – umstritten. Beide Faktoren werden von Laien (im Sinne von Nicht-Risikoexperten) mit individuellen Präferenzen getrennt betrachtet; insbesondere wird eine Begrenzung des maximal möglichen Schadens als eigenständiger Wert gesehen und zunehmend gefordert.

Hinzu kommt, dass der den üblichen Risikoabschätzungen zugrunde liegende Risikobegriff in der Regel nur wenige Schadensdimensionen umfasst, oft nur Todesfälle und Gesundheitsschäden, seltener auch Vermögensverluste. Die Öffentlichkeit sieht meist noch andere Dimensionen und Begleitumstände von Risiken, wie etwa den Zeitraum schädigender Wirkungen, die räumliche Schadensausdehnung, Unausweichlichkeit des Risikos, Evakuierungen, Umsiedlungen und andere ausgeprägt soziale Aspekte. Hier muss jeweils im Rahmen konkreter CRA-Verfahren zwischen Experten einerseits und Stakeholdern andererseits geklärt werden, welche Bewertungsaspekte bei der CRA berücksichtigt werden sollen.

Eine konsensfähige und erfolgreiche CRA kann somit nur auf einer Kooperation von Experten (für Risikoermittlung und -management) und gesellschaftlichen Stakeholdern als Repräsentanten der öffentlichen Meinung beruhen.

5. Die CRA kann auf eine Reihe von ausgearbeiteten und bewährten Instrumenten zurückgreifen

Die vergleichende Risikobewertung kann als multiattributes Bewertungsverfahren konzeptualisiert werden. Sie kann damit auf dem Ansatz der multiattributen Entscheidungsfindung aufbauen, der ein theoretisch fundiertes und strukturiertes Vorgehen in überschaubaren Einzelschritten ermöglicht. Für die einzelnen Schritte (Problemstrukturierung, Attributstrukturierung und -gewichtung, Sensitivitätsanalyse etc.) gibt es eine Reihe von praktisch erprobten Verfahren.

Eine Stärke dieses Ansatzes ist, dass er das Explizitmachen von Vorannahmen und Wertungen fördert und deswegen eine transparente vergleichende Risikobewertung unterstützt. Deshalb ist dieser Ansatz auch gerade für CRA-Verfahren geeignet, an denen verschiedene Bewerter(gruppen) – Experten und Stakeholder – teilnehmen. Für die Organisation und Durchführung solcher Multistakeholderbewertungen gibt es eine Vielzahl von Beteiligungsverfahren und Modellen, die im Detail ausgearbeitet sind.

Darüber hinaus kann – vor allem in Bezug auf die praktische Organisation und Durchführung von CRAs – auf einen reichhaltigen Erfahrungsschatz zurückgegriffen werden. Insbesondere in den USA wurden bereits mehr als 100 CRAs durchgeführt, für die auch Zusammenfassungen der dabei gemachten Erfahrungen verfügbar sind.

6. CRA erfordert den Einbezug bislang unbeachteter wissenschaftlicher Erkenntnisse

Die naturwissenschaftlichen Grundlagen der Risikoabschätzung sind eine notwendige Voraussetzung der CRA, aber das Prozessmanagement entscheidet über deren Erfolg. Dazu ist Verfahrenskompetenz nötig. Es sind Erkenntnisse aus den Verhaltenswissenschaften, den Sozialwissenschaften und insbesondere aus der Entscheidungsanalyse nutzbar zu machen, um Prozesse der Kooperation zwischen Experten und gesellschaftlichen Stakeholdern und der kooperativen Entscheidungsfindung vorausschauend zu planen und rational zu gestalten.

Das Prozessmanagement einer CRA beginnt mit der Auswahl der Teilnehmer und der Festlegung der Ziele für die CRA. Anschließend müssen die zu bewertenden Risiken ausgewählt und kategorisiert werden; es müssen Bewertungskriterien bestimmt und ggf. gewichtet und Beurteilungen der Risiken auf diesen Attributen vorgenommen werden. Aus diesen Bewertungen ergibt sich dann das Ranking.

Diese Prozesse sind – soweit es geht – einvernehmlich zu gestalten. Es wäre jedoch naiv, hier divergente Positionen, Interessen und Machtpotenziale auszublenden. Die CRA ist deshalb nicht von Verhandlungsprozessen zu trennen.

Eines der wichtigsten Merkmale erfolgreicher kollektiver Bewertungs- und Abwägungsprozesse ist es, die Perspektive des Partners einnehmen zu können, d.h. sich in die Belange, Zwänge, Interessen und Ziele des Gegenübers hineinversetzen zu können. Gerade für eine CRA, an der Vertreter unterschiedlicher gesellschaftlicher Gruppen teilnehmen, gelten nicht nur die Kriterien der Effizienz und Rationalität guten Handelns, sondern vor allem der Grundsatz: Der Umgang miteinander sollte die gegenseitigen Beziehungen verbessern oder zumindest nicht beeinträchtigen.

7. CRA braucht wissenschaftliche Expertise und muss weiterentwickelt werden

Eine wichtige Voraussetzung für eine CRA ist die Harmonisierung der Verfahren zur Abschätzung und zur Bewertung von Risiken in unterschiedlichen Bereichen. Dabei sollte die Charakterisierung der Risiken ein eigenständiger Schritt sein. Der gesamte Prozess erfordert wissenschaftliche Expertise und muss noch weiterentwickelt werden, um effizient anwendbar zu sein.

Die *Risikoabschätzung* beginnt mit der *Identifikation der Gefährdungspotenziale*. Hier sind drei Problemfelder von Bedeutung: (1) der erforderliche Evidenzgrad zur Begründung einer Kausalität zwischen den fraglichen Ursachen und Wirkungen, (2) die Einstufung einer Wirkung als advers (bzw. unerwünscht) sowie (3) eine mögliche Exposition von Schutzgütern.

Ein wesentliches Problem ist die Evidenzbewertung. Hier müssen Kategorien („wie stark ist der Verdacht?“) entwickelt werden, die in eindeutiger Weise operationalisiert sind. Worst-case-Szenarien sind wegen ihrer Beliebigkeitskomponente – es gibt immer noch gravierendere Szenarien – keine tragfähige Basis zur Abschätzung und Bewertung von Risiken. *Dosis-Wirkungs-Abschätzungen* sollten nach standardisierten und harmonisierten Verfahren festgelegt werden. Mit Blick auf die Relevanz von Gefährdungspotenzialen kommt der *Expositionsabschätzung* eine große Bedeutung zu.

Mit Risikocharakterisierung ist die Zusammenfassung der Ergebnisse von Gefährdungspotenzial-Identifikation, Dosis-Wirkungs-Abschätzung sowie Expositionsabschätzung gemeint. Als Aufbereitung von Informationen ist sie auch eine inhaltliche Voraussetzung für vergleichende Analysen. Daher sollte die Risikocharakterisierung als ein besonderer Verfahrensschritt ausgewiesen werden.

Die Risikobewertung bildet das Bindeglied zwischen der vorwiegend naturwissenschaftlichen Risikoabschätzung und einer gesellschaftlich-politisch ausgerichteten

Beurteilung des Risikos. Die Basis zur Beurteilung quantifizierbarer Risiken kann ein durch gesellschaftliche Diskussion erzielter Konsens über tolerierbare Risiken sein. Da sich viele Betrachtungen jedoch auf (noch) nicht quantifizierbare Risiken erstrecken, müssen – wiederum auf der Basis wissenschaftlicher Expertise – Kriterien zur Differenzierung zwischen Gefahrenabwehr (mit eindeutigem Regulierungsbedarf, meist über Grenzwerte) und Vorsorge entwickelt werden. Weiterhin ist fraglich, welcher Risikoverdacht zur Anwendung des Vorsorgeprinzips ausreicht, um auf diese Weise eine erste Basis für einen Vergleich mit anderen Risiken zu erhalten. Hier fehlen jedoch noch Qualitätsanforderungen an Einzelstudien sowie an die Erstellung des wissenschaftlichen Gesamtbildes zum Risikoverdacht.

Ein weiteres Problemfeld betrifft Risikoeinschätzungen zu neuen, sich entwickelnden Technologien (etwa die Nano-Technologie oder die Gentechnik). Hier sind spezielle – auch prognostisch orientierte – Verfahren zur Früherkennung von Gefährdungspotenzialen und deren Relevanz zu entwickeln. Dabei sollte sichergestellt sein, dass solche Verfahren konsistent angelegt sind und auch Vergleiche unterschiedlicher Technologien ermöglichen.

8. CRA benötigt ein Risikokommunikationsprogramm

Eine wichtige Voraussetzung für den Erfolg eines CRA-Projekts ist eine gute Kommunikation. Wesentlich sind hier sowohl eindeutige Zielsetzungen, explizite und verständliche Definitionen der zu vergleichenden Risiken sowie der Bewertungskriterien und Messgrößen als auch eine verständliche Risikocharakterisierung.

Jede CRA setzt voraus, dass die Beteiligten ausreichend informiert sind, um kompetent an der vergleichenden Bewertung der Risiken mitwirken zu können. Weiterhin sind aber auch prozessuale Aspekte bei der Kommunikation zu beachten. Dabei geht es vor allem um die Umsetzung von Fairness, Kompetenz und Vertrauen.

Für die Kommunikation von Ergebnissen der CRA in die Öffentlichkeit sind die folgenden Aspekte zu unterscheiden (1) Akzeptanz von CRA-Verfahren, (2) Verdeutlichung des Beitrags der CRA für das Verständnis der Größe von Risiken und (3) Akzeptanz der Ergebnisse der CRA.

Akzeptanz findet eine CRA nur dann, wenn es gelingt, im Sinne der Risikomündigkeit Verständnis für ein rationales Abwägen von Risiken zu schaffen. Die Verdeutlichung des Beitrags der CRA für das Verständnis von Risiken hängt davon ab, ob und wie vorhandene Informationsdefizite ausgeglichen werden. Dabei spielt der Austausch von Standpunkten und Ansichten eine wichtige Rolle. Dies ist auch wesentlich für die Akzeptanz von CRA-Ergebnissen.

9. In Deutschland sollte ein Kompetenzzentrum für CRA aufgebaut werden

Wegen der uneinheitlichen Regulierung von Gesundheits- und Umweltrisiken haben im Oktober 2000 die Bundesgesundheitsministerin und der Bundesumweltminister im Rahmen des Aktionsprogramms Umwelt und Gesundheit (APUG) die Risikokommission gegründet. Die Kommission erhielt den Auftrag, Vorschläge und Empfehlungen zu einer Neuordnung der Verfahren und Strukturen zur Risikobewertung und Stan-

dardsetzung im gesundheitlichen Umweltschutz zu erarbeiten. In ihrem Abschlussbericht empfiehlt die Kommission unter anderem die Einrichtung eines hochrangigen, interdisziplinär zusammengesetzten Risikorats, dessen Aufgabe die Harmonisierung und Koordination der Risikoabschätzung ist.

Vor diesem Hintergrund sollte in Deutschland ein Kompetenzzentrum für CRA aufgebaut werden. Dessen Aufgaben sollen hier nur skizziert werden; die wichtigsten Punkte sind

- eine Sammlung und umfassende Dokumentation aller relevanten CRA-Projekte (dabei kann der vorliegende Bericht als Ausgangspunkt verwendet werden);
- der Aufbau eines Arsenal an CRA-Methoden einschließlich einer bewertenden Zusammenstellung von Leistungsspektrum, Defiziten sowie Vor- und Nachteilen;
- eine Unterstützung bei der Durchführung von CRA-Vorhaben einschließlich des Angebots grundlegender Kurse und Trainingseinheiten;
- Initiativen und Planungen für wissenschaftliche Weiterentwicklungen.

Die Anwendung verfügbarer CRA-Methoden auf konkrete Fragestellungen ist die beste Möglichkeit, ihren Entwicklungsstand und praktischen Nutzen zu beurteilen. Daher sollten möglichst bald mittels einer Test-CRA praktische Erfahrungen gesammelt werden, um zu klären, welche Basis für Weiterentwicklungen existiert und welche Ansätze hierfür erfolgversprechend sind.

1. Einleitung

Entscheidungen über den Umgang mit Gesundheits- und Umweltrisiken müssen in den unterschiedlichsten Zusammenhängen getroffen werden. Diese reichen von Entscheidungen im privaten Bereich über administrative bis hin zu Entscheidungen auf einer globalen politischen Ebene. Während im privaten Bereich solche risikobezogenen Entscheidungen (z.B. über die Ausübung riskanter Sportarten oder den Abschluss einer Unfallversicherung) meist ohne eine genaue Analyse der Risiken und der Vor- und Nachteile der verfügbaren Handlungsalternativen getroffen werden, sollten administrative und politische risikobezogene Entscheidungen auf der Basis einer eingehenden Analyse der Risiken und der Entscheidungsoptionen erfolgen.

Der Blick auf einige aktuelle Probleme zeigt, dass das Spektrum risikobezogener Entscheidungen in der Gesellschaft sehr groß ist:

- Muss die Verwendung eines bestimmten Stoffes reguliert werden (z.B. durch Grenzwertsetzung)?
- Kann ein als schädigend erkannter Stoff durch einen anderen ersetzt werden (z.B. Substitution von Chemikalien)?
- Wo sollen in der Gesundheitspolitik die Schwerpunkte gesetzt werden (z.B. Krebsvorsorge, AIDS-Vorsorge)?
- Welche Energiesysteme sollen genutzt werden (z.B. fossile Energieträger, regenerative Energie)?
- Welche landwirtschaftliche Produktionsweise sollte vorzugsweise genutzt werden (konventionelle Landwirtschaft, grüne Gentechnik)?

Aufgaben der Risikoabschätzung, der Risikobewertung und des Risikomanagements werden in Deutschland von unterschiedlichen Institutionen wahrgenommen.¹ So sind auf der Bundesebene verschiedene Ministerien (z.B. BMU, BVEL, BfG) verantwortlich, die dabei auf den Sachverstand der ihnen zugeordneten Bundesoberbehörden (z.B. BfS, UBA, RKI), von privatrechtlichen Einrichtungen (z.B. Senatskommission zur Bewertung Maximaler Arbeitsplatzkonzentrationen der DFG) und von Expertengremien (z.B. SSK) zurückgreifen. Auch auf der Länderebene sind zahlreiche Gremien involviert (z.B. LAI, LAUG). Diese Vielfalt von Zuständigkeiten und Bearbeitungsebenen führt teilweise zu einer unübersichtlichen Situation. In ihrem Abschlussbericht zur Neuordnung der Verfahren und Strukturen zur Risikobewertung und Standardsetzung im gesundheitlichen Umweltschutz der Bundesrepublik Deutschland kommt die Risikokommission zu der Einschätzung:

Das alles führt dazu, dass der Umgang mit möglichen Gesundheitsgefahren meist von einer zufälligen oder interessengeleiteten Problemwahrnehmung durch die Öffentlichkeit oder die Medien bestimmt wird. Dass ein effektives Krisenmanagement fehlt, wurde in der jüngeren Vergangenheit an einer Vielzahl von Risikokonflikten deutlich. Eine Beteiligung der Öffentlichkeit und von Interessengruppen an der Risikoregulierung findet nur ausnahmsweise statt; sie ist unsystematisch und schwerfällig. Zögerliche und widersprüchliche Regulierungen durch die beteiligten staatlichen Stellen, selektive Information der Betroffenen und Kommunikationsdefizite haben zur

¹ Siehe dazu die Analyse der Risikokommission (2003).

Folge, dass auf der einen Seite relativ unbedeutende Risiken einen erheblichen Platz in der öffentlichen Wahrnehmung einnehmen, während andererseits z.T. schwerwiegende Risiken unterschätzt oder sogar verdrängt werden.

(Risikokommission 2003, 20)

Diese Einschätzung macht deutlich, dass vergleichende Risikobewertungen bei der Regulierung von Risiken eine wichtige Rolle spielen können.

Vergleiche von Risiken sind an und für sich nichts Neues. Denn jede Risikobewertung impliziert immer einen Vergleich: entweder einen Vergleich des betrachteten Risikos in bezug auf einen Umweltstandard (Grenzwert, Richtwert etc.) oder den Vergleich mit anderen Risiken.

Beispiele für den Vergleich mit Umweltstandards sind die Maximale Arbeitsplatzkonzentration (MAK-Werte), die akzeptable tägliche Aufnahmemenge für Substanzen (*Acceptable Daily Intake*; ADI-Werte) oder technische Richtkonzentrationen (TRK-Werte). Hier werden Stoffemissionen oder Expositionssituationen mit Standards verglichen, um über die Notwendigkeit von Risikomanagementmaßnahmen zu entscheiden. Die Zahl der in Deutschland geltenden Umweltstandards ist sehr groß und in ihrer Begründung und Zielsetzung sehr heterogen.² Das spricht nicht gegen diese Vergleichspraxis, verweist aber darauf, dass eine solche Bewertungspraxis intransparent und auch inkonsistent sein kann.

Neben den Vergleichen mit Umweltstandards finden sich vergleichende Bewertungen verschiedener Risiken mit dem Ziel des Risikorankings. In den USA ist dieser Typ von vergleichender Risikobewertung unter dem Namen *Comparative Risk Assessment* (CRA) schon seit Ende der 80er Jahre diskutiert und auch praktisch angewendet worden. Das bekannteste und wohl auch ambitionierteste CRA-Projekt ist das von der amerikanischen Umweltschutzbehörde EPA durchgeführte Ranking von Umweltrisiken. Darüber hinaus gibt es in den USA zahlreiche andere CRA-Projekte, die meist für bestimmte Bundesstaaten oder auf der lokalen Ebene durchgeführt wurden bzw. werden. Zielstellung dieser Ansätze war und ist, die Grundlage für eine rationalere Prioritätensetzung in der Gesundheits- und Umweltpolitik zu liefern. Auf der Basis wissenschaftlicher Risikoabschätzungen sollen Risiken aus ganz unterschiedlichen Bereichen bewertet und verglichen werden, um so zu einer Einschätzung ihrer relativen Bedeutsamkeit zu kommen.

Anders als in den USA sind vergleichende Risikobewertungen in Deutschland (und Europa) bislang nur selten genutzt worden – weder im Bereich der Risikoabschätzung und -bewertung, noch im Bereich des Risikomanagements.³ Die Gründe dafür sind vielfältig: Zum einen wird auf die methodischen Schwierigkeiten der CRA hinge-

² In seiner Analyse von Umweltstandards erfasst der Sachverständigenrat für Umweltfragen 154 Listen von Umweltstandards mit insgesamt ca. 10.000 einzelnen Umweltstandards; siehe SRU (1996).

³ Morgenstern, Shih & Sessions (2000) nennen in ihrem internationalen Vergleich von CRA-Methoden und Ergebnissen eine Reihe von Entwicklungsländern, finden jedoch in Europa bzw. den OECD Staaten keine Anwendungen der CRA mit dem Ziel eines Risikorankings. Vergleichende Risikobewertungen für spezifische Themen werden derzeit in einigen Forschungsprojekten untersucht, z.B. im GaBE Projekt zu Risiken von Energieversorgungssystemen; vgl. Hirschberg, Spiekernan & Dones (1998). Siehe dazu Kapitel 3.2.5 dieses Berichts.

wiesen, zum anderen gibt es ganz unterschiedliche Einschätzungen des Nutzens von CRA für die Umweltpolitik.

Die Abstinenz bezüglich der CRA hängt wohl auch mit der Regeldichte zusammen und mit der in Deutschland verfolgten Regulationsphilosophie, welche Risiken aus verschiedenen Regulationsbereichen (nicht-ionisierende Strahlung, ionisierende Strahlung und Chemikalien, Verkehr etc.) und aus verschiedenen Umweltpolitikbereichen (Naturschutz, Bodenschutz, Gewässerschutz, Klimaschutz und Luftreinerhaltung, Abfallwirtschaft sowie Gefahrenstoffe) nicht gemeinsam betrachtet. Diese sektorale Betrachtung hat aber auch Nachteile. Nicht zuletzt verunsichert sie die Bürger, wenn in einem Bereich Risiken hingenommen werden, die in einem anderen Bereich nicht toleriert werden. Dies ist auch die Folge des in Deutschland verfolgten Grundprinzips, bei technischen Anlagen die Regulierung über Sicherheitsnachweise zu operationalisieren, ohne einen quantitativen Bezug zu Risiken herzustellen. Insbesondere die separate Behandlung von Einzelsektoren ohne hinreichende Berücksichtigung des übergreifenden Sicherheits- und Risiko-Kontextes gilt es zu ändern. Hierbei kann die CRA, wenn sie in den Prozess der Risikoregulation eingebaut wird, einen Beitrag leisten.

Der Anspruch auf eine rationale Begründung gesundheits- und umweltbezogener Entscheidungen mit Hilfe vergleichender Risikobewertung hat – vor allem in den USA – zu einer kritischen Diskussion geführt, in der eine Reihe von Argumenten gegen die CRA vorgebracht wurden. Ein Teil dieser Kritiken bezieht sich auf die Nutzung der CRA im Kontext der US-amerikanischen Regulierung von Gesundheits- und Umwelttrisiken (vor allem durch die EPA). So wird etwa argumentiert, dass mit der Anwendung der CRA eine „Tyrannei des Rationalen“ einher gehen würde, welche die Flexibilität des bestehenden US-amerikanischen Regulierungssystems einschränken würde, die gerade daraus resultiere, dass Risikoabschätzung und Risikomanagement nicht von politischen Prozessen getrennt seien (Silbergeld 1995, 422).

Andere Kritikpunkte sind grundsätzlicher Natur. Neben dem Vorwurf der Trivialisierung von Risiken – insbesondere von Umwelttrisiken, denen im Vergleich zu sogenannten Lebensstiltrisiken (z.B. Rauchen, Ernährungsgewohnheiten) häufig nur ein geringer Anteil an den Risiken für die menschliche Gesundheit zugeschrieben wird – werden vor allem drei Argumente gegen die CRA vorgebracht (z.B. Hornstein 1992; Shrader-Frechette 1995; Silbergeld 1995):

Das erste Argument behauptet, dass Risiken häufig qualitativ so unterschiedlich seien – etwa hinsichtlich ihrer Entstehung oder der Art der Schädigung –, dass man sie nicht sinnvoll vergleichen könne. Dies ist das bekannte „Man kann Äpfel nicht mit Birnen vergleichen“-Problem. Tatsächlich aber kann man Äpfel und Birnen sehr wohl miteinander vergleichen – etwa hinsichtlich ihres Geschmacks, Vitamingehalts oder Preises. Voraussetzung ist nur, dass es mindestens eine gemeinsame Eigenschaft der Vergleichsobjekte gibt, in Bezug auf die hin verglichen werden kann. Bedingung ist aber auch, dass diese Eigenschaft als bedeutsam für den Zweck des Vergleichs angesehen wird. Und in dieser Hinsicht macht das Inkommensurabilitäts-Argument auf einen wichtigen Aspekt für die vergleichende Bewertung von Risiken aufmerksam: Um sinnvoll zu sein, müssen solche Vergleiche die Aspekte der zu betrachtenden Risiken berücksichtigen, die für die Bewerter relevant sind. Man kann also sehr wohl das Risiko des Rauchens mit dem des Autofahrens vergleichen, wenn man etwa den statistischen Erwartungswert von Todesfällen, die pro Jahr durch Rauchen

und Autofahren verursacht werden, als bewertungsrelevantes Vergleichsmerkmal akzeptiert.

Das zweite Argument verweist auf die großen Lücken in unserem Wissen über Gesundheits- und Umweltrisiken und folgert daraus, dass vergleichende Bewertungen dieser Risiken nicht möglich oder zumindest nicht sinnvoll sind. In der Tat sind Risikoabschätzungen häufig mit großen Unsicherheiten behaftet, die sowohl im theoretischen Verständnis wie in den verfügbaren empirischen Daten begründet sind. Dies ist ein gravierendes Problem – allerdings nicht nur für eine CRA, sondern für jede Form der Bewertung und Regulation von Risiken. Trotzdem müssen Entscheidungen über den Umgang mit Risiken getroffen werden – und sie werden auch getroffen. Bei jeder Form der Bewertung und Regulation von Risiken müssen die Unsicherheiten aber angemessen berücksichtigt werden. Wie dies im Rahmen einer CRA geschehen kann, ist ein zentrales Thema des vorliegenden Berichts.

Das dritte Argument schließlich sieht in der CRA ein Hindernis für Partizipationsmöglichkeiten der Öffentlichkeit bei Entscheidungen über Gesundheits- und Umweltrisiken. In der Fundierung solcher Entscheidungen durch vergleichende Risikobewertungen wird die Gefahr einer Expertokratie gesehen. Denn die Risikoabschätzungen, die für eine CRA herangezogen werden müssen, sind häufig sehr kompliziert und für Nicht-Experten kaum nachzuvollziehen. Und auch die Durchführung einer vergleichenden Risikobewertung erfordert methodisches Know-How. Aber gerade eine CRA bietet die Möglichkeit, wissenschaftliche Expertise einerseits und Wertungen gesellschaftlicher Gruppen (Stakeholder) andererseits systematisch und in transparenter Weise miteinander zu verbinden. Den Experten kommt dabei die Aufgabe zu, Informationen über Art, Ausmaß und Wahrscheinlichkeit von Risiken (sowie die dabei vorhandenen Unsicherheiten) zu bereitzustellen, während die Bewertung dieser Information, wie auch die Auswahl und Gewichtung von Bewertungskriterien, durch die gesellschaftlichen Gruppen (Stakeholder) erfolgen kann.

Streffler et al. (2000, 15) weisen darauf hin, dass man drei Arten von Wissen haben muss, um Risiken in rationaler Weise vergleichen zu können: *Erstens* müssen Art und Wahrscheinlichkeit von Handlungsfolgen abgeschätzt werden können. Das heißt, man muss über Wissen zu Kausal- oder zumindest Konditionalverhältnissen verfügen. *Zweitens* muss man wissen, welche Handlungsfolgen die mit dem Risikovergleich betrauten gesellschaftliche Akteure (z.B. Experten, politisch Verantwortliche, von den Risiken Betroffene) als schädlich einschätzen. Dies mag im Falle offensichtlicher Schäden (Tod, Krankheit) unstrittig sein, bei anderen Folgen (z.B. Befindlichkeitsstörungen oder erst in ferner Zukunft zu erwartende Schäden) können sich aber durchaus unterschiedliche Beurteilungen zum „Schadensstatus“ zwischen verschiedenen Akteuren ergeben. *Drittens* ist Wissen darüber erforderlich, wie man Risikovergleiche durchführt. Dabei geht es nicht um die Ermittlung von Risikowerten, d.h. um die Risikoabschätzung, sondern es geht um die Methodik vergleichender Risikobewertung und deren Prämissen.

Vor allem um diese dritte Art von Wissen geht es in dem vorliegenden Abschlussbericht zum BfS-Projekt „Ausarbeitung von Konzepten eines integrierten und vergleichenden Risikoansatzes“. Ziel des Projekts war es, die methodischen Grundlagen für Verfahren der integrierten und vergleichenden Risikobewertung zu erarbeiten, die den verschiedenen Anforderungen, die sich im Rahmen der Risikoregulation für vergleichende Risikobewertungen ergeben können, gerecht werden. Wir beziehen uns

dabei vor allem auf das Problem der vergleichenden Bewertung von Risiken mit dem Ziel des Risikorankings. Es geht nicht um die vergleichende Bewertung von Handlungsoptionen im Rahmen des Risikomanagements mit dem Ziel der Priorisierung – dafür sind weitere Aspekte einzubeziehen, die jenseits der Risikoproblematik liegen (z.B. die Kosteneffizienz von Maßnahmen). Die vergleichende Risikobewertung liefert aber die Grundlage für diese weitergehenden Bewertungen.

Der Begriff „integriert“ verweist darauf, dass Risiko – und damit auch die vergleichende Risikobewertung – ein multidimensionales Konzept ist, in das verschiedene Aspekte eingehen, die explizit behandelt werden müssen. Risiko kann sich auf verschiedene Gefährdungen beziehen, die unterschiedliche Folgen (für die Gesundheit, Umwelt, Ökonomie etc.) haben können und für deren Bewertung unterschiedliche Kriterien herangezogen werden können. Die Risikoabschätzung erfolgt durch unterschiedliche Experten(gruppen), die unter Umständen zu divergierenden Einschätzungen kommen. Solche Einschätzungen werden von verschiedenen Stakeholdern (möglicherweise) unterschiedlich bewertet.

Diese Ausgangslage macht es für die Konzeption einer CRA erforderlich, neben den naturwissenschaftlichen Grundlagen der Risikoabschätzung auch das Problem der Umsetzung von CRA-Verfahren zu berücksichtigen. Damit kommen sowohl methodische Probleme als auch Fragen des Prozessmanagements ins Blickfeld. Zur wissenschaftlichen Fundierung der für das Prozessmanagement erforderlichen Verfahrenskompetenz können Erkenntnisse aus den Verhaltenswissenschaften, den Sozialwissenschaften und der Entscheidungsanalyse nutzbar gemacht werden.

Zielsetzungen der Arbeit

Ziel der Studie ist es, vorliegende Ansätze zur CRA zu bewerten sowie Möglichkeiten zu erkunden, derartige Vergleiche zu verbessern, Partizipationschancen zu erörtern und derartige Verfahren so zu gestalten, dass sie bei politischen Entscheidungen über gesundheitsbezogene Umweltrisiken mehr Akzeptanz finden.

Zu klären sind die folgenden Fragen:

- Welche Vorarbeiten zu vergleichenden Risikobewertungen gibt es? Welche Lücken und welche Probleme weisen diese auf?
- Welche Typen von vergleichenden Risikobewertungen lassen sich unterscheiden und welche sind - unter den gegebenen administrativen Randbedingungen - sinnvoll und sollten ausgearbeitet werden?
- Auf welche formale Struktur kann eine vergleichende Risikobewertung aufbauen?
- Wie sind die einzelnen Schritte einer vergleichenden Risikobewertung zu gestalten und welche Hinweise ergeben sich aus der umfangreichen Literatur der sozialwissenschaftlichen Risikoforschung?
- Welche besonderen Anforderungen stellt eine integrative vergleichende Risikobewertung, die insbesondere verschiedene Bewerter aus Wissenschaft und Gesellschaft einbezieht?
- Welche Erfolgchancen haben vergleichende Risikobewertungen und wo sind ihre Grenzen?

Aus diesen Zielstellungen ergibt sich der Aufbau des vorliegenden Berichts:

- Kapitel 2 klärt zentrale Konzepte und Definitionen (Risiko, Gefahr, Risikoabschätzung etc.).
- Kapitel 3 erläutert an Fallbeispielen die Bandbreite der bislang durchgeführten CRA-Verfahren.
- Kapitel 4 gibt einen Überblick zu den Ergebnissen sozialwissenschaftlicher Risiko- und Entscheidungsforschung, welche für die Organisation und Durchführung von CRA-Verfahren und die dabei erforderliche Risikokommunikation genutzt werden können.
- In Kapitel 5 werden methodische Probleme der CRA behandelt und der Ansatz der multiattributen Entscheidungsfindung als Methode für die vergleichende Risikobewertung mit verschiedenen Bewertungskriterien und Beteiligten dargestellt.
- Kapitel 6 diskutiert die praktische Organisation und Durchführung einer CRA.
- Im Anhang finden sich vertiefende Informationen zu Risikometriken sowie Erläuterungen und Beispiele zur multiattributen vergleichenden Risikobewertung sowie zur vergleichenden Bewertung von undeutlichen Risiken und von Störfallrisiken.

2. Konzepte und Definitionen

Im folgenden werden die wesentlichen Konzepte für die Diskussion und Ausarbeitung eines vergleichenden Risikoansatzes erläutert. Weitere Erläuterungen und Definitionen finden sich im Glossar am Ende dieses Berichts.

2.1 Risiko

In den Begriffen Chance und Risiko drückt sich jeweils aus, dass aus einer bekannten Möglichkeit Realität werden kann. Meist verbindet sich mit diesem Begriffsinhalt auch die Vorstellung, dass sich der Wechsel von einer Möglichkeit zur Realität in irgendeiner Weise bewerten lässt. Ob sich eine solche Möglichkeit tatsächlich auch realisiert, bleibt jedoch ungewiss. Ungewissheit ist eine konstitutive Eigenschaft von Chancen und Risiken. Wenn wir das, was möglich ist, nur ungenau kennen, ergibt sich daraus zusätzlich eine Unsicherheit. Auch diese Unsicherheit kann für sich genommen bewirken, dass aus einem bestimmten Sachverhalt ein Risiko wird.

Mit Risiken werden meist negative Erwartungen verbunden. Verallgemeinert bedeutet ein Risiko ein potentielles Ereignis mit unerwünschten Wirkungen oder – anders ausgedrückt – die Möglichkeit des Eintritts eines Schadensereignisses. Die spezifische Konzeptualisierung dieser allgemeinen Vorstellung fällt in verschiedenen Wissenschaftsdisziplinen unterschiedlich aus.

Jungermann & Slovic (1993, 169f.) nennen sechs verschiedene Definitionen von Risiko, von denen die Definitionen (c) bis (f) vor allem im Bereich der Finanz- und Versicherungsmathematik genutzt werden:

- (a) Risiko als Wahrscheinlichkeit eines Schadens;
- (b) Risiko als Ausmaß des möglichen Schadens;
- (c) Risiko als Funktion (meist als Produkt) von Wahrscheinlichkeit und Ausmaß des Schadens;
- (d) Risiko als Varianz der Wahrscheinlichkeitsverteilung aller möglichen Konsequenzen einer Entscheidung;
- (e) Risiko als Semivarianz der Verteilung aller (negativen) Konsequenzen mit einem bestimmten Bezugspunkt;
- (f) Risiko als gewichtete lineare Kombination der Varianz und des Erwartungswertes der Verteilung aller möglichen Konsequenzen.

Technische Risikoanalysen bzw. Systemanalysen werden in der Regel für Anlagen durchgeführt, die eine erhebliche Gefährdung ihrer Umwelt hervorrufen können, meist aufgrund verwendeter *"Materialien, die durch toxische oder radioaktive Wirkung oder durch Energiefreisetzung gesundheitliche oder materielle Schäden verursachen können"* (Hauptmanns, Herttrich & Werner 1987). Derartige Anlagen verfügen über Sicherheitssysteme, die solche Materialien einschließen oder Energiefreisetzungen verhindern sollen. Technische Risikoanalysen bzw. Systemanalysen betonen daher die möglichen und zunächst unbekannten Abläufe von Störfällen und

den Aspekt des Systemversagens: „*The concept of risk is used to assess and evaluate uncertainties associated with an event. Risk can be defined as the potential for loss as a result of a system failure*“ (Ayyub & Bender 1999, S. 7).

Auch im Gesundheitsbereich selbst werden – je nach Wissenschaftsfeld – noch unterschiedliche Varianten des Risikobegriffs genutzt. Für die Krebsforschung definieren Williams & Paustenbach (2002, 368f.): „*Risk is a unitless probability of an individual developing cancer*“. In der Epidemiologie wird Risiko üblicherweise als Relatives Risiko, Odds Ratio oder Attribuierbares Risiko ausgedrückt. Risiko ist hier die Relation von Beobachtungen von Schadensfällen in einer Population in bezug zu einer anderen Population. In der Toxikologie wird Risiko verstanden als „*the likelihood, or probability, that the toxic properties of a chemical will be produced in populations of individuals under their actual conditions of exposure*“ (Rodricks 1992, 48). Ähnlich haben die WHO und der *United Nations Conference on the Human Environment* Risiko definiert als „die zu erwartende Häufigkeit unerwünschter Effekte, ausgelöst durch die Exposition gegenüber einem Fremdstoff“ (zitiert nach Neubert 1994, 848).

Eine bereichsunabhängige Definition haben Kaplan und Garrick gegeben, die die drei wesentlichen Elemente von Risiko hervorhebt. Sie haben Risiko (R) als Menge von Tripeln definiert (Kaplan & Garrick 1981, 13):

$$R = \{(s_i, p_i, x_i)\}, i = 1, 2, \dots, N.$$

Wobei s_i die Beschreibung eines Szenarios ist;

p_i die Wahrscheinlichkeit des Szenarios angibt;

x_i das Ausmaß des Schadens in dem Szenario angibt.

Mit dieser Konzeptualisierung wird Risiko als die Menge der betrachteten Folgen mit deren Wahrscheinlichkeit und Ausmaß beschrieben.⁴ Das *Scientific Steering Committee's Working Group on Harmonisation of Risk Assessment Procedures* der Europäischen Kommission nimmt eine ähnliche Konzeptualisierung von Risiko vor:

The probability and severity of an adverse effect / event occurring to man or the environment following exposure, under defined conditions, to a risk source(s).

(EC 2000b, 5)

Dies ist die Definition von Risiko, die in diesem Bericht zugrunde gelegt wird.

Verschiedene Autoren haben eine Erweiterung des wissenschaftlich-technischen Risikobegriffs durch sogenannte „qualitative“ Risikofaktoren vorgeschlagen (Fischhoff, Watson & Hope 1984; Florig et al. 2001).⁵ Diese „qualitativen“ Risikofaktoren basieren auf den Ergebnissen der Forschung zur Risikowahrnehmung (s.u.). Dort hatten sich zwei Aspekte als bedeutsam für die Risikowahrnehmung erwiesen: Das Ausmaß, in dem eine Risikoquelle (Technologie, Produkt, Stoff etc.) als furchterre-

⁴ Kaplan & Garrick (1981) und Kaplan (1997) differenzieren diese Konzeptualisierung noch weiter aus. Für den vorliegenden Zusammenhang ist diese Definition ausreichend.

⁵ Auch der WBGU (1998) nimmt bei seiner Klassifizierung von Risiken in verschiedene Typen eine solche Erweiterung vor, indem u.a. Ubiquität und Mobilisierungspotenzial als Klassifizierungskriterien für Risiken verwendet werden.

gend wahrgenommen wird, und das Ausmaß, in dem die Risikoquelle als etwas für die Wissenschaft bzw. Öffentlichkeit Unbekanntes wahrgenommen wird. Fischhoff, Watson & Hope (1984) beispielsweise unterscheiden drei Dimensionen von Risiko: Mortalität (wobei sie weiter zwischen Arbeitsplatz und Öffentlichkeit differenzieren), Morbidität und Besorgnis. Letztere ist der Oberbegriff für die beiden qualitativen Risikodimensionen „furchterregend“ und „unbekannt“ (siehe Abbildung 1).

Ein anderes Verständnis liegt der Verwendung des Risikokonzepts im juristischen Bereich zugrunde.⁶ Risiko wird dort im Rahmen der Begriffstrias Gefahr, Risiko und Restrisiko betrachtet. Exakte juristische Definitionen liegen für die Begriffe Risiko und Gefahr bislang allerdings noch nicht vor. Von Gefahr wird gesprochen, „wenn eine Sachlage bei ungehindertem Ablauf des objektiv zu erwartenden Geschehens mit hinreichender Wahrscheinlichkeit zu einem Schaden führt, d.h. zu einer nicht unerheblichen Beeinträchtigung eines rechtlich geschützten Gutes“ (SRU 1999, 39). Gefahren müssen nach deutschem Recht abgewehrt werden.

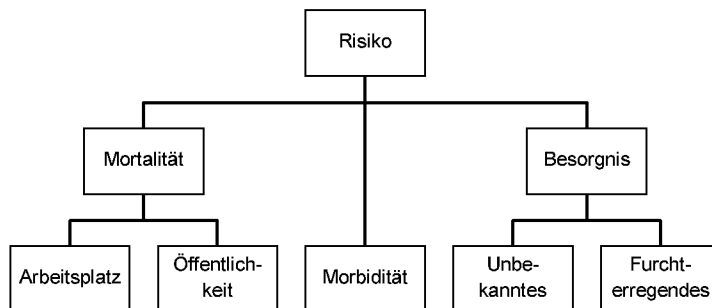


Abbildung 1: Multidimensionales Risikokonzept von Fischhoff, Watson & Hope (1984, 129)

Risiko wird dagegen verstanden als jede Möglichkeit, dass ein Schaden lediglich mit einer Gewissheit eintritt, die nicht ausreicht, um das Vorhandensein einer Gefahr zu begründen. Restrisiken schließlich bezeichnen die von der Gesellschaft hinzunehmenden Risiken. Diese beruhen, wie das Bundesverfassungsgericht in seiner berühmten „Kalkar-Entscheidung“ sinngemäß formuliert hat, auf Ungewissheiten jenseits der Schwelle der praktischen Vernunft, die ihre Ursache in der Begrenztheit des menschlichen Erkenntnisvermögens haben.

Dementsprechend existieren prinzipiell auch unbekannte Risiken und unbekannte Gefahren. Denn sowohl für den Bereich des Risikos wie für den der Gefahr sind Fälle denkbar, von denen man bislang zwar noch nichts weiß, die aber rein logisch auch nicht ausgeschlossen werden können. Abbildung 2 zeigt den Zusammenhang zwischen Risiko, Gefahr und den zugehörigen umweltrechtlichen Begriffen.

⁶ Die Ausführungen zum juristischen Risikobegriff basieren im wesentlichen auf SRU (1999).

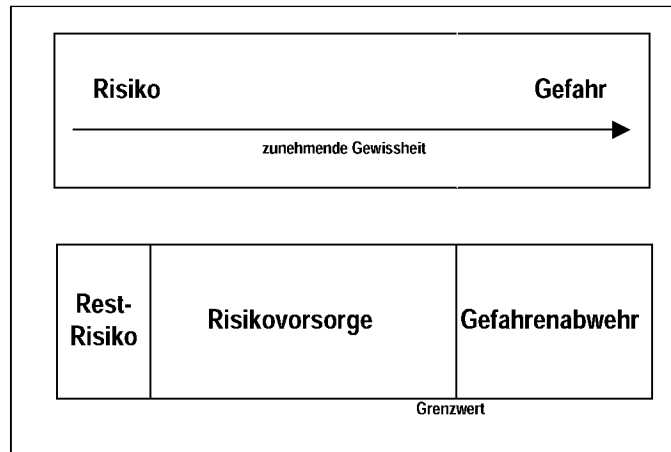


Abbildung 2: Risiko, Gefahr und die zugehörigen umweltrechtlichen Begriffe

2.2 Gefahr / Schädigungspotenzial – Gefährdungspotenzial

Anders als im alltäglichen Sprachgebrauch, wo Gefahr und Risiko häufig synonym verwendet werden, wird im Kontext der wissenschaftlichen Risikoabschätzung zwischen den beiden Konzepten unterschieden. Gefahr meint hier das Potenzial einer Risikoquelle (Technologie, Produkt, Substanz, Situation, Aktivität etc.), adverse Effekte⁷ zu verursachen. Auch die soziologische Risikoforschung unterscheidet zwischen Risiko und Gefahr, und zwar bezüglich des Zustandekommens einer negativer Erwartung (Beck 1986). Gefahr steht für etwas Bedrohliches oder auch Fremdartiges, das sich ereignet, etwa in Form einer Naturkatastrophe. Risiko dagegen wird überwiegend als mögliche zukünftige Folge der Wahl einer bestimmten Handlungsoption und daher als grundsätzlich beeinflussbar angesehen.

Es ist deshalb im vorliegenden Zusammenhang angemessener, von „Schädigungspotenzial“ oder „Gefährdungspotenzial“ zu sprechen, um dieses Potenzial näher zu bestimmen und so zu einer Risikoeinschätzung zu kommen. Gefahr im Sinne von Schädigungspotenzial oder Gefährdungspotenzial ist damit auch zu unterscheiden vom Gefahrenbegriff, wie er im rechtlichen Bereich als quantitativ definierte Größe verwendet wird (s.o.).

2.3 Ungewissheit und Unsicherheit

Ungewissheit und Unsicherheit gehören konstitutiv zum Risikobegriff. Häufig werden Ungewissheit und Unsicherheit als Synonyme oder nicht klar gegeneinander abge-

⁷ Die WHO (1994, 20) gibt folgende Definition für den Begriff „adverser Effekt“: „change in morphology, physiology, growth, development or life span of an organism which results in impairment of functional capacity or impairment of capacity to compensate for additional stress or increase in susceptibility to the harmful effects of other environmental influences. Decisions on whether or not any effect is adverse require expert judgement.“

grenzt verwendet. Im Kontext einer vergleichenden Risikobewertung ist es jedoch nützlich, diese Begriffe deutlich zu unterscheiden.

Dabei soll *Ungewissheit* den Sachverhalt beschreiben, dass wir nicht wissen können, ob etwas, das wir als Möglichkeit identifiziert haben, Realität wird. Ungewissheiten dieser Art können auch dann bestehen, wenn die Möglichkeiten als solche gut bekannt sind; Beispiele sind die einzelnen Resultate beim Würfeln. Dagegen soll *Unsicherheit* als Maß dafür verwendet werden, wie genau die jeweiligen Möglichkeiten – und gegebenenfalls auch ihre statistisch zu erwartende Häufigkeit – bekannt sind.

Gelegentlich werden diese beiden Sachverhalte dadurch unterschieden, dass man Unsicherheiten als aleatorisch (im Sinne der obigen Ungewissheit) oder als epistemische (durch mangelndes Wissen verursacht) bezeichnet. Bei Schlussfolgerungen aus vergleichenden Risikobewertungen ist es von erheblicher Bedeutung, ob die einzelnen Risiken aleatorisch oder epistemisch „verursacht“ sind.

2.4 Risikoabschätzung

Unter Risikoabschätzung wird der „Prozess von der Identifizierung des Gefährdungspotentials bis zur quantitativen Charakterisierung von Risiken“ (Risikokommission 2002, 9) verstanden. Je nach dem, ob es um die Abschätzung von Gesundheits- oder Umweltrisiken oder von Risiken technischer Systeme geht, werden unterschiedliche Verfahren eingesetzt.

Bei der Abschätzung von Gesundheitsrisiken werden meist vier Schritte unterschieden (vgl. NRC 1983): Am Anfang steht die Identifizierung einer Gefährdung, d.h. die qualitative Charakterisierung der gesundheitsschädigenden Eigenschaften eines Stoffes. Hier geht es beispielsweise um die Frage, ob eine Substanz Krebs erzeugen kann. Der zweite Schritt ist die quantitative Beschreibung der Dosis-Wirkungs-Beziehung. Untersucht wird, bei welcher Dosis welche Wirkung auftritt und vor allen Dingen, ob es einen Schwellenwert für eine schädliche Wirkung gibt oder ob auch bei geringsten Dosen Risiken bestehen. Ein weiterer Schritt besteht in der Abschätzung der Exposition, d.h. der Stärke und zeitlichen Dauer, mit der eine Population dem Schadstoff ausgesetzt ist. Schließlich wird im vierten Schritt eine Risikocharakterisierung vorgenommen, d.h. eine zusammenfassende Bewertung des Risikos, in der Art und Häufigkeit der zu erwartenden Gesundheitsschäden für die exponierte Population dargestellt werden.

Bei der Abschätzung ökologischer Risiken geht es um die Abschätzung möglicher adverser Effekte auf das Ökosystem als Resultat der Exposition mit einem oder mehreren Stressoren. Die ökologische Risikoabschätzung lässt sich in drei Phasen unterteilen (EPA 1998): In der ersten Phase (Problemformulierung) wird der zu betrachtende Ausschnitt des Ökosystems festgelegt und es werden die ökologischen Effekte (Endpunkte) und die als Verursacher angenommenen Stressoren in einem (meist qualitativen) Modell (*conceptual model*) beschrieben. Darauf aufbauend wird in der zweiten Phase (Analyse) die Wahrscheinlichkeit der Exposition untersucht und Ausmaß und Art der für eine bestimmte Expositionstärke zu erwartenden ökologischen Effekte bestimmt. Auf der Basis der dabei resultierenden Expositions- und Stressor-Effekt-Profile wird in der dritten Phase eine zusammenfassende Risikocharakterisierung vorgenommen.

rakterisierung vorgenommen. Diese Risikocharakterisierung umfasst auch die zugrunde gelegten Annahmen und Abschätzungsunsicherheiten der Analyse.

Bei der Bestimmung von Risiken technischer Anlagen ist zwischen dem bestimmungsgemäßen Betrieb ("Normalbetrieb") und dem möglichen Auftreten von Störfällen zu unterscheiden. Falls im bestimmungsgemäßen Betrieb mit der Emission risikorelevanter Stoffe zu rechnen ist, können die entsprechenden Risiken wie oben beschrieben abgeschätzt werden. Eine Abschätzung von Störfallrisiken erfordert zusätzliche spezifische Vorgehensweisen. Dabei geht es zunächst um eine statistische Erfassung und Modellierung des Verhaltens der einzelnen Komponenten eines technischen Systems (vgl. Mertens 1993). Zum Einsatz kommen hier Verfahren wie die Fehlerbaum- oder Ereignisbaumanalyse. Auf dieser Basis wird das Ausfallverhalten des Gesamtsystems (Störfall) mit den entsprechenden Wahrscheinlichkeiten berechnet. Schließlich können dann für die einzelnen Ausfallsituationen die resultierenden Risiken (etwa Gesundheitsrisiken oder ökologische Risiken aufgrund störfallbedingter Emissionen), die ebenfalls wieder über Wahrscheinlichkeitsverteilungen charakterisiert werden, angegeben werden.

2.5 Risikobewertung

Die Risikobewertung baut auf den Ergebnissen der wissenschaftlichen Risikoabschätzung auf. Während die Risikoabschätzung die Ermittlung der Höhe bzw. Größe eines Risiko zum Ziel hat, geht es bei der Risikobewertung darum, dieses Risiko in bezug auf seine Bedeutsamkeit, Schwere oder Ernsthaftigkeit zu bewerten. Die Größe bzw. Höhe eines Risikos allein gibt keinen Aufschluss darüber, wie „schlimm“ dieses Risiko ist. Ein relatives Risiko von 2 (also eine Verdopplung des Risikos für die exponierten gegenüber den nicht exponierten Personen), eine bestimmte Krankheit zu bekommen, sagt nichts darüber aus, wie bedeutsam das ist. Diese Bedeutsamkeit erschließt sich erst, wenn das Risiko zu einem Referenzwert (z.B. einem Umweltstandard) oder zu einem anderen Risiko in Beziehung gesetzt wird. Eine Risikobewertung impliziert also immer einen Risikovergleich. Damit stellt sich die Frage, wer die Risikobewertung vornehmen soll. Für den Fall individuellen Risikoverhaltens, also für Risiken, von denen nur diejenigen betroffen sind, die diese Risiken „hervorbringen“ (dabei geht es im wesentlichen um die sogenannten *life-style* Risiken), ist die Antwort einfach: Die Bewertung sollen die Betroffenen selbst vornehmen.

Problematisch wird es, wenn es um die gesellschaftliche Regulierung von Risiken geht. Hier geht es um die Frage der Zumutung bzw. Akzeptabilität⁸ von Risiken. In der Vergangenheit hat sich gezeigt, dass sich hierzu eine allgemeingültige Lösung nicht finden lässt. Wie groß ein Risiko sein darf – oder: wie gering ein Risiko sein muss – damit es für eine Gesellschaft akzeptabel ist, den Bürgern zugemutet werden darf, darüber hat sich bislang kein gesellschaftlicher Konsens erzielen lassen.⁹ Bevor man im Rahmen der gesellschaftlichen Risikobewertung allerdings überhaupt vor der

⁸ Hier geht es um Akzeptabilität und nicht um Akzeptanz von Risiken. Während letztere sich als deskriptiver Begriff auf die empirisch feststellbare Billigung von Risiken durch Personen, Gruppen oder einer Gesellschaft bezieht, meint erstere die Forderung, die jeweiligen Risiken zu billigen. Akzeptabilität ist also ein normatives Konzept.

⁹ Vgl. hierzu die ausführliche Diskussion in WBGU 1998, Kap. C 1

Frage steht, wie die Akzeptabilität eines Risikos begründet werden kann, müssen zwei andere Probleme gelöst werden:

Erstens, die Dimensionen, die bei der wissenschaftlichen Abschätzung von Risiken betrachtet werden (Schaden und Wahrscheinlichkeit), sind nicht die einzigen Aspekte, die für die Bewertung von Risiken durch verschiedene gesellschaftliche Akteure von Bedeutung sein können. Faktisch sind zwar für Risikobewertungen in der Vergangenheit meist die Dimensionen herangezogen worden, die auch bei der wissenschaftlichen Abschätzung von Risiken herangezogen werden. Dies sind die in den verschiedenen Disziplinen verwendeten Risikometriken: Todesfälle, spezifische Endpunkte (Krankheiten) sowie toxikologische oder epidemiologische Parameter.¹⁰ Die psychologische und sozialwissenschaftliche Risikowahrnehmungsforschung hat aber gezeigt, dass für die intuitive Risikobeurteilung durch Laien eine ganze Reihe sogenannter „qualitativer“ Beurteilungsaspekte eine Rolle spielen können. Diese beziehen sich auf die Art („Qualität“) des Schadens und auf den situativen Kontext, in dem ein Risiko steht (Jungermann & Slovic 1994). Wie oben zum Begriff Risiko ausgeführt, wurde als Lösung hierfür vorgeschlagen, diese qualitativen Aspekte explizit in die Konzeptualisierung des Risikobegriffs und damit auch in die Risikoabschätzung einzubeziehen. Der WBGU hat in seinem Gutachten „Welt im Wandel – Strategien zur Bewältigung globaler Umweltrisiken“ einen ähnlichen Ansatz verfolgt. Er schlägt vor, „die wissenschaftlichen Abschätzungen gemeinsam mit den durch empirische Studien erfassten Risikowahrnehmungen als Informationsbasis für eine rationale Abwägung heranzuziehen (Fiorino, 1989). Beide Typen von Informationen sind wichtige Bestandteile der Risikobewertung“ (WBGU 1998, 40). Im Gegensatz zum WBGU halten wir eine solche Einbeziehung der Risikowahrnehmung in die Risikobewertung nicht für sinnvoll; diese sollte unseres Erachtens ausschließlich auf den wissenschaftlichen Risikoabschätzungen beruhen. Bei der Priorisierung von Risikomanagementmaßnahmen kann die Risikowahrnehmung einzelner Gruppen bzw. der Öffentlichkeit dagegen durchaus ein relevantes Beurteilungskriterium sein.

Zweitens, Risikoabschätzungen liefern nur selten eindeutige Resultate. In der Regel sind sie mit mehr oder weniger großen Unsicherheiten behaftet.¹¹ Damit ergibt sich das Problem, wie diese Unsicherheiten bei der Risikobewertung berücksichtigt werden können. Zunächst müssen diese Unsicherheiten überhaupt bei der Risikobewertung bekannt sein. Das ist durchaus nicht selbstverständlich. Ergebnisse der Risikoabschätzung werden häufig als Punktschätzungen angegeben, ohne dass Charakterisierungen der mit den Schätzungen verbundenen Unsicherheiten berichtet werden. Diese Praxis ist in den letzten Jahren vielfach kritisiert worden (z.B. NAS 1994). Es wird argumentiert, dass bei der Punktschätzung meist der *Worst-Case* zugrunde gelegt und die Risiken deshalb überschätzt würden. Dies könne eine Fehlallokation von Ressourcen im Gesundheits- bzw. Umweltschutz zur Folge haben (Finkel 1989). Unsicherheiten können in allen Schritten im Prozess der Risikoabschätzung bestehen. Für die Abschätzung der Dosis-Wirkungs-Beziehung und der Exposition ist es möglich, quantitative Angaben zur Unsicherheit zu machen und diese zum Beispiel in Form von Konfidenzintervallen, Perzentilen oder Wahrscheinlich-

¹⁰ Für die Toxikologie sind dies zum Beispiel: *Unit Risk*, *Benchmark Dose*, *Margin of Exposure*; für die Epidemiologie: *Relatives Risiko*, *Odds Ratio*, *Attribuierbares Risiko*.

¹¹ Für eine zusammenfassende Diskussion dieser in der Literatur zur Risikoabschätzung immer wieder konstatierten Problematik siehe z.B. Bailer & Bailer (1999) sowie EPA (1996).

keitsverteilungen darzustellen (vgl. Thompson & Bloom 2000). Dagegen lassen sich die Unsicherheiten bei der Identifizierung des Schädigungspotenzials nur begrenzt quantitativ beschreiben. Das Kernproblem ist hier, dass sich diese Unsicherheiten aus der begrenzten oder widersprüchlichen Befundlage ergeben. Zwar kann hier versucht werden, mit Hilfe von Meta-Analysen die Befundlage quantitativ zusammenzufassen. Allerdings stellt dies hohe Anforderungen an die Qualität und Vergleichbarkeit der verfügbaren Daten.¹²

Diese beiden Probleme, das der Multidimensionalität der Risikobewertung und das des Umgangs mit Unsicherheit, sind zentral für die Risikobewertung und werden im Rahmen dieser Arbeit immer wieder aufgegriffen.

2.6 Risikovergleich

Ein Risikovergleich hat zum Ziel, auf der Basis von Risikoabschätzungen zwei oder mehr Risiken in Beziehung zueinander zu setzen. Dabei müssen klare Zielsetzungen formuliert werden, um zu sinnvollen und nützlichen Ergebnissen zu kommen. Derartige Zielsetzungen ergeben sich meist aus Problemen mit zugehörigen Fragestellungen.

Ein solches Problem ist beispielsweise die einstufende Verdeutlichung eines Risikos mit dem Ziel, seine Bedeutung - etwa in Bezug auf bestimmte Gesundheitsschäden - zu vermitteln. Hier geht es primär um das Risikoverständnis und um Risikokommunikation. Eine andere Zielsetzung für einen Risikovergleich kann darin bestehen, Risiken in eine Reihenfolge zu bringen, um eine begründete Prioritätensetzung bei Entscheidungen zum Risikomanagement vor dem Hintergrund begrenzter Ressourcen oder unter Kosten-Nutzen-Gesichtspunkten zu erreichen. Schließlich kann ein Risikovergleich dazu dienen, unterschiedliche Optionen (etwa die Auswahl unter mehreren technischen Lösungen oder zwischen Handlungsmöglichkeiten) auch im Hinblick auf ihre Risiken zu beurteilen. Hier ist das Risiko einer von mehreren Gesichtspunkten, wodurch eine Gewichtung und/oder Aggregation erforderlich wird.

In systematischer Hinsicht kann man die vergleichende Risikobewertung zwischen Risikomanagement und Risikoabschätzung ansiedeln. Abbildung 3 stellt diese Beziehung in Form einer Pyramide dar.

Die Spitze dieser Pyramide bildet das Risikomanagement. Die hier bestehenden Fragen über die relative Bedeutsamkeit von Risiken bestimmen den Kontext für die vergleichende Risikobewertung. Für deren Durchführung werden dann die erforderlichen Informationen über die zu betrachtenden Risiken aus vorhandenen Risikoabschätzungen extrahiert bzw. noch erforderliche Risikoabschätzungen initiiert. Auf dieser Datenbasis wird die vergleichende Risikobewertung durchgeführt und als Ergebnis ein Risikoranking bereitgestellt. Die Methodik der vergleichenden Risikobewertung sowie Fragen der Organisation und Durchführung werden ausführlich in den Kapiteln 5 und 6 behandelt.

¹² Zum Problem der zusammenfassenden Charakterisierung der Befundlage bei der Identifizierung des Schädigungspotenzials siehe: Wiedemann, Schütz & Thalmann 2002, Kap. 7.

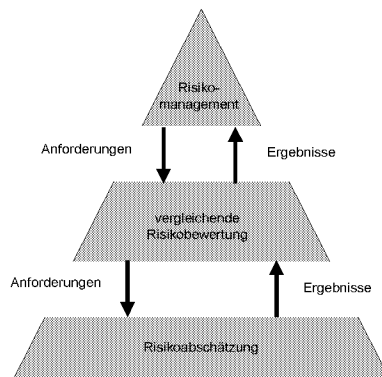


Abbildung 3: CRA Pyramide

Das Ranking wiederum liefert die Grundlage für die Priorisierung von Risiken. Die Priorisierung von Risiken ist also vom Risikoranking zu unterscheiden. Die Priorisierung erfordert weitergehende Überlegungen, z.B. ob überhaupt Maßnahmen (unter den gegebenen Randbedingungen) möglich sind oder nach welchen Kriterien über die Auswahl von Risikomanagementmaßnahmen entschieden werden soll (beispielsweise Kosteneffizienz, Gerechtigkeit u.ä.). Die vergleichende Bewertung von Risiken ist also nur ein – wenn auch grundlegender – Aspekt für die Priorisierung. Die Auswahl der Kriterien für die Beurteilung von Risikomanagementmaßnahmen und die Beurteilung selbst sind gesonderte Probleme, die ggf. die Einbeziehung verschiedener Stakeholder / Interessengruppen erfordern.

2.7 Risikomanagement

Das Risikomanagement bezieht sich auf diejenigen Aktivitäten, die mit der Bewältigung von Risiken verbunden sind. Sie setzt damit nach der Risikobewertung ein, bei der geklärt wird, wie bedeutsam ein Risiko ist. Die Entscheidung darüber, ob und wie ein Risiko reguliert oder kontrolliert werden soll, ist dann Aufgabe des Risikomanagements, denn für diese Entscheidung sind neben der Risikobewertung noch andere Aspekte relevant. Risikomanagement und Risikobewertung sollten deshalb als eigenständige Konzepte getrennt werden, die unterschiedliche Funktionen im Prozess der Risikoregulierung (s.u.) erfüllen.

Zunächst muss die Entscheidung getroffen werden, ob überhaupt Maßnahmen zum Risikomanagement erforderlich sind. Von grundlegender Bedeutung ist dabei die Auswahl des Schutzziels: geht es um Gefahrenabwehr oder um vorsorgliche Maßnahmen. Wie oben schon ausgeführt, sollen durch die Gefahrenabwehr Schäden verhindert werden, mit deren Eintreten man ohne solche Abwehrmaßnahmen mit hinreichender Gewissheit rechnen müsste. Gefahrenabwehr setzt also zumindest voraus, dass im Prozess der Risikoabschätzung eine wissenschaftlich eindeutige Feststellung eines Schädigungspotenzials vorgenommen werden konnte. Für die Gefahrenabwehr sind Kriterien und Maßnahmen in Gesetzen und Verordnungen, die den Gesundheits- bzw. Umweltschutz und die Sicherheit technischer Systeme betreffen, festgelegt. Dagegen geht es bei der Vorsorge um die Frage, ob und wie mit möglichen Risiken umgegangen werden soll, bei deren wissenschaftlicher Risikoabschätzung noch große Unsicherheiten bestehen. Diese Unsicherheiten betreffen vor

allein die Frage der Identifizierung des Schädigungspotenzials, also die Frage, ob ein bestimmter Stoff eine gesundheitsschädigende Wirkung hat. Auch das Schutzziel Vorsorge ist schon in einigen Gesetzen enthalten (z.B. Bundes-Immissionsschutzgesetz, Chemikaliengesetz, Gentechnikgesetz).¹³ Seine Bedeutung für das Risikomanagement liegt zur Zeit allerdings noch vornehmlich auf der Ebene gesellschaftlicher bzw. politischer Auseinandersetzungen zwischen verschiedenen Interessengruppen, von denen die einen die Anwendung des Vorsorgeprinzips fordern und die anderen dies zurückweisen.

Nach der Entscheidung, dass Maßnahmen in bezug auf ein Risiko ergriffen werden sollen, müssen hierzu Handlungsoptionen entwickelt werden. Dabei müssen gesundheitliche, ökologische, ökonomische, technische, soziale und politische Gesichtspunkte berücksichtigt werden. Der nächste Schritt ist die Bewertung der Handlungsoptionen anhand von Kriterien, die sich zum Beispiel auf die Effizienz der Optionen zur Risikominimierung, auf Kosten-Nutzen-Aspekte, aber auch auf die gesellschaftliche Durchsetzbarkeit der Maßnahmen beziehen können. Ein weiterer wesentlicher Bewertungsaspekt sind mögliche unerwünschte Folgen von Handlungsoptionen (s.u.). Diese Informationen bilden die Basis für die Entscheidung über die zu implementierende Risikomanagementstrategie. Zum Risikomanagement gehört aber auch, dass die Implementierung der Maßnahmen überwacht wird und diese hinsichtlich ihrer Effektivität evaluiert werden.

Grundsätzlich lassen sich für die Bewältigung von Risiken vier Strategien unterscheiden: Vermeiden, Vermindern, Nachsorgen und Absichern (Wiedemann 1993, 209).

Beim *Vermeiden* von Risiken geht es um das Ausschalten der Möglichkeit von Schädigungen durch die Nutzung von Technologien, Produkten, Substanzen etc.. Auf der gesellschaftlichen Ebene wird dies im Regelfall durch Gesetze oder Verordnungen erreicht. Beispiele sind Verbote von Chemikalien oder der Setzung von Grenzwerten. Durch Grenzwertsetzung vermeiden lassen sich Gesundheitsrisiken nur, wenn für die jeweiligen Noxen in der Dosis-Wirkungs-Beziehung ein Schwellenwert existiert. Dies ist für die meisten nicht kanzerogenen Stoffe der Fall. Dagegen wird bei kanzerogenen Stoffen von einer (meist linearen) Dosis-Wirkungs-Beziehung ohne Schwellenwert ausgegangen. Risiken durch solche Noxen können durch Grenzwerte prinzipiell nicht vermieden, sondern nur vermindert werden.

Die Strategie des *Verminderns* zielt nicht auf die Beseitigung von Risiken, sondern auf deren Reduzierung. Das kann zum einen durch Maßnahmen geschehen, die die Eintrittswahrscheinlichkeit von Schadensereignissen verringern. Bei kanzerogenen Stoffen bedeutet das, die Exposition so weit wie möglich zu verringern,¹⁴ bei technischen Anlagen kann dies durch entsprechende Sicherheitsvorkehrungen geschehen. Zum anderen können Maßnahmen zur Begrenzung des Schadensumfangs getroffen werden, zum Beispiel durch Begrenzung der Exposition auf den Arbeitsbereich.

Nachsorgen bedeutet, Maßnahmen für den Eintritt eines Schadensfalls zu ergreifen, die bei der Bewältigung der Folgen helfen. Hierzu gehört ein breites Spektrum ganz

¹³ Vgl. Rehinder (1988) sowie Wiedemann et al. (2001), Kap. 4.

¹⁴ Für diese Problemlage wurde das ALARA-Prinzip (*As Low As Reasonably Achievable*) ursprünglich entwickelt.

unterschiedlicher Maßnahmen, von der Bereitstellung der für eine Schadensbewältigung notwendigen Infrastruktur bis hin zu finanzieller Notfallunterstützung und psychosozialer Betreuung der Betroffenen. Im wesentlichen bezieht sich diese Strategie auf Extremereignisse, wie Natur- oder Technikkatastrophen, die im Zusammenhang mit dem Management umweltbezogener Gesundheitsrisiken bislang wenig beachtet wurden. In den letzten Jahren allerdings wird diesen Extremereignissen auch in diesem Zusammenhang zunehmend Aufmerksamkeit geschenkt.¹⁵

In eine ganz andere Richtung zielt die Strategie des *Absicherns*. Hier geht es um die Entschädigung von Betroffenen für entstandene Schäden. Üblicherweise geschieht dies durch das Abschließen von Versicherungen, mitunter auch durch die Übernahme von Bürgschaften durch den Staat. Beispiele für letzteres sind die staatliche Absicherung gegen Schäden durch Kernkraftunfälle, welche die Versicherungssumme der Betreiber überschreiten, oder (für eine gewisse Zeit nach dem 11. September 2001) gegen die Risiken von Terroranschlägen auf die Luftfahrt.

Vor allem für die Strategien *Vermeiden* und *Vermindern* ergibt sich das Problem des *risk-risk trade-offs* (Graham & Wiener 1997). Damit ist gemeint, dass die Vermeidung bzw. Verminderung von Risiken selbst wieder Risiken zur Folge haben kann. Beispielsweise zielt die Sanierung von Altlasten auf die Vermeidung bzw. Verringerung des Krebsrisikos von Anwohnern durch die dort vorhandenen kanzerogenen Stoffe; solche Sanierungsmaßnahmen bringen allerdings auch Unfallrisiken für die damit beschäftigten Arbeiter mit sich, die die Reduzierung des Gesundheitsrisikos für die Anwohner unter Umständen aufwiegen können (Cohen, Beck & Rudel 1997).¹⁶ Die üblichen Analyseinstrumente für die Entscheidung über Risikomanagement Optionen: Kosten-Nutzen-Analyse bzw. Kosteneffizienz-Analyse, zielen auf die Beurteilung des Kosten-Nutzen-Verhältnisses bzw. der Kosteneffizienz von Risikomanagement Optionen. Ob durch Risikomanagementmaßnahmen unbeabsichtigt andere Risiken entstehen können, bleibt bei dieser Betrachtungsweise meist unberücksichtigt. Solche Risikoverschiebungen sind vermutlich keine Einzelfälle und sollten bei der Beurteilung und Auswahl von Risikomanagement Optionen im Rahmen einer Risk-Tradeoff-Analyse berücksichtigt werden (Graham & Wiener 1997; Presidential Commission 1997a, 35).

2.8 Risikoregulierung

Im gesellschaftlichen Umgang mit Risiken sind Risikoabschätzung, Risikobewertung und Risikomanagement in vielfältiger Weise aufeinander bezogen. Zugleich stellen sie aber auch eine Abfolge notwendiger Schritte für die rationale Regulierung von Risiken dar. Die Risikokommission (2003) bezeichnet diesen Gesamtprozess von der Risikoabschätzung über die Risikobewertung bis zum Risikomanagement als Risikoregulierung (siehe Abbildung 4).

An den Beginn dieses Prozesses setzt die Risikokommission als ersten Schritt ein Vorverfahren. Hier sollen eine Reihe von Entscheidungen zur Problemeingrenzung

¹⁵ Vgl. dazu Kunreuther (2002).

¹⁶ Graham und Wiener (1997) nennen diese unbeabsichtigten Risiken, die die Risikoreduzierung durch Risikomanagementmaßnahmen entgegen wirken können, deshalb „*countervailing risks*“.

und zur Steuerung der folgenden Prozessabschnitte getroffen werden. Dies umfasst die Organisation der Früherkennung von Risiken, die Festlegung der Rahmenbedingungen für die spätere Risikoabschätzung, eine Setzung von Prioritäten – etwa für besonders dringende Fälle – sowie die eventuelle Wahl eines verkürzten Regulierungsverfahrens, beispielsweise zur Ressourceneffizienz, wenn größere Risiken von vorneherein ausgeschlossen werden können.

Der zweite Schritt des Prozesses besteht in der Risikoabschätzung. Die Aufgabe dieser naturwissenschaftlichen Analyse besteht in der Identifizierung des Gefährdungspotenzials einer Noxe, in der quantitativen Gefährdungsabschätzung, in der Expositionsabschätzung sowie in einer Angabe der jeweiligen Unsicherheiten.

Als nächsten Schritt sieht die Risikokommission eine Risikobewertung vor. Hier sollen die Ergebnisse der Risikoabschätzung im Hinblick auf Konsequenzen für das Risikomanagement ausgewertet, diskutiert und bewertet werden. Damit wird das naturwissenschaftlich bestimmte Risiko in gesellschaftliche Wertkategorien „übersetzt“ (Risikokommission 2003).

Schließlich wird im vierten Schritt, dem Risikomanagement, ein etwaiger Handlungsbedarf festgestellt, Maßnahmeoptionen zur Risikoreduzierung werden identifiziert und bewertet, und in einer vergleichenden Bewertung dieser Optionen werden deren Auswirkungen differenziert – etwa in der Wirkung auf unterschiedliche Schutzgüter – analysiert. Am Schluss steht eine Empfehlung an die behördlichen und politischen Entscheidungsträger.

Nach Ansicht der Risikokommission sollten in diesen gesamten Prozess nicht nur Fachleute involviert sein, sondern auch andere gesellschaftliche Gruppen:

Risikobewertungen stoßen in einer Gesellschaft, in der Wertepluralismus herrscht und politische Handlungen stets unter hohem Rechtfertigungsdruck stehen, oft auf Skepsis oder Misstrauen. Aussagen über Risiken sind daher mehr als andere Aussagen auf Plausibilität und Vertrauen in die Regulierungsgremien angewiesen. Je mehr Individuen und Gruppen die Möglichkeit haben, aktiv an der Risikoregulierung mitzuwirken, desto größer ist die Chance, dass sie Vertrauen in die Institutionen entwickeln und auch selbst Verantwortung übernehmen.

[...]

Insofern ist eine frühzeitige und gegenseitige Beteiligung der Betroffenen und der organisierten gesellschaftlichen Gruppen an der Entscheidungsfindung sachlich angemessen sowie rechtsstaatlich und demokratisch geboten.

(Risikokommission 2003, 53f.)

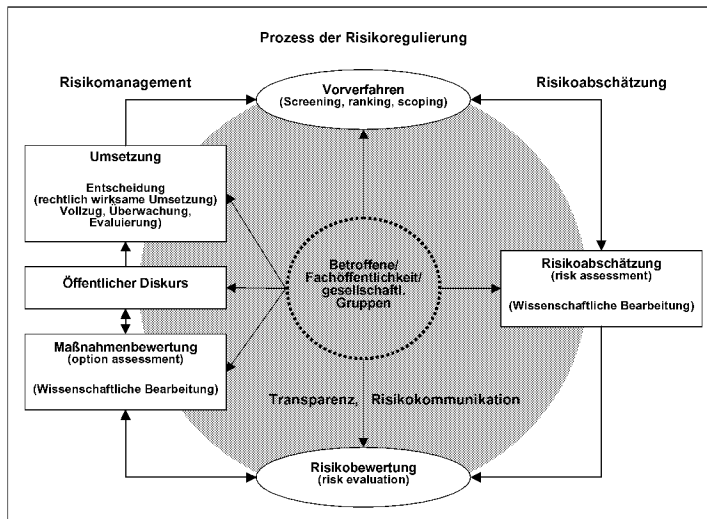


Abbildung 4: Prozess der Risikoregulierung (Risikokommission 2003, 28)

2.9 Risikokommunikation

Eine zentrale Rolle für eine transparente Darstellung der Annahmen, Erwägungen und Ergebnisse der einzelnen Schritte im Prozess der Risikoregulierung spielt die Risikokommunikation. Dies gilt nicht nur in Bezug auf die Kommunikation zwischen den am Prozess der Risikoregulierung direkt Beteiligten (Experten, gesellschaftliche Gruppen), sondern auch für die Öffentlichkeit insgesamt. Dabei ist Risikokommunikation mehr als der bloße Informationsaustausch zwischen einzelnen Beteiligten. Sie umfasst vielmehr alle

Kommunikationsprozesse, die die Identifikation, Analyse, Bewertung und das Management von Risiken sowie die dafür notwendigen Voraussetzungen und Beziehungen zwischen den beteiligten Personen, Gruppen und Institutionen zum Gegenstand haben.

(Wiedemann 1993, 197)¹⁷

In erster Näherung kann zwischen zwei Zielen der Risikokommunikation unterschieden werden: Das Herstellen von Verständigung und das Sichern von Verständnis. Damit sind auch zwei zu unterscheidende Aufgabenfelder der Risikokommunikation genannt. *Verständigung* bezieht sich auf Prozesse zur Herstellung und Aufrechterhaltung von Kommunikation. Hier geht es um zwei-Wege-Kommunikation, um Beteiligung sowie um das Sichern von Fairness, Respekt sowie von Vertrauen. Erst unter solchen Randbedingungen wird es möglich, *Verständnis* bei den Rezipienten für die fachlichen Grundlagen der Risikoabschätzung zu fördern und somit zu deren Risikomündigkeit beizutragen.

¹⁷ Das US amerikanische National Research Council definiert: "Risk communication is an interactive process of exchange of information and opinion among individuals, groups, and institutions. It involves multiple messages about the nature of risk or other messages, not strictly about risk, that express concerns, opinions, or reactions to risk messages or to legal and institutional arrangements for risk management." (National Research Council 1989, 21).

Um der Gefahr der Inflationierung des Begriffs „Risikokommunikation“ zu begegnen, soll im weiteren von Risikokommunikation nur dann die Rede sein, wenn bei Sender und Empfänger beträchtliche kognitive, motivationale oder emotionale Unterschiede in bezug auf das thematische Feld „Risikoabschätzung“ vorherrschen. Das ist insbesondere der Fall, wenn Experten mit Laien kommunizieren. Solche Differenzen können aber auch unter Umständen bei der Kommunikation unter Experten angenommen werden, wenn diese von ganz verschiedenen professionalen Wissenbasen ausgehen.

3. CRA in der Praxis

Ansätze der integrierten und vergleichenden Risikobewertung können nur dann richtig bewertet werden, wenn man ihre Nutzenwahrnehmung einbezieht und berücksichtigt, welche Varianten der CRA bislang in der Praxis existieren. Deshalb sollen in diesem Kapitel zuerst die verschiedenen Einschätzungen und Positionen zur CRA in den USA, in Europa und in Deutschland behandelt werden. Im Anschluss daran werden vier Beispiele vergleichender Bewertungen dargestellt und kommentiert.

3.1 Das politische Umfeld der CRA

3.1.1 CRA in den USA

Vergleichende Risikobewertungen zu einer Grundlage der Umwelt- und in gewisser Weise auch der Gesundheitspolitik zu machen, war eine Erfindung der Reagan / Bush-Administration in den USA der 80er Jahre. Seit dieser Zeit steht das Thema CRA in den USA auf der politischen Agenda.

Der Kongress und die Bundesbehörden, aber auch die Länder- und Kommunalpolitik haben sich intensiv mit dem Problem der Allokation von Ressourcen für die Sicherung von Umwelt und Gesundheit befasst. Ursprünglich ging es im Zusammenhang mit der Deregulation der Umweltpolitik um eine rationalere Prioritätensetzung ("the biggest bang for the buck"). Die CRA wurde als ein Instrument für die strategische Orientierung angesehen, d.h. als entscheidende Steuerungsgröße zwischen Risikoabschätzung und Risikomanagement.¹⁸

Die in den 70er Jahren gegründeten Einrichtungen – vor allem die Environmental Protection Agency (EPA), die Occupational Safety and Health Administration (OSHA) und die Consumer Product Safety Commission (CPSC) – wurden dafür kritisiert, dass sie große Anteile ihrer Budgets auf marginale Risiken verwendeten, die im Vergleich mit anderen Risiken kaum Bedeutung besäßen (vgl. Breyer 1993). Ihnen wurde vorgehalten, dass politische Erwägungen und eine zu starke Gewichtung der öffentlichen Risikowahrnehmung dazu geführt hätten, dass die objektive Risikolage (gemeint waren die Risiko-Abschätzungen von Experten) bei der Allokation der Mittel zur Gestaltung der Umweltpolitik weitgehend unberücksichtigt blieb.¹⁹

Das erste Projekt der vergleichenden Risikobewertung „Unfinished Business“ – initiiert von der EPA – diente dem Ziel: „to develop a ranking of the relative risks associated with major environmental problems that could be used as one of several important bases on which EPA could set priorities“ (U.S. EPA 1987).

Nach dem „Unfinished Business Project“ wurden in den USA eine ganze Reihe ähnlicher Projekte durchgeführt, die zugrundeliegenden Probleme solcher Vergleiche in-

¹⁸ "Comparative risk analysis is intended principally as a policy-development and broad resource-allocation tool." SAB 2000, S. 5.

¹⁹ Für England zeigen Untersuchungen von Hood et al. (2001) vom Centre of Risk Regulation der London School of Economics ähnliche Tendenzen wie in den USA. Strenge und Umfang der Risikoregulation wird wesentlich von der öffentlichen Risikowahrnehmung bestimmt.

tensiv diskutiert (vgl. Finkel & Golding 1994) und verschiedene Vorschläge zur Verbesserung gemacht (Konisky 1999), z.B. im Rahmen des Integrated Risk Project der EPA. Im Bericht „Toward Integrated Environmental Decision-Making“ des US-EPA Science Advisory Board (SAB 2000) wird dazu eine Vorgehensweise vorgeschlagen. Aufgrund der Merkmale: (1) Ist das Risiko niedrig, hoch oder ungewiss?, (2) Ist das Potenzial für eine Risikoreduktion hoch?, (3) Ist Risikoreduktion ökonomisch wirtschaftlich, gerecht und von der Öffentlichkeit unterstützt? soll eine von folgenden Maßnahmen ausgewählt werden:

- Monitoring (Kontrolle, Überwachung) des Risikos auf Veränderung,
- Forschung oder Aufklärung, und
- Risikomanagement mit Monitoring der Wirksamkeit.

Die Beteiligung aller gesellschaftlicher Gruppen wird ausdrücklich gefordert, allerdings in einer angemessenen Weise: „Obwohl Integrated Decision-Making die Beteiligung eines breiten Spektrums von Teilnehmern erfordert (z.B. von Wissenschaftlern, Ingenieuren, Wirtschaftsfachleuten, Entscheidungsträgern sowie der Öffentlichkeit), haben die verschiedenen Gruppen verschiedene Rollen. Wie von Wissenschaftlern nicht erwartet wird, die Ansicht der Öffentlichkeit einzubringen, wird von der Öffentlichkeit nicht erwartet, die technischen Analysen der Wissenschaftler durchzuführen.“

In den USA wurden bislang etwa 100 CRA durchgeführt. Praktisch in jedem Bundesstaat finden sich entsprechende Projekte.²⁰

Die Resultate der ersten CRA-Projekte in den USA zeigten zwar auch Mängel der CRA auf, dennoch wurde an dem Verfahren festgehalten. Das Ziel, die knappen Ressourcen nach der Bedeutung der Risiken zu verteilen, blieb unangetastet. Aus heutiger Sicht ist jedoch festzuhalten, dass die politische Auseinandersetzung über CRA anhält und insbesondere durch die Diskussion um das Vorsorgeprinzip neue Facetten aufweist.²¹ Es scheint, als ob sich die Positionen in den USA bzgl. der Anwendbarkeit von vergleichenden Risikobewertungen polarisieren.

Zwischen 1993 und 1998 befasste sich der US Kongress (Abgeordnetenhaus und Senat) in mehr als einem Dutzend legislativer Initiativen mit dem Thema Risikobewertung. Zweifelsohne waren die Bemühungen der Politik, umfassende Gesetzesstrukturen für die Risikobewertung und die CRA zu schaffen, vorteilhaft für die Weiterentwicklung der CRA. Nur die wenigsten dieser Initiativen wurden allerdings von beiden Kammern als Gesetz verabschiedet.

Die Fortschritte in der Anwendung und der daraus resultierenden politischen Legitimierung von CRA wurden durch *Executive Orders* erreicht. Die erste Initiative stammt von Reagan 1981, der mit der Executive Order (E.O.) 12291 (Regulatory

²⁰ Die beste Übersicht bietet der Environmental Defense Fund. CRA-Projekte finden sich aber auch in Kanada, Neuseeland, Ungarn, Polen, Tschechien und Südostasien.

²¹ Politisch wird ihr vor allem vorgeworfen, zu technokratisch ausgerichtet zu sein und die Partizipation der Öffentlichkeit an risikobezogenen Entscheidungen zu verhindern, oder zumindest zu erschweren. Als bedenklich wird außerdem angesehen, dass die CRA die Flexibilität des umweltpolitischen Handelns einschränken könnte, zu sehr Lebensstil-Risiken priorisiere und damit umweltbezogene Risiken von der Agenda absetze (vgl. Commoner 1992, Hornstein 1992, McCloskey 1994, siehe aber Minard 1993).

Planning Process, 50 Federal Register 1036) die Überwachung der Kosten-Nutzen-Analysen von gesetzlichen Regelungen an das Office for Management and Budget übertrug. Unter Clinton wurden die E.O. 12291 und E.O.12498 durch E.O.12866 und E.O.12875 ersetzt (Regulatory Planning and Review, 58 Federal Register 51735, 1993), wobei CRA explizit genannt wird, entsprechende Verfahren aber nicht definitiv vorgeschrieben werden. E.O. 12866 besagt, dass

„...each agency shall consider, to the extent reasonable, the degree and nature of the risks posed by various substances or activities within its jurisdiction....“

Darüber hinaus werden Kosten-Nutzen-Bewertungen als Teil der behördlichen Verpflichtungen angesehen:

„...[assess] all costs and benefits and select those approaches that maximize net benefits, incorporating economic and distributive impacts and equity considerations into the analysis. Agencies should adhere to the principle that the regulations should be pursued in the most cost effective manner, the benefits should justify the costs, and the regulation should be based on the best scientific, technical, economic and other information....“

Man kann also davon ausgehen, dass sich die CRA als Instrument der Prioritätensetzung in den USA etabliert hat.

3.1.2 CRA in Europa

In Europa hat die CRA bei weitem nicht die Bedeutung wie in den USA. Im ersten Bericht der EU-Kommission (2000) zur Harmonisierung der Risikoabschätzungsverfahren wird auf die methodischen Schwierigkeiten der CRA hingewiesen und festgehalten, dass es ganz unterschiedliche Einschätzungen des Nutzens von CRA für die Umweltpolitik gibt. Im Bericht selbst wird der Nutzen vorsichtig positiv eingeschätzt und eine weitere Befassung mit diesem Thema gefordert.

Trotzdem gibt es auch in Europa Ansätze für eine vergleichende Risikobewertung. In einigen Bereichen, wie z.B. bei der Prioritätensetzung im Bereich der Chemikalien wurden Risiko-Ranking Verfahren vorgeschlagen (vgl. Hansen et al. 1999). Das Risiko aufgrund einer Exposition des Menschen mit einer Chemikalie wird über die Exposition (näherungsweise über die Tonnage, die Nutzungsweise sowie physikalische Eigenschaften der emittierten Chemikalien) und über die Toxizität ausgedrückt und in R-Sätzen erfasst.²² Ähnliches findet sich im Chemie-Weißbuch der EU Kommission. Hier wird ein neues Bewertungssystem REACH („REACH“ für *Registration, Evaluation, Authorisation of Chemicals*) vorgeschlagen: Je nach Tonnage und Eigenschaften der Stoffe wird ein Registrierungs-, Evaluierungs- und Zulassungsprozess durchlaufen. Eine prioritäre Bewertung ist für die Chemikalien vorgesehen, die in Mengen von mehr als 100 t pro Jahr vorliegen. Gleiches gilt für Stoffe, die als gefährlich aufgrund von inhärenten Eigenschaften eingeschätzt werden (Stoffe mit erbgutverändernden Eigenschaften, sehr giftige Stoffe, Stoffe mit hoher Persistenz oder

²² The EURAM Rankings are intended to provide a reference and a focus for the discussions leading to the preparation of priority lists. They should be seen as a convenient tool for summarising the vast amount of information in IUCLID in such a fashion that an expert can use this data summary to identify those substances which are expected to exhibit the greatest risk to man or the environment. Siehe <http://ecb.jrc.it/existing-chemicals/>

Bioakkumulationsfähigkeit und Stoffe, deren Molekularstruktur Anlass zur Besorgnis gibt). Für besonders gefährliche Stoffe, die zu „großer Besorgnis Anlass geben“, soll ein Zulassungsverfahren durchgeführt werden. Zulassungspflichtig sollen persistente organische Schadstoffe (POPs) sowie Stoffe sein, die nachweislich kanzerogene, mutagene oder reproduktionstoxische Eigenschaften besitzen.

Ein gesamteuropäisches Projekt stammt von der Europäischen Umweltbehörde EEA, die 1995 das Dobris-Assessment²³ über den Zustand der europäischen Umwelt veröffentlichte. Hier wurden 56 verschiedene Umweltprobleme bewertet. Die EEA verwendete dabei die folgenden Attribute: (1) threat to sustainability; (2) impact of global problems in the European area; (3) prominence of a European aspect; (4) trans-boundary aspect; (5) long-term character; (6) risk for human health; (7) social or cultural impacts; (8) ecological damage und (9) economic loss.

Der Report zeigt nicht, wie diese Attribute angewandt und gewichtet wurden. Es wird eine Prioritätenliste vorgestellt, die selbst keine Abstufungen mehr enthält. Die EEA weist auf diese Weise die folgenden 12 Umweltprobleme als besonders dringlich aus: Klimaveränderung, Ozonloch, Verlust der Biodiversität, technische und natürliche Risiken, Versauerung, Bodenzon, Binnengewässer, Waldschäden, Meeres- und Küstenumwelt, Abfall, städtische Umwelt und chemische Stoffe. Ein zweiter Lagebericht wurde 2001 veröffentlicht; darin werden die Fortschritte bezüglich dieser Umweltprobleme bewertet.

Die EEA unterscheidet nicht klar genug Risiken – im Sinne von möglichen Ereignissen – von bereits vorhandenen Schäden. Ein Blick auf die Liste der Umweltprobleme macht darüber hinaus deutlich, dass die Probleme nicht nur in ihrem Begriffsumfang differieren, sondern sich auch verschiedene Positionen innerhalb der Kausalkette des Risikogeschehens – von der Verursachung bis hin zum Effekt – aufweisen. So finden sich in der Liste sowohl Ursachen von Risiken (z.B. Biotechnologie, Transport und Lagerung von gefährlichen Abfällen) als auch Auswirkungen (z.B. Bodendegradation, Ozonloch).²⁴

In jüngster Zeit haben sich verschiedene europäische Behörden mit der CRA auseinandergesetzt. So berief sich z.B. die European Medicines Evaluation Agency in 2001

²³ Englisch: <http://reports.eea.eu.int/92-827-5122-8/en/>

Deutsch: <http://reports.eea.eu.int/92-827-5122-8/de/>

²⁴ Nuclear accidents, Stratospheric ozone, depletion, increasing UV, Loss of biodiversity and genetic resources, Resources and quality of groundwater, Acidification, Hazardous waste (transport and storage), Climate change, Forest degradation, Waste disposal, Nuclear waste, Urban air quality, Nature conservation and sensitive ecosystems, Persistent air toxics, Industrial accidents, Direct pollution (discharges, dumping) to sea, Soil contamination from waste disposal, Conservation of threatened species, Tropospheric ozone increase and episodes, Waste production, Soil and resource contamination, Fragmentation and destruction of habitats, Eutrophication of surface waters, Riverine inputs into seas, Changes in hydrological regime, Management of large rivers and lakes, Desertification, Lack of water supply, Stress and degradation due to tourism, Food security, Urban waste, Growing vulnerability of complex systems, Bioaccumulation (metals, POCs), Energy security, Soil erosion, Risks of biotechnology, Microbiological pollution of surface water, Oil spills, Sea-level rise, Intensification of land-use, Introduction of new organisms, Shortage of industrial raw materials, Coastal erosion, Natural radioactivity (radon), Loss of agricultural land, Shift of biogeographical zones, Draining wetlands, Floods, droughts and storms, Non-ionising radiation, Societal health, Landscape modification, Noise, Occupational health, Loss of cultural heritage, Seismic activity volcanoes, Pests and locusts, Thermal pollution of waters.

auf ein „comparative risk/benefit reassessment“, um die Autorisierung für das Medikament Orlaam teilweise zu widerrufen.²⁵

In Großbritannien befasste sich das Advisory Committee on Pesticides mit der komparativen Risikobewertung und nahm dieses Thema in ihre Sitzungen von 2001 auf.²⁶

Die schwedische Regierung, genauer das Ministry of Education and Science, verwies in ihrer Stellungnahme zu dem 5. Rahmenprogramm der EU auf die fehlende Auseinandersetzung mit komparativer Risikobewertung²⁷

„...Furthermore, there is an important need for environmentally oriented socio-economic and behavioural research. A comprehensive perspective, comparative risk assessment and risk management as well as developing legal instruments, financial mechanisms and incentives are important aspects of environmental research....“.

Das Standing Committee on the Food Chain and Animal Health der Europäischen Kommission (General Health & Consumer Protection – Direktorat E) befürwortete im Oktober 2002 die Anwendung der komparativen Risikobewertung.²⁸

Die European Food Safety Authority fördert bei den Forschungsaktivitäten zu Acrylamid auch die komparative Risikobewertung über „Health risks from heat treated foods and food products“.²⁹

Im jüngsten Bericht (2003) der Europäischen Kommission (General Health & Consumer Protection - Direktorat C) über Risikobewertung für Tierpopulationen bejaht die Scientific Steering Committee Task Force die komparative Risikobewertung.³⁰

Insgesamt deutet sich in Europa ein Trend zu einer positiven Bewertung der CRA an.

3.1.3 CRA in Deutschland

In Deutschland ist die CRA bislang unüblich. Eine der Ausnahmen ist der Bericht der Arbeitsgruppe des Länderausschusses für Immissionsschutz (LAI 1992). Die Gruppe hat in mehrjähriger Arbeit „Beurteilungsmaßstäbe zur Begrenzung des Krebsrisikos durch Luftverunreinigungen“ erarbeitet und 1992 veröffentlicht. In dieser Studie wurden folgende Stoffe behandelt: Arsen und anorganische Arsenverbindungen, Asbest, Benzol, Cadmium und Cadmiumverbindungen, Dieselrußpartikel, PAK (Benzo[a]pyren) und 2,3,7,8-Tetrachlordibenzo-p-dioxin (TCDD). Die Bewertung basiert auf den „unit risk“-Werten, die das geschätzte zusätzliche Krebsrisiko eines Menschen nach konstanter Exposition über 70 Jahre gegenüber einer Konzentration von 1 µg Schadstoff pro m³ Luft beschreiben.

²⁵ <http://www.emea.eu.int/pdfs/human/press/pus/877601en.pdf>

²⁶ http://www.pesticides.gov.uk/committees/ACP/ACP-288_mins.htm

²⁷ <http://www.cordis.lu/fifth/src/ms-s-1.htm>

²⁸ http://europa.eu.int/comm/food/fs/rc/scfcah/plants/rap12_en.pdf

²⁹ http://europa.eu.int/comm/food/fs/sfp/fcr/acrylamide/study_area10.pdf

³⁰ http://europa.eu.int/comm/food/fs/sc/ssc/out325_en.pdf

CRA in Form programmatischer Risikovergleiche – wie sie in den USA durchgeführt wird – stößt jedoch in Deutschland auf erhebliche Bedenken.

Als Alternative wird eine Risikotaxonomie vorgeschlagen. So hat z.B. der WBGU 1998 einen Ansatz zur Risikoklassifikation vorgelegt, der nach sechs verschiedenen Risikotypen differenziert. Die Typen unterscheiden sich nach den Attributen: (1) Wahrscheinlichkeit, (2) Schadensausmaß, (3) Abschätzungssicherheit der Wahrscheinlichkeit, (3) Abschätzungssicherheit des Schadens, (4) Persistenz, (5) Ubiquität, (6) Verzögerungswirkung, (7) Irreversibilität und (8) Mobilisierungspotenzial. Eine Priorisierung der verschiedenen Typen wurde vom WBGU jedoch nicht angestrebt, vielmehr wurden aus der Risikoklassifikation verschiedene Risikomanagement-Strategien abgeleitet.

Ein interessanter Aspekt des WBGU-Verfahrens besteht darin, dass die vorgeschlagene Vorgehensweise erlaubt, unterschiedlichste Risikoquellen zu bewerten. Insofern ähnelt das Verfahren programmatischen Risikovergleichen. Im Unterschied dazu werden die Risikoquellen jedoch nicht „gerankt“, sondern verschiedenen Handlungsklassen zugeordnet, die jedoch keine Abstufungen hinsichtlich der Prioritätensetzung beinhalten. Das WBGU-Verfahren ist jedoch nicht ohne Schwierigkeiten anzuwenden. Denn konkrete Risiken lassen sich den Risikotypen des WBGU nicht so einfach zuordnen, wie es auf den ersten Blick scheint, und die Managementstrategien für die unterschiedlichen Risikotypen bleiben keineswegs so unterschiedlich, wenn man ihre Operationalisierung versucht (wie kann zum Beispiel zwischen Vorsorge und Langfristverantwortung unterschieden werden, die als zwei verschiedene Optionen gesehen werden?).

Ein zweiter Ansatz, der zumindest fragmentarische Vergleiche beinhaltet, wurde vom Rat von Sachverständigen für Umweltfragen (SRU 1999) vorgelegt. Danach werden Gesundheitsbeeinträchtigungen durch Allergien, durch ultraviolette Strahlung oder durch Lärm in der Öffentlichkeit und in der Politik eher unterschätzt. Hingegen stellt der SRU eher eine Überschätzung der Gesundheitsrisiken durch hormonähnlich wirkende Stoffe und bei Ursachen des Krankheitsbildes der Multiplen Chemikalien-Überempfindlichkeit (in der öffentlichen Diskussion) fest. Diese Beurteilungen werden von Expertenseite (den Mitgliedern des SRU) getroffen, eine explizite Vergleichsmethodik findet sich in dem Bericht nicht.

Auch im Aktionsprogramm Umwelt und Gesundheit (APUG 1999) des Bundesministeriums für Gesundheit und des Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit werden die wichtigsten gesundheitsrelevanten Umweltprobleme (Lärm, Luft, Strahlung, Wasser, Boden und Altlasten, Siedlungshygiene, Lebensmittel und Bedarfsgegenstände, Produkte, Störfälle und ausgewählte Umweltgifte) diskutiert. Die Autoren des APUG entschließen sich aber nicht zu einem Risiko-Ranking.

Die Risikokommission mahnt an, Risiken in Zukunft systematisch zu bewerten. Die Vielzahl an Zuständigkeiten und Bearbeitungsebenen in den verschiedenen Institutionen führe dazu, „dass der Umgang mit möglichen Gesundheitsgefahren meist von einer zufälligen oder interessengeleiteten Problemwahrnehmung durch die Öffentlichkeit oder die Medien bestimmt wird. [...] Zögerliche und widersprüchliche Regulierungen durch die beteiligten staatlichen Stellen, selektive Information der Betroffenen und Kommunikationsdefizite haben zur Folge, dass auf der einen Seite relativ unbedeutende Risiken einen erheblichen Platz in der öffentlichen Wahrnehmung einnehmen.“

men, während andererseits z.T. schwerwiegende Risiken unterschätzt oder sogar verdrängt werden“ (Risikokommission 2003, 20).

Die Funktion von Risikovergleichen sieht die Kommission zum einen bei der Risikoabschätzung und zum anderen bei der Risikokommunikation. Sie äußert sich wie folgt: „Risikovergleiche in der öffentlichen Risikokommunikation sind fragwürdig. Für eine rationale Prioritätensetzung beim Einsatz öffentlicher und privater Mittel zur Risikominderung oder -vermeidung sind sie aber essentiell. Die naturwissenschaftliche Risikoabschätzung kann hierzu lediglich quantitative Schätzungen der Eintrittswahrscheinlichkeit und (quantitative) Angaben zur Vertrauenswürdigkeit dieser Schätzung beitragen. Schon die Frage der Vergleichbarkeit unterschiedlicher adverser Effekte (Qualität des Risikos) übersteigt den Bereich des Empirisch-Kognitiven und ist überwiegend durch Wertevorstellungen bestimmt.“

Trotz aller Skepsis gegenüber der CRA scheint es fast unvermeidlich zu sein, Risiken zu vergleichen und sie zu priorisieren. Dennoch gibt es in Deutschland bislang kein einziges methodisch ausgearbeitetes Projekt zur CRA, das Risiken über verschiedene Medien und Endpunkte vergleicht, wie das in den USA der Fall ist.

Diese Abstinenz hängt unserer Auffassung nach einerseits mit der Regelungsichte zusammen, andererseits aber auch mit der hier verfolgten Regulationsphilosophie, die Risiken aus verschiedenen Regulationsbereichen (wie nicht-ionisierende Strahlung, ionisierende Strahlung und Chemikalien, Verkehr etc.) und aus verschiedenen Umweltpolitikbereichen (Naturschutz, Bodenschutz, Gewässerschutz, Klimaschutz und Luftreinerhaltung, Abfallwirtschaft sowie Gefahrenstoffe) nicht zusammen betrachtet. Im Vordergrund der deutschen Regulierungspraxis steht außerdem die Bewertung der Sicherheit, nicht die des Risikos. Eine CRA würde in Deutschland nur Sinn machen, wenn sie unter den Gesichtspunkten der Vorsorge bzw. der Risikominimierung durchgeführt würde.

3.2 Projekte als Fallbeispiele

3.2.1 US EPA „Unfinished Business“ Projekt

Ziele des Projekts

Die komparative Risikobewertung entstand 1987 aus dem Bedürfnis, die Risiken verschiedener Regulationsbereiche der US EPA (Environmental Protection Agency) für Ressourcen-Allokationszwecke zu priorisieren. Dies ging einher mit einer Überprüfung der Angemessenheit der bisherigen Ausgaben der EPA für die verschiedenen Umweltprogramme.

Vorgehensweise

EPA-Administrator Lee M. Thomas berief 1986 eine interne Task Force von 75 EPA Wissenschaftlern und Managern, um die „relative risks to human health and the environment posed by various environmental problems“ zu untersuchen. Die Arbeiten der Task Force wurden wie folgt aufgliedert und eingegrenzt:

1. Die gesamte Umweltproblematik (auch die Themen, die EPA nicht unbedingt kontrollierte) wurden in 31 Themenblöcke klassifiziert. Die Unterteilung wurde entsprechend der bestehenden Programmzugehörigkeiten und Regelungen durchgeführt. So sind z.B. die Themenblöcke „...criteria air pollutants, hazardous air pollutants, contaminants in drinking water, abandoned hazardous waste (e.g., Superfund) sites, pesticide residues on food, and worker exposures to toxic chemicals...“ aus den bestehenden Programmen ausgewählt worden.³¹
2. Vier Risikobereiche wurden differenziert: Krebsrisiko, Nicht-Krebs-Risiko, Ökologische Effekte und gesellschaftliche („Wohlfahrt“) Effekte. Es sei angemerkt, dass in „Unfinished Business“ die gesellschaftlichen Effekte nicht unbedingt die gesamte soziale Umwelt berücksichtigen, sondern vorwiegend die materiellen Schäden, wie z.B. „...damages to property, goods and services or activities to which a monetary value can often be assigned. These include natural resources (e.g., crops, forests and fisheries), recreation (e.g., tourism, boating), materials damage and soiling (e.g., building materials), aesthetic values (e.g., visibility) and various public and commercial activities....“ Die Risikobereiche wurden von vier verschiedenen Arbeitsgruppen separat evaluiert. Eine Zusammenfassung der Risikobereiche blieb aus.
3. Im Projekt wurden die folgenden Aspekte nicht berücksichtigt: (1) die ökonomische oder technische Kontrollierbarkeit des Risikos, (2) die qualitativen Aspekte der Risiken, die von der Öffentlichkeit oft als besonders wichtig empfunden werden, wie z.B. der Grad der Freiwilligkeit des Risikos, der Bekanntheit oder der Fairness sowie (3) der gesellschaftliche Nutzen der Aktivitäten, die mit den Risiken zusammenhängen.

Die weiteren Schritte in „Unfinished Business“ befassten sich mit der Problematik der Integration von quantitativen Daten und qualitativen Informationen („expert judgment“) sowie mit der Festlegung einer gemeinschaftlichen analytischen Richtlinie für die vier Arbeitsgruppen. Innerhalb der Arbeitsgruppen wurde jeweils eine Metrik ausgewählt, wobei für Krebsrisiko (Anzahl der Fälle) und gesellschaftliche Effekte (monetäre Verluste) bereits etablierte Vorgehensweisen bestanden, für Nicht-Krebs-Risiken und ökologische Effekte aber neue Metriken erst entwickelt werden mussten.

Für jeden der 31 Themenblöcke wurde in jeder der 4 Arbeitsgruppen ein „Summary Sheet“ angelegt, in dem die bestehenden Risikoinformationen sowie deren Quellen und Unsicherheiten festgehalten wurden. Die summary sheets wurden von den zuständigen Programmeinheiten der EPA ausgefüllt und innerhalb der Arbeitsgruppe evaluiert. Obwohl die Arbeitsgruppen nach Objektivität strebten, („successfully developed a spirit of striving for objectivity in this project“), waren subjektive Bewertungen – auch der Experten – nicht zu vermeiden.

Diese Einschränkungen im „comparative risk project“ von „Unfinished Business“ wurden aber nicht als gravierende Qualitätsminderungen betrachtet, die die Ergebnisse des Rankings in Frage stellen:

³¹ Aus diesem Grunde wurde auch der Begriff „programmatic comparative risk assessment“ benutzt, um eine komparative Risikobewertung zu beschreiben, welche die Risiken eines Programms oder auch über verschiedene Programme berücksichtigt.

„...The method we used to compare environmental problem areas can best be described as systematically generating informed judgments among Agency experts. About 75 career managers and experts representing all EPA programs participated over a period of about nine months. The participants assembled and analyzed masses of existing data on pollutants, exposures, and effects but ultimately had to fill substantial gaps in available data by using their collective judgment. In this sense the project represents expert opinion rather than objective and quantitative analysis. But despite the difficulties caused by lack of data and lack of accepted risk assessment methods in some areas, the participants feel relatively confident in their final relative rankings....”

Die Vorgehensweise im Unfinished-Business-Projekt wird in Abbildung 5 noch einmal zusammengefasst.

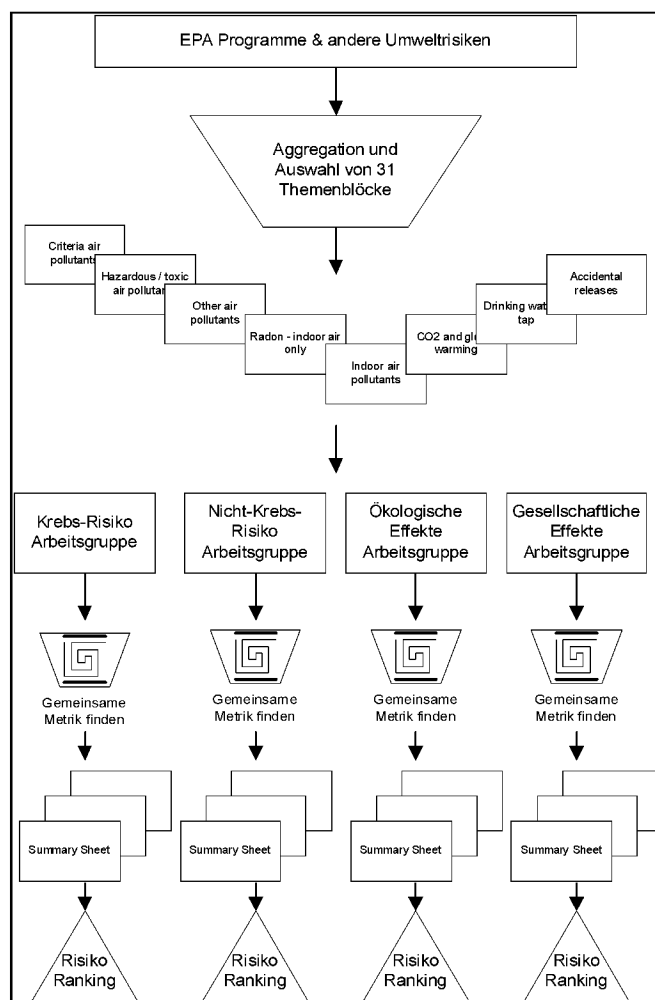


Abbildung 5: Vorgehensweise im Unfinished Business Projekt

Auswahl der zu betrachtenden Schutzgüter, Einwirkungen und Auswirkungen

Bei der Entwicklung der 31 Themenblöcke und bei der Auswahl der Hauptattribute für die CRA in den 4 Arbeitsgruppen spielten Konventionen eine Rolle. Laut „Unfinished Business“ hätten es genauso gut 11 oder 111 Themenblöcke sein können. Das Projektteam entschied sich für ca. 30, um das Bestreben nach Detailliertheit mit der Notwendigkeit der Handhabbarkeit der Analyse auszubalancieren. Allerdings ist es nicht nebensächlich, wie kategorisiert wird. Denn je größer der Themenblock, d.h. je mehr einzelne Einwirkungen dazu zählen, desto größer ist die Chance, dass der Themenblock in der Rangliste einen vorderen Platz einnimmt.

Zusätzlich kann die Zusammenstellung zu den Themenblöcken zu Überschneidungen in den Einwirkungen und Auswirkungen führen, so dass es zu einem „double-counting“ von Risiken kommen kann (siehe auch Morgan et al. 2000).

„...Some of these problems are diffuse mixtures of sources effects and a variety of different problems. This has led to some "double counting" of risks in more than one problem area. For example health risks from inactive hazardous waste sites (problem area #17) often result when people either drink contaminated water or inhale volatile toxic chemicals. Such health risks will be double counted by the drinking water problem area (#15) or by the hazardous air pollutant problem area (#2)..."

Eine Eigenschaft umwelt-relevanter Stoffe kann sein, sich rasch über verschiedene Umweltmedien zu verbreiten. Auch diese Eigenschaft kann zu einem „double-counting“ der Risiken führen.

„...when we were able to, we tried to take account of intermedia transfers and secondary effects. Some pollutants easily cross the boundaries of environmental media (e.g., air, water). Sulfur dioxide (SO₂) emissions ... covers risks from SO₂ in the air ... as well as damages from the deposit of sulfates on structures ... and from their eventual arrival in water ... Essentially, we attempted to follow the pollutant "from the cradle to the grave." Secondary effects occur when pollutants such as SO₂ are chemically transformed ... and do damage in a new form...."

Der Bericht betont, dass die Auswahl der Umweltprobleme auch anhand von konsistenteren Kriterien – wie Quellen, Kontaminanten, Pfade oder Rezeptoren – hätte vorgenommen werden können. Es wird außerdem erwähnt, dass die „intangible“-Effekt-Dimensionen, z.B. Grad der Freiwilligkeit oder Fairness, aufgrund mangelnder Konzepte, Daten und Metriken nicht in die Analysen integriert werden konnten.

Diese Einschränkungen zählen zu den bewusst in Kauf genommenen Problemen des Prozesses. Andere beruhen auf den mangelnden Datensätzen, die nur durch Expertenschätzungen – mit ihren bekannten Problemen – behoben werden konnten.

Dosis-Wirkungs-Beziehungen

Eine konsistente Berücksichtigung von Dosis-Wirkungs-Beziehungen sowie von Expositionen erwies sich als schwierig. Dazu heißt es im Bericht:

„...EPA analyses have made differing assumptions about emissions of various pollutants, the exposure that people, ecosystems and objects receive and dose/response relationships. Comparing estimates generated under incompatible assumptions was difficult. In some instances where

there were sufficient data, we tried to coordinate assumptions. In other instances, we simply tried to judge the magnitude of the bias produced by the incompatible assumptions and compared risks without reworking the underlying quantitative estimates..."

In manchen Fällen wurde versucht, die unterschiedlichen Dosis-Wirkungs-Beziehungen zu berücksichtigen. Arbeitsgruppen-übergreifende Abstimmungen fehlen jedoch, so dass die Risiken, die von den verschiedenen Arbeitsgruppen zu Krebs-Risiken, Nicht-Krebs-Risiken sowie zu ökologischen und gesellschaftlichen Effekten berechnet wurden, nicht auf einheitlichen Annahmen zu Dosis-Wirkungs-Beziehungen beruhen.

Auswahl der betroffenen Regionen und Bevölkerungsgruppen

In den Schlussfolgerungen des Berichts wird darauf hingewiesen, dass diese national angelegte Studie nicht unbedingt den örtlichen Gegebenheiten entspricht und lokale komparative Risikobewertungen erforderlich sind.

"...This analysis is not a guide to what may be the most serious problems in a particular area or for particular individuals. Any attempts to set local priorities should take into account local conditions..."

Weiterhin wurden Unterschiede in der Vulnerabilität verschiedener Bevölkerungsgruppen oder Ökosysteme nicht einbezogen.

Auswahl der zu betrachtenden Zeiträume

Die EPA hat in „Unfinished Business“ den Zeitrahmen auf den damaligen Status quo gesetzt:

"...To facilitate a fair representation of the damages from all environmental hazards, we tried to estimate all the damages that will occur from the problem as it exists now..."

Zukünftige Entwicklungen in den Gefahrenpotenzialen wurden demnach nicht berücksichtigt, ebenso nicht unterschiedliche Latenzzeiten bis zum Eintreten der adversen Wirkungen:

"...many of the effects associated with various air pollutants take place at essentially the same time as they are emitted into the environment ground water contamination from hazardous waste sites may not affect human health for years, even decades, after contaminants have leached from a site.... Instead of discounting or using some other method of indicating time preference, we simply presented damages in terms of their magnitude, sometimes with a notation about the time at which they occur...."

Wertung der Schäden

Während die Metrik für das Krebs-Risiko (Anzahl der Fälle) und die Metrik für die gesellschaftlichen Effekte (monetäre Verluste) sich auf etablierte Verfahren stützen konnten, mussten neue gemeinsame Metriken der komparativen Risikobewertung für ökologische Effekte und für Nicht-Krebs Risiken entwickelt werden.

Die sogenannten „intangiblen“ Effekte, wie z.B. „options values“ und „existence values“, die in den Nutzwert einer Ressource eingehen, wurden nur z.T. berücksichtigt.

„...In many instances, people value environmental quality to a greater degree than the tangible benefits they derive from a clean environment would seem to suggest. This is because of the intangible characteristics of the risks. People place intrinsic value on a clean environment. The public's vehement insistence on cleaning up contaminated ground-water aquifers even when they are not currently being used is an example of these important intangible values.

The fact that we did not fully consider these intangible characteristics in this project should in no way suggest that they are unimportant. In fact a number of studies have found that they are a significant component of what the public perceives as the total value derived from environmental protection. Rather we did not fully consider them here merely because not enough work has been done in this area to enable the project participants to assess them fairly for all the environmental problems examined in the project...."

Beteiligte / Stakeholder

In dem „Unfinished Business“ Projekt waren ausschließlich Wissenschaftler und Manager aus den Reihen der EPA beteiligt. Andere Stakeholder, insbesondere die allgemeine Öffentlichkeit, wurden nicht einbezogen.

Gesamtbewertung

Dem „Unfinished Business“ Projekt der US EPA kommt das Verdienst zu, sich erstmals gezielt mit den Fragestellungen komparativer Risikobewertung befassen zu haben. Dabei wurden die aus heutiger Sicht zu erkennenden Defizite zum Teil auch schon von den damals Beteiligten gesehen und bewusst als bestehende Probleme in Kauf genommen.

Von besonderer Bedeutung ist die Einteilung von Risiken in Themenblöcke, die sich an damaligen EPA-Strukturen orientierten; kombiniert mit den gewählten Risikobereichen führte sie zum Teil zur Doppelzählung von Risiken. Hier wird deutlich, dass entlang der gesamten Kausalkette des Risikogeschehens – von der Verursachung über die Exposition bis hin zum Effekt – Querbeziehungen (auch latente) zwischen Sachverhalten bestehen können, die für das Untersuchungsdesign komparativer Bewertungen und für die Beurteilung von Ergebnissen von entscheidender Bedeutung sein können.

Eine Analyse des Projekts verdeutlicht zudem die Wichtigkeit einheitlicher Annahmen zu Dosis-Wirkungs-Beziehungen, die Einbeziehung des Nutzens in Risikobetrachtungen, die Problematik der konsistenten Bewertung unterschiedlicher Schäden und dabei insbesondere die Frage, ob eine ausschließlich monetäre Bewertung unter gesellschaftlichen Aspekten ausreichend ist.

Bei der Vorgehensweise fallen folgende Besonderheiten auf:

- Die Risiken (Themenblöcke) wurden in vier Bereiche (Krebsrisiko, Risiko anderer Erkrankungen, ökologische Effekte, materielle gesellschaftliche Effekte) aufgeteilt und diese jeweils separat behandelt.
- Bei Krebsrisiken und materiellen gesellschaftlichen Effekten standen bekannte Metriken zur Verfügung, mit deren Hilfe die Risiken quantitativ bewertet werden konnten.
- Bei Risiken anderer Erkrankungen und bei ökologischen Effekten mussten Metriken erst entwickelt werden. In Anbetracht unzureichender Daten wurden diese Risiken nur qualitativ in 3 bzw. 6 Kategorien eingestuft.

In der Gesamtbewertung wurde keiner der vier Bereiche für wichtiger als die anderen angesehen, und es wurde auch nicht versucht, aus den vier Bereichen aggregierte Rangwerte zu ermitteln. Bei getrennter Betrachtung der vier Bereiche werden vier Cluster identifiziert (EPA 1987; Morgenstern & Sessions 1988):

- hoher Rang in drei der vier Bereichen oder mindestens mittlerer Rang in allen vier Bereichen;
- hoher Rang in Gesundheitsrisiken, aber niedriger Rang in ökologischen und materiellen Risiken;
- hoher Rang in ökologischen und materiellen Risiken, aber niedriger Rang in Gesundheitsrisiken;
- überall nur mittlerer oder niedriger Rang.

Davon werden die drei ersten Cluster als wichtig angesehen, der vierte als weniger wichtig.

Die folgende Tabelle gibt einen Überblick über die Vorgehensweise.

Tabelle 1: Vorgehensweise im US-EPA "Unfinished Business Project"

US-EPA "Unfinished Business Project"	
Zielsetzung	Prioritäten für Zuordnung von Ressourcen (Personal und Mittel): „public health“
Risikobereiche	Krebsrisiko Risiko anderer Erkrankungen ökologische Effekte materielle gesellschaftliche Effekte
Vergleichskategorien	„Themenblöcke“ entsprechend der internen Organisationsstruktur der EPA
Zeitlicher Rahmen	Künftige Schäden (z. B. Krebserkrankungen, die erst nach einer Latenzzeit auftreten) werden so behandelt, als ob sie sofort eintreten, d. h. sie werden nicht diskontiert.
räumlicher Rahmen	Regionale Besonderheiten werden ausdrücklich nicht berücksichtigt.
Attribute	Todesfälle, Erkrankungen, ökologische Effekte, materielle Effekte
Aggregation / Metrik	4 Risikobereiche (s.o.) getrennt; innerhalb jedes Bereichs Einstufung in Kategorien: Krebsrisiko: 26 quantitative Stufen Risiko anderer Erkrankungen: 3 Kategorien ökologische Effekte: 6 Kategorien materielle gesellschaftliche Effekte: 24 quantitative Stufen „informed judgments“, d. h. die Risiken wurden von Experten bewertet unter Berücksichtigung der zusammengestellten Fakten einschließlich der Unsicherheiten. Bei (in vielen Fällen) fehlenden Daten und Methoden wurde der Rang des Risikos subjektiv geschätzt.
Gesamtbewertung	vier Cluster: - hoher Rang in drei der vier Bereiche oder mindestens mittlerer Rang in allen vier; - hoher Rang in Gesundheitsrisiken, aber niedriger Rang in ökologischen und materiellen Risiken; - hoher Rang in ökologischen und materiellen Risiken, aber niedriger Rang in Gesundheitsrisiken; - überall nur mittlerer oder niedriger Rang.
Unsicherheiten	Unsicherheiten wurden prinzipiell berücksichtigt, konnten aber vielfach wegen mangelnder Daten nicht quantifiziert werden. Sie wurden qualitativ bewertet.
Ergebnisdarstellung	Ranglisten

3.2.2 Comparative Risk Framework Methodology and Case Study (US-EPA)

Hintergrund

Der Bericht (US-EPA, *Comparative Risk Framework Methodology and Case Study* 1998) wird ausdrücklich als Entwurf bezeichnet und stellt noch keine ausgereifte Methodik dar.³² Dies kommt auch im Review zu diesem Bericht (US-EPA Science Advisory Board 1999) zum Ausdruck.

³² »INTERNAL DRAFT: DO NOT CITE OR QUOTE – This document is a preliminary draft. It has not been formally released by the U.S. Environmental Protection Agency and should not at this stage be construed to represent Agency policy. It is being circulated for comments on its technical accuracy and policy implications.«

Ziel der Studie

Anhand eines hypothetischen Fallbeispiels wird die Methodik der EPA zur komparativen Risikobewertung vorgeführt. Im Fallbeispiel geht es um den Vergleich verschiedener Verfahren zur Trinkwasser-Entkeimung. Hintergrund ist die Erkenntnis, dass bei der bisher vorherrschenden Chlorierung des Wassers organische Chlorverbindungen gebildet werden, die ihrerseits Gesundheitsrisiken verursachen können. Daher sind sowohl die Risiken der Erkrankung durch bakterielle Infektionen bei zu schwacher Entkeimung als auch die Risiken der Erkrankung durch Entkeimungs-Beiprodukte (disinfection byproducts, DBP) zu berücksichtigen und für etablierte und alternative Entkeimungsverfahren zu vergleichen.

Modellansatz

Der prinzipielle Modellansatz berechnet einerseits die Keime, die – abhängig von der Rohwasser-Qualität – nach der Entkeimung noch im Trinkwasser vorhanden sind, andererseits die durch die Entkeimungsverfahren entstehenden chemischen Verbindungen. Für beide werden Gesundheitseffekte berechnet und in eine gemeinsame Gesundheits-Metrik umgerechnet. Die hierfür verwendete Metrik ist „Quality Adjusted Life Year“ (QALY). Abbildung 6 zeigt ein Schema des Modellansatzes.

Auswahl der Technologie

Es geht laut Aufgabenstellung um eine Bewertung von Verfahren und Standards zur Trinkwasser-Entkeimung. Entsprechend wird das zur Zeit standardmäßig eingesetzte Entkeimungsverfahren mit zwei modifizierten Verfahren verglichen. Einflüsse des Wasserverteilungssystems werden vernachlässigt (dies wird kritisiert).

Auswahl der zu betrachtenden Schutzgüter, Einwirkungen und Auswirkungen

Es wird von bekannten schädlichen Folgen von Keimen und Substanzen ausgegangen. An betroffenen Schutzgütern wird ausschließlich die menschliche Gesundheit betrachtet.

In der Fallstudie werden berücksichtigt:

- nur der direkte Pfad durch unmittelbaren Genuss des Trinkwassers (und nicht z.B. beim Waschen, Duschen und Baden über Haut oder Inhalation);
- nur die direkten Folgen für die das Trinkwasser genießende Person; Infektionen Dritter durch bereits Infizierte werden angesprochen, (hier) aber nicht berücksichtigt; diese Einschränkung wird vom US-EPA Science Advisory Board (SAB) (1999) kritisiert;
- nur die Auswirkungen des Normalbetriebs, aber nicht die Auswirkungen von Störfällen; diese Einschränkung wird vom US-EPA SAB (1999) kritisiert, da Störfälle, d. h. Fehler in den Entkeimungsanlagen, in der Vergangenheit mehrfach zu epidemischen Infektionen geführt hatten.

Auf Grund der Aufgabenstellung ergeben sich außerdem folgende Einschränkungen:

- keine Wirkungen, die über die direkt betroffenen Personen hinaus gehen, insbesondere keine Umweltauswirkungen;
- nur direkte Auswirkungen der Entkeimungsverfahren, keine Auswirkungen, die sich z.B. aus dem Bau der Entkeimungsanlagen oder der Herstellung der Entkeimungsanlagen ergeben können.

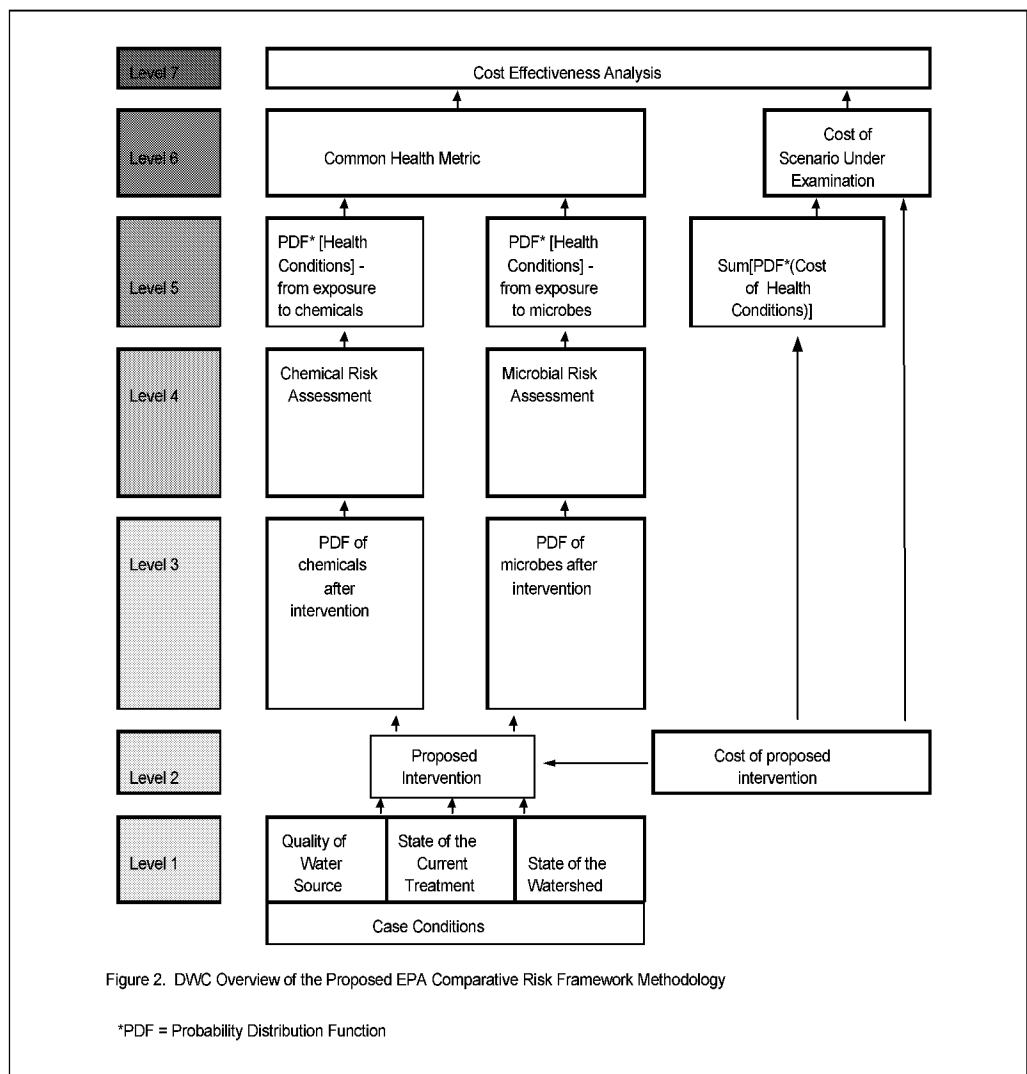


Abbildung 6: Schematische Darstellung des Modellansatzes (nach US-EPA Science Advisory Board 1999)

Auswahl der betroffenen Regionen und Bevölkerungsgruppen

Die Berechnungen werden für einen bestimmten Standort durchgeführt. Es wird zum einen die Normalbevölkerung, zum anderen eine besonders sensitive Gruppe immungeschwächter Menschen betrachtet.

Dies wird vom US-EPA SAB (1999) kritisiert: Um eine Standardsetzung begründen zu können, müssten ein repräsentatives Kollektiv von Standorten sowie repräsentative Bevölkerungsgruppen einbezogen werden. Die in der Fallstudie als besonders sensitiv berücksichtigte Bevölkerungsgruppe habe einen zu geringen Anteil an der Bevölkerung, so dass, wenn besondere Sensitivität nur bei dieser Bevölkerungsgruppe berücksichtigt werde, die Effektivität besserer Wasseraufbereitung unterschätzt werden könne. Diese Kritik bezieht sich auf folgende Eigenschaft der Vorgehensweise: Für jede der im Vergleich untersuchten Alternativen werden die QALYs der Gesamtbevölkerung ermittelt als Summe der QALYs der nicht besonders sensitiven Personen und der QALYs der besonders sensitiven Personen. Ist die Gruppe der besonders sensitiven Personen klein, beispielsweise 1% der Gesamtbevölkerung, dann würde eine Alternative, bei der die QALYs dieser Gruppe um z.B. 20% abnehmen würden, während die übrige Bevölkerung nicht betroffen wäre, in der Gesamtbeurteilung nur zu einer Abnahme von $1\% \times 20\% = 0,2\%$ führen.

Dosis-Wirkungs-Beziehungen

Die Berechnungen in der Fallstudie gehen von bekannten (oder als nachgewiesen angesehenen) Dosis-Wirkungs-Beziehungen aus. Allerdings sind diese mit erheblichen Unsicherheiten verbunden, z. B. hinsichtlich der Übertragbarkeit von Tierversuchen auf den Menschen. Näheres hierzu siehe Abschnitt „Unsicherheiten“ (S.52).

Wertung der Schäden

Um einen Vergleich unterschiedlicher Verfahren zu ermöglichen, müssen unterschiedliche Schäden auf einen einheitlichen Vergleichsmaßstab umgerechnet werden. Dies ist ein notwendiger und fundamentaler Bewertungsschritt, der besonderer Transparenz bedarf, weil er zwangsläufig subjektive Wertungen einschließt.

In *Comparative Risk Framework Methodology and Case Study* ist dies das „Quality Adjusted Life Year“ (QALY). Diesem Ansatz liegt der Gedanke zu Grunde, dass es einen Unterschied macht, ob ein gesunder Mensch oder ein durch Krankheit geschwächter Mensch durch eine Einwirkung stirbt. Der Ansatz versucht dies dadurch zu berücksichtigen, dass jedem Lebenszeitabschnitt ein Gewichtungsfaktor zwischen 0 und 1 zugeordnet wird, wobei 1 „perfekte Gesundheit“ und 0 „Tod oder krankheitsbedingter vollständiger Verlust an Lebensqualität“ bedeutet. Mit diesem Ansatz können Erkrankungen und Todesfälle auf ein gemeinsames Maß umgerechnet werden.

Bei diesem Vorgehen treten die folgenden Probleme auf (Hubbel 2002):

- Wer definiert den Lebensqualität-Faktor für eine gegebene Erkrankung? Was ist perfekte Gesundheit? (Gibt es das überhaupt?)
- Gibt es nicht auch Erkrankungen, die als „schlimmer als der Tod“ gewertet werden? Müsste diesen ein negativer Gewichtungsfaktor zugeordnet werden?

- Die zweite Frage zeigt, dass das QALY-Konzept implizit eine WTP (willingness to pay) - Frage enthält: „Wie viel Lebenszeit wärest du bereit, zu opfern, wenn dir diese Krankheit erspart bliebe?“ Hierbei ist allerdings unklar, ob damit Lebenszeit aus dem gleichen Lebensabschnitt gemeint ist (was eine entsprechende Vorverlegung des Alterungsprozesses bedeuten würde) oder eine Vorverlegung des Todeszeitpunkts bei ansonsten gleichem Lebensverlauf. Der Bewerter muss also eine konkret vorstellbare Situation (Erkrankung) mit einer eher theoretischen Situation (ungenau spezifizierter Lebenszeitverlust) vergleichen.
- Die WTP-Frage zieht eine Diskontierungs-Frage nach sich: „Würde es einen Unterschied machen, in welchem Lebensalter die Krankheit eintritt?“
- Wie müssen akute Symptome bewertet werden? Diese würden wegen ihrer Kürze u. U. geringer bewertet als dies der subjektiven Einschätzung der Betroffenen entspricht (US-EPA SAB 1999; Hubbel 2002, S.7).
- Bei der Gewichtung des möglichen Lebensdauer-Verlusts infolge akuter Belastungen muss berücksichtigt werden, dass möglicherweise die besonders empfindlichen Personen genau deswegen besonders empfindlich sind, weil sie zuvor langfristig einer chronischen Belastung ausgesetzt waren. Reduziert man die durch akute Belastung verlorenen Lebensjahre wegen des bestehenden schlechten Gesundheitszustandes, ohne zu berücksichtigen, dass der schlechte Gesundheitszustand durch chronische Belastung aus derselben Quelle verursacht wurde, so wird nicht nur die chronische Belastung ignoriert, sondern auch noch die akute Belastung unzulässigerweise geringer gewichtet. Dies ist kein Fehler des QALY-Ansatzes, sondern zeigt die Notwendigkeit, Erkrankungen und Todesfälle im Zusammenhang und nicht als separate Endpunkte zu betrachten (Hubbel 2002, S. 9-10). „Verursacht die Belastung den Tod von bereits erkrankten Menschen, oder verursacht die Belastung die Erkrankung, die zum Tod führt?“ (S. 21).
- Die QALY-Metrik setzt einige restriktive Annahmen voraus (Hubbel 2002, S. 5):
[...] QALYs are consistent with von Neumann-Morgenstern utility theory only if one imposes several restrictive assumptions, including independence between longevity and quality of life in the utility function, risk neutrality with respect to years of life, and constant proportionality in tradeoffs between quality and quantity of life.

Unsicherheiten

Die Fallstudie berücksichtigt im wesentlichen Unsicherheiten numerischer Parameter. (Deren Werte können relativ leicht bestimmt und ihr Einfluss auf die Ergebnisse kann relativ leicht ermittelt werden.) Das US-EPA Science Advisory Board (1999, S.29) kritisiert, dass keine Modellunsicherheiten berücksichtigt wurden. Der Bericht solle aber zumindest eine qualitative Diskussion der möglichen Größenordnung der Modellunsicherheiten enthalten. Außerdem wird empfohlen, bei der Analyse von Unsicherheiten grundsätzlich zwischen Variabilität (in Form einer Verteilung z. B. über eine Population) und eigentlicher Unsicherheit zu unterscheiden, und diese Unterscheidung auch bis in die Ergebnisdarstellung beizubehalten.

Ein Beispiel für eine Modellunsicherheit, die nicht nur hier, sondern generell bei Risikoabschätzungen eine Rolle spielt: Die Extrapolation der Dosis-Wirkungs-Beziehungen von Tierversuchen mit hohen Dosen auf niedrige Dosen beim Menschen ist mit hohen Unsicherheiten verbunden.

Weiter besteht noch Forschungsbedarf hinsichtlich Methoden, um aus berechneten Risiken („Level 4“ in Abbildung 6) die Wahrscheinlichkeitsverteilung von Gesundheitszuständen („Level 5“ in Abbildung 6) zu berechnen (US-EPA SAB 1999, S.11).

Gesamtbewertung

Etliche der bei komparativen Risikobewertungen auftretenden Probleme sind wegen der Aufgabenstellung in dieser Fallstudie nicht vorhanden: Es sind Alternativen zu vergleichen, die einen *gleichartigen Nutzen* erzielen sollen und mit *gleichartigen Risiken* behaftet sind. So lassen sich die bakteriellen Risiken zweier Alternativen direkt und problemlos miteinander vergleichen; dies gilt ebenfalls für die chemischen Risiken. Probleme treten erst auf, wenn die eine Alternative mit einem höheren chemischen, die andere mit einem höheren bakteriellen Risiko verbunden ist. Dann müssen verschiedenartige Risiken verglichen werden, die u. U. auch noch für verschiedene Bevölkerungsgruppen unterschiedlich groß sind. Hier kommt der gemeinsamen Metrik die entscheidende Rolle zu.

Das US-EPA Science Advisory Board (1999) bewertet die „Comparative Risk Framework Methodology“ als ein „potentially powerful tool that is well suited to helping identify important variables that are presented by complex environmental problems“ (S.10). Das QALY-Konzept wird grundsätzlich befürwortet, jedoch müssten auch alternative Metriken verglichen werden, um die bestgeeignete auszuwählen, und für die ausgewählte Metrik verteidigbare Gewichte entwickelt werden. Die Wahl der Metrik müsse klar begründet werden.³³

Diese Forderungen entspringen der Tatsache, dass die Basis des Risikovergleichs in der Metrik liegt und auf Wertungen basiert, die nicht naturwissenschaftlich hergeleitet werden können, sondern anders begründet werden müssen.

Schließlich muss die Unsicherheit der Ergebnisse bei einer Entscheidungsfindung unbedingt berücksichtigt werden. Es ist nicht auszuschließen, dass bei einem Vergleich zweier Risikowerte die Unsicherheit eines oder beider Werte größer ist als der Unterschied zwischen diesen Werten. Ist dies der Fall, kann daraus keine Entscheidung abgeleitet werden.

Die folgende Tabelle gibt einen Überblick über die Vorgehensweise:

³³ »Even though the Committee is not questioning whether the QALY is the most suitable metric, it believes that alternatives should be considered, adequately discussed in the document, and a clear case made for the metric selected.« (US-EPA Science Advisory Board, 1999, S.10)

Tabelle 2: Vorgehensweise in US-EPA "Comparative Risk Framework Methodology and Case Study"

US-EPA "Comparative Risk Framework Methodology and Case Study"	
Zielsetzung	Vergleich alternativer technischer Verfahren „public health“ und Individualschutz
Risikobereiche Vergleichskategorien	Gesundheitsrisiken durch Trinkwasser-Aufbereitung verschiedene technische Verfahren (in Form von Szenarien)
zeitlicher Rahmen	hinsichtlich der Emissionen die Lebensdauer der betrachteten Anlage, hier auf 20 Jahre festgesetzt; hinsichtlich der Gesundheitseffekte die Lebenszeit der exponierten Personen
räumlicher Rahmen	Versorgungsbereich der Anlage
Attribute	Todesfälle, Erkrankungen
Aggregation / Metrik	Verschiedene Erkrankungen und Todesfälle werden mit der QALY-Metrik aggregiert.
Gesamtbewertung	implizit multiattribut durch Verwendung der QALY-Metrik
Unsicherheiten	Für die in die Rechnungen eingehenden numerischen Parameter wurden sowohl Variabilität als auch Unsicherheit berücksichtigt. Modell- Unsicherheiten wurden nicht berücksichtigt. Die resultierenden Unsicherheiten der Ergebnisse wurden über Monte-Carlo-Simulation berechnet. Die Unsicherheiten sind jedoch nur für die Detail-Ergebnisse angegeben, nicht in der zusammengefassten Ergebnisdarstellung.
Ergebnisdarstellung	verschiedene Detailergebnisse, aufgeschlüsselt nach den - zu vergleichenden Alternativen, - verschiedenen Erkrankungen und - verschiedenen Bevölkerungsgruppen und mit Varianten für unterschiedliche Diskontierung Gesamtergebnis als Gegenüberstellung der aggregierten QALY-Werte für die Alternativen

3.2.3 Projekt "Global Burden of Disease (GBD) Study"

Ziele des Projekts

Von der WHO wurde 1992 zusammen mit der *Harvard School of Public Health* und der Weltbank das Projekt „Global Burden of Disease (GBD)“ initiiert (GBD 1996). Ziel war eine umfassende Analyse der Belastung der Weltbevölkerung durch Krankheiten und Verletzungen. Die betrachteten Endpunkte waren vorzeitiger Tod und Behinderungen, jeweils unterschieden nach Geschlecht und Altersgruppen der Betroffenen und getrennt für acht geographische Weltregionen. Unter anderem sollten für mehr als 100 Krankheiten und Verletzungen die Sterblichkeitsraten, fast 500 Erscheinungsformen (Verläufe und Folgen) sowie die Beiträge der zehn wichtigsten Risikofaktoren ermittelt werden. Als Blick in die Zukunft wurden Szenarien für die Entwicklung von vorzeitigen Todesfällen und Behinderungen entwickelt. In einem Folgepro-

jekt, GBD2000, wurden die Methoden anhand der gewonnenen Erkenntnisse z.T. weiterentwickelt und auch für nationale Zwecke nutzbar gemacht (GBD 2000).

Ziel des Projektes war es, verlässliche Grundlagen zu schaffen, damit Gesundheitsbehörden Datensätze unterschiedlicher Güte und Vollständigkeit für eine globale Bewertung des Gesundheitszustandes und der Ursachen von Krankheit und Behinderung und für eine zukunftsweisende Gesundheitspolitik nutzen können. Lopez und Murray (1998) beschrieben das Ziel als:

„... to weave these disparate data sets of variable reliability and completeness into a global assessment of health conditions and the causes of disease and injury to guide public policy into the next century....“

Aus diesem Gesamtziel wurden drei spezifische Zielstellungen für das GBD-Projekt abgeleitet:

1. Eine systematische Einbeziehung von weiteren Informationen – außer der reinen Anzahl von Todesfällen – in die Bewertung des Gesundheitsstatus. Hierzu sollte als zeitbasiertes Maß die Anzahl verlorener gesunder Lebensjahre verwendet werden, entweder verursacht durch vorzeitigen Tod oder verstanden als die Lebensjahre, die mit einer Behinderung – aufgrund einer Krankheit oder aufgrund einer Verletzung – verbracht werden. Dabei wird das Ausmaß der Behinderung durch einen Gewichtungsfaktor justiert.
2. Alle Schätzungen und Hochrechnungen sollten auf möglichst objektiven epidemiologischen und demographischen Methoden basieren.
3. Die Messung der Belastung durch Krankheiten oder Verletzungen sollte mittels einer Metrik erfolgen, die es auch ermöglicht, die Kosten-Effektivität von Maßnahmen oder Eingriffen abzuschätzen.³⁴ Die gewählte Metrik ist die Anzahl der Behinderungs-justierten Lebensjahre, besser bekannt als DALY (Disability-Adjusted Life Years).

Vorgehensweise

Für die Studie wurde ein neuartiger quantitativer Indikator entwickelt und standardisiert, der die Belastung durch vorzeitige Sterblichkeit und nicht-tödliche Krankheitsbilder quantitativ erfasst. Dieser Indikator beruht auf dem Prinzip der verlorenen Lebensjahre, justiert nach der Lebensqualität sowie nach der Beeinträchtigung und Behinderung durch Krankheit oder Unfälle. Der als DALY bezeichnete Indikator ist die Summe der aufgrund vorzeitiger Sterblichkeit verlorenen Jahre (years of life lost, YLL) sowie der mit einer Beeinträchtigung oder Behinderung zu lebenden Jahre (years of life lived with disability, YLD), jeweils bezogen auf die gesellschaftlichen Verhältnisse:

$$DALY_i = YLL_i + YLD_i,$$

wobei i für den jeweiligen Krankheitszustand steht.

³⁴ Der Erfolg dieser Zielstellung resultierte jedoch in einer Übernahme und breiten Verwendung des DALY Indikators in großem Ausmaß, so dass dieser „objektive“ Indikator sich rasch zu einem normativen Maßstab entwickelte.

Ein DALY bedeutet ein verlorenes Jahr an Lebenszeit in Gesundheit. Die Vorzeitigkeit eines Todes wird ermittelt aus der jeweiligen Differenz zur statistischen Lebensdauer in der Gesellschaft mit der höchsten Lebenserwartung (Japan)³⁵. Die DALY-Werte wurden für 108 unterschiedliche, Tod oder Behinderungen verursachende Krankheiten und Verletzungsursachen gemäß den Kodierungen der ICD-9 (Ninth Revision of the International Classification of Diseases) berechnet. Diese Krankheiten und Verletzungsursachen – im Folgenden GBD-Ursachen genannt – wurden in drei Hauptkategorien gruppiert (ansteckende Krankheiten, nicht-ansteckende Krankheiten, Verletzungen), die wiederum in 9 Unterkategorien³⁶ gruppiert und in insgesamt 483 Kategorien von Folgeerscheinungen unterteilt sind.

Diese Vorgehensweise wird in Abbildung 7 verdeutlicht. Der linke Teil der Abbildung enthält eine schematische Darstellung des Vorgehens anhand der einzelnen Schritte. Der rechte Teil zeigt die quantitative Ausgestaltung der Schritte im GBD-Projekt.

Murray und Lopez (1996) beschrieben vier Ziele, die sie bei der Entwicklung des DALY Indikators verfolgten:

1. Die Entwicklung von in sich konsistenten Schätzungen der Sterblichkeit durch 107 Haupt-Todesursachen^{37,38}, getrennt nach Alter und Geschlecht sowie nach acht geographischen Regionen der Erde.
2. Die Entwicklung von in sich konsistenten Schätzungen des Auftretens, des Vorrhorrhens, der Dauer und der spezifischen Sterblichkeit für 483 behindernde Folgeerscheinungen, die aus den oben genannten Ursachen resultieren, getrennt nach Alter, Geschlecht und geographischer Region.
3. Die Abschätzung des Anteils der Sterblichkeit und Behinderung, die den 10 Haupt-Risikofaktoren zuzuschreiben sind, ebenfalls getrennt nach Alter, Geschlecht und geographischer Region.
4. Die Entwicklung von Szenarien für Sterblichkeit und Behinderung, getrennt nach Ursache und wiederum nach Alter, Geschlecht und geographischer Region.

³⁵ In der Anfangsphase des GBD Projekts wurde die japanische Lebenserwartung als ein real auftretender maximaler Wert genutzt. Es lassen sich jedoch auch eine theoretische Lebenserwartung für den Menschen und speziell eine praktische Lebenserwartung für eine bestimmte Bevölkerung aus den dynamischen Lebens- bzw. Todesstatistiken ableiten. Dies wurde z.T. in GBD2000 und für nationale Burden of Disease Projekte genutzt.

³⁶ Die Anzahl der Cluster für Todesursachen wurde auf 21 Unterkategorien für GBD2000 erhöht. http://www3.who.int/whosis/menu.cfm?path=whosis,burden,burden_manual&language=english, S. 17.

³⁷ Gemeint sind die GBD-Ursachen.

³⁸ Die Anzahl der untersuchten Krankheiten und Verletzungen als Todesursache wird unterschiedlich als 107 oder als 108 in der Originalliteratur wiedergegeben, wobei 108 meist in den späteren Referenzen genannt wird.

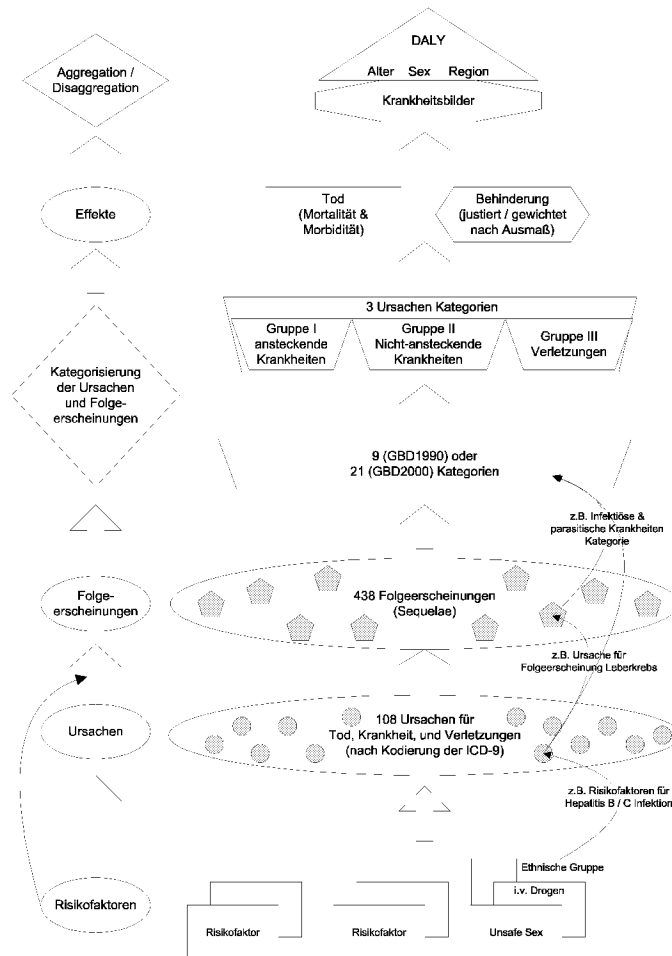


Abbildung 7: Vorgehensweise bei der Ermittlung von DALYs

Die Risikofaktoren lassen sich nach den Aspekten Ernährung, Konsum- und sonstige Gewohnheiten, Umwelteinflüsse, Arbeitsbedingungen und ähnliche Gesichtspunkte kategorisieren. Beispiele sind Vitaminmangel, Rauchen und Alkohol, Luftverschmutzung, Lärm, aber auch körperliche Inaktivität (Ezzati et al. 2002).

Wichtig ist, dass die Indikatoren von sozialen Werten abhängig sind. Benötigt werden daher Gewichtungen, beispielsweise für die Frage, wie schwerwiegend eine bestimmte Behinderung ist oder ob ein Lebensjahr, verbracht als junger Erwachsener, mehr oder weniger wert ist als ein Lebensjahr eines Kleinkindes. Diese Gewichtungen wurden durch Konsensverfahren ermittelt und transparent dargelegt.

Berücksichtigt man die Faktoren Alter, Geschlecht und zeitliche Diskontierung, so transformieren sich die zunächst einfachen Formeln für YLL und YLD,

$$YLL = N * L$$

$$YLD = I * DW * L,$$

mit

- N = Anzahl der Todesfälle
- L = standardisierte Lebenserwartung in Jahren einer Altersgruppe
- I = Anzahl der Inzidenzfälle
- DW = Behinderungsgewichtung (von 0 = gesund bis 1 = denkbar schlechter Gesundheitszustand),

sehr schnell in komplexere Gleichungen, die im Rahmen dieses Berichts nicht weiter vertieft werden. Es soll nur erwähnt werden, dass sie abhängig sind von der Diskontrate, von altersbedingten Gewichtungen und von anderen impliziten Annahmen (GBD 2001):

$$YLL = N C e^{(ra)} / (\beta + r)^2 \left[e^{-(\beta + r)(L+a)} [-(\beta + r)(L+a) - 1] - e^{-(\beta + r)a} [-(\beta + r)a - 1] \right]$$

mit

- YLL = Verlorene Lebensjahre
- N = Anzahl der Todesfälle
- C = altersgewichtete Korrektur-Konstante (GBD Standard ist 0.1658)
- r = Diskontrate (GBD Standard ist 0.03)
- a = Alter beim Eintritt des Todesfalls
- β = Parameter der Alters-Gewichtungs-Funktion (GBD Standard ist 0.04)
- L = Dauer der Behinderung oder der verlorenen Zeit aufgrund frühzeitigen Todes

und

$$YLD = I DW C e^{(ra)} / (\beta + r)^2 \left[e^{-(\beta + r)(L+a)} [-(\beta + r)(L+a) - 1] - e^{-(\beta + r)a} [-(\beta + r)a - 1] \right]$$

mit

- YLD = Lebensjahre mit einer Behinderung
- I = Anzahl der Vorfälle (incidents) während der Referenz Periode
- DW = Gewichtungsfaktor der Behinderung (Wert von 0 (keine Behinderung) bis 1 (denkbar schlechter Gesundheitszustand))
- C = altersgewichtete Korrektur-Konstante (GBD Standard ist 0.1658)
- r = Diskontrate (GBD Standard ist 0.03)
- a = Alter beim Eintritt des Todesfalls
- β = Parameter der Alters-Gewichtungs-Funktion (GBD Standard ist 0.04)
- L = Dauer der Behinderung oder der verlorenen Zeit aufgrund frühzeitigen Todes

Komparative Risikobewertung

Im GBD-Projekt bedeutet „Komparative Risikobewertung“ eine systematische Untersuchung der potenziellen Veränderungen im Gesundheitsstatus einer Bevölkerungsgruppe, wenn sich die Expositionsverteilung bezüglich eines Risikofaktors verändert. Dabei werden solche Veränderungen in Relation zu den Beiträgen durch andere Risikofaktoren betrachtet.

Um eine solche komparative Risikobewertung durchführen zu können, wurde das Konzept des „Attribuierbaren Anteils“ eines Risikofaktors, z.B. Rauchverhalten, an der Belastung durch eine bestimmte Krankheit, z.B. Lungenkrebs gewählt. Dieser „Attribuierbare Anteil“ (*attributable fraction*, AF) ist definiert als die prozentuale Redu-

zierung des Auftretens einer Behinderung oder der Sterblichkeit aufgrund eines Risikofaktors, wenn dieser Faktor eliminiert wird.

Die folgende Formel beschreibt diesen „Attribuierbaren Anteil“

$$AF = \frac{P(RR - 1)}{P(RR - 1) + 1} \quad (1)$$

mit

AF = attribuierbarer Anteil
P = Prävalenz oder Wahrscheinlichkeit einer Exposition
RR = relatives Risiko.

Alternativ lässt sich der Beitrag eines Risikofaktors auch ermitteln durch Vergleich der Konsequenzen aus der vorliegenden Expositionsverteilung (bezüglich dieses Risikofaktors) mit denjenigen Konsequenzen aus einer fiktiven, postulierten Verteilung, die einem bestimmten Szenario entspricht ("counterfactual distribution").

Für eine solche *counterfactual distribution* gilt:

$$AF = \frac{\sum_{i=1}^n P_i RR_i - \sum_{i=1}^n P'_i RR_i}{\sum_{i=1}^n P_i RR_i} \quad (2)$$

mit

AF = attribuierbarer Anteil
P_i = Anteil der Bevölkerung exponiert in Expositions-kategorie i
RR_i = relatives Risiko der Expositions-kategorie i
n = Anzahl der Expositions-kategorien oder -ebenen.

Falls die Verteilung kontinuierlich ist, wie z.B. bei Bluthochdruck in der Bevölkerung oder bei Luftverschmutzung, dann gilt

$$AF = \frac{\int_{x=0}^m RR(x)P(x) - \int_{x=0}^m RR(x)P'(x)}{\int_{x=0}^m RR(x)P(x)}$$

mit

AF = attribuierbarer Anteil
P(x) = Verteilung der Bevölkerung in Expositionsebene x
P'(x) = Counterfactual Verteilung der Bevölkerung für Expositionsebene x
RR(x) = relatives Risiko für Expositionsebene x
m = maximale Expositionsebene.

Kombiniert man den attribuierbaren Anteil eines Risikofaktors für eine bestimmte Ursache für Krankheit, Verletzung und Tod mit der gesamten Belastung für diese Ursache, so erhält man die attribuierbare Belastung für den Risikofaktor.

$$AB = AF \times B \quad (4)$$

mit

AB	= attribuierbare Belastung
AF	= attribuierbarer Anteil
B	= gesamte Belastung für eine Ursache.

Mit Hilfe dieser Metriken lassen sich verschiedene Risikofaktoren vergleichen. Dadurch können die wichtigsten Risikofaktoren identifiziert und ggf. verringert werden.

Der vollständige Ablauf der komparativen Risikobewertung besteht aus folgenden Schritten:

1. Auswahl des (zu bewertenden) Risikofaktors,
2. Auswahl der relevanten Krankheiten und Verletzungen, die durch den Risikofaktor verursacht (oder beeinflusst) werden,
3. Auswahl einer geeigneten Variable für die Exposition,
4. Zusammenstellung von Daten zur Bevölkerungsverteilung, bezogen auf eine Exposition hinsichtlich des fraglichen Risikofaktors (1),
5. Spezifikation des Risikofaktors (Bezüge zu (2), beispielsweise Dosis-Wirkungs-Beziehungen),
6. Auswahl der Bezugsverteilung der Exposition (Basis für Risikovergleich),
7. Berechnung der Belastungen (aufgrund (2)),
8. Analyse von Unsicherheiten.

Die Schritte 1 - 6 ergeben zusammen mit Gleichung (2) bzw. (3) den Beitrag des jeweiligen Risikofaktors zu einer bestimmten Krankheit oder Verletzung. Hieraus und aus dem nachfolgenden Schritt 7 ergibt sich die dem Risikofaktor zuzuordnende Belastung aus Gleichung (4).

Für die Auswahl von Risikofaktoren einschließlich ihrer kausalen Zusammenhänge mit bestimmten Ursachen für Tod, Krankheit oder Verletzung muss eine durchdachte Strategie fallspezifisch entwickelt werden. Zu berücksichtigende Aspekte sind (Mathers et al. 2001):

- Criteria for identifying relevant studies
- Search strategy for studies
- Characteristics of excluded studies
- Description of studies, including methodological qualities
- Characteristics of included studies such as sample size and region
- Standard errors reported by studies
- Steps to assess and reduce random error
- Reported standard errors and other measures of uncertainty due to variability
- Meta-analysis if appropriate
- Steps to assess and reduce bias and to assess causality
- Confounding
- Selection bias
- Information bias (exposure and outcome measurement error)

Die Abbildung 8 zeigt die wesentlichen Eingabegrößen und die Ergebnisse einer solchen vergleichenden Risikobetrachtung. Durch Verwendung aktueller oder prognostizierter Verteilungen für einen Risikofaktor beim Vergleich mit Bezugsverteilungen ("counterfactual distributions") lassen sich sowohl die Einflüsse früherer oder aktueller Expositionen auf aktuelle Belastungen durch Krankheit als auch Einflüsse aktueller und künftiger Expositionen auf künftige Belastungen ermitteln.

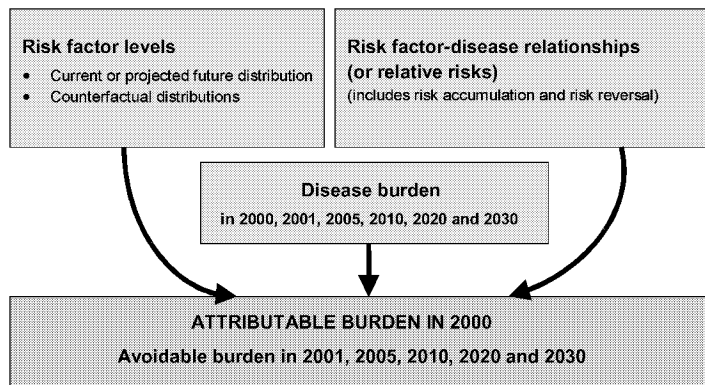


Abbildung 8: Input/Output der vergleichenden Risikobetrachtung in GBD (GBD 2001)

Auswahl der betroffenen Regionen und Bevölkerungsgruppen

Die Vorgehensweise, wie der Weltgesundheitsstatus anhand von geopolitischen Gesichtspunkten in Weltregionen aufgebrochen wurde, ist eher willkürlich, entspricht aber den traditionellen Auffassungen vom Entwicklungsstand einer Weltregion. Dass jedoch innerhalb einer Weltregion und sogar innerhalb eines Landes sehr unterschiedliche Krankheitsbilder etabliert sein können, ist selbstverständlich. Man denke z.B. nur an die Differenzen der Krankheitsbilder zwischen urbanen und ländlichen Räumen. Andere Einflussfaktoren bestehen auch in rein sozialen Unterschieden, wie z.B. der Religion, der ethnischen Zugehörigkeit sowie kulturellen Gewohnheiten. Diese Heterogenität fällt jedoch der globalisierten Betrachtungsweise der Statistiken zum Opfer, da hier ein hohes Niveau von Aggregation verlangt wird; die Heterogenität wird selbst in der Weltregion-spezifischen Disaggregation nur unzureichend ersichtlich. Aus diesem Grunde entwickelte die WHO Hilfestellungen, um nationale und subnationale Gesundheitsbelastungsstudien durchzuführen.

Auswahl der zu betrachtenden Zeiträume und Wahl der Diskontierung

Die Auswahl der Zeiträume in den GBD Projekten ist von grundlegender Bedeutung. Hier geht es nicht nur um die Nutzung der Zeit als Messindikator und um die damit verbundene, willkürliche Auswahl des Standards, wie groß eine Lebenserwartung sein sollte, sondern auch um die Frage, wie eine Altersgruppe definiert werden soll, insbesondere in Bezug auf Annahmen zu der sich verändernden und zeitlich-dynamischen Sterblichkeitsrate.

Unsicherheiten

In der ersten Studie (GBD1996) wurden nur wenige Unsicherheiten diskutiert: Nennenswerten Einfluss auf die Ergebnisse haben die Diskontrate und die altersabhängige Gewichtung, während die Unterschiede der Gewichtungen verschiedener Behinderungen in der DALY-Metrik gering sind. Demgegenüber würde die (nicht quantifizierte) Unsicherheit der epidemiologischen Daten die Ergebnisse erheblich stärker beeinflussen (Murray & Lopez 1996, GBD Executive Summary). In GBD2000 sollen die Unsicherheiten der Daten berücksichtigt werden (Mathers et al. 2002) sowie die Unsicherheiten der Gewichtungen in der DALY-Metrik (Salomon et al. 2001).

Beteiligte / Stakeholder

Die erste GBD Studie wurde bereits als eine breitangelegte Zusammenarbeit zwischen mehr als 100 Wissenschaftlern in 20 Nationen angelegt. Die darauf folgenden Diskussionen, auch z.T. um GBD2000, zeigten jedoch, dass es Lücken in der Beteiligungsstrategie gab. So fehlten nicht nur bestimmte Expertenrepräsentanten, sondern auch Beteiligte außerhalb der Expertengemeinde. Letzteres ist insbesondere deshalb problematisch, da Einschätzungen über Soll-Werte für fundamentale Kriterien des Gesundheitszustandes einer Bevölkerung, z.B. die Messung der Präferenz für einen Gesundheitszustand durch die betroffene Person, eng verknüpft sind mit sozialen Wertschätzungen. Die Operationalisierung solcher sozialen Wertschätzungen bedarf einer breiteren Einbeziehung von Beteiligten. Die Strategie der WHO, Hilfestellungen für nationale und regionale Gesundheitsbelastungsstudien zu geben, ist ein bedeutender Schritt in diese Richtung.

Gesamtbewertung

Die komparative Betrachtung von Risikofaktoren für eine Gesundheitsbelastung wurde im GBD-Projekt anhand einer hoch-aggregierten und hinsichtlich Krankheitsbildern integrativen Belastungs-Metrik (DALY) durchgeführt. Obwohl die DALY-Metrik eine intuitive und vereinfachte Wertdimension repräsentiert, ist ihre Abschätzung (bzw. Berechnung) in der Praxis ein hochkompliziertes Unterfangen, das meist nur mit speziell entwickelten Software-Lösungen zu bewältigen ist.

Als Maß für die Bedeutung von Risikofaktoren bietet sich der Gewinn bzw. der Verlust an Lebenszeit an. Im GBD-Projekt wird deutlich, dass ein solches Risikomaß mit Überlegung definiert und transparent dargestellt werden muss.

Die folgende Tabelle gibt einen Überblick über die Vorgehensweise:

Tabelle 3: Vorgehensweise in der "Global Burden of Disease (GBD) Study"

"Global Burden of Disease (GBD) Study"	
Zielsetzung	konsistente Abschätzungen der Hauptursachen für Todesfälle, für Krankheitsfolgen, die zu Behinderungen führen, und der Todesfälle und Behinderungen durch 10 wichtige Risikofaktoren, sowie Projektionen von Sterblichkeit und Behinderungen „public health“, Prioritätensetzung in der Gesundheitspolitik
Risikobereiche	Gruppe I: ansteckende Krankheiten, Einflüsse vor und während der Schwangerschaft, Mangelernährung; Gruppe II: nichtansteckende Krankheiten; Gruppe III: Verletzungen (absichtliche und unabsichtliche)
Vergleichskategorien	- Liste verschiedener Krankheiten, Krankheitsursachen und Krankheitsfolgen; wurde über eine Baumstruktur entwickelt, - Werte für verschiedene Kalenderjahre.
zeitlicher Rahmen	jeweils 1 Kalenderjahr
räumlicher Rahmen	global und nach 8 geographischen Regionen unterschieden
Attribute	Todesfall, Krankheit, Behinderung
Aggregierung / Metrik	DALY; Gewichtungen wurden durch Befragung kleiner Gruppen ermittelt, Unterschiede der Gewichtungen auch zwischen verschiedenen Ländern nur gering
Gesamtbewertung	implizit multiattribut
Unsicherheiten	in GBD1996 einige Unsicherheiten qualitativ diskutiert, keine quantitative Behandlung; in GBD2000 Angabe von Unsicherheiten vorgesehen
Ergebnisdarstellung	Tabellen und Grafiken in absoluten Fallzahlen und in Metrik-Einheiten (DALYs)

3.2.4 ExternE-Projekt

Vorgehensweise

Das ExternE-Projekt (ExternE - Summary 1995) wird hier als Beispiel einer vergleichenden Risikobewertung für Optionen der Energiegewinnung vorgestellt. Anhand der Vorgehensweise und Ergebnisse dieses Projekts lassen sich typische Verfahren, aber auch typische offene Fragen aufzeigen.

Mit dem ExternE-Projekt sind erhebliche Fortschritte bei der konsistenten und umfassenden Ermittlung der externen Kosten der Energieerzeugung erzielt worden. Es werden damit Vergleiche zwischen unterschiedlichen Energiequellen ermöglicht. Das Projekt

- berücksichtigt alle Lebensphasen einer Technologie;

- berücksichtigt die direkten Folgen aller durch eine Technologie verursachten Emissionen und Belastungen (zumindest im Prinzip; Einschränkungen bestehen auf Grund beschränkter Kenntnis), jedoch keine indirekten Folgen wie z.B. Auswirkungen auf das Arbeitsplatzangebot;
- berücksichtigt auch Folgen für künftige Generationen;
- basiert auf der „Wirkungspfad-Vorgehensweise“ (Abb. 9 auf Seite 65);
- rechnet alle Schäden auf eine einheitliche Maßeinheit um (in diesem Fall: Geldwert), um Vergleiche zu ermöglichen.

Wichtig hinsichtlich des Risikovergleichs ist, dass nicht erst durch Betrachtung der so ermittelten Ergebnisse für verschiedene Energiesysteme ein Vergleich stattfindet, sondern dass Vergleiche bereits in die beschriebene Ermittlung einfließen. So ist durch die Wahl der Maßeinheiten beispielsweise festgelegt, wie ein plötzlicher Unfalltod im Vergleich zu einer zusätzlichen, erst nach langer Latenzzeit auftretenden Krebserkrankung bewertet wird, wie im Vergleich dazu vorübergehende und bleibende, aber nicht tödliche Erkrankungen bewertet werden, und wie diese alle im Vergleich z.B. zu ökologischen Schäden oder Schäden an immateriellen Gütern bewertet werden. Außerdem beeinflussen weitere, teils normative, teils willkürliche Festsetzungen in hohem Maße die berechneten „externen Kosten“ und damit auch das Ergebnis eines damit vorgenommenen Vergleichs.

Im folgenden wird die Vorgehensweise näher erläutert. Die problematischen Vorgaben werden bei den Analyseschritten angesprochen, für die sie benötigt werden.

Ziel des Projekts

Ziel des Projekts ist die Ermittlung der externen Kosten der Energieerzeugung. Begründet wird dies wie folgt:

»In Europe, policy analysts are being required to take account of environmental aspects in their decision making and to undertake cost-benefit analysis of available options. The 5th Environmental Action Programme 'Towards Sustainability', clearly indicates the need for the assessment of externalities and monetary evaluation.« (*ExternE Overview*, o. D.)

Modellansatz

Der prinzipielle Modellansatz basiert auf der Analyse von „Schadensfunktion“ oder „Wirkungspfad“. In *ExternE Overview* wird dies anhand einer Abbildung³⁹ veranschaulicht, in der allerdings ein wichtiger Schritt fehlt: Die Aggregation der Schäden. Die folgende Abbildung 9 basiert auf dieser Abbildung, ist aber um die Aggregation der Schäden ergänzt.

³⁹ Die gleiche Abbildung wie in *ExternE - Summary* (1995), S. 152.

Auswahl der Technologie

Wie im Ziel des Projekts genannt, geht es um die Energieerzeugung. Entsprechend werden nur Energieerzeugungstechnologien behandelt, und davon nur die, die derzeit im Einsatz sind (fossil: Steinkohle, Braunkohle, Öl, Gas; Kernenergie; erneuerbar: Wasserkraft, Photovoltaik, Wind, Biomasse; sowie Energieeinsparung (*ExternE - Summary* 1995, S. 12)).

Auswahl der zu betrachtenden Schutzgüter, Einwirkungen und Auswirkungen

Es wird ausgegangen von bekannten schädlichen Folgen von Substanzen, die bekanntermaßen bei der betrachteten Art der Energiegewinnung freigesetzt werden. Dabei werden prinzipiell alle Phasen des Brennstoffzyklus (z.B. Bergbau, Abraumentsorgung, Brennstoffaufbereitung, Transport, Verbrennung, Entsorgung der Verbrennungsprodukte) und alle Lebensphasen der hierzu benötigten Anlagen (Errichtung, Betrieb, Demontage, Entsorgung) betrachtet.

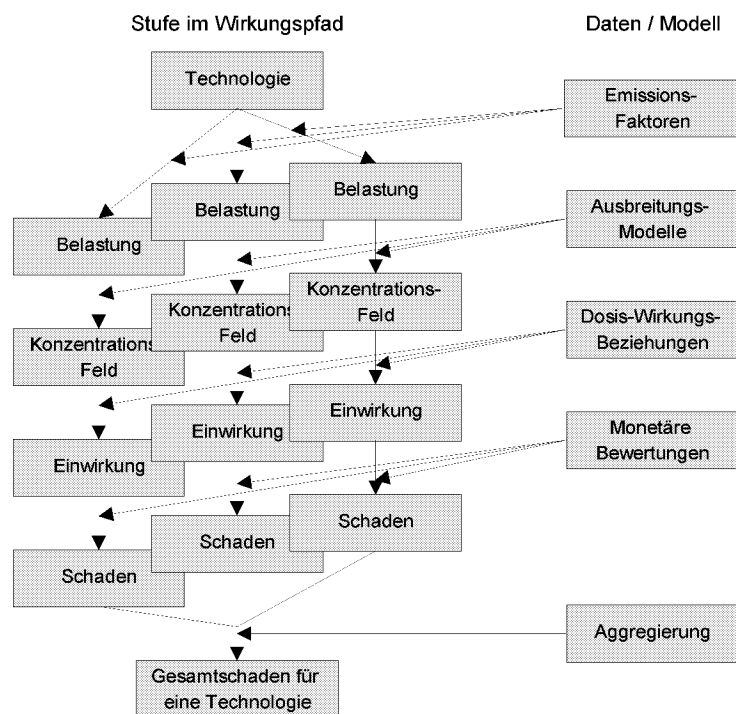


Abb. 9: „Typische Wirkungspfad-Vorgehensweise“ (nach *ExternE Overview*, ergänzt um die Aggregation mehrerer Wirkungspfade)

In *ExternE Overview* werden die betrachteten Einwirkungen und Auswirkungen wie folgt beschrieben:

„Externe Kosten von Energie sind [...] Kosten für Gesellschaft und Umwelt, die von Produzenten und Konsumenten der Energie nicht berücksichtigt werden, d. h. die nicht im Marktpreis enthalten sind. Dazu gehören Schäden an der natürlichen und der erbauten Umwelt wie z.B.

- Auswirkungen von Luftverschmutzung auf
 - Gesundheit,
 - Gebäude,
 - landwirtschaftliche Erträge,
 - Wälder und
 - globale Erwärmung;
- Berufskrankheiten und Arbeitsunfälle; und
- Beeinträchtigungen durch den Anblick einer Anlage oder Geräuschemissionen.“

Weiter unten werden noch genannt:

- Süßwasserfischerei und
- Biodiversität.

Auf der Basis erster grober Abschätzungen werden in die eigentliche Quantifizierung nur die Ursachen und Auswirkungen einbezogen, die einen nennenswerten Anteil zum berechneten Schaden beitragen. Es wird aber immer betont, dass prinzipiell alles zu berücksichtigen sei. (*ExternE Methodology Annexes* 1997)

Auswahl der betroffenen Regionen und Bevölkerungsgruppen

Die Berechnungen werden standortspezifisch durchgeführt. Da davon ausgegangen wird, dass die Schadstoffkonzentration mit zunehmender Entfernung von der emittierenden Anlage abnimmt, sind die ermittelten Schäden umso größer, je größer die Bevölkerungsdichte in der Nähe der emittierenden Anlage ist. Ebenso werden für küstennahe Standorte geringere Schäden ausgewiesen als für küstenferne Standorte, da der Anteil der Emissionen, der aufs Meer hinaus transportiert wird, keine Gesundheitsschäden verursachen kann, weil dort keine Bevölkerung ist. (Mayerhofer et al. 1997, S. 105)

»The system boundary also has spatial and temporal dimensions. Both should be designed to capture impacts as fully as possible.

This has major implications for the analysis of the effects of air pollution in particular. It necessitates extension of the analysis to a distance of hundreds of kilometres for many air pollutants operating at the 'regional' scale, such as ozone, secondary particles, and SO₂. For greenhouse gases the appropriate range for the analysis is obviously global. Consideration of these ranges is in marked contrast to the standard procedure employed in environmental impact assessment which considers pollutant transport over a distance of only a few kilometres and is further restricted to primary pollutants. The importance of this issue in externalities analysis is that in many cases in the ExternE Project it has been found that regional effects of air pollutants like SO₂, NO_x and associated secondary pollutants are far greater than effects on the local scale.« (Mayerhofer et al. 1997, S. 9-10).

Dies ist gleichbedeutend mit der ausschließlichen Betrachtung des Kollektivrisikos. Das durch eine Anlage verursachte Individualrisiko eines in der Nähe dieser Anlage lebenden Menschen ist natürlich unabhängig davon, wie viele Menschen sonst noch in der Umgebung leben.

Auswahl der zu betrachtenden Zeiträume und Wahl der Diskontierung

Im Prinzip sollten alle Zeiträume, für die Auswirkungen zu erwarten sind, berücksichtigt werden:

»Despite the uncertainties involved it is informative to conduct analysis of impacts that take effect over periods of many years. By doing so it is at least possible to gain some idea of how important these effects might be in comparison to effects experienced over shorter time scales. The chief methodological and ethical issues that need to be addressed can also be identified. To ignore them would suggest that they are unlikely to be of any importance.« (Mayerhofer et al. 1997, S. 10-11).

»[...] the collective dose resulting from uranium mining and milling [...] is mainly caused by the release of Radon from abandoned tailing piles assumed to continue over 10000 years. [...] Because of the time horizon considered, the discounted present value of the damage is rather small, while the undiscounted damage from the uranium mining and milling stage dominates the total external costs of the fuel cycle.« (Krewitt et al. 1997, S. 80).

Allerdings sind darin große Unsicherheiten enthalten:

1. Die Wahl des Zeitraums von 10 000 Jahren ist willkürlich. Sie ist nicht etwa damit begründet, dass bis dahin die Belastungen auf einen vernachlässigbaren Wert abgeklungen seien, vielmehr heißt es bei Krewitt et al. (1997) auf S. 38: »[...] an exhalation rate of 3 Bq/m²/s from reasonably covered mill tailings that is expected to remain unchanged over at least 10 000 years.« Würde man über eine längere Zeit rechnen, wäre der errechnete Schaden größer, über einen kürzeren Zeitraum wäre er geringer.

2. Die langfristigen Auswirkungen finden in einer Zukunft statt, von der wir nicht wissen, wie sie aussehen wird:

»These impacts will be imposed on societies with very different characteristics from those of today. The magnitude of the impact can vary by orders of magnitude depending upon the scenario assumptions. These uncertainties are far larger than those caused by scientific uncertainty.« (ExternE - Summary 1995, S. 153)

Bei derart langen Betrachtungszeiträumen spielt natürlich eine entscheidende Rolle, ob und in welchem Maß zukünftige Schäden diskontiert werden sollen (ExternE - Summary 1995, S. 43-46). Auch dies ist keine Frage, die naturwissenschaftlich beantwortet werden kann, sondern eine Frage der Bewertung.

Auch die Folgen seltener Unfälle werden über die Zeit gemittelt (ExternE Methodology Annexes 1997, S. 34; Krewitt et al. 1997, S. ES-4 und S. 38). Würde man allerdings davon ausgehen, dass die gravierendste Kategorie von Kernkraftwerksunfällen für ein Land einem wirtschaftlichen Totalschaden gleichkommen würde, dann wäre die Zeitmittelung der Folgen eines solchen Unfalls nicht sinnvoll.

Dosis-Wirkungs-Beziehungen

Die Berechnungen im ExternE-Projekt gehen von bekannten und anerkannten Dosis-Wirkungs-Beziehungen aus. Dennoch ist die Ermittlung der Folgen selbst bei be-

kannter Belastung schwierig. Im Projekt werden die beiden folgenden problematischen Aspekte genannt:

Zum einen gibt es bei einigen Schadstoffen einen Schwellenwert und/oder eine nichtlineare Dosis-Wirkungs-Beziehung, so dass die Wirkung der vom betrachteten Energieerzeugungssystem verursachten *zusätzlichen* Emission von den bereits vorliegenden Emissionen anderer Systeme abhängt; zum anderen kann die Wirkung eines Schadstoffs von der Konzentration anderer Stoffe abhängen, von denen der größte Teil nicht vom betrachteten Energieerzeugungssystem, sondern von anderen Systemen emittiert wird. Dadurch hängen die durch das Energieerzeugungssystem verursachten Schäden auch davon ab, wie hoch die Emissionen anderer Systeme sind:

»One of the main findings concerning the aggregation methodology is that damage costs per tonne of pollutant emitted might differ significantly even within a country. The main reason for such differences is the spatial variation in SO₂, NO_x, and NH₃ emissions. As ammonia is mainly emitted from agricultural activities, and the availability of free ammonia in the atmosphere also depends on the level of SO₂ and NO_x emissions from other sources, the damage costs we refer to as 'energy' externalities in fact strongly depend on emissions from various industrial activities.«
(Krewitt et al. 1997, S. ES-6; S. 76)

Offen bzw. dem Entscheidungsträger überlassen bleibt die Frage, welchem „Verursacher“ solche multikausalen Schäden zuzuordnen sind.

Auf jeden Fall werden die bereits vorhandenen Belastungen berücksichtigt. Die durch eine zusätzliche Anlage verursachten, zusätzlichen Schäden werden auf der Basis des Abschnitts der Dosis-Wirkungs-Beziehung zwischen der bereits vorhandenen Belastung und der Gesamtbelastung unter Einbezug der zusätzlichen Anlage berechnet. Dadurch ergibt sich auch bei Schadstoffen, die eine Dosis-Wirkungs-Beziehung *mit* Schwellenwert aufweisen, eine differentielle Zunahme der Schäden, als ob eine (näherungsweise lineare) Dosis-Wirkungs-Beziehung *ohne* Schwellenwert vorläge. Das führt rechnerisch zur gleichen Zunahme von Gesundheitsschäden, unabhängig davon, ob 100 Personen mit einem um 10% erhöhten Schadstoffwert belastet werden oder 100 000 Personen mit einem um 0,01% erhöhten Schadstoffwert. Ob eine solche Gleichsetzung berechtigt ist, ist jedoch fraglich.

Außerdem hängt bei dieser Betrachtungsweise der berechnete Schaden von der – je nach Fragestellung vorgegebenen oder willkürlichen – Reihenfolge der Betrachtung ab. Nimmt man ausgehend vom unbelasteten Zustand jeweils einen Emittenten nach dem anderen hinzu, so ergibt sich für die als erste berücksichtigten Emittenten überhaupt kein Schaden, solange die Belastung unter dem Schwellenwert bleibt. Erst für die Emittenten, bei deren Berücksichtigung der Schwellenwert überschritten ist, ergeben sich Schäden.

Schließlich ermittelt das Projekt ausdrücklich die *inkrementellen* Folgen z.B. eines neuen Kraftwerksprojekts oder von Änderungen in der Energiepolitik (*ExternE - Methodology* 1995, S. 11). Dies bedeutet andererseits, dass die ermittelten Werte umso weniger verlässlich sind, je mehr die tatsächliche Gesamt-Belastung von dem im Projekt berücksichtigten Ist-Zustand abweicht.

Wertung der Schäden

Um einen Vergleich unterschiedlicher Energiesysteme zu ermöglichen, werden die Schäden auf einen einheitlichen Vergleichsmaßstab umgerechnet. In ExternE ist dies der Geldwert (die Schäden werden „monetarisiert“). Die folgenden Probleme treten aber auch dann auf, wenn als Vergleichsmaßstab nicht der Geldwert, sondern eine andere Maßeinheit gewählt wird:

Die Bewertung immaterieller Schäden:

Wie bewertet man z.B. den Wert des Waldes als Erholungsraum? „Der Analytiker sollte Informationen für die Entscheidungsträger sammeln und nicht willkürliche Entscheidungen treffen, welche Aspekte einer weiteren Diskussion Wert seien“:

»Many receptors that may be affected by fuel chain activities are valued in a number of different ways. For example, forests are valued not just for the timber that they produce, but also for providing recreational resources, habitats for wildlife, their interactions (direct and indirect) with climate and the hydrological cycle, protection of buildings and people in areas subject to avalanche, etc. Externalities analysis should include all such aspects in its valuation. Again, the fact that a full quantitative valuation along these lines is rarely possible is besides the point when seeking to define what a study should seek to address: the analyst has the responsibility of gathering information on behalf of decision makers and should not make arbitrary decisions as to what may be worthy of further debate.« (Mayerhofer et al. 1997, S. 9)

Die Bewertung verschiedener Todesfälle:

Die erste Phase des ExternE-Projekts ging vom 'value of a statistical life' (VSL) aus. Aber: Sind alle Todesfälle gleich zu werten, egal in welchem Alter und unter welchen Randbedingungen der/die Betroffene stirbt? Dies ist einerseits besonders fraglich bei den Krebserkrankungen, die erst in höherem Alter auftreten, andererseits bei Todesfällen durch Luftverschmutzung, die hauptsächlich gesundheitlich geschwächte Personen betreffen, die möglicherweise ohnehin nicht mehr lange gelebt hätten. In der Fortschreibung der Methodik wird daher die Größe 'years of life lost' (YOLL) eingeführt. Hierzu heißt es:

»[...] in earlier phases of the project a number of questions were raised regarding the use of the VSL for every case of mortality considered. These originally related to the fact that many people whose deaths were linked to air pollution were suspected of having only a short life expectancy even in the absence of air pollution. Was it logical to ascribe the same value to someone with a day to live as someone with tens of years of remaining life expectancy? Furthermore, is it logical to ascribe the full VSL to cases where air pollution is only one factor of perhaps several that determines the time of death, with air pollution playing perhaps only a minor role in the timing of mortality? In view of this the project team explored valuation on the basis of life years lost. [...] Within ExternE it has been concluded that VSL estimates should be restricted to valuing fatal accidents, mortality impacts in climate change modelling, and similar cases where the impact is sudden and where the affected population is similar to the general population for which the VSL applies. The view of the project team is that the VSL should not be used in cases where the hazard has a significant latency period before impact, or where the probability of survival after impact is altered over a prolonged period. In such cases the value of life years (YOLL) lost approach is recommended.« (Mayerhofer et al. 1997, S. 260)

Die Wahl des einen (VSL) oder anderen (YOLL) Maßes ist bereits eine Entscheidung für eine bestimmte Bewertung von Todesfällen und eine Entscheidung gegen andere mögliche Bewertungen.

Die Bewertung verschiedener Erkrankungen in Abhängigkeit von ihrer Dauer und Schwere:

Hier treten ähnliche Probleme auf wie bei der Bewertung unterschiedlicher Todesfälle: Der Maßstab, mit dem Zeitpunkt des Krankheitsbeginns und Krankheitsdauer umgerechnet werden, ist bereits eine bestimmte Bewertung der Krankheiten. Ein weiterer Bewertungsmaßstab bestimmt, wie die Schwere der Krankheit eingerechnet wird.

Die Monetarisierung von Umwelt- und Gesundheitsschäden sowie Todesfällen:

Diese Vorgehensweise stößt häufig auf Kritik, weil dadurch Leben und Gesundheit einerseits und materielle Werte andererseits gleichgesetzt werden. (Diese Kritik gilt aber für *jeden* Vergleich unterschiedlicher Schadensarten, wie weiter unten ausgeführt wird.)

»Although valuation of environmental impacts in monetary terms is widespread, there are still many people who find the idea strange at best, and distasteful and unacceptable at worst. [...] The essence of economic valuation of environmental goods is that it attempts to identify the preference of individuals for the allocation of resources by, for example, quantifying people's 'Willingness to Pay' (WTP) for environmental improvement. It does not pretend to quantify the intrinsic value of a human life or biodiversity. [...]

It is, however, acknowledged that there are a range of problems with the use of valuation on the base of WTP:

1. The concept is 'income constrained'. Since you cannot pay what you do not have, a poorer person's WTP is less than that of a richer person, other things being equal. [...]
2. Putting money values on health and the environment and pricing accordingly does not necessarily guarantee sustainability, and therefore taken alone it is insufficient as a prescription for economic policy on the environment.
3. At the most fundamental level, it is disputed whether monetary valuation of health and the environment is philosophically defensible. This issue is outside the scope of our study« (*ExternE - Summary* 1995, S. 41-42)

Unabhängig davon, ob als einheitlicher Maßstab für Schäden der Geldwert oder eine andere Größe verwendet wird, bleibt das Problem, dass durch Umrechnung auf den einheitlichen Maßstab *zwangsläufig* ein Vergleich vorgenommen wird zwischen unterschiedlichen Schäden wie z. B. dem Verlust eines Menschenlebens, der Beeinträchtigung der Gesundheit, der Beeinträchtigung der Umwelt und ggf. der Beeinträchtigung immaterieller Güter. Dies wäre nur dadurch zu umgehen, dass auf die Umrechnung auf einen einheitlichen Maßstab verzichtet würde und statt dessen die unterschiedlichen Schäden im Ergebnis getrennt ausgewiesen würden. Letztlich kommt man aber, wenn man Risiken vergleichen will, nicht umhin, solche unterschiedlichen Schäden gegeneinander abzuwägen.

Unsicherheiten

Unsicherheiten werden systematisch berücksichtigt, soweit dies möglich ist. Dabei wird unterschieden zwischen statistischen Unsicherheiten, Modellunsicherheiten, unsicheren Annahmen, und Fehlern (*ExternE Methodology Annexes* 1997, S. 87):

»It is appropriate to group the main contributions to the uncertainty into qualitatively different categories:

- statistical uncertainty - deriving from technical and scientific studies, e.g. dose-response functions and results of valuation studies,
- model uncertainty - deriving from judgements about which models are the best to use, processes and areas excluded from them, extension of them to issues for which they are not calibrated or designed. Obvious examples are the use of models with and without thresholds, use of rural models for urban areas, neglecting areas outside dispersion models and transfer of dose-response and valuation results to other countries,
- uncertainty due to policy and ethical choices - deriving from essentially arbitrary decisions about contentious social, economic and political questions, for example decisions on discount rate and how to aggregate damages to population groups with different incomes and preferences,
- uncertainty about the future - deriving from assumptions which have to be made about future underlying trends in health, environmental protection, economic and social development, which affect damage calculations, e.g. the potential for reducing crop losses by the development of more resistant species, and
- human error.

For human error, little can be done other than by attempting to minimise it. The ExternE Project uses well reviewed results and models wherever available and calculations are checked. The use of standardised software (EcoSense) has greatly assisted this.

Uncertainties of the first type (statistical) are amenable to analysis treatment by statistical methods, allowing the calculation of formal confidence intervals around a best estimate. Uncertainties in the other categories are not amenable to this approach, because there is no sensible way of attaching probabilities to judgements, scenarios of the future, the 'correctness' of ethical choices or the chances of error. There is no reason to expect that a statistical distribution has any meaning when attempting to take into account the possible variability in these parameters. In addition, our best estimate in these cases may not be a median value, thus the uncertainty induced may be systematic. Nevertheless the uncertainty associated with these issues is important and needs to be addressed.«

Die Unsicherheiten der Ergebnisse werden aus den Unsicherheiten der Parameter mit einem einfachen arithmetischen Verfahren ermittelt. Dieses Verfahren geht davon aus, dass die Ergebnisse im wesentlichen durch Multiplikation der Parameter berechnet werden⁴⁰. Wenn für jeden Parameter x_i ($i = 1 \dots n$) die Unsicherheit durch eine Lognormalverteilung mit der geometrischen Standardabweichung $s_{g,i}$ ($i = 1 \dots n$) approximiert wird, lässt sich die Unsicherheit des *Produkts* dieser Parameter durch die Formel

$$\left[\log(s_{g,prod}) \right]^2 = \prod_{i=1}^n \left[\log(s_{g,i}) \right]^2$$

berechnen. Diese Formel gilt *nicht* für die Unsicherheit einer *Summe* unsicherer Werte. Für die aggregierten (summierten) Risiken werden daher keine Unsicherheiten angegeben. Bei einem Kosten-Nutzen-Vergleich werden jedoch Unsicherheiten berücksichtigt: Die Auswirkungen (Attribute) werden in der Reihenfolge ihrer „confidence“ in die Betrachtung einbezogen und mit den Vermeidungskosten verglichen,

⁴⁰ z. B.: »Damage = pollution concentration x population x exposure-response function x valuation« (ExternE Methodology Annexes 1997, S. 88).

wie im folgenden beschrieben (Krewitt et al. 1997, S. 89-90; *ExternE - Vol. 10: National Implementation 1999*, S. 153):

»In order for an abatement technology to be considered worthwhile it is necessary for benefits to exceed costs. Comparison of the private costs with the estimated damage costs per tonne of pollutant emitted suggests that the abatement measures reviewed by ERM (1996) are justified for both UK and German plant. However, the externality estimates are subject to significant uncertainty. To take into account these uncertainties, we rank the results for each impact of the three pollutants in terms of the confidence attached to each (see e.g. Figure 9.7). For the German plant, current best available emissions control techniques as defined in ERM (1996) is justified without reference to mortality effects. In fact for SO₂ it is only effects on materials and crops that need to be included before the lower end of the range of private costs is passed. For the UK plant, where overall damages are lower because of the location of the power plant site close to the North Sea, the combined effect of damage to materials and crops (where appropriate) and acute morbidity are sufficient to take results for all three pollutants past the lower end of the range for the private costs of abatement, and also past the upper end of the range for particulates. [...]

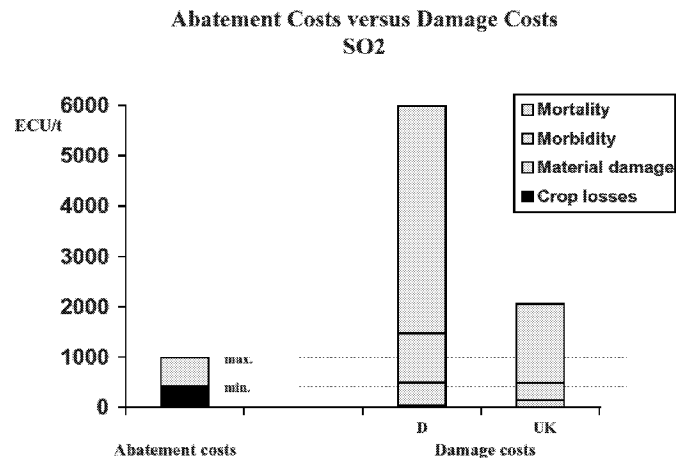


Figure 9.7 Comparison of abatement costs and damage costs per tonne of SO₂ emitted from a coal fired power plant in Germany and the UK«

Beteiligte / Stakeholder

Am Projekt beteiligt waren über 30 Teams aus Forschungsinstituten und Beraterfirmen in neun Mitgliedsstaaten der EU und anderen europäischen Ländern. Ob und welche Stakeholder beteiligt waren, konnte anhand der verfügbaren Informationen nicht geklärt werden.

Gesamtbewertung

Das Projekt behandelt streng genommen nicht Risiken, sondern externe Kosten. Da jedoch ein großer Teil der externen Kosten aus Risiken resultiert, kann man aus diesem Projekt viele Schlussfolgerungen für Risikovergleiche ziehen. Dennoch ist zu beachten, dass das Ziel des Projekts nicht ein fertiger Vergleich ist, sondern die Bereitstellung von Daten und Methoden, um Vergleiche zu ermöglichen. Da die exter-

nen Kosten u. a. von der Bevölkerungsdichte um den Standort einer Anlage und von der Schadstoffbelastung aus anderen Quellen abhängen, sind solche Vergleiche nur standortspezifisch und für ein jeweils festgelegtes Szenario sinnvoll.

Ein besonderes Verdienst des Projekts ist, Annahmen zu dokumentieren und teilweise auch zu diskutieren, die sonst oft stillschweigend vorausgesetzt werden. Insbesondere sind hier zu nennen:

- Die Monetarisierung immaterieller Schutzgüter ist umstritten, entweder prinzipiell (d. h. immaterielle Schutzgüter sollten überhaupt nicht mit materiellen Gütern gleichgesetzt werden) oder bezüglich der Gewichtung (wieviel Geld ist ein Ökosystem wert).
- Die Wirkung einer geringfügigen Dosis wird möglicherweise überbewertet: Aus einer geringfügigen Erhöhung der Radon-Belastung wird nur deshalb eine nennenswerte Zahl zusätzlicher Krebsfälle ermittelt, weil über eine (willkürlich gewählte) Betrachtungsdauer von 10 000 Jahren und die in diesem Zeitraum lebende große Zahl von Menschen aggregiert wird. Bei einer Beschränkung auf einen Betrachtungszeitraum von 100 Jahren (oder Diskontierung künftiger Schäden) wäre diese Schadstoffbelastung nicht mehr signifikant.
- Schwere, aber sehr seltene Unfälle werden über die Zeit gemittelt. (Nur das Produkt „Häufigkeit mal Schaden“ geht in das Ergebnis ein.)
- Die Diskontierung zukünftiger Schäden ist umstritten. Im Projekt werden daher Ergebnisse für verschiedene Annahmen der Diskontrate angegeben.
- Neu ist die Vorgehensweise, Unsicherheiten der Ergebnisse dadurch zu berücksichtigen, dass die Einzelbeiträge nicht direkt aggregiert werden, sondern, beginnend mit den „vertrauenswürdigsten“, nach und nach in den Vergleich einbezogen werden, bis eine Aussage möglich ist. Dadurch sind vielfach relativ sichere Aussagen möglich, ohne auf die besonders unsicheren Teilbeiträge zurückgreifen zu müssen.

Die folgende Tabelle gibt einen Überblick über die Vorgehensweise:

Tabelle 4: Vorgehensweise im ExternE-Projekt

ExternE-Projekt	
Zielsetzung	Vergleich alternativer Technologien „public health“, „public wealth“
Risikobereiche	Risiken durch Anlagen zur Erzeugung elektrischer Energie
zeitlicher Rahmen	im Prinzip unbegrenzt (10 000 Jahre bei Radon); künftige Schäden werden diskontiert
räumlicher Rahmen	nur begrenzt durch Bereiche der Schadstoffausbreitung. Regionale Besonderheiten werden berücksichtigt.
Attribute	Schäden an Gesundheit, Gebäuden, landwirtschaftlichen und Fischerei-Erträgen, Wäldern und Biodiversität und durch globale Erwärmung; Berufskrankheiten und Arbeitsunfälle; Beeinträchtigungen durch den Anblick einer Anlage oder Geräuschemissionen.
Aggregierung / Metrik	VSL oder YOLL, Monetarisierung
Gesamtbewertung	Vergleich der aggregierten Werte, implizit multiattribut durch die verwendeten Metriken
problematische Annahmen	Monetarisierung immaterieller Schutzgüter, Aggregierung von Schäden durch extrem kleine Dosen in extrem großer Population, Bewertung extrem seltener Großunfälle allein nach dem Produkt „Häufigkeit mal Schadensumfang“, Diskontrate für zukünftige Schäden.
Unsicherheiten	„Statistische“ Unsicherheiten werden systematisch berücksichtigt und in den detaillierten Ergebnistabellen angegeben. Beim Vergleich werden die Teilbeiträge in der Reihenfolge ihrer „Vertrauenswürdigkeit“ hinzugezogen. Für einige (aber nicht alle) <i>unsichere Annahmen</i> werden Sensitivitätsanalysen durchgeführt.
Ergebnisdarstellung	für jede Technologie detaillierte Tabelle, aufgeschlüsselt nach Attributen; keine zusammenfassende Gegenüberstellung, aber Beispiele für Anwendung in Kosten-Nutzen-Vergleichen, hier auch Berücksichtigung von Unsicherheiten.

3.2.5 Projekt „Ganzheitliche Betrachtung von Energiesystemen (GaBE)“

Das aktuelle, seit einigen Jahren laufende Projekt „Ganzheitliche Betrachtung von Energiesystemen“ (GaBE) des Paul Scherrer Instituts (PSI) gehört in den Bereich der Vergleiche von Risiken, Kosten und Nutzen mit dem Ergebnis einer Wahlmöglichkeit unter Optionen (Hirschberg 1999). Damit verfolgt es im Hinblick auf Risikovergleiche die dritte der in Abschnitt 1.5 genannten Zielsetzungen. Das Projekt hat zum Ziel, eine umfassende Methodik zur konsistenten und detaillierten Bewertung von Energiesystemen anwendungsreif zu entwickeln. Dabei war Nachhaltigkeit ursprünglich kein Forschungsthema; im Laufe der Zeit hat sich jedoch aus diesem Leitbild ein

Rahmenkonzept zur Bewertung von Energiesystemen entwickelt, nicht zuletzt durch die Aufnahme der nachhaltigen Entwicklung als Leitprinzip in die Verfassung der Schweiz.

Vorläufer Poly-Projekt

Schon zu Beginn der neunziger Jahre wurde im Poly-Projekt der ETH Zürich (Berg et al. 1995) die Thematik von Risikovergleichen aufgegriffen. Eine wesentliche Zielsetzung des Projektes bestand darin, Beiträge zu einer verallgemeinerungsfähigen Methode zur Risikobewertung technischer Systeme zu erarbeiten. Als Basis für die Entwicklung einer entsprechenden Risikobewertung im Energiebereich wurde die Schweizer Störfallverordnung (StFV) gewählt. Mit dem für diese Verordnung entwickelten Instrumentarium strebt man an, diejenigen Störfallrisiken weitgehend zu erfassen und zu beurteilen, die durch den Umgang mit chemischen Stoffen und Mikroorganismen in stationären Anlagen sowie beim Transport gefährlicher Güter entstehen. Die Verordnung erstreckt sich nicht auf Risiken durch den Betrieb energieerzeugender – oder genauer: energieumwandelnder – Systeme. Da außerdem bekannt ist, dass bei solchen Systemen auch der bestimmungsgemäße, also störfallfreie Betrieb erhebliche Umweltauswirkungen hat, waren drei Fragestellungen zu bearbeiten (Berg et al. 1995):

- Ist der Risikoansatz der StFV zur Bewertung von Umweltrisiken in technischen Systemen geeignet?
- Kann die Methodik der StFV auf Störfallrisiken im Energiebereich angewendet werden?
- Können Begriffe und Methoden der Risikowissenschaft, die störfallorientiert sind, auf risikobezogene Aspekte des Normalbetriebs übertragen werden?

Der Risikoansatz der StFV zielt auf die Ermittlung von Störfallszenarien ab, denen ein Risiko in Form der Kombination von Schadensausmaß und Eintrittswahrscheinlichkeit zugeordnet wird. Dabei werden unterschiedliche Verfahren zur Szenarioentwicklung eingesetzt, je nach zu untersuchendem System. Die Autoren (Berg et al. 1995) kommen bei ihrer Analyse zu dem Schluss, dass die als System zu verstehende Umwelt der analysierten Anlagen zu komplex ist, um die zur Schadensanalyse erforderlichen Endpunkte zu definieren und Ursache-Wirkungs-Zusammenhänge zu formulieren. Sie weisen ferner darauf hin, dass der Risikoansatz der StFV nur den Betrieb der Anlage erfasst und durch ein life-cycle-assessment ergänzt werden muss. Je weiter der Betrachtungsrahmen gefasst wird, so die Autoren, desto eher sind mit dem Risikoansatz nur Teilaspekte dessen zugänglich, was für eine umfassende Bewertung berücksichtigt werden muss.

Die genannte Ausarbeitung zur Risikobewertung im Energiebereich (Berg et al. 1995) enthält eine Zusammenstellung der Umwelteingriffe, die im Zuge von Gewinnung, Transport und Aufbereitung der Primärenergieträger, Bereitstellung und Nutzung der Endenergie sowie Abfallentsorgung stattfinden. Zu diesen Eingriffen zählt auch die Schaffung von Potenzialen für Störfälle und Unfälle. Eine Abschätzung der Auswirkungen dieser Umwelteingriffe ergab, dass langfristige Umweltveränderungen in der Regel die Folge von betrieblichen, bestimmungsgemäßen Vorgängen sind, wobei Störfälle besonders auffällig und lokal schwerwiegend sein können. Nur bei der

Kernenergie und bei Erdöltransporten dominiert nicht eindeutig der Betrieb die Umwelteinwirkungen.

Bei der Anwendung des Risikoansatzes der StFV auf den Normalbetrieb und auf Störfälle treten spezifische Schwierigkeiten auf. Bei Störfällen entsteht die Frage nach der Aussagekraft von Häufigkeits- und Wahrscheinlichkeitsbetrachtungen. Störfälle mit sehr großem Schadensausmaß haben in der Regel als sehr niedrig ermittelte Eintrittswahrscheinlichkeiten. Diese sind nur schwer in überzeugender Weise zu bestimmen. Zudem wird in solchen Fällen das Schadensausmaß zum dominierenden Bewertungsfaktor. Die Risiken des Normalbetriebs werden dagegen durch ständig auftretende Einwirkungen verursacht, die meist Beiträge zu einer ohnehin bestehenden Belastung sind. Hier ist es schwierig, eine Kausalbeziehung zwischen Emission (aus einer bestimmten Anlage) und Wirkung herzustellen.

Die Autoren (Berg et al. 1995) schlagen vor, an Stelle von Schäden eine Umweltgefährdung als Bewertungsgrundlage zu verwenden. Sie wählen hierzu Indikatoren, die auf der Ebene von Expositionen drei Schlüsselgrößen für die Umweltgefährdung berücksichtigen: die räumliche und zeitliche Reichweite sowie das Ausmaß. Als weiteres Kriterium verwenden sie den Aspekt, ob die beschriebene Umwelteinwirkung Gegenstand der aktuellen öffentlichen Diskussion ist. Dies führt zu den Indikatoren CO₂-Emissionen, radioaktive Abfälle und Flächenbeanspruchung. Die Ermittlung der Zahlenwerte erfolgt auf der Basis einer Ausarbeitung zu Ökoinventaren von Energiesystemen (Frischknecht 1996). In diesen Ökoinventaren sind sehr seltene Störfälle unterhalb einer Grenzhäufigkeit – orientiert an Ereignissen pro Energieumsatz – nicht berücksichtigt; dies betrifft Kernenergieunfälle und große Öltankerkatastrophen. Kennzeichnend für die Analyse unterschiedlicher Energieträger ist, dass je nach Indikator ihre Reihenfolge unterschiedlich ausfällt. Eine Skala für die „Gesamtgefährdung“ steht jedoch in Berg et al. (1995) nicht zur Verfügung.

Hier wird deutlich, dass schon auf der Ebene eines Risikovergleiches die Zusammenfassung von Einzelaspekten zu einer Gesamtaussage schwierig ist. Derartige Fragen der Aggregation sind auch ein Thema des GaBE-Projektes.

GaBE-Projekt

Das GaBE-Projekt (Hirschberg 1998; Hirschberg 1999; Kröger 2001) greift Ansätze des Poly-Projektes auf; es analysiert und bewertet interdisziplinär und nach einem integralen Konzept die

- Umwelteinflüsse des Normalbetriebs,
- gesundheitlichen Auswirkungen des Normalbetriebs,
- Risiken schwerer Störfälle und
- ökonomischen Aspekte

von Systemen zur Energieversorgung.

Prinzipiell erlaubt die im Rahmen des Projektes am PSI entwickelte Methodik die Untersuchung folgender Fragestellungen:

- In welcher Hinsicht kann eine bestimmte Technik verbessert werden (Vergleich nicht-aggregierter Indikatoren)?
- Wie lassen sich mehrere Technik-Optionen (heutige/künftige Anlagen) vergleichen, beispielsweise anhand bestimmter Indikatoren oder anhand aggregierter Werte?
- Wie sieht ein effizienter Vergleich unterschiedlicher Szenarien der Energieversorgung aus?

Das Projekt verfolgt damit explizit als ein Ziel, Risiken zu Vergleichszwecken zu bestimmen, wobei Fragestellungen der Auswahl unter verschiedenen Energieversorgungs-Optionen eine wesentliche Rolle spielen.

Ansatz und Vorgehensweise

Zentraler Bestandteil der verwendeten Methodik ist das Life Cycle Assessment (LCA), das den Normalbetrieb von Anlagen umfassend bilanziert (Frischknecht 1996). Ein ausdrücklich an der Vorgehensweise des Life Cycle Assessment und am Drei-Säulen-Konzept der Nachhaltigkeit orientiertes und quantitativ angelegtes Bewertungsraster für Energiesysteme („Rahmenkonzept“) stellt Kröger im Kontext des GaBE-Projektes (Kröger 1998; Kröger 2000) vor. Drei Prinzipien (Vermeidung von Ressourcenerschöpfung, Vermeidung von nichtabbaubaren Abfällen sowie Empfindlichkeit gegenüber dem Umfeld) vereinen ökologische, ökonomische und soziale Aspekte und werden durch Kriterien und Indikatoren konkretisiert (siehe Tabelle 5). Entscheidendes Charakteristikum dieses Ansatzes sind Quantifizierbarkeit und geeignete Maßeinheiten der Indikatoren („Nachhaltigkeit ist messbar“). Dabei werden Aspekte, die üblicherweise bei der Ermittlung von Umweltrisiken untersucht werden, unter der Kriterienkategorie „kein Ressourcenabbau“ behandelt (s. u.).

Mit den verwendeten Indikatoren lassen sich sowohl die derzeitigen als auch weiterentwickelte Systeme der elektrischen Energieversorgung erfassen. Wie Vergleiche (Kröger 2001) zeigen, decken diese Indikatoren auch die relevanten Attribute ab, die von Institutionen wie der OECD/NEA und der UN/IAEA zur Beurteilung von Energiesystemen vorgeschlagen werden. Anhand eines Vergleichs unterschiedlicher Energie-Ketten wurde am PSI untersucht, ob die Datenlage eine verlässliche Quantifizierung der Indikatoren zulässt (Kröger et al. 2001). Ferner wurde getestet, ob eine derartige Bewertung auch der breiten Öffentlichkeit in wissenschaftlich einwandfreier Weise die grundlegenden Unterschiede zwischen Energieversorgungsoptionen demonstrieren kann. Eine entsprechende Datentabelle wurde vom PSI im Vorfeld einer Volksabstimmung in der Schweiz zu erneuerbaren Energien publiziert. Sie enthält Indikatoren zu Versorgungssystemen, die auf den Energiequellen Steinkohle, Erdgas, Kernspaltung, Wasser und Photovoltaik beruhen. In Tabelle 6 (Kröger et al. 2001) sind auch Emissionen von SO-Verbindungen aufgeführt; sie stehen repräsentativ für negative Gesundheitseffekte. Die Autoren weisen ausdrücklich darauf hin, dass die Quellen der Daten und die getroffenen Annahmen transparent zu machen sind, da bereits geringfügig erscheinende Änderungen beträchtlichen Einfluss auf die Ergebnisse haben können.

Tabelle 5: Konkretisierung des Begriffs „Nachhaltigkeit“ im Zusammenhang mit Energiesystemen (nach Kröger 2000)

Kriterien-Kategorien	Spezifische Kriterien	Indikatoren		Maßeinheit
„Kein“ Ressourcenabbau	Brennstoff- und Materialnutzung	Abbauzeit		Jahre
	Landverluste	Nichtnutzbare Fläche x Dauer	Betrieb	km ² x Jahre
			Schwere Unfälle	km ² x Jahre
	Umwelteffekte	Klimarelevante Gase		Tonnen CO ₂ äquiv.
		Saurer Regen		Tonnen SO ₂ äquiv.
	Gesundheits-Effekte	Akute und späte Todes- fälle	Betrieb	Anzahl Todes- fälle
			Schwere Unfälle	Anzahl Todes- fälle
			Objektive Risiko- abschätzung	Anzahl Todes- fälle
	Soziale Aspekte	Akzeptanz		Δ %
		Proliferationsrisiko		qualitativ
Ökonomische Wettbewerbsfähigkeit		Währungseinheit / kWh		
„keine“ Produktion nicht abbaubarer Abfälle	Abfallmenge			m ³ oder Tonnen
	Einschlusszeit			Jahre
Hohes Potential für Robustheit, langfristige Stabilität	Versorgungssicherheit	Abhängigkeit vom Ausland		qualitativ
		Technologieverfügbarkeit		Währungseinheit
	Resistenz gegen Versagen	Karenzzeit		Stunden
		Soziopolitische Bedingungen		qualitativ

Tabelle 6: Ausgewählte (exemplarische) Indikatoren für aktuelle und zukünftige Energiesysteme (Kröger et al. 2001)

Energy chain		Hard coal	Natural gas	Nuclear	Hydro (storage)	Photovoltaic
Indicator	Unit					
Fuel reserves	years	160 - 2 300	70 - 170	120 - 400	–	–
Material consumption (copper ore)	kg/GWh _e	14 - 19 59	16 8	7 - 9 4	< 1 < 1	270 - 1 600 350
Greenhouse gases	t(CO ₂ -äq)/GWh _e	950 - 1 200 770	530 390	8 - 29 6	4 4	110 - 260 44
Sulphur dioxide	kg(SO _x) / GWh _e	920 - 25 000 520	260 150	56 - 150 33	8 - 10 7	700 - 3 600 160
Inorganic waste in repository	kg/GWh _e	5 800 - 54 000 4 000	1 500 1 100	650 - 1 200 600	30 30	4 900 - 10 000 1 600
High- and middle radioactive waste	m ³ /GWh _e	0.13 - 0.20 0.04	0.04 0.004	9.0 - 11.0 2.4	0.006 0.002	0.6 - 1.2 0.06
Production costs	Rp./kWh _e	5.7 - 7.4 6.3	4.7 - 5.8 4.7 - 8.2	5.1 - 7.5 5.7 - 7.2	4 - 21 12 - 16	70 - 140 45
External (environmental) costs	Rp./kWh _e	3.1 - 15.8 5.1 - 8.6	0.8 - 5.5 2.5 - 4.2	0.2 - 1.3 0.3 - 0.4	0 - 1.2 0.1	0.1 - 1.5 0.5 - 0.7

(Die obere Zahlenreihe stellt ein Bandbreite von Werten für heutige Systeme dar; die untere Reihe enthält die entsprechenden Werte für die besten zukünftigen Systeme (spezifische schweizerische Randbedingungen).)

Behandlung von Risiken

Das GaBE-Projekt räumt dem Themenfeld Risiko einen deutlichen Stellenwert ein. Dabei geht es um Umweltrisiken; sie werden im nachhaltigkeitsorientierten Ansatz (Kröger 2000) dadurch berücksichtigt, dass entsprechende Schutzgüter als Teil von Ressourcen verstanden werden. Ressourcen sind hier in einem weiten Sinn zu verstehen und umfassen auch die Gesundheit von Mensch und Natur, die Wirtschaftlichkeit und den sozialen Frieden. Dabei werden Risiken als Folgen möglicher Störfälle oder Unfälle – insbesondere schwerer Unfälle – als ein spezieller und wichtiger Aspekt der Bewertung von Energiesystemen verstanden; gesundheitliche Folgen („Gesundheitsrisiken“) des Normalbetriebs sind begrifflich von diesen Störfall-Risiken getrennt und Bestandteil der Analysen im Verlauf eines „klassischen“ LCA.

Risiken durch gravierende Störfälle werden daher in einer methodischen Erweiterung des LCA im Rahmen des GaBE-Projekts (Hirschberg 1998; Kröger 2001) durch Auswertung von Berichten über aufgetretene Ereignisse oder mittels Prognoseverfahren – etwa probabilistische Risikoanalysen – einbezogen. Ein breites Spektrum an Schäden wird erfasst (Tote, ernsthaft Verletzte, Evakuierungen, Kontamination von Boden/Wasser, finanzielle Verluste). Eine Aversion gegenüber Schadensereignissen mit Katastrophenausmaß wird durch den Verlust an nutzbarer Bodenfläche und durch die Zahl der Todesfälle gemessen.

Zu den Ergebnissen des GaBE-Projekts gehört die wichtige Feststellung, dass ein Vergleich zwischen verschiedenen Energie-Ketten dadurch erschwert wird, dass zur Ermittlung der Risiken konventioneller Energiequellen statistische Daten vorliegen, wogegen im Fall der Kernenergie – jedenfalls für westliche Reaktoren – die Ergebnisse probabilistischer Risikoanalysen – also von Prognoseverfahren – verwendet werden müssen. Allerdings sind auch statistische Daten nicht ohne Probleme: historische Daten erfordern eine sorgfältige Prüfung auf Konsistenz und Anwendbarkeit.

Aggregieren von Einzelaspekten

Für die Gewichtung und Aggregation der vorgeschlagenen Indikatoren sowie für deren anschließende Nutzung bei Entscheidungsfindungsprozessen existieren Ideen und auch erste Erfahrungen. Ein tragfähiges Konzept ist jedoch noch zu entwickeln (Kröger et al. 2001). Einige Resultate sind von Stromversorgungsunternehmen in der Schweiz zur Strategieentwicklung und in öffentlichen Debatten (Referendum zur Frage der Energiepolitik) verwendet worden.

Eine Möglichkeit zur Aggregation ist die Ermittlung von Kosten. Dabei können Vollkosten, die außer den etablierten Komponenten („interne Kosten“) auch die Kosten für Umwelteinflüsse (als Nutzung der Ressource Umwelt, „externe Kosten“) enthalten, als Maßstab für die Ressourcenintensität von Energiesystemen verwendet werden. Umweltauswirkungen einschließlich Gesundheitsschäden werden als externe Kosten ermittelt (vgl. EcoSense). Externe Kosten werden als ein hochaggregierter Indikator für Umwelteinflüsse verwendet; die Gesamtkosten (interne und externe) werden als ein Nachhaltigkeits-Indikator angesehen. In den letzten Jahren wurden, so Hirschberg (1999), große Fortschritte bei der Berechnung externer Kosten gemacht. Auch wenn die Monetarisierung der Folgen des Treibhauseffektes oder von Schäden an Ökosystemen noch nicht belastbar möglich ist, liegt der prinzipielle Vorteil eines solchen Ansatzes in der Verwendung empirisch ermittelter, individueller Präferenzen, die „den Wertansätzen einzelner Entscheidungsträger oder Gruppen in der Regel vorzuziehen ist“ (Hirschberg 1999).

Ein anderer Ansatz zur Aggregation basiert auf multikriteriellen Entscheidungsverfahren, der auch die soziale Dimension – charakteristisch für die Nachhaltigkeitsdiskussion – berücksichtigt. Die folgende Tabelle 7 zeigt die von Hirschberg (1999) verwendeten Bewertungskriterien und relativen Wichtungen.

Ziel hierbei war die vergleichende Bewertung der Umwelt- und Gesundheitseffekte verschiedener Stromerzeugungssysteme (Erdgas, Wasserkraft, Windenergie, Kernenergie und Photovoltaik), also – im Sinne des vorliegenden Berichts – ein reiner Risikovergleich. Die angegebenen Gewichtungen sind exemplarisch und können fall-spezifisch gewählt werden. Die Quantifizierung der einzelnen Kriterien stammte aus LCA und aus speziellen Risikobetrachtungen zu schweren Störfällen. Unter Verwendung der – wie gesagt exemplarischen – Gewichtungsfaktoren aus der Tabelle ergaben sich die besten aggregierten Werte für die Windenergie und für die Wasserkraft. Versuchsweise wurde außerdem bei der multikriteriellen Bewertung die Gewichtung der drei Hauptkriterien verändert. Dabei wurde unter anderem sichtbar, dass insbesondere die Bewertung der Kernenergie vom Stellenwert schwerer Störfälle abhängt (Hirschberg 1999).

An der Tabelle fällt auf, dass als Belastungen durch den Normalbetrieb Emissionen angegeben sind (ohne Quantifizierung ihrer gesundheitlichen Wirkungen), während die Belastungen durch schwere Störfälle anhand von Todesfällen eingestuft werden. Die aus einem LCA abgeleiteten Emissionen werden dabei als Indikator für Umwelt- und Gesundheitseffekte angesehen (Hirschberg 1999). Damit bleibt eine Aggregation unter dem Aspekt „Gesundheitsschäden“ letztlich offen.

Tabelle 7: Bewertungskriterien und relative Wichtungen für Umwelt- und Gesundheitsschäden (Hirschberg 1999)

Hauptkriterium und Wichtungen	Kriterium (2.Stufe) und Wichtungen	Kriterium (3.Stufe) und Wichtungen
Belastungen durch Normalbetrieb50	Lokal10	CO40
		Lärm40
		Beeinträchtigung des Landschaftsbilds20
	Regional30	SOx30
		NOx40
		NMVOG10
Global (CO-2-Äquivalent)60	Schwebstoffe15	
	Strahlung5	
Belastungen durch schwere Störfälle25	Umwelt25	Kontamination von Boden und Binnengewässern70
		Kontamination von Meerwasser30
	Gesundheit / Soziales75	Auftreten akuter Todesfälle40
		Maximale Anzahl von Todesfällen20
		Auftreten verzögerter Todesfälle20
		Verletzte5
	Evakuierte15	
Abfälle25	Menge50	
	Notwendige Einschlusszeit50	

Ebenfalls als Ergebnis des GaBE-Projekts wird in (Kröger 2001) ein weiteres Anwendungsbeispiel multikriterieller Analyseverfahren vorgestellt. In diesem Fall werden die Hauptkriterien aus dem Drei-Säulen-Modell der Nachhaltigkeit übernommen. Dementsprechend werden ökonomische, ökologische und soziale Aspekte entweder gleich gewichtet oder jeweils besonders betont.

Dabei zeigt sich beispielsweise, dass die Kernenergie unter einer Betonung der ökonomischen oder ökologischen Aspekte relativ gut abschneidet – besser als etwa Erdgas –, dass sie aber ungünstig eingestuft wird, wenn soziale Gesichtspunkte im Vordergrund stehen. Allerdings wurden auch hier die Gesundheits- und Umweltfolgen des Normalbetriebs nicht mit denen von Störfällen bzw. Unfällen zusammengefasst. Untersucht wurden auch die Bewertungschancen einer zukünftigen Kernenergie, bei der die möglichen Unfallfolgen und die erforderliche Einschlusszeit der Abfälle minimiert sind. Unter diesen Randbedingungen und bei einer Gleichgewichtung der Nachhaltigkeitskriterien wird die Kernspaltung als günstigste Energiequelle eingestuft.

Gesamtbewertung

Die Arbeiten zum Poly-Projekt und – insbesondere – des GaBE-Projekts zeigen, dass in letzter Zeit ein umfangreiches und leistungsfähiges Instrumentarium zur vergleichenden Bewertung von Energieversorgungs-Systemen entwickelt worden ist. Die Lebenszyklusanalyse (LCA) bildet dabei die Basis. Die Risiken von Energieversorgungs-Systemen – gesundheitliche und umweltbezogene Risiken – bilden dabei einen Aspekt des Vergleiches. Die fundierte Wahl unter verschiedenen Optionen zur Energieversorgung ist eine wesentliche Zielsetzung der vergleichenden Bewertungen. In Teilaspekten ist das Instrumentarium jedoch noch zu verbessern und abzusichern. Dies betrifft einerseits die Daten zu Störfällen, bei denen – sofern sie aus retrospektiven Erhebungen stammen – ihre fallspezifische Verwendbarkeit und ihre Aussagekraft sorgfältig zu prüfen sind. Bei Störfalldaten aus Prognoseverfahren (etwa probabilistische Risikoanalysen) sollte erkennbar sein, wie sie ermittelt wurden; optimal wären hier standardisierte und qualitätsgesicherte Vorgehensweisen.

Besondere Schwierigkeiten liegen in der Aggregierung unterschiedlicher Facetten der Ergebnisse aus Analysen eines Systems zu einer griffigen Gesamtaussage. Dies betrifft auch eine integrierende Sicht von betrieblichen und störfallbedingten Gesundheits- und Umweltrisiken von Energieversorgungs-Systemen. Generell müssen bei Aggregierungen auch gesellschaftliche Wertentscheidungen einfließen. Dabei können Ansätze aus der Nachhaltigkeitsdiskussion in dem Sinne verwendet werden, dass Nachhaltigkeit mit ihren drei in weitgehendem Konsens gesehenen Dimensionen (Ökologie, Ökonomie, Soziales) ein Rahmenkonzept für Bewertungen darstellen kann.

Die folgende Tabelle gibt einen Überblick über die Vorgehensweise:

Tabelle 8: Vorgehensweise im GaBE-Projekt

GaBE-Projekt	
Zielsetzung	Vergleich alternativer technischer Verfahren „public health“, „public wealth“
Risikobereiche	Risiken durch Anlagen zur Erzeugung elektrischer Energie
zeitlicher Rahmen	im Prinzip unbegrenzt
räumlicher Rahmen	im Prinzip unbegrenzt
Attribute	Mengen ausgewählter Schadstoffe, Landschaftsverbrauch, Lärm; Abfallmenge und Einschlusszeit; Verletzte und Todesfälle durch Unfälle
Aggregierung / Metrik	entweder Monetarisierung oder separate Gewichtung der Attribute
Gesamtbewertung	implizit multiattribut (Monetarisierung) oder explizit multiattribut
Unsicherheiten	Unklar, inwieweit berücksichtigt. Für Indikatoren werden Bandbreiten angegeben.
Ergebnisdarstellung	Tabelle mit Werten der Attribute

3.2.6 Klassierung krebserzeugender Luftschadstoffe für die TA-Luft-Novelle

Dieses Projekt (Schneider et al. 2001; Schuhmacher-Wolz 2002) wird hier als einer der wenigen Anwendungsfälle von Risikovergleichen in Deutschland vorgestellt.

Zielsetzung ist die Einordnung krebserregender Stoffe in eine der drei Klassen der TA-Luft-Novelle an Hand ihrer Wirkstärke, da aus Gründen der Verhältnismäßigkeit Emissionen besonders stark wirkender Kanzerogene strikt begrenzt werden sollen.

Die Beurteilung soll allein an Hand der Wirkstärke erfolgen, d. h. es wird keine Abschätzung der tatsächlichen Exposition oder des tatsächlichen Risikos vorgenommen. Für die Einordnung eines Schadstoffs wird der unit-risk-Wert herangezogen, falls dieser „gut geeignet“ oder „geeignet“ ist; anderenfalls wird auf Basis der vorliegenden Informationen mit Methoden, die nur eine grobe Einschätzung der krebserzeugenden Potenz im Bereich höherer Dosen zulassen, eine Zuordnung vorgenommen.

Hervorzuheben ist, dass hier neben der Höhe des Risikos auch die Abschätzungssicherheit in die Bewertung eingeht: Vorliegende unit-risk-Werte werden nach ihrer Qualität (Aussagesicherheit) beurteilt und einer der Kategorien „gut geeignet“, „geeignet“ und „nicht geeignet“ zugeordnet. Hierfür wird ein detaillierter Kriterienkatalog herangezogen (Schneider 2000).

Tabelle 9: Vorgehensweise bei der Klassierung krebserzeugender Luftschadstoffe für die TA-Luft-Novelle

Klassierung krebserzeugender Luftschadstoffe für die TA-Luft-Novelle	
Zielsetzung	Begründung für Grenzwertsetzungen für kanzerogene Noxen
Risikobereiche	Krebsrisiko
zeitlicher Rahmen	Expositionsdauer, Lebensdauer des Exponierten
räumlicher Rahmen	spielt keine Rolle, da Schadstoffkonzentrationen betrachtet werden
Attribute	Krebserkrankung
Aggregation / Metrik	unit risk
Gesamtbewertung	an Hand des unit risk; Zuordnung zu einer von drei Klassen
Unsicherheiten	Angabe und Berücksichtigung von Unsicherheitsbandbreiten der unit risk-Werte
Ergebnisdarstellung	Tabelle mit unit risk-Werten und Klassen-Zuordnungen

3.2.7 Zusammenfassung

Zusammenfassend lässt sich für die sechs diskutierten Fallbeispiele feststellen, dass Risikovergleiche sich im Hinblick auf die Zielsetzung, die Bildung von Risikokategorien, den zeitlichen Rahmen, die Vergleichsmetriken und die Randbedingungen unterscheiden.

Zielsetzung: Typische Zielsetzungen von Risikovergleichen sind das Festlegen von Prioritäten in der Gesundheits- oder Umweltpolitik, die Wahl zwischen technischen oder administrativen Alternativen und die Begründung für Grenzwertsetzungen.

Bildung von Kategorien: Ein typisches Beispiel für Probleme der Kategorienbildung ist das US EPA „Unfinished Business“ Projekt. Hier orientierten sich die zu vergleichenden Kategorien von Risiken an der internen Organisationsstruktur der EPA, die sich ihrerseits an ihrem durch Gesetze vorgegebenen Auftrag orientieren muss. Es ist dabei nicht nebensächlich, wie kategorisiert wird, denn je umfangreicher eine Kategorie ist, desto größer ist die Chance, dass sie in der Rangliste einen vorderen Platz einnimmt. Die im „Unfinished Business“ Projekt definierten Kategorien haben außerdem den Nachteil, dass Emissionsquellen und Wirkungspfade und auch Wirkungen als Kategorien verwendet wurden, wodurch einige Risiken doppelt gezählt wurden.

Zeitlicher Rahmen: Die zeitliche Mittelung der Folgen seltener Unfälle mit hohem Schaden, um sie dann mit betrieblichen Risiken zusammenfassen zu können, ist problematisch (siehe ExternE-Projekt). Wir schlagen daher vor, solche Risikobeiträge separat zu vergleichen.⁴¹

Vergleichsmetrik: Bei den Projekten mit den Zielsetzungen „Prioritäten in der Gesundheits- oder Umweltpolitik“ und „technische oder administrative Alternativen“ waren Risiken zu vergleichen, die sich z. B. hinsichtlich Schadensart, Zeitpunkt des Schadenseintritts (Frage der Diskontierung) und Gruppen unterschiedlich sensibler Personen unterscheiden. Um diese Beiträge aggregieren und vergleichen zu können, müssen sie auf eine gemeinsame Metrik umgerechnet werden. Da hierbei zwangsläufig Wertungen vorgenommen werden, solche Wertungen aber grundsätzlich subjektiv sind, wird die verwendete Metrik oft kritisiert. Die Metrik bedarf daher der Legitimation.

Bei dem Projekt mit dem Ziel „Grenzwertsetzung“ war ein gemeinsames Vergleichsmerkmal (Krebsrisiko) vorhanden, für das eine etablierte Metrik (*unit risk*) verfügbar ist.

Randbedingungen: Das ermittelte Risiko ist abhängig von vielen Randbedingungen, wie z. B. räumliche Verteilung der Schutzgüter (Bevölkerung, Ökosysteme, Sachwerte) in der Umgebung der Risikoquelle, bereits vorhandene Belastungen aus anderen Risikoquellen (dies ist insbesondere relevant bei Schadstoffen mit nichtlinearer Dosis-Wirkungs-Beziehung), Wechselwirkungen, die zu Abbau oder Aufbau von Schadstoffen führen, wie sie z. B. in ExternE zwischen SO₂, NO_x und NH₃ festgestellt wurden.

⁴¹ Siehe dazu im Einzelnen Kapitel 5 und 6.

Daraus folgt, dass Risikoangaben immer nur für das bei der Analyse vorausgesetzte oder bei der statistischen Erhebung vorhandene Gesamtszenario gelten.

Im folgenden Kapitel werden die hier aufgezeigten Fragestellungen und Probleme der vergleichenden Risikobewertung auf der Basis der empirischen Risiko- und Entscheidungsforschung genauer untersucht.

4. Empirische Grundlagen der CRA

4.1 Ausgangspunkt für Risikovergleiche

Die technischen Voraussetzungen für einen Risikovergleich sind simpel: Für die vergleichenden Risiken müssen sich Ausprägungen auf mindest einem gemeinsamen Attribut angeben lassen, auf die hin der Vergleich durchgeführt werden kann.⁴² Eine weitere notwendige Randbedingung ist, dass die Varianz der Ausprägungen nicht so groß sein darf, dass sie eine eindeutige Ordnung ($<$, $>$, $=$) der Risiken unmöglich macht. Zum Beispiel lassen sich Risiken von genetisch veränderten Organismen mit dem Risiko von Meteoriteneinschlägen dann vergleichen, wenn in beiden Fällen Todesfälle betrachtet werden und diese sich mit einer Genauigkeit abschätzen lassen, die eine eindeutige Rangbildung ermöglicht. Hinzu kommt, dass die entsprechenden Daten auch verfügbar sein müssen.

Welcher Vergleich jedoch sinnvoll ist, steht auf einem anderen Blatt. Die Frage, welche Risiken sich vergleichen lassen, müsste deshalb eher lauten: Warum sollten die ausgewählten Risiken verglichen werden? Die Klärung dieser entscheidungs- und handlungsrelevanten Voraussetzungen von Risikovergleichen ist von zentraler Bedeutung und soll im weiteren genauer erörtert werden.

Weitgehend unstrittig ist, Risiken von Entscheidungsalternativen zu vergleichen. Solche Vergleiche haben im Feld der Technikbewertung eine lange Tradition. Ebenso geläufig ist es, verschiedene Therapien nach ihren Risiken vergleichend zu bewerten. In der Technik beispielsweise werden Standorte oder technische Optionen für die Energiewirtschaft (Keeney 1980, siehe aber auch Kunreuther und Linnerooth 1983), für die Abfallwirtschaft (Powell 1996) sowie Optionen von Gefahrguttransporten (Schiene versus Straße) miteinander verglichen. Risikovergleiche zwischen verschiedenen technologischen Optionen (zum Beispiel Kernkraft versus fossile Brennstoffe) stehen aber vor dem Problem, wie sich Risiken vergleichen lassen, die einerseits durch eine Kombination von sehr kleinen Häufigkeiten mit sehr großen Schäden und andererseits durch relativ große Häufigkeiten bei kleineren Schäden entstehen. Anders formuliert stellt sich die Frage, wie Katastrophenpotenziale zu gewichten sind.⁴³

Auch ganz verschiedene Risiken lassen sich vergleichen. So könnten der Straßenbau, die Verstädterung, die Umweltverschmutzung und die Anwendung der Biotechnologie im Pflanzenbau auf das Kriterium „Verlust der Biodiversität“ hin verglichen werden, um Prioritäten zu setzen oder um Schieflagen zu erkennen.

⁴² Beim Vergleichen werden mindestens zwei Objekte gedanklich in Beziehung gesetzt, um hinsichtlich ihrer Merkmale und Relationen Ähnlichkeiten, Unterschiede oder Gemeinsamkeiten nach bestimmten Aspekten zu bestimmen. Auf dieser Grundlage lassen sich z.B. Ordnungen bilden, d.h. Gruppierungen oder Rang- bzw. Reihenfolgen nach einem oder mehreren inhaltlichen oder formalen Merkmalen. Fachlexikon der Psychologie, Deutscher Verlag, 1995.

⁴³ Dabei wäre konsequenterweise vorauszusetzen, dass bei fossilen Energieträgern ebenfalls eine Störfallanalyse über hypothetische worst case Fälle vorzunehmen ist.

Außerdem können Risiken von medizinischen Untersuchungen mit ionisierender Strahlung mit den Risiken von Radonbelastungen verglichen werden. Ziel könnte hier die Verdeutlichung der relativen Bedeutsamkeit der Risiken sein.

Die Beispiele zeigen, dass es darauf ankommt, den Handlungskontext genau zu klären und den Fokus der Risikovergleiche in diesem Handlungskontext zu bestimmen.

Im wesentlichen lassen sich drei Handlungskontexte unterscheiden, die in Tabelle 10 aufgeführt sind.

Tabelle 10: Kontexte von Risikovergleichen

Fokus	Handlungskontext	Ziel
Kommunikationsproblem	Risikoverständnis	Vermittlung der Bedeutsamkeit von Risiken
Priorisierungsproblem	Risikobewertung	Ranking / Kategorisierung
Entscheidungsproblem	Risikomanagement	Beste Option / Variante / Handlungsbegründung

Im weiteren soll auf diese unterschiedlichen Zielstellungen von Vergleichen genauer eingegangen werden. Hier werden wir uns auf zwei Bereiche konzentrieren: auf Risikovergleiche als Mittel der Risikokommunikation sowie auf Verfahrensprobleme der CRA bei der Prioritätensetzung. Darüber hinaus sollen organisatorische Probleme bei der CRA und kommunikative Probleme bei der Ergebnisvermittlung in Politik und Öffentlichkeit erörtert werden.

4.2 Risikovergleiche als Mittel der Risikokommunikation

Unter Risikovergleich versteht man hier den Vergleich eines unbekannten mit einem bekannten Risiko. Solche Vergleiche dienen primär der Verständnissicherung. Wie bereits ausgeführt, setzen Vergleiche voraus, dass die zu vergleichenden Objekte zumindest ein Attribut gemeinsam haben, auf das hin sie verglichen werden können.

In der Risikokommunikation haben Vergleiche immer eine besondere Rolle gespielt. Dabei werden zwei oder mehr Risiken gegenübergestellt, um ihre Art, ihr Ausmaß und damit ihre relative Bedeutung aufzuzeigen. Im Kern ging es dabei um eine angemessene Beurteilung der "Riskantheit" eines Risikos und um die Minimierung von Aufregungsschäden. Allerdings wurde schnell klar, dass sich diese Erwartungen nicht ohne weiteres realisieren lassen.

Die ersten methodisch reflektierten Vorschläge zu Risikovergleichen stammen aus dem Jahr 1988. Damals veröffentlichte die Chemical Manufacturers Association der USA einen Leitfaden von Covello, Sandman und Slovic zu „Risikokommunikation, Risikostatistik und Risikovergleichen“, in dem sich auch ein Kapitel über die Erfolgsbedingungen von Risikovergleichen findet.

Nach Auffassung der Autoren helfen Vergleiche, Risiken in Perspektiven zu setzen. Sie helfen auch, die Bedeutung eines Risikos besser zu verstehen. Andererseits wird darauf hingewiesen, dass Risikovergleiche heikel sind, weil sie als Versuch wahrgenommen werden können, einer besorgten Öffentlichkeit klar zu machen, dass sie sich um die falschen Risiken sorgt. Dass solche Vorhaltungen schlecht ankommen

können, versteht sich fast von selbst. Es war deshalb die Intention von Covello, Sandman und Slovic (1988) zu zeigen, welche Risikovergleiche Sinn machen und welche nicht.

Sie fokussieren dabei auf den Typ der entwarnenden Risikovergleiche, d.h. solche, die zeigen, dass ein bestimmtes Risiko – verglichen mit anderen – weit weniger bedeutsam ist. Warnende Risikovergleiche haben es ihrer Auffassung nach einfacher: sie treffen weniger auf Vorbehalte. Die Autoren ordnen Risikovergleiche in 5 Ränge, die deren Akzeptanz ausdrücken sollen:⁴⁴ Je niedriger der Rang, desto geringer wird die Akzeptanz eingeschätzt (siehe Tabelle 11).

Tabelle 11: Akzeptanzränge der Risikovergleiche nach Covello, Sandman und Slovic

<i>Risikovergleiche</i> <u>Rang 1: (höchste Akzeptanz)</u> • Vergleiche von gleichen Risiken zu verschiedenen Zeitpunkten • Vergleiche mit einem Standard (Grenzwert) <u>Rang 2:</u> • Vergleich von Risiken bei Durchführung gegenüber Risiken bei Unterlassung einer Handlung • Vergleich von Risiken von verschiedenen Optionen für ein Problem • Vergleich mit dem gleichen Risiko an einem anderen Ort <u>Rang 3:</u> • Vergleich des durchschnittlichen Risikos mit dem Exzess-Risiko • Vergleich eines Risikos für einen adversen Effekt von einer Quelle mit den Risiken von allen anderen Quellen für den gleichen adversen Effekt <u>Rang 4:</u> • Vergleich von Risiken und Kosten bzw. Risiko-Kosten-Verhältnissen • Vergleich von Risiken mit Nutzen • Vergleich von beruflichen mit umweltbezogenen Gesundheitsrisiken • Vergleich mit anderen Risiken von der gleichen Quelle • Vergleich mit anderen Ursachen für den gleichen Schaden <u>Rang 5: (niedrigste Akzeptanz)</u> • Vergleich von Risiken, die in keiner Beziehung zueinander stehen
--

Das der Tabelle zugrunde liegende Modell ist relativ einfach: Risikovergleiche sollen um so akzeptabler sein, je plausibler und eindeutiger dem Vergleichenden der gemeinsame Aspekt erscheint, der einen sinnvollen Vergleich von zwei Sachverhalten erst möglich macht. Schon der Vergleich eines Risikos mit einem Standard wird beispielsweise für angemessener gehalten als der Vergleich eines bestimmten Risikos mit einem Risiko aus einer anderen Quelle, das einen gleichartigen Schaden verursacht. Insbesondere würde aus dem Modell von Covello folgen, dass eine CRA, die dem Rang 4 oder 5 zuzuordnen wäre, auf große Schwierigkeiten stoßen müsste. Dem ist aber nicht so. Ein Survey zur Akzeptanz von CRA-Verfahren aus den USA (Charlton Research Company 2001), in dem 800 Personen befragt wurden, ob sie

⁴⁴ "... It is possible to rank different kinds of risk comparisons in terms of their acceptability to people in the community." (Covello, Sandman und Slovic, 1988, S. 17).

die CRA für eine gute Idee halten, erbrachte folgendes Ergebnis: 39% der Befragten halten die CRA für eine gute Idee (starke Zustimmung) und immerhin 41% stimmen dieser Einschätzung schwach zu (Bewertung: Etwas). Nur 11% lehnen das Konzept der CRA ab.

Aber auch sonst ist das Modell von Covello et al. problematisch. Das soll an folgendem Beispiel verdeutlicht werden. Nehmen wir an, ein Hausbesitzer lässt eine Radonmessung machen, bei der eine Radonkonzentration von 800 Bq/m^3 gemessen wird. Die SSK hat einen Sanierungswert für Radon festgelegt (SSK 1994); danach ist bei einer Radonkonzentration von mehr als 1000 Bq/m^3 eine Sanierung erforderlich. Der Vergleich bestünde also hier in dem Gegenüberstellen von Mess- und Aktionswert. Dieser 'Rang 1'-Vergleich ist nach Covello et al. besser als der nachstehende 'Rang 4'-Vergleich, der von einem Schweizer Merkblatt stammt: „Aufgrund heutiger Erkenntnis, ist Radon nach dem Rauchen die häufigste Ursache für Lungenkrebs.“

Unseres Erachtens macht das Beispiel deutlich, dass die Akzeptanzränge nicht ohne weiteres nach der Ähnlichkeit der Risiken vergeben werden können. Es spielt sicher auch eine Rolle, wie bekannt das Referenzrisiko ist und ob die Vergleiche Handlungsoptionen nahelegen.

Dennoch hat das Modell von Covello et al. eine überraschend hohe Resonanz gefunden und ist noch heute in Übersichten und Empfehlungen zur Risikokommunikation zu finden (vgl. WHO 2002). Dabei wird jedoch übersehen, dass diese Akzeptanzordnung nicht durch empirische Untersuchungen gestützt wird. Anzumerken ist außerdem, dass bei Risikovergleichen vier unterschiedliche Aspekte zu unterscheiden sind:

- die Akzeptanz des Risikovergleichs,
- der Beitrag, den der Vergleich für das Verständnis des Risikos liefert,
- der Effekt von Vergleichen auf die Risikowahrnehmung und
- der Einfluss der Risikovergleiche auf die Akzeptanz eines Risikos.

Diese Aspekte können zu einem Kaskadenmodell verknüpft werden: Zuerst muss überhaupt der Vergleich akzeptiert werden. Erst dann ist es möglich, dass er Wirkungen entfalten kann. Eine erste kognitive Wirkung kann in der Verbesserung des Verständnisses liegen. Hier sind allerdings verschiedene Möglichkeiten denkbar: Größenordnungen werden verständlicher, die Bedeutung des Risikos kann besser eingeschätzt werden, oder das Ausmaß der Unsicherheit des Risikos wird verständlich. Erst wenn diese kognitiven Voraussetzungen vorliegen, kann mit einem Einfluss auf die Risikowahrnehmung gerechnet werden. Eine Veränderung der Risikowahrnehmung kann - muss aber nicht - zu einer veränderten Risikoakzeptanz führen.

Eine erste Studie, die die Wirkung von Risikovergleichen im Bezug auf das Radon-Risiko belegt, wurde von Weinstein et al. (1989) durchgeführt. Sie konnte zeigen, dass Vergleiche das Risikoverständnis verbessern.

Eine weitere Studie stammt von Roth et al. (1990). Sie konstruierten Risikokommunikations-Szenarien, die die Probanden⁴⁵ vor die Aufgabe stellten, einen Manager bei

⁴⁵ Die Untersuchungsteilnehmer in psychologischen Experimenten werden Probanden genannt.

der Auswahl von Risikovergleichen für eine Präsentation zu beraten. Insgesamt wurden 14 verschiedene Risikovergleiche vorgegeben, die die Probanden auf 7 Skalen zu bewerten hatten (Klarheit, Verständnishilfe, Nützlichkeit der Information, Über/Unterbewerten des Risikos, Beruhigung, Vertrauensbildung sowie die Skala „sollte genutzt werden“). Die 14 Risikovergleiche stammen dabei aus den fünf verschiedenen Rängen der Kategorisierung von Covello, Sandman und Slovic.

Auf der Skala „sollte genutzt werden“ wurden alle Risikovergleiche ähnlich gut bewertet (Werte zwischen 1.82 und 2.92, auf einer 7-stufigen Skala mit dem besten Wert 1). Eine Korrelationsanalyse zwischen den Punktwerten auf den 7 Beurteilungsskalen und den Risikovergleichsrängen nach Covello, Sandman und Slovic erbrachte durchweg negative Resultate, bis auf die Skala „Vertrauensbildung“. Hier gab es praktisch eine Nullkorrelation ($r = 0.1$). Offenbar spielen die Akzeptanzränge nicht die postulierte Rolle.

Roth et al. (1990) kommen zu dem Schluss, dass die Akzeptanz von Risikovergleichen nicht von der Ähnlichkeit der Risiken abhängt. Experimente von Johnson (2003) stützen diese Einschätzung. Auch in dieser Studie findet sich kein Zusammenhang zwischen den Bewertungen auf den 7 Skalen (Klarheit, Verständnishilfe, Nützlichkeit der Information, Über/Unterbewerten des Risikos, Beruhigung, Vertrauensbildung sowie auf der Skala „sollte genutzt werden“) und den Risikovergleichsrängen. Darüber hinaus konstruierte Johnson auch ein Ablehnungsszenario. Hier wurde den Probanden mitgeteilt, dass der Risikomanager auf Misstrauen und Ablehnung treffen wird. In diesem Kontext zeigt sich ebenfalls kein Zusammenhang. Das Modell von Covello et al. (1988) steht somit empirisch auf schwachen Beinen.

Slovic et al. (1990) berichten über ein weiteres Experiment, in dem drei Varianten vorgegeben wurden. Unter Bedingung A erhielten die Probanden ein einfaches Risikoszenario (Asbestbelastung in einer Schule, Kontrollbedingung), unter Bedingung B dazu noch Risikovergleiche und schließlich unter Bedingung C neben den Risikovergleichen einen kritischen Kommentar, der die Risikovergleiche in Frage stellte. Dabei zeigte sich unter Bedingung B eine verringerte Risikowahrnehmung (als Effekt des Risikovergleichs), unter Bedingung C wurde diese Wirkung wieder aufgehoben. Das zur Verteidigung des Ansatzes von Covello, Sandman und Slovic gedachte Experiment stützt jedoch nicht die angenommenen Akzeptanzränge. Vielmehr macht es deutlich, wie komplex die Bedingungen sind, von denen die Wirkung von Risikovergleichen abhängt.⁴⁶ Offenbar spielen Kontextfaktoren eine wesentliche Rolle.

Auch dieses Experiment wurde von Johnson (2002) wiederholt. In einer umfangreichen Studie überprüfte er, ob verschiedene Bedingungen der Informationsvorgabe einen Einfluss auf die Wirkung von Risikovergleichen besitzen. Dabei konnte er die Befunde von Slovic et al. replizieren. Außerdem zeigen seine Resultate, dass eine Inokulation (d.h. eine vorwegnehmende Auseinandersetzung mit Kritik an Risikovergleichen) die Wirkung der Kritik einschränkt.

Ein drittes Experiment stammt von Freudenburg und Rursch (1994). Die Autoren gaben ein Risikoszenario vor (Bau einer Sondermüllverbrennungsanlage) und die Probanden sollten sich in die Rolle der Anwohner versetzen. Zuerst wurde nach der Haltung zum Bau der Anlage gefragt (63% der Befragten waren dagegen), dann

⁴⁶ Leider führen Slovic et al. (1990) keine Signifikanztests durch. Das schränkt ihre Aussagen ein.

wurde ihnen ein Statement vorgelegt (eine unabhängige Studie, beauftragt von der Industrie, findet ein Risiko von 1 : 1 Million). 36% der Probanden gaben dann an, dass sie den Bau der Anlage eher unterstützen würden (keine Änderung: 54,9%, weniger Unterstützung: 8,5%). Danach wurde ein Risikovergleich vorgestellt (1 : 1 Million ist ein geringeres Risiko als das Rauchen einiger Dutzend Zigaretten). Hier sagen 62,1%, dass dieser Vergleich sie nicht umstimmt, 18,2% würden dann den Bau der Anlage eher unterstützen und 19,3% würden ihn weniger unterstützen.

Johnson und Chess (2003) untersuchten die Wirkung von Vergleichen mit einem Standard auf die Risikobewertung. Vorab: Ungefähr ein Drittel bis zur Hälfte der Befragten wies eine negative Einstellung gegenüber Grenzwerten auf. Sie glaubten z. B. nicht, dass die Grenzwerte wirklich die Gesundheit schützen. Oder sie nahmen an, dass bei der Festlegung von Grenzwerten die Kosten für die Industrie stärker beachtet werden als der Gesundheitsschutz. In dem Experiment wurden verschiedene Expositionsszenarien vorgegeben, die den Grenzwert unterschiedlich ausschöpften (95% oder 50% vom Grenzwert). In allen Treatmentbedingungen waren die Probanden verunsichert und nahmen an, dass sie einem ernsthaften Risiko ausgesetzt sind. Risikovergleiche mit einem Grenzwert haben demnach keine beruhigende Wirkung.

Die hier beschriebenen Experimente zeigen, dass die Akzeptanz von Risikovergleichen vergleichsweise hoch ist und nicht von der Akzeptanzordnung nach Covello, Sandman und Slovic (1988) abhängt. Die Probanden von Roth et al. (1990) und Johnson (2002) bewerten die Akzeptanz der Vergleiche bis auf eine Ausnahme mit Werten zwischen 2 und 3 auf einer 7er Skala (deren bester Wert 1 ist). Der Beitrag der Risikovergleiche für das Risikoverständnis ist – siehe die Daten von Roth et al. (1990) und Johnson (2002) – ebenfalls gut (Werte zwischen 2 und 3 auf einer 7er Skala mit dem besten Wert 1), unabhängig vom situativen Kontext (d.h. Empörung oder neutral-positives Klima). Die Wirkung der Risikovergleiche auf die Risikowahrnehmung ist jedoch eher marginal. Bis auf einige Ausnahmen ergaben sich nur nicht-signifikante Effekte. Das erklärt z.T. auch die (vermutete und bislang nicht durch empirische Studien belegte) negative oder ausbleibende positive Wirkung auf Akzeptanzbeurteilungen von Risiken. Denn Akzeptanzveränderungen können auf zweierlei Weise zustande kommen: Zum einen durch eine veränderte Risikowahrnehmung, zum anderen durch einen veränderten Akzeptanzstandard. Die Wirkungen eines Risikovergleichs hängen also keinesfalls nur von der Ähnlichkeit der Risiken ab. Vielmehr spielen offenbar die wahrgenommene Absicht des Vergleichs sowie die Konsistenz bzw. Inkonsistenz der vorhandenen Informationen über Größe und Bedeutsamkeit des Risikos eine Rolle.

Insgesamt zeigt sich also:

- Risikovergleiche werden akzeptiert.
- Sie haben eine klare Wirkung auf das Risikoverständnis.
- Sie haben kaum eine Wirkung auf die Risikowahrnehmung.
- Sie haben offenbar fast keine Wirkung auf die Risikoakzeptanz.

Daraus folgt, dass es keine wissenschaftlichen Evidenzen gibt, die explizit gegen die Durchführung einer CRA sprechen. Zwar ist die Wirkung von Risikovergleichen auf die Risikowahrnehmung gering, es muss aber bedacht werden, dass eine integrative CRA auch die Beteiligung der Bürger einschließt – sie sind nicht nur Konsumenten

von Risikovergleichen. Die Ergebnisse legen außerdem nahe, die Aspekte Akzeptanz von CRA-Verfahren, Beitrag der CRA für das Verständnis der Größe von Risiken, Einfluss der CRA auf die Risikowahrnehmung und Akzeptanz der Ergebnisse der CRA ebenfalls zu unterscheiden und als Aufgabenfelder bei der Entwicklung und Durchführung einer CRA zu beachten.

4.3 Verfahrensprobleme

Florig et al. (2001) weisen darauf hin, dass eine Vielzahl der bislang durchgeführten CRAs methodischen Standards nicht genügen und so nicht ausreichend valide sind. Sie fordern eine stärkere Berücksichtigung von Erkenntnissen der Entscheidungsforschung, der Einsichten aus der Forschung zur Risikowahrnehmung und Risikokommunikation sowie der psychologischen Untersuchungen zur menschlichen Informationsverarbeitung.

Wir wollen im weiteren dazu einen Überblick geben. Im Mittelpunkt stehen: Die Festlegung des Rahmenkonzepts, die Entwicklung von Risikokategorien, die Auswahl der Attribute, die Bewertung der Optionen auf den Attributen und die Durchführung der Vergleiche. Hier sollen - ausgehend von empirischen Befunden - Fehler und Fallen bei der Durchführung einer partizipativen vergleichenden Risikobewertung aufgezeigt werden. Das heißt, es wird immer davon ausgegangen, dass neben Experten auch Vertreter gesellschaftlicher Gruppen und Laien an der CRA teilhaben.

Rahmenkonzepte: Was ist bei der Festlegung von Systemgrenzen für die CRA zu beachten?

Risikovergleiche setzen Grenzziehungen voraus: Es sind geographische, zeitliche und inhaltliche Abgrenzungen nötig. Geographisch sind Regionen zu bestimmen (z.B. Deutschland, NRW, Kreis Düren), für die die Risikovergleiche gemacht werden sollen. Gleiches gilt für den zeitlichen Rahmen. Es macht einen Unterschied, ob nach bislang existenten Risiken (retrospektive Blickrichtung) oder nach künftigen Risiken (prospektive Blickrichtung) gefragt wird.

Das Rahmenkonzept für die Risikovergleiche ist nicht natürlich vorgegeben, sondern es wird - je nach Zielstellung - konstruiert bzw. ausgewählt. Damit werden unterschiedliche Grenzen gezogen, je verschiedene Risikoklassen fokussiert sowie auch unterschiedliche Referenzpunkte und verschiedene Messgrößen⁴⁷ nahe gelegt. So ist vorstellbar, dass einmal von eher vorsichtigen Annahmen ausgegangen wird, zum anderen aber von eher optimistischen. Tabelle 12 zeigt dafür ein Beispiel:

⁴⁷ Ein Beispiel für die Anfälligkeit für Framing-Effekte auf der Ebene von Messgrößen wäre z.B., dass Risiken auf der Basis von relativen Risiken verglichen werden, und damit Inzidenzen, die eigentlich wichtig sind, ausgeblendet werden.

Tabelle 12: Beispiel für Wirkungen von Rahmenkonzepten

	Rahmen A	Rahmen B
Einstellung	Vorsichtig	Optimistisch
Verursachung	Fokus auf multiple Stressoren	Fokus auf einen Stressor
Konsequenzen	Analyse von Folgen von Folgen	Analyse von Folgen erster Ordnung
Toxizität	Fokus auf positive Fälle	Fokus auf wissenschaftliches Gesamtbild
Exposition	Einbeziehen verschiedener Expositionspfade	Einbeziehen des angenommenen Hauptpfades

Das Rahmenkonzept hat aber einen beträchtlichen Einfluss auf nachfolgende Schritte, insbesondere auf das Bewerten und Vergleichen, wie die psychologische Entscheidungsforschung zeigt, die in diesem Zusammenhang von Framing-Effekten spricht (siehe Russo und Schoemaker 1998, Schoemaker und Russo 2001).

Schoemaker und Russo (2001) betonen hierbei die Framing-Fallen. Dazu zählen: (1) die Illusion der vollständigen Problemrepräsentation, (2) Übervertrauen in das vorhandene Wissen und damit die Dominanz des eigenen Frames, (3) die Tendenz, nur solche Informationen zu betrachten, die den eigenen Frame unterstützen.

In Zusammenhang mit der ersten Framing-Falle ist eine Untersuchung von Fischhoff, Slovic und Lichtenstein (1978) aufschlussreich. Sie zeigt, dass Menschen die Vollständigkeit von Problemrepräsentationen überschätzen - sie übersehen Auslassungen. Es ist leicht vorstellbar, dass diese Falle auch bei der Zusammenstellung von Risiken für eine vergleichende Bewertung auftreten kann. Ob eine Repräsentation eines Bereichs, z.B. aller Risiken für ein Ökosystem oder aller gesundheitlichen Risiken in einer Region vollständig oder jedenfalls ausreichend ist, kann falsch eingeschätzt werden. Dabei spielt das Übervertrauen eine Rolle: Übervertrauen ist die Tendenz, sich zu sicher zu sein, dass die eigenen Annahmen und Beurteilungen richtig sind. In einer Reihe von Experimenten wurde nachgewiesen, dass auch bei Experten mit diesem Effekt zu rechnen ist, wenn sie Schätzungen abgeben (vgl. Bazerman 1990). Der Effekt, vorzugsweise bestätigende Informationen zu betrachten, zeigt sich ebenfalls auch bei Experten (Einhorn und Hogarth 1978).

Als wesentliche Konsequenz bleibt festzuhalten: Mit der Zusammenstellung der Problemliste für die vergleichende Risikobewertung werden wichtige Weichenstellungen vorgenommen. Es ist deshalb viel Sorgfalt auf eine adäquate Problemrepräsentation zu legen. Über verschiedene Anläufe mit unterschiedlichen Leitfragen sollte versucht werden, zu einer vollständigeren, nach konsistenten Kriterien ausgewählten und strukturierten Liste von Risikoquellen zu kommen. Es empfiehlt sich, theoriegeleitet und fragestellungsbezogen vorzugehen und homogene Klassen zu bilden, die von einem Kettenglied aus der Risikokette ausgehen. In der Regel werden das initiierte Ereignisse (Emissionen) sein, die sich einer Risikoquelle zuordnen lassen (Stoffklassen: z.B. Emissionen von Verkehr).

Risikokategorien: Welchen Einfluss haben sie auf das Vergleichen?

Es ist beinahe trivial festzustellen, dass die Vergleichsobjekte bei Vergleichen eine zentrale Rolle spielen. Die Tücken liegen allerdings im Detail.

Ein Blick auf die Risikolisten durchgeführter CRA lässt erkennen, dass in der Regel mit sehr heterogenen Risikokategorien gearbeitet wird. Neben Umwelt- und Gesundheits- sowie Wohlfahrtsproblemen finden sich Innenraumluftverschmutzung, Abfall, Mangel an Umweltbewusstsein, globale Klimaveränderungen, Verlust innerstädtischer Grünflächen u.a.. Es werden aber auch Technologien und Endpunkte aufgeführt (z.B. Müllverbrennung und Aids). Der SAB (1990) – in seiner Kritik an dem Unfinished Business Projekt – hat bereits ausgeführt, welche Messprobleme und Vergleichsprobleme man sich einhandelt, wenn derart heterogene Referenzsysteme ausgewählt werden. Das Gesagte soll an Hand zweier Beispiele verdeutlicht werden.

Das Risiko-Ranking des „Environmental Strategies Project for Metro Denver“⁴⁸ ist das erste Beispiel. Hier fällt auf, dass die Risikokategorien als Risikoquellen (z.B. Mülldeponien), als Kontaminanten (z.B. Luftschadstoffe), aber auch als zu schützende Güter (Drinking water, Surface water) angegeben werden. Damit entstehen aber Überlappungen. So kann beispielsweise eine Risikoquelle wie eine Mülldeponie (Superfund Site) auch Auswirkungen auf das Trinkwasser haben. Ebenso ist es möglich, dass die unterirdischen Tankbehälter (Underground Storage Tanks) die Qualität des Grundwassers beeinträchtigen. Durch den Wechsel der Referenzsysteme schleichen sich in das Ranking Fehlerquellen ein.

In dem Risiko-Ranking in Florida kommen dagegen ganz andere Risikokategorien zur Anwendung (vgl. Florida Center for Public Management 1995). Hier dominieren zu schützende Güter wie Innenraumluft-Qualität, Nahrungsmittel-Qualität sowie die landschaftlichen (*scenic*), historischen und kulturellen Ressourcen. Daneben finden sich Risikotreiber wie „Patterns of Delevopment“ und „Use and Management of public lands“. Auf diese Weise entstehen Ungereimtheiten: Nicht klar ist, welche Risiken z.B. die Nahrungsmittelqualität beeinträchtigen oder welche Auswirkungen die angegebenen Risikotreiber haben.

Im folgenden werden diese Kategorisierungsprobleme aus zwei Perspektiven betrachtet. Zum einen geht es um den Einfluss kognitiver Ordnungsbildungen, zum anderen um affektive Einflüsse, die aus den Bezeichnungen bzw. Beschreibungen der Risikoklassen resultieren können.

Für einen Vergleich sind Objekte, d.h. in unserem Fall Risikokategorien, erforderlich. Denn Vergleiche auf der Grundlage von einzelnen Stoffen sind nicht immer möglich, da dann zehntausende und mehr Stoffe verglichen werden müssten. Deshalb müssen Risiken gruppiert werden. Was aber ist eine adäquate Gruppierung? Diese scheinbar einfache Frage ist nicht leicht zu beantworten.

Morgan et al. (2000) schlagen ein Klassifikationsschema vor, das sich an der Risikokette von Hohenemser, Kasperson und Kates (1985) orientiert. Ihrer Auffassung nach kann die Kette, die von riskanten Aktivitäten bis hin zu Endpunkten einschließ-

⁴⁸ Comparison of Human Health Risks and Economic Damages from Major Sources of Environmental Pollution in Denver. Industrial Economics, Inc. June 1998.

lich deren Bedeutung für Betroffene reicht, benutzt werden, um ein Set von Risiken für die CRA auszuwählen. Beispielsweise kann damit begonnen werden, Aktivitäten in einer Region zusammen zu stellen, die zu Luftverschmutzungen führen: Abgase von Verkehr (PKW, Lastkraftwagen), Kohlekraftwerke, Chemiewerke, usw. Diese können dann weiter spezifiziert werden anhand der emittierten Schadstoffe, die wiederum zu endpunktspezifischen Schadstoffgruppen zusammengefasst werden können. Ähnliches wird auch von der EPA (1993) vorgeschlagen.

Im Ergebnis solcher Klassenbildungen sollten dann Kategorien vorliegen, die

- logisch konsistent sind (erschöpfend, nicht überlappend, und weitgehend homogen, so dass mit den ausgewählten Attributen auch bewertet werden kann);
- kognitive Biases möglichst klein halten (Verständnis erleichtern, Framing-Effekte minimieren) und Grenzen der Informationsverarbeitung berücksichtigen (z.B. überschaubare Anzahl von Kategorien);
- regulative Belange berücksichtigen (Verfügbarkeit von Daten, Zuständigkeit und Umsetzbarkeit in Management-Optionen);
- Fairness sichern (Berücksichtigung der Belange aller Stakeholder und der Öffentlichkeit).

Solche Vorschläge lassen sich unter Bezugnahme von empirischen Untersuchungen zur Risikowahrnehmung weiter verfeinern. Dazu sollen im weiteren einige Untersuchungen vorgestellt werden.

In einer Studie zur Wahrnehmung der Dringlichkeit von Umweltproblemen haben Balderjahn und Wiedemann (1999) – ausgehend von den Kategorisierungen der Umweltprobleme – eine Korrespondenzanalyse durchgeführt. Diese Analyse ergab, dass Umweltprobleme mit hohem Abstraktionsgrad wie beispielsweise Luftverschmutzung, Wasserverschmutzung etc. als 'dringlicher' zu lösen, betrachtet werden als solche mit mittlerem bzw. geringem Abstraktionsgrad. Umweltprobleme mit niedrigem Abstraktionsgrad (Verschmutzung des Rheins, Aussterben von Störchen etc.) wurden 'unsicher hinsichtlich der Dringlichkeit' eingeordnet. Dies entspricht den Ergebnissen, die für die Risikobewertung (s. Karger & Wiedemann 1998) erzielt wurden: abstrakte Probleme werden eher als 'große Gefährdung für den Menschen' betrachtet als Probleme mit geringem Abstraktionsgrad.

Neben solchen kognitiven Gruppierungsproblemen sind auch affektive Aspekte bei der Kategorisierung zu beachten, die die Risikowahrnehmung beeinflussen können. Die Nennung von Schlüsselbegriffen wie Gen, Atom, Chemie kann z.B. eine Reihe von negativen Assoziationen auslösen. Solche Stigmatisierungen beeinflussen dann die Bewertung des Risikos (vgl. auch Gregory, Flynn & Slovic 1995, Murphy 2001, Wilson und Crandon 1998).

Um ein Beispiel zu nennen: Lindell und Earle (1980) haben gezeigt, dass das gleiche Risiko, einmal anonym und einmal unter dem Namen „Kernkraft“ vorgegeben, zu unterschiedlichen Bewertungen führt. Einen ähnlichen Befund erbrachte ein Survey in den UK (1999). Er zeigt, dass die Substitution von "organic farming" durch "Biotechnologie auf Kohlenstoffbasis" in Befragungen zu strikteren Regulationsforderungen führt. Wir konnten diesen Effekt replizieren (Wiedemann und Brüggemann 2001).

Eigene Untersuchungen weisen darauf hin, dass solche emotionalen Framing-Effekte nicht nur auf der Ebene von Worten zu finden sind, sondern durch Einbettungen in verschiedene Geschichten bewirkt werden können (Wiedemann und Schütz 2002). In einer Reihe von Experimenten konnte demonstriert werden, dass die Darstellung eines identischen Risikos in emotional verschieden getönten Geschichten zu unterschiedlichen Risikobewertungen führt (siehe Abbildung 10).

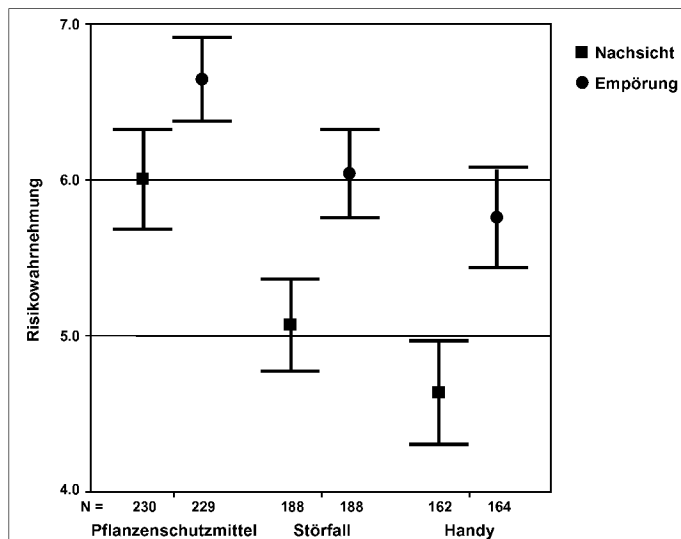


Abbildung 10: Wirkung von Risiko-Stories auf die Risikowahrnehmung

In solchen Geschichten werden z.B. die Personen (Risikoerzeuger, Opfer), ihr Handeln sowie ihre Motive und Reaktionen charakterisiert. Damit können Emotionalisierungen⁴⁹ (z.B. Nachsicht versus Empörung) transportiert werden, die Einfluss auf die Risikowahrnehmung nehmen. Diese Effekte sind stabil über verschiedene Risikoszenarien (Risiken aufgrund eines Störfalls, von Pflanzenschutzmitteln, durch Mobilfunk). Offenbar tritt eine Art Halo-Effekt⁵⁰ ein. Die negative Darstellung des Risikoerzeugers führt zu Emotionen, die sich wiederum auf die Risikobewertung auswirken.

Bei Risikovergleichen hätte das zur Folge, dass das Ranking nicht den Risikoattributen folgt, sondern durch Kontexteigenschaften (die z.B. nur mitgedacht werden) beeinflusst wird. Dass solche Ergänzungen durch Kontexteigenschaften bei Befragungen durchaus der Fall sind, haben wir in früheren Untersuchungen gezeigt (Schütz, Wiedemann & Gray 2000).

Die vorhandenen Befunde weisen deutlich darauf hin, dass die Kategorisierung von Risiken und deren Etikettierung einen Einfluss auf die Risikobewertung und damit

⁴⁹ Für die Untersuchung wurden *stories* konstruiert, die ein Risiko beschreiben. Dazu wurden einzelne Elemente der Geschichten in unterschiedlichen Facetten dargestellt. Zum Beispiel wurden die Größe des betroffenen Unternehmens variiert, ebenso der Umgang mit dem Risiko und den Anwohnern. Solche *stories* wurden den Untersuchungsteilnehmern zur Beurteilung vorgelegt.

⁵⁰ Halo-Effekt: Bezeichnet die Tendenz, dass bei Beurteilungen die Bewertung eines markanten Merkmals auf die Bewertung anderer Merkmale ausstrahlt.

auch auf Risikovergleiche ausübt. Die Etikettierung ist möglichst neutral vorzunehmen, stigmatisierende Begriffe sind zu vermeiden. Dennoch kann nie ganz ausgeschlossen werden, dass affektive Dimensionen den Risikovergleich beeinflussen.

Attribute: Welchen Einfluss haben sie auf das Vergleichen?

Ähnlich wie bei der Erstellung von Risikokategorien gibt es im Hinblick auf die Risikoattribute eine Vielzahl von Möglichkeiten. Die durchgeführten CRA unterscheiden sich deshalb beträchtlich. So finden sich neben ökologischen und gesundheitlichen Endpunkten Attribute wie „Peace of Mind“, Fairness, und Effekte auf künftige Generationen (vgl. den Scorecard-Ansatz des Environmental Defense Fonds).

Aus entscheidungstheoretischer Sicht müssen bei der Attributwahl mindestens drei Forderungen erfüllt sein: (1) Die Zahl der Attribute sollte übersichtlich sein. Bei Risikovergleichen ist immer nur eine begrenzte Anzahl von Attributen heranzuziehen. (2) Die Attribute müssen in ihrer Bewertung unabhängig voneinander sein (siehe Kap. 6.2.3). (3) Die gewählten Attribute sollten für die Beurteiler relevant sein.

Eine besondere Herausforderung für eine partizipative CRA stellt die Forderung nach Relevanz dar. Denn die Befunde zur Risikowahrnehmung legen nahe, dass für Laien und Experten unterschiedliche Attribute relevant sein können. Daraus wird abgeleitet, neben wissenschaftlichen Kriterien auch Laienkriterien zu verwenden. Ob dies so zutrifft, und ob daraus auch die obige Forderung nach Einbezug von Laien-Attributen in die CRA schlüssig abgeleitet werden kann, ist umstritten (Sjöberg 1996, Schütz und Wiedemann 2003).

Wir sind deshalb dieser Frage nachgegangen und haben Daten aus früheren Untersuchungen (Wiedemann und Kresser 1997, Wiedemann und Balderjahn 1997) neu analysiert. Dabei zeigte sich, dass die Attribute, die seitens der psychometrischen Risikoforschung als relevant herausgestellt werden, und die etwa Morgan et al. (2001) in ihrem Ansatz zur CRA nutzen (Katastrophenpotenzial, Kontrollierbarkeit, Unsicherheit des wissenschaftlichen Wissens und Anzahl der Betroffenen) nicht die sind, die Laien nutzen, wenn es ihnen freigestellt wird, selbst Attribute zu wählen.

In einem weiteren von uns durchgeführten Experiment wurde eine Liste von Umweltproblemen vorgegeben, die von den Probanden in Klassen nach Handlungsdringlichkeit (dringlich, weniger dringlich, unsicher) zu ordnen waren⁵¹. Ausgehend von diesen Ordnungen wurden die Attribute erfragt, nach denen die Probanden ihre Einordnung vorgenommen hatten. Hier dominierten Attribute wie Problemmuster, Schadensmerkmale und Merkmale des Problemmanagements.⁵²

⁵¹ Die Liste umfasst 30 Umweltprobleme und besteht aus abstrakten Umweltproblemen (z.B. Luftverschmutzung, Wasserverschmutzung) sowie konkreten Schädigungen (z.B. Schädigung der Eichen, Verschmutzung der Nordsee.).

⁵² Im Bereich der Problemmuster spielen hauptsächlich die Ausdehnung ("global - lokal") und das Ausmaß ("allgemein - spezifisch") des Umweltproblems eine Rolle, ebenso das Wissen über das Problem. Außerdem ist die Verursachung bedeutsam. Wenn die Probanden Folgenmerkmale nennen, beziehen sie sich am stärksten auf die "Folgen für den Menschen", danach auf die "Folgen für die Umwelt". Wird das Problemmanagement thematisiert, liegt die Betonung auf der "Kontrollierbarkeit" und den "Lösungsmöglichkeiten" für ein Umweltproblem.

Ein Experiment von Morgan et al. (2001) erhärtet unsere Auffassung, dass die sogenannten qualitativen Faktoren keine besondere Rolle beim Risiko-Ranking spielen. Werden Risikoattribute nach ihrer Wichtigkeit bewertet, so spielen klassische Attribute wie Mortalität und Morbidität eine größere Rolle als Kontrollierbarkeit des Risikos, Unsicherheit, Qualität des wissenschaftlichen Wissens und Latenzzeiten.

In einer weiteren Untersuchung, die an Balderjahn und Wiedemann (1999) anschließt, haben wir Attribute für die Risikobewertung vorgegeben. Befragt wurden Experten, Laien, Manager und Verwaltungsangehörige. Die Aufgabe der Befragten bestand nun darin, für jede der Risikoquellen anzugeben, wie wichtig bzw. unwichtig die Attribute für die Beurteilung des jeweiligen Risikos sind.

Für die Bewertungskriterien ergab sich über alle Risikoquellen und Gruppen zusammengefasst die folgende Rangordnung nach der Wichtigkeit (siehe Box 1).

1. Folgen für Gesundheit und Leben der Menschen
2. Irreversibilität möglicher Umweltschäden
3. Art und Stärke möglicher Umweltschäden
4. Sicherheit der Beherrschbarkeit dieses Risikos
5. Eintrittswahrscheinlichkeit eines Schadens
6. Kenntnisstand zum Risiko
7. Persönliche Betroffenheit im Schadensfall
8. Wirtschaftliche Notwendigkeit der Akzeptanz des Risikos
9. Ausmaß öffentlicher und sozialer Konflikte (Medienwirkung) im Schadensfall
10. Art und Umfang ökonomischer Risiken

Box 1: Rangordnung der Bewertungskriterien nach der Wichtigkeit

Die Unterschiede hinsichtlich der Einschätzung der Bedeutung der einzelnen Kriterien sind zwischen den Gruppen gering und in keinem Fall signifikant. Auch zwischen den Risikoquellen variiert die Bedeutung der Bewertungskriterien nur wenig.

Wir haben in Experimenten, basierend auf den Arbeiten von Balderjahn und Wiedemann (1999) ebenfalls ermittelt, welche Bedeutung unterschiedliche Attributausprägungen für verschiedene Gruppen (Experten, Laien, Manager und Verwaltungsmitarbeiter) haben. Die Probanden hatten ein hypothetisches Umweltrisiko zu beurteilen, mit Bezug auf eine Reihe von Beurteilungskriterien, die in verschiedenen Merkmalsausprägungen dargeboten wurden. Variiert wurden Gesundheitsschäden, Umweltschäden und Schadenswahrscheinlichkeiten sowie die Bekanntheit des Problems in der Öffentlichkeit und der Beschäftigungszuwachs. Aufgabe der Probanden war es, für die möglichen Merkmalskombinationen anzugeben, inwieweit die mit den Kombinationen verbundenen Risiken für sie akzeptabel sind.

Die Bedeutung der Attribute unterschied sich kaum zwischen den einzelnen Gruppen. Die *Schadenswahrscheinlichkeit* dient allen Gruppen als das wichtigste Entscheidungskriterium. Während bei Verwaltung, Experten und Laien dieses Kriterium ca. ein Drittel der Entscheidung erklärt, ist es bei den Managern nur gut ein Viertel. Zweitwichtigstes Attribut sind die Folgen für den Menschen (ca. 25%), gefolgt vom Beschäftigungszuwachs und den Folgen für die Umwelt. Allein für die Experten sind

die Folgen für die Umwelt etwas wichtiger als der erwartete Beschäftigungszuwachs. Relativ unwichtig ist das Kriterium der öffentlichen Auseinandersetzung. Nur die Manager beachten auch dieses Kriterium etwas stärker.

Jüngstens haben Lion et al. (2003) untersucht, welche Informationen Laien bezüglich unbekannter Risiken präferieren. Folgende Risiken wurden verwendet: Genetisch modifizierte Lebensmittel, Radon-Konzentration in Wohnungen, neue Anti-Blutgerinnungsmittel, Dioxin-Emissionen von Müllverbrennungsanlagen und EMF. Es zeigte sich, dass Laien vor allem an den Aspekten „Wie ist man exponiert?“, „Welche Konsequenzen hat das Risiko?“, „Worin besteht das Risiko?“ und „Wie groß ist die Wahrscheinlichkeit?“ interessiert sind.

Diese empirischen Ergebnisse demonstrieren, wie schwierig es ist, die Frage zu beantworten, welche Risikoattribute „natürlich“ präferiert werden. Schon geringfügige Änderungen im Frageformat sowie im Untersuchungsdesign legen verschiedene Risikoattribute nahe.

Eine Inklusion der qualitativen Risikofaktoren, die seitens der psychometrischen Risikoforschung empfohlen werden (Katastrophenpotenzial, Schrecklichkeit, Bekanntheit des Risikos, Wissen über das Risiko in den Wissenschaften, Freiwilligkeit etc.), wie es unter anderem von Fischhoff, Watson & Hope (1984) vorgeschlagen und u.a. von Morgan et al. (2000, 2001) versucht wurde, empfiehlt sich aber nicht. Dagegen sprechen die vorhandenen Untersuchungsbefunde.

Attribute müssen aus Zielstellungen abgeleitet werden (siehe Keeney 1992). Dabei sollten die grundsätzlichen Schutzgüter menschliche Gesundheit, intakte Öko-Systeme und Lebensqualität beachtet werden. Auf der Ebene der Attribute spielt die Eintrittswahrscheinlichkeit - entgegen der Annahmen der psychometrischen Risikoforschung - eine wichtige Rolle, auch bei Laien. Weiterhin ist ratsam, auch den Problemgrad (Schweregrad des Effekts), Problemumfang (Welchen Umfang hat das Risiko: Wie viele Menschen sind betroffen? Wie verbreitet sind die Umweltprobleme?) und gegebenenfalls die Problemqualität (die Reversibilität versus Irreversibilität der möglichen Schäden) zu betrachten.

Bewertungen auf den Attributen: Welchen Einfluss haben Risikomaße aus Vergleichen?⁵³

Für die Beschreibung von Risiken auf den Attributen lassen sich unterschiedliche Risikomaße nutzen. Beispielsweise kann man ein technologisches Risiko mit der Anzahl der vorhandenen Anlagen angeben, von denen dieses Risiko ausgeht. Weiterhin kann die Wahrscheinlichkeit pro Betriebszeitraum betrachtet werden, mit der Schadstoffe durch einen Störfall freigesetzt werden. Risikobeschreibungen können sich auch auf andere Angaben zu Schadstoffen beziehen. So lassen sich Mengen-, Konzentrations- oder Intensitätsangaben machen oder es kann die Ausbreitung (welche Fläche ist von Schadstoffen betroffen?) charakterisiert werden. Schließlich finden sich auch Beschreibungen der Persistenz – wie lange bleibt der Schadstoff ein Schadstoff – in Zeitmaßen (z.B. Halbwertszeit).

⁵³ Dieser Abschnitt entstand in Zusammenarbeit mit Lic. phil. Andrea Thalmann, der wir für ihre Unterstützung danken.

Bezüglich der Exposition lässt sich z.B. charakterisieren, wie viele Menschen betroffen sind. Hier kann weiterhin die Akkumulation über die Zeit dargestellt werden: In welchem Maß reichert sich der Schadstoff im Körper oder in der Umwelt – bezogen auf eine bestimmte Zeitspanne – an? Auch die adversen Effekte können unterschiedlich beschrieben werden. Es können Angaben zu akuten oder chronischen Effekten, zur Erkrankungshäufigkeit und zur Art des möglichen Gesundheitsschadens, zu Effekten auf künftige Generationen oder zu Umweltschäden gemacht werden. Natürlich hängt die Auswahl dieser Maße auch davon ab, ob sich das jeweilige Risiko danach überhaupt sinnvoll beschreiben lässt, und ob dazu Daten vorliegen.

Experimentelle Untersuchungen demonstrieren, dass die Auswahl der Maße die Risikocharakterisierung und -wahrnehmung beeinflusst (vgl. Femers 1993, Gray 1996). Ein Beispiel von Crouch und Wilson (1982) veranschaulicht das: Will man wissen, ob die amerikanische Kohleproduktion in den Jahren von 1950 bis 1970 sicherer geworden ist, so kann man als Basis die Zahl der Unfalltoden pro 1 Mio. Tonnen geförderter Kohle nutzen. Es zeigt sich, dass die Kohleproduktion über den betrachteten Zeitraum sicherer geworden ist. Man kann sich aber auch auf die Zahl der Toten pro tausend Arbeiter in der Kohleproduktion beziehen. In diesem Fall ergibt sich, dass die Kohleproduktion über die zwanzig Jahre unsicherer geworden ist (siehe Abbildung 11).

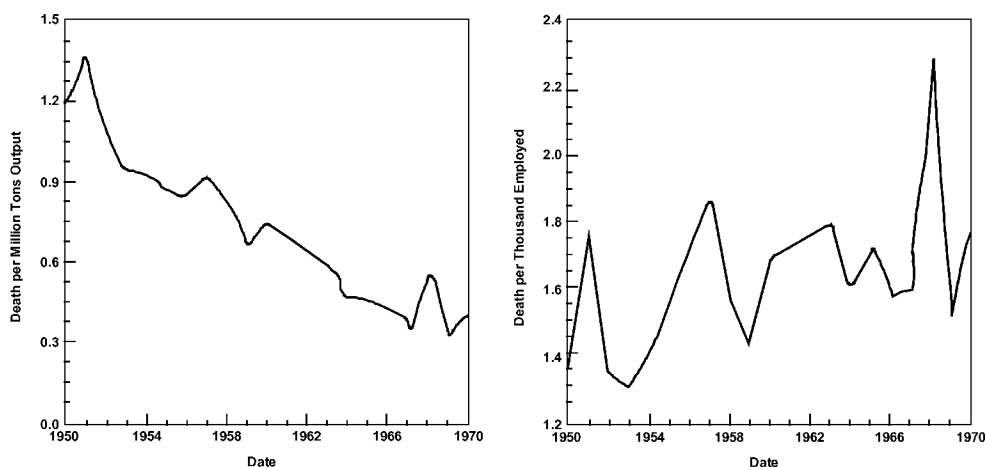


Abbildung 11: Einfluss gewählter Indikatoren auf Risikodarstellungen (aus: Crouch & Wilson 1982, 99)

Ein anderes Beispiel ist der häufig angeführte Vergleich vom Risiko des Fliegens mit dem des Autofahrens. Legt man als Referenzgröße für die Berechnung des Risikos die Zahl der Todesfälle (durch Flugzeugabstürze bzw. Autounfälle) pro zurückgelegtem Kilometer (bzw. Meilen) zugrunde, so ergibt sich ein deutlich höheres Risiko für das Autofahren: 13 Tote pro 10^9 gefahrener Meilen für das Autofahren versus 0,6 Tote pro 10^9 geflogener Meilen für das Fliegen. Betrachtet man jedoch nur Reisen, die kürzer als 600 Meilen sind (dies macht – zumindest in den USA – einen großen

Teil der Flüge und einen noch größeren Teil der Autofahrten aus), so ergibt sich ein geringeres Risiko für das Autofahren.⁵⁴

Tabelle 13 fasst die Effekte von verschiedenen Risikomaßen auf die Risikowahrnehmung zusammen. Beispielsweise fanden Stone, Yates und Parker (1994), dass Risikoinformationen, welche relativ angegeben werden (z.B. Verdopplung des vorliegenden Risikos), einen stärkeren Einfluss auf das Urteil ausüben als die Beschreibung in Form der Inzidenzrate. Dies trifft vor allem bei sehr kleinen Wahrscheinlichkeiten zu. So werden extrem kleine Inzidenzraten als „fast Null“ wahrgenommen, während dies bei der Angabe in Form eines Relativen Risikos nicht der Fall ist. In den Experimenten von Stone et al. (1994) sind die Personen bereit, mehr für ihre Sicherheit zu bezahlen, wenn die Risikoreduktion relativ präsentiert wurde im Vergleich zu einer Darstellung mit Hilfe der Inzidenzrate. In die gleiche Richtung gehen die Ergebnisse von Magat, Viscusi und Huber (1987).

Tabelle 13: Effekte von Risikobeschreibungen auf die Risikowahrnehmung

Risikobeschreibung	Effekt auf die Risikowahrnehmung
Wahrscheinlichkeit • Relatives Risiko • Basisrate (bei kleinen Risiken) • positive verbale Wahrscheinlichkeitsangabe • negative verbale Wahrscheinlichkeitsangabe	→ verstärkt eher → schwächt eher ab → verstärkt eher → schwächt eher ab
Umgang mit Unsicherheit • Konfidenzintervallangaben /Thematisierung von Unsicherheit	→ verstärkt eher
Konsequenzen • Ernsthaftigkeit / Schweregrad	→ verstärkt eher

Auch Halpern, Blackman und Salzman (1989) untersuchen den Einfluss verschiedener Informationspräsentationen auf Risikourteile. Die Autoren stellen fest, dass die Wahrscheinlichkeitsdarstellung in Form des relativen Risikos zu einer Erhöhung der Risikobewertung führte, verglichen mit einer Darstellung der gleichen Wahrscheinlichkeit mit Hilfe der Häufigkeit von Todesfällen. Halpern und Mitarbeiter erklären diesen Unterschied damit, dass die Befragten, welchen die Häufigkeit bzw. die absolute Zahl der Todesfälle bekannt war, durch das Wissen um die Basisrate das Risiko anders beurteilen.

In einer Reihe von Studien (v.a. Weber und Hilton, 1986; Wallsten, Fillenbaum und Cox, 1986; Fischer und Jungermann, 1996) wurde die Basisrate als eine zentrale Variable identifiziert, welche die Interpretation von verbalen Wahrscheinlichkeitsbeschreibungen beeinflusst. So zeigten Weber und Hilton (1990), dass die Interpretation der verbalen Wahrscheinlichkeitsangaben durch das Wissen um die Basisrate der

⁵⁴ Das Beispiel stammt von Evans, Frick & Schwing (1990). Die Daten beziehen sich auf den Zeitraum 1978 bis 1987. Die Autoren kritisieren die Verwendung von Mittelwerten aus Unfallstatistiken, die vor allem für die Autofahrer die große Variabilität in den Unfallhäufigkeiten für Personen und Altersgruppen verdecken. Tatsächlich können solche Berechnungen nicht ohne eine ganze Reihe weiterer Annahmen gemacht werden, die von den Autoren ausführlich diskutiert und hinsichtlich ihrer Auswirkungen auf den Risikovergleich dargestellt werden.

Ereignisse beeinflusst wird. Dies stimmt auch mit den Ergebnissen von Wallsten und Mitarbeiter (1986) überein, die ebenfalls die wahrgenommene Basisrate als Einflussgröße auf die Interpretation von Wahrscheinlichkeitsaussagen identifizieren konnten.

Verbale Wahrscheinlichkeitsangaben werden oft als zu vage eingeschätzt (vgl. Budescu und Wallsten, 1997; 1995; Fillenbaum, Wallsten, Cohen, und Cox, 1991; Wallsten, Budescu, Rapoport, Zwick, Forsyth, 1986; Zimmer, 1983).

Kontextuelle Effekte auf die numerische Interpretation von verbalen Beschreibungen der Wahrscheinlichkeit wie z.B. „einige“, „wenige“, „viele“ oder bei der Häufigkeit wie z.B. „selten“ sind in einer Reihe von Studien festgestellt worden (Beyth-Marom, 1982; Brun und Teigen, 1988; Budescu und Wallsten, 1985; Fillenbaum, Wallsten, Cohen und Cox, 1991; Gonzales und Frenck-Mestre, 1993; Hamm, 1991; Teigen, 1988; Weber und Hilton, 1990). So wird beispielsweise dem Satz „häufig in das Kino gehen“ ein anderer Zahlenwert zugeordnet als „häufig die USA besuchen“ (Newstead und Collis 1987). Ebenfalls wird die Bedeutung von „selten“, „gelegentlich“, „häufig“ in verschiedenen Kontexten unterschiedlich interpretiert.

Verbale Ausdrücke sind weniger neutral als numerische Angaben und können deswegen auf subtile Weise Entscheidungen beeinflussen (vgl. Budescu und Wallsten, 1995; Moxey und Sanford, 1993; Champaud und Bassano, 1987). So gibt es Ausdrücke, die das Vorhandensein eines Ereignisses hervorheben wie z.B. „wahrscheinlich“ oder „möglicherweise“ und solche die das Nicht-Vorhandensein betonen wie z.B.: „zweifelhaft“ (vgl. Teigen und Brun 1999; 2000).

Im Zusammenhang mit dem Gebrauch verbaler und numerischer Wahrscheinlichkeitsangaben beschreiben Erev und Cohen (1990) sowie Wallsten, Budescu, Zwick und Kemp (1993) ein Paradoxon, das „communication mode preference (CMP) paradox“. Das besagt, dass die meisten Personen bevorzugen, Informationen über Wahrscheinlichkeiten von Ereignissen in Form numerischer Angaben zu erhalten – aber selber lieber verbale Wahrscheinlichkeitsangaben zur Kommunikation verwenden.

Die Studie von Erev und Cohen (1990) weist darauf hin, dass Personen verbale Termini bevorzugen, um unpräzise Wahrscheinlichkeiten bzgl. des Eintretens von Ereignissen zu bewerten. Die Präferenz basiert möglicherweise auf der angenommenen Kongruenz zwischen Präzision eines Ausdrucks und der darunter liegenden Unsicherheit (vgl. Teigen und Brun 1999). Dabei darf allerdings die große interindividuelle Variabilität in der Bedeutung von Wahrscheinlichkeitswörtern nicht vergessen werden, die bei Laien (Thalmann, in Vorbereitung) und bei Experten (vgl. Jablonowski 1994) zu finden ist.

In eigenen Untersuchungen (Thalmann, in Vorbereitung) sind wir der Bedeutung verbaler Wahrscheinlichkeitsangaben bei Laien nachgegangen. Es handelte sich dabei um die Evidenzmaße der SSK, die zur Beschreibung der vorhandenen Evidenz bei gesundheitlichen Effekten von EMF benutzt werden.

Es stellt sich heraus, dass die Probanden Hinweis, Verdacht und Nachweis eines Risikos⁵⁵ anders einschätzen als es die SSK vorsah. Entgegen der Originalreihenfolge Nachweis – Verdacht – Hinweis der SSK, bildeten rund die Hälfte (55%) der Be-

⁵⁵ Beurteilt wurden die Evidenzmaße der SSK (2001) für gesundheitliche Effekte von EMF.

fragten die Reihenfolge Nachweis – Hinweis – Verdacht. Die Probanden interpretierten die Evidenzkategorie „Hinweis“ als stärker bzw. beweiskräftiger als „Verdacht“. Außerdem zeigen sich enorme Differenzen zwischen den Probanden.

Solche unterschiedlichen Bedeutungen von verbalen Risikomaßen – in diesem Falle verbale Etiketten zur Darstellung der Beweiskraft – sind nicht nur bei Laien festzustellen. In einer Studie von Javlonowski (1994) wurden Experten (Risikomanager) nach der Bedeutung verschiedener Häufigkeitsangaben gefragt.⁵⁶ Auch hier fand sich eine sehr große Variation in der Einschätzung der einzelnen Wahrscheinlichkeits-Etiketten.

Die Thematisierung von Unsicherheiten führt zu unterschiedlichen Bewertungen. Zum Teil führt sie zu einer Steigerung des Vertrauens in die jeweilige Informationsquelle, zum Teil wird sie aber auch als Zeichen von Inkompetenz und Unehrlichkeit gewertet (vgl. Johnson und Slovic, 1995).

In einem Telexperiment der gleichen Studie zeigt sich, dass Intervallangaben von Risikoschätzungen zu einer höheren Risikowahrnehmung führen, als wenn Punktschätzungen gemacht werden. Es erwies sich auch, dass Unsicherheiten in Form von Intervallangaben z.T. zu höherer Besorgnis führten.

In einer weiteren Studie zur Wahrnehmung und Bewertung von Unsicherheit in Risikoabschätzungen (Johnson und Slovic 1998) fanden die Autoren, dass bei Intervallangaben die obere Grenze als der glaubwürdigste Schätzwert angesehen wurde. Außerdem wurde die Ursache von Unsicherheit nicht in der Natur der Sache gesehen, sondern zumeist sozialen Faktoren zugeschrieben (Eigeninteressen der Experten, Inkompetenz). Zudem glaubte nahezu die Hälfte der Befragten, dass die Behörden nur dann ein Risiko kommunizieren, wenn dies beträchtlich ist. Ungefähr die Hälfte der Befragten präferierte eine „Entweder oder Kommunikation“, d.h. eine klare Botschaft über das Vorliegen oder Nichtvorliegen eines Risikos.

Die Befunde von Johnson (2003) replizieren diese Ergebnisse im wesentlichen. Dabei zeigte sich, dass generelle Einstellungen und Bewertungen gegenüber dem Informationsgeber einen Einfluss auf die Bewertung der Unsicherheit haben. Solche, die misstrauisch sind, werten die Angabe von Unsicherheiten eher als Zeichen von Unehrlichkeit und Inkompetenz.

Als eine Einflussgröße auf die Risikowahrnehmung wurde in der Studie von Weber und Hilton (1990) die Ernsthaftigkeit der Schadensereignisse untersucht. Schwerwiegendere Ereignisse führen zu einer Erhöhung in der Wahrscheinlichkeits-Einstufung. Die Ernsthaftigkeit von Ereignissen beeinflusst die Wahrscheinlichkeitsbewertung möglicherweise im Sinne eines „worry effect“, d.h. ernstere bzw. schwerwiegendere Ereignisse ziehen die Aufmerksamkeit in potenziell höhere Wahrscheinlichkeitsgrade. Diese Hypothese wird durch die Studie von Fischer und Jungermann (1996) gestützt. Hier ordneten die Befragten den gleichen verbalen Ausdrücken in fiktiven Beipackzetteln von Medikamenten unterschiedliche numerische Wahrscheinlichkeitswerte zu, je nachdem ob es sich um schwere oder um leichte Nebenwirkungen handelte.

⁵⁶ rare – very unlikely – unlikely – somewhat unlikely – likely – frequent – extremely likely

Als Konsequenz drängt sich auf, verbale Repräsentationen von Quantitäten durch quantitative Angaben zu ergänzen. Wenn zusammenfassende Bewertungen wie Nachweis, Verdacht usw. benutzt werden, sollte ihre Bedeutung in einem Vortest geprüft werden. Unsicherheiten sollten in einfacher Weise kommuniziert werden, beispielsweise mit Hilfe einer verbalen Charakterisierung.

Vergleiche: Welchen Einfluss hat die Art des Vergleichens auf den Vergleich?

Das Vergleichen von Objekten (auch Risiken), um eine Rangordnung zu erstellen, kann auf unterschiedliche Weise vorgenommen werden, z.B. durch direktes Ranking, Paarvergleiche, Tripelvergleiche etc. (Winterfeld und Edwards 1986). Diese Verfahren unterscheiden sich hinsichtlich des kognitiven Aufwands und damit nach den erforderlichen Voraussetzungen. Sie sind aber auch abhängig von Akzeptanz- und anderen Randbedingungen.

In der Literatur zu Risikovergleichen finden sich eine Reihe von Untersuchungen, die Aufschluss darüber geben, auf welche Probleme CRA bei der Vergleichsbildung treffen können. Darüber hinaus gibt es zu einzelnen Verfahren von Vergleichen eine Vielzahl von - empirischen und eher theoretischen - Bewertungen, die hier nicht aufgeführt werden können. Es sei hier auf die Spezialliteratur, z.B. Keeney und Raiffa (1976) und Winterfeld und Edwards (1986) verwiesen.

Ein erstes Vergleichsproblem resultiert aus der Abneigung gegenüber Abwägungen von Alternativen. Einschlägig hierzu sind Baron und Sprance (1997) sowie Ritov und Baron (1999), die den Einfluss von Protected Values (PV) auf die Präferenzbildung und das Abwägen von Entscheidungsalternativen untersucht haben.⁵⁷ Darunter verstehen sie absolute Werthaltungen, die so stark sind, dass sie Abwägungen verbieten. Solche Protected Values sind moralische Bewertungen von Aktivitäten, die Konsequenzen ausblenden (Man darf X nicht tun, gleichgültig, wie gering die Konsequenzen sind). Damit wird auch die Größenordnung der Konsequenzen ausgeblendet (Ein kleines Schadenspotenzial ist ebenso schlimm wie ein großes Schadenspotenzial). Außerdem spielen solche Protected Values eine größere Rolle bei der Bewertung von aktivem Tun (Veränderung des Status quo) als bei Unterlassungen des Tuns (Beibehalten des Status quo). Zwar ist der Entdeckungskontext der PV nicht identisch mit den Umständen des Risiko-Ranking im Rahmen einer CRA, dennoch können solche Protected Values zu Konflikten führen, zum Beispiel, wenn Risiken für bedrohte Spezies und für Menschenleben gegeneinander abzuwägen sind. Damit werden Vergleiche, insbesondere Kosten-Nutzen-Vergleiche schwierig, wenn nicht unmöglich (Fiske & Tetlock 1997).

Außerdem zeigt sich, dass die Präsentation der Vergleichsoptionen einen Einfluss hat. Werden Optionen separat nacheinander bewertet, so werden eher Abwägungen gemacht, als wenn die Optionen gleichzeitig bewertet werden (Bazerman et al. 1999; Hsee 1996). Es scheint, dass die Probanden im Falle simultaner Vergleiche danach bewerten, wie sie meinen, dass sie bewerten sollten; im Fall sequenzieller Vorge-

⁵⁷ In der Regel werden in den Experimenten zu PV die Probanden in eine Art moralisches Dilemma gebracht. Sie müssen entscheiden ob eine Therapie eingesetzt werden kann, die 5 Personen das Leben kostet, um etwa 500 Personen das Leben zu retten.

hensweise bewerten sie jedoch nach ihren faktischen Präferenzen (vgl. Irwin und Baron 2001).

Es macht ebenfalls einen Unterschied, wie die Optionen verglichen werden. Ein direktes Rating orientiert mehr auf Wertbezüge als die Methode der Zahlungsbereitschaften (Irwin und Baron 2001). Beide Methoden führen zu unterschiedlichen Präferenzordnungen. Wird die Methode der Zahlungsbereitschaft zu Vergleichen herangezogen, so spielt der Vergleichsmodus eine Rolle. Wird die Verkaufsmethode herangezogen (Für wieviel Geld würden Sie x veräußern?) so werden eher wertbezogene Entscheidungen getroffen. Im Gegensatz dazu dominieren bei der Kaufmethode (Wieviel ist ihnen X wert?) monetäre Aspekte (Chapman und Johnson 1995; Irwin 1994).

Schließlich ist auch zu beachten, dass bei Risikobewertungen (und damit auch bei Risikovergleichen) Risiken nicht unabhängig von ihrem Nutzen beurteilt werden. Experimente stützen eine Beziehung zwischen Nutzen- und Risikowahrnehmung (z.B. Fischhoff et al. 1978; Harding & Eiser 1984; Alhakami 1991; Alhakami & Slovic 1994). Dazu haben Finucane et al. (2000) unter dem Stichwort „Affektheuristik“ interessante Untersuchungen gemacht. Sie konnten zeigen, dass Risiken geringer eingeschätzt werden, wenn hoher Nutzen gesehen wird und umgekehrt. Außerdem waren die Risiko- und Nutzenurteile, wenn sie unter hohem Zeitdruck abgegeben werden mussten, stärker negativ korreliert, als bei Urteilen ohne Zeitbegrenzung. Nach Auffassung der Autoren sind dafür die Affekte in bezug auf die Risikoquelle verantwortlich.

Die bislang einzige empirische Untersuchung zur Durchführung einer CRA stammt von der Arbeitsgruppe um Granger Morgan von der Carnegie Mellon Universität (Morgan et al. 2000, 2001; Florig et al. 2001). Die Probanden waren Risikomanager aus Behörden und Industrie. Sie hatten Rankings durchzuführen. Die Risikoliste umfasste 9 - 11 Risiken. Vorab (D1) hatten sie eine holistische Risikobewertung vorzunehmen, danach waren sie aufgefordert, jeweils die Attribute, auf denen die Risiken beschrieben waren, zu gewichten. Daraus wurde der multiattribute Rang des Risikos berechnet (D2). Nach einer Instruktion und Einführung wurden die Probanden in Gruppen zusammengefasst und hatten ein holistisches Ranking (D3) und ein multiattributes Ranking (D4) vorzunehmen. Danach wurden in D5 und D6 wieder individuelle Rankings vorgenommen (D5: holistisch, D6: multiattribut).

Abbildung 12 zeigt die Resultate der Korrelationsanalyse. Zwischen dem holistischen und dem multiattributen Ranking zeigt sich eine durchaus beachtliche Korrelation (zwischen 0.595 und 0.686). Das spricht nach Auffassung der Autoren für eine ausreichende konvergente Validität. Auch die Übereinstimmung zwischen den Probanden in den 6 Sessions war gut (Korrelationen zwischen 0.595 und 0.913).

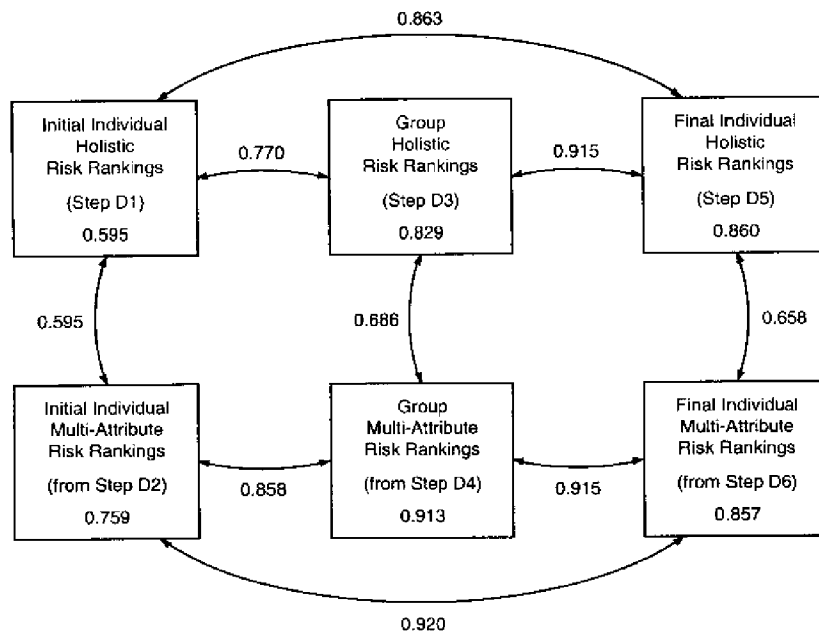


Abbildung 12: Korrelationen zwischen Risk-Rankings (Quelle: Morgan et al. 2001)

Die Probanden waren mit dem Gruppen-Diskussionsprozess vor D3 und D4 auch zufrieden ($x = 6,12$ auf einer 7-stufigen Skala) sowie mit den Rankings in D3 und D4 ($x = 5,8$ und $5,41$, ebenfalls auf einer 7-stufigen Skala).

Die hier aufgeführten Untersuchungen deuten darauf hin, dass es nicht unerheblich ist, welche Personen die Vergleiche durchführen sollen, und welche Vergleichsmethode eingesetzt wird. So können Abweichungen gegenüber Rankings aus stärkeren Werthaltungen resultieren. Allerdings weisen unsere eigenen Untersuchungen darauf hin, dass unterschiedliche Gruppen in ähnlicher Weise – bei relativ einfachen Vergleichsanforderungen – Attribute bewerten und zu weitgehend konsistenten Präferenzen kommen. Es spielt allerdings eine Rolle, ob die Vergleiche nach Risiken (jedes Risiko wird für sich auf allen Attributen bewertet) oder nach Attributen (pro Attribut werden alle Risiken bewertet) durchgeführt werden. Weiterhin unterscheiden sich die Ergebnisse in Abhängigkeit von den gewählten Verfahren. Zahlungsbereitschaften und direkte Ratings führen zu unterschiedlichen Präferenzordnungen. Schließlich können vorhandene emotionale Einstellungen die Risiko- und Nutzenbewertung beeinflussen, was sich auf die Vergleiche auswirkt.

Welches Vergleichsverfahren auch immer gewählt wird, die Vorgehensweise sollte methodisch konsistent, fair und gut dokumentiert sein. Außerdem ist darauf zu achten, dass das gewählte Verfahren verständlich und nachvollziehbar ist. Dabei kommt es auf verständliche und eindeutige Risikocharakterisierungen an. Wo immer möglich, sollten mehrere und unterschiedliche Ranking-Verfahren genutzt werden, um die Stabilität der Rankings überprüfen zu können.

4.4 Verhandlungskonflikte bei der CRA

Die Bereitschaft, sich auf Verfahren wie die CRA einzulassen, ist nicht immer gegeben. Mitunter werden Widerstände oder Bedenken formuliert. Minard (o.J.) beschreibt eine Reihe von Beispielen.⁵⁸

Eine CRA, an der ganz unterschiedliche Personengruppen beteiligt sind, erfordert Einigungen. Das beginnt mit der Festlegung der Ziele, setzt sich mit der Auswahl der Teilnehmer fort, bezieht sich auf die Festlegung der zu bewertenden Risiken, deren Kategorisierung, die Bestimmung der Attribute, die Beschreibung der Risiken auf den Attributen und betrifft schließlich das Ranking. Weiterer Einigungsbedarf ergibt sich bei der Entwicklung von Aktionsplänen und deren politischer Durchsetzung. Hier können Dissense und Konflikte auftreten.

Dazu gehören:

- das Beharren der Entscheider, unabhängig von allen anderen Gruppierungen eigenständig alle Probleme zu bewerten,
- die Angst der Entscheider, durch ein solches Verfahren Macht abgeben zu müssen,
- die Bedenken, durch das Verfahren möglicherweise eine andere, nicht akzeptierte Position zu legitimieren und
- fehlendes Wissen über Design und Werkzeuge der konsensorientierten CRA.

Im weiteren werden solche Fallstricke bei der CRA dargestellt. Sie betreffen alle Prozesse des Abwägens und Bewertens in einer CRA; darunter fallen die Beurteilung der eigenen Position und Interessen, die Beurteilung der anderen Parteien, die Beurteilung von Vorschlägen und von Ergebnissen. Bei der Darstellung folgen wir der Literatur zum Verhandeln (Bazerman 1986, Bazerman und Lewicki 1983, Neale und Bazerman 1985, Pruitt und Rubin 1986, Raiffa 1982, Thomson 1990, Thompson und Hastie 1990).

Fallstricke bei der Beurteilung der eigenen Position und Interessen

Die wichtigsten Fallen bei der Beurteilung der eigenen Position und Interessen sind (a) Übervertrauen, (b) interessengetriebene Wahrnehmungen und (c) Kontrollillusion (siehe Tabelle 14).

Das Überschätzen des eigenen Einflusses (Kontrollillusion) ist dabei der gravierendste Fallstrick. Man ist sich zu sicher, dass die Ergebnisse den eigenen Erwartungen entsprechen werden. Das Übervertrauen macht sich dann besonders bemerkbar, wenn bereits erzielte Zwischenergebnisse den eigenen Wünschen entsprechen. Die-

⁵⁸ Er weist darauf hin, dass ein CRA-Projekt im Schnitt 2-3 Jahre dauert. Ungefähr 1 Jahr lang braucht es, um eine CRA vorzubereiten und einen entsprechenden politischen Beschluss zu erwirken. Danach sind noch einmal 6-10 Monate nötig, um die Prozedur durchzuführen. Ein weiteres Jahr ist für die Rückmeldung in die Politik und die Beschlussfassung für einen Aktionsplan erforderlich. Während dieser Zeit sind vielfältige Konsultationen und Abstimmungen durchzuführen und Entscheidungen zu treffen. Siehe www.epa.gov/comp_risk/tool6/part1_2.htm

ser Effekt ist besonders kritisch, da Informationen, die die eigenen Hypothesen und Bewertungen untermauern, selektiv wahrgenommen und in ihrer Bedeutung überschätzt werden. Selbst wenn Hinweise auf die Angemessenheit des eigenen Standpunktes fehlen, ist man geneigt, nur bestätigende Information zu suchen.

Tabelle 14: Befangenheiten in der Beurteilung der eigenen Partei

Eigene Partei		
Befangenheit	Beschreibung	Bedeutung bei der CRA
Unrealistischer Optimismus	Die eigenen Chancen bei der Durchsetzung der Interessen werden zu hoch angesetzt.	Geringe Bereitschaft zu Konzessionen und Enttäuschungen.
Kontrollillusion	Tendenz, selbst alle Festlegungen und Vereinbarungen den eigenen Fähigkeiten zuzuschreiben.	Überschätzen der eigenen Verhandlungsmacht.
Wunschdenken	Die Wahrscheinlichkeit, dass ein erwünschtes Ereignis tatsächlich eintritt, wird überschätzt.	Geringe Bereitschaft zu Kompromissen.
Interessengeleitete Perspektive	Informationen, die den eigenen Standpunkt bestätigen, werden selektiv wahrgenommen.	Überschätzen der "Angemessenheit" des eigenen Standpunktes.
Suche nach Bestätigung	Tendenz, Informationen, die die eigenen Annahmen und Hypothesen bestätigen, aktiv zu suchen, jedoch nicht falsifizierende Informationen.	Ausblenden von Information, die ein Überdenken des eigenen Standpunktes notwendig macht.

Übervertrauen kann dazu führen, dass (a) eine mögliche Einigung nicht erreicht wird, weil man an seinem eigenen Standpunkt festhält und (b) man annimmt, die anderen Teilnehmer werden den eigenen Standpunkt teilen und eine den eigenen Interessen entsprechende Bewertung mittragen. Gerade bei starken Wertkonflikten oder Interessenbindungen kann diese Einstellung zu Positions- und Grabenkämpfen führen, die integrative Lösungen vereiteln. Überzeugt vom eigenen Erfolg wird erwartet, Konzessionen müssten vom anderen ausgehen. Sitzen sich z.B. Interessenvertreter von Industrie und Umweltverbänden gegenüber, ist die Gefahr groß, dass über Positionen anstatt über Präferenzen verhandelt wird.

Fallstricke bei der Beurteilung der anderen Parteien

Eines der wichtigsten Merkmale erfolgreicher kollektiver Bewertungs- und Abwägungsprozesse ist es, die Perspektive der anderen Beteiligten einnehmen zu können, d.h. sich in die Belange, Zwänge, Interessen und Ziele des Gegenübers hineinversetzen zu können. Gerade für die CRA, an der Vertreter unterschiedlicher gesellschaftlicher Gruppen teilnehmen, die auch in anderen Situationen zusammenarbeiten und so auf dauerhaft tragfähige Beziehungen angewiesen sind, gelten nicht nur die Kriterien der Effizienz und Rationalität guten Handelns, sondern vor allem der Grundsatz: der Umgang miteinander sollte die gegenseitigen Beziehungen verbessern oder zumindest nicht beeinträchtigen (Fisher & Ury 1981).

Die wichtigsten Fehleinschätzungen und Befangenheiten, die einer flexiblen und offenen Beziehung der Beteiligten entgegenstehen, sind die Stereotypisierung, die fehlgeleitete Zuschreibung der Ursachen des Verhaltens und das Überbewerten negativer Eindrücke (siehe Tabelle 15).

Tabelle 15: Fallstricke in der Beurteilung der anderen Partei

Andere Partei		
Befangenheit	Beschreibung	Bedeutung für die CRA
Stereotypes Denken	Aus der Beurteilung des Persönlichkeits-Typs werden Hypothesen über Eigenschaften, Ziele oder Präferenzen der Person abgeleitet.	Fehlbewertung von Interessen, Zielen und Absichten
Kontrast-Effekt	Je größer die wahrgenommenen Unterschiede zur eigenen Einstellung sind, desto mehr verstärkt sich die Tendenz, Statements oder Personen als konträr zur eigenen Einstellung wahrzunehmen, konträrer als sie es eigentlich sind.	Freund-Feind-Stereotypisierung
Konsistenz-Effekt	Erwartungen darüber, was man wahrnehmen wird, beeinflussen das, was man tatsächlich wahrnimmt. Eigenschaften des anderen, die nicht ins Bild passen, werden unterschätzt bzw. ignoriert.	Inflexibler Umgang
Falscher Konsenseffekt	Ist das Entscheidungs- und Urteilsverhalten des anderen dem eigenen ähnlich, wird es als angemessen wahrgenommen. Weicht das Verhalten des anderen vom eigenen ab, wird es als abweichend von der Norm eingeschätzt.	Schwarz-Weiß Malerei. Fehlbewertung und Ausblenden der Ursachen für das Verhalten des anderen.
Negativitätsverzerrung	Negative Informationen determinieren stärker das Gesamturteil eines anderen als positive Informationen.	Überbewerten negativer Eindrücke von den Interessen, Zielen und Absichten des Anderen
Effekt des ersten Eindrucks	Der erste Eindruck oder das erste Statement werden am besten in Erinnerung behalten.	Überbewerten des ersten Eindrucks

Bei kooperativen Handlungen bildet sich jeder der Beteiligten eine Vorstellung von der Persönlichkeit, den Werten, Zielen und Interessen der anderen Parteien. Dabei wird auf Vorerfahrungen zurückgegriffen, die helfen, konkrete Situationen einzuordnen, zu strukturieren und dadurch Komplexität zu reduzieren. Einer dieser Mechanismen zur Einschätzung anderer Personen ist das Denken in Stereotypen. Solches Wissen über Typen von Menschen erleichtert es, selbst ungenaue Informationen über andere zu verstehen, und bereits aus dem ersten Eindruck Hypothesen über Einzelheiten der Persönlichkeit, über Charaktereigenschaften oder Präferenzen abzuleiten. Es kann aber auch zu einer Fehleinschätzung des anderen führen. Zudem wird häufig der erste Eindruck überschätzt. Auf der anderen Seite, wenn sich die Parteien bereits kennen, ist es besonders schwierig, solche Annahmen über die intendierten Strategien und Interessen der anderen Beteiligten zu flexibilisieren, da man annimmt, dass andere sich immer in gleicher Weise verhalten. Der Konsistenz-Effekt trägt zusätzlich dazu bei, die Stereotypisierung aufrechtzuerhalten. Denn Eigenschaften, die nicht in das erwartete Bild passen, werden ausgeblendet und ignoriert. Die Gefahr einer stereotypen Beurteilung liegt dann im gegenseitigen Nicht-Verstehen.

Je mehr Nähe zur eigenen Einstellung wahrgenommen wird, desto größer ist die Tendenz, Unterschiede zu verharmlosen. Hingegen erfolgt beim Kontrasteffekt eine Polarisierung. Je größer die wahrgenommenen Unterschiede, desto stärker werden sie als konträre Positionen aufgebauscht. Diese Tendenz zur Freund-Feind-Stereotypisierung kann bewirken, dass ein bestehender Dissens überschätzt wird und Konsensmöglichkeiten nicht aufgedeckt werden. Solche Fehleinschätzungen

können zu Missverständnissen führen, die Positionskämpfe fördern, zur „Lagerbildung“ führen und integrative Lösungen verhindern.

Ebenso gravierend können sich falsche Hypothesen über die Gründe des Verhaltens auswirken. Sich in einen anderen hineinzuversetzen bedeutet, auch situative Einflüsse wahrzunehmen, die ein bestimmtes Verhalten provozieren. Eine „harte“ Kommunikationsstrategie zu führen, muss nicht notwendigerweise auf einen unnachgiebigen Charakter zurückzuführen sein, mit dem man auch in anderen Situationen rechnen muss. Geht man von einer stabilen Persönlichkeitseigenschaft aus, unterschätzt man die Möglichkeiten, den Standpunkt der anderen Partei beeinflussen zu können oder übersieht Handlungszwänge, in denen sich andere Interessenvertreter befinden. Nur Verständnis für solche situativen Gegebenheiten kann zu einem fairen Dialog führen, der dann auch effiziente Lösungen verspricht.

Fallstricke bei der Beurteilung von Vorschlägen

Die wichtigsten Fallstricke bei der Beurteilung von Vorschlägen sind (a) der Einfluss einer Nullsummenannahme, (b) der Umgang mit und die Bewertung von Verlusten und Gewinnen und (c) die Gefahr des positionalen Handelns, um das eigene Gesicht zu wahren (siehe Tabelle 16).

Untersuchungen zu Verhandlungsverhalten haben gezeigt, dass die Beteiligten mit Vorannahmen in Verhandlungen gehen. Hier ist die "Nullsummenannahme" von besonderer Bedeutung. Die Parteien glauben, dass sie verschiedene, miteinander konkurrierende Ziele anstreben. Sie definieren so die CRA als Wettbewerb, in dem es Gewinner und Verlierer gibt.

Ist ein Vorschlag auf dem Tisch, steht man vor zwei Alternativen: entweder diese sichere Alternative anzunehmen oder sich für die unsichere Alternative zu entscheiden, d.h. weiterzuverhandeln – mit dem Risiko, dass die Einigung scheitern kann oder nur ein schlechteres Ergebnis als das bisherige erzielt wird. In dieser Situation spielt die Bewertung des bisher Erreichten eine erhebliche Rolle. Denn es ist von Bedeutung, ob man das bisher Erreichte als Verlust oder als Gewinn betrachtet. Wird der Vorschlag als Gewinn für die eigene Position gewertet, ist man eher geneigt, ihn zu akzeptieren. Beurteilt man den Vorschlag im Sinne eines Verlustes, ist man risikobereiter und setzt das bislang Erreichte für ein unsicheres, aber vielleicht besseres Ergebnis aufs Spiel. Dieser Bewertungseffekt tritt unabhängig vom Inhalt des Vorschlages auf, auf den man sich zu einigen sucht. Die gleiche Ausgangssituation kann zu einer unterschiedlichen Entscheidung führen, je nach dem, ob man das Angebot als positiv oder negativ einschätzt.

Tabelle 16: Fallstricke in der Beurteilung des Verhandlungsverlaufs

Vorschläge		
Befangenheit	Beschreibung	Bedeutung für die CRA
Nullsummenannahme	Annahme, dass die Gewinne der einen Partei Verluste der anderen Partei sind.	Dominanz einer Gewinner-Verlierer-Strategie.
Ankereffekt	Anker werden als Ausgangspunkt für die Beurteilung und Bewertung herangezogen.	Durch gesetzte bzw. vermutete Anker werden Bewertungen bestimmt.
Orientierung an den gesunkenen Kosten	Vorangegangene Kosten anstatt zukünftige Gewinne werden als Grundlage für Entscheidungen herangezogen.	Kompromisse, die nach Stand der Verhandlungen rational sind, werden ausgeschlagen.
Kognitive Dissonanz	Vorangegangene Positionen zu revidieren, ist nicht mit dem eigenen Selbstbild vereinbar. Unvereinbarkeiten mit dem eigenen Selbstbild werden ausgeblendet oder uminterpretiert.	Der Spielraum der Kompromissbereitschaft hängt davon ab, wie sehr man das Gesicht zu wahren sucht.

Dieser Effekt schlägt insbesondere dann durch, wenn die beteiligten Parteien positional diskutieren. Mit dem Verkünden einer bestimmten Position wird eine Ergebnismarge festgelegt. Um sein Gesicht nicht zu verlieren, ist man gezwungen, jeden Vorschlag, der davon abweicht, als Verlust zu werten. Eine Einigung fällt dann besonders schwer.

Ähnlich gravierend kann es sich auswirken, wenn die Beteiligten bei ihrer Entscheidung, auf einen Vorschlag einzugehen oder nicht, ihr Augenmerk auf den vorangegangenen Verhandlungsverlauf richten. Rechnen die Parteien in diese Abwägung vorangegangene Aufwendungen mit ein, z.B. bereits investierte Zeit, Mühe und Kosten, kann dies zu einem Festhalten an den eigenen Zielvorstellungen führen. Eine mögliche Einigung wird verworfen, obwohl es bei einer Abwägung, die zukünftige Kosten und Nutzen betrachtet, rational ist, Konzessionen zuzugestehen. Und: Je mehr man sich an eine vorher eingenommene Position gebunden und darauf verpflichtet hat, sei es durch öffentliche Bekundung oder eine extreme Anfangsposition, die wenig Raum für Zugeständnisse öffnet, desto eher können Konflikte eskalieren. Es verstärkt sich die Tendenz, die eigene Position zu verteidigen. Daher ist es besonders wichtig, Tendenzen zur Selbstbindung an bestimmte Positionen immer wieder zu thematisieren, Alternativen aufzuzeigen und Flexibilität zu fördern.

Fallstricke bei der Beurteilung von Ergebnissen

Die Bewertung des Zustandekommens des Ergebnisses und die Identifikation der Beteiligten damit ist wichtig, da (a) die erzielte oder verfehlte Einigung nach außen (in der Öffentlichkeit) und nach innen (in die eigene Partei hinein) vertreten werden muss, und (b) Lerneffekte nur dann erreicht werden, wenn das Erreichte adäquat beurteilt wird. Die Fallstricke des Überschätzens der eigenen Rolle, besonders in Abhängigkeit vom Erfolg der Verhandlungen und der Rekonstruktionsfehler, können dabei hinderlich sein (siehe Tabelle 17).

Scheitert eine CRA, liegt es an den beteiligten Parteien, die Gründe dafür offen zu legen. Werden Konsense im Risiko-Ranking erzielt, ist die Bewertung dieser Ergebnisse maßgeblich dafür, ob man sich als Gewinner oder Verlierer sehen kann. Da die Tendenz besteht, Erfolge den eigenen Bemühungen und Fähigkeiten zuzuschreiben, besteht die Möglichkeit, dass bei Misserfolgen versucht wird, die Verantwortung ab-

zuwälzen, und dass jede der Parteien versuchen wird, Erfolge für sich zu beanspruchen. Die Neigung, sich Erfolge auf die eigene Fahne zu schreiben, wird dadurch verstärkt, dass man den eigenen Beitrag beim Zustandekommen eines Konsenses häufig überschätzt. Negative Ergebnisse für die eigene Partei z.B. der mangelnden Kompromissbereitschaft der anderen Parteien zuzuschreiben, gelingt nicht zuletzt deshalb, da im Rückblick Geschehnisse, die die eigene Darstellung stützen, besser erinnert werden; davon abweichende Situationen werden ausgeblendet oder als bedeutungslos betrachtet.

Tabelle 17: Fallstricke in der Beurteilung des Verhandlungsergebnisses

Ergebnis		
Befangenheit	Beschreibung	Bedeutung für die CRA
Erinnerungsverzerrung	Die Unfähigkeit, Details von Ereignissen zu erinnern, führt zu einer "logischen" Rekonstruktion, die ungenau sein kann.	Inkonsistenzen, Brüche und Ambiguitäten des Verhandlungsverlaufes werden im Nachhinein in ein stimmiges Bild gebracht.
Effekt des Selbstbild-Bewahrens	Bewahrung des Selbstbildes durch die Tendenz, Erfolge den eigenen Fähigkeiten zuzuschreiben, Misserfolge externen Faktoren.	Die Verantwortung für das Scheitern von Verhandlungen wird abgewälzt.
Egozentrik-Effekt	Tendenz, den eigenen Anteil an einem gemeinsam erarbeiteten Ergebnis zu überschätzen.	Übertriebene Darstellung einer Gewinner-Position.

Diese Tendenz, Widersprüchlichkeiten und Brüche im Nachhinein auszublenden, ist wenig hilfreich für eine Klärung der Abläufe einer CRA mit dem Zweck, daraus zu lernen. Gerade dafür ist es notwendig, sich die eigenen Anteile klar vor Augen zu halten, auch wenn sie sich im Laufe der Verhandlungen und im Hinblick auf das Verhandlungsergebnis als Fehleinschätzungen erwiesen haben. Außerdem gilt: Im Nachhinein ist man häufig über das eingetretene Ergebnis nicht überrascht. Plausible Erklärungen werden ohne Mühe gefunden. Dies verhindert ebenfalls, nach den eigenen Anteilen zu fragen, diese aufzuarbeiten und darauf aufbauend den weiteren Umgang mit den anderen Parteien zu gestalten.

Die hier beschriebenen Fehler und Fallen müssen nicht, sie können aber im Prozess der CRA auftreten. Damit ist insbesondere dann zu rechnen, wenn das Risiko-Ranking mit der Prioritätensetzung gleichgesetzt wird und es im Interesse einzelner Beteiligter liegt, dass bestimmte Risiken eine hohe (oder gegebenenfalls auch niedrige) Position erreichen.

5. Konzeptioneller Rahmen für eine integrierte und vergleichende Risikobewertung

5.1 Methodische Probleme bei einer CRA

Die in Kapitel 3 beschriebenen CRA-Beispiele sowie die Diskussion der empirischen Forschung zum Urteils- und Entscheidungsverhalten in Kapitel 4 zeigen, dass es für die Durchführung von CRA-Vorhaben vor allem zwei große Problemfelder gibt: (1) der Umgang mit den *Unsicherheiten*, die bei den zugrundeliegenden Risikoabschätzungen bestehen können, und (2) die Auswahl der *Bewertungskriterien*, die bei der CRA berücksichtigt werden sollen. Diese beiden Problemfelder sollen im folgenden diskutiert werden.

5.1.1 Problem: Unsicherheit und Ungewissheit

Das größte Problem für eine vergleichende Risikobewertung sind die Unsicherheiten, die bei der Abschätzung von Risiken bestehen können. So finden beispielsweise Schuhmacher-Wolz, Konietzka und Schneider (2002) in einer Analyse von Risikoabschätzungen zu kanzerogenen Luftschadstoffen, dass selbst bei qualitativ guten Studien sich noch Unterschiede in den Abschätzungen für das Unit Risk von bis zu zwei Größenordnungen ergeben. Auch Expositionsabschätzungen können, abhängig von den Mess- und Schätzverfahren, erheblich variieren (vgl. Swartjes 2002; WHO 2000).⁵⁹ In einer zusammenfassenden Analyse dieser Problematik kommen Neus et al. (1998, 79) zu der Einschätzung,

daß bei ungünstigen Konstellationen Risikoabschätzungen im Einzelfall stark von den „wahren“ Verhältnissen abweichen können, und daß Risikoabschätzungen nicht im Sinne von exakten Risikoprognosen verstanden werden können. Soweit die exemplarisch ausgewählten Zusammenstellungen eine Verallgemeinerung erlauben, dürfte die Toleranzbreite der Risikoschätzer im allgemeinen in der Größenordnung von zwei Zehnerpotenzen liegen.

Ein weiteres Problem beim Vergleich von Risiken, die auf einer Exposition mit Schadstoffen beruhen, liegt in der Charakteristik dieser Exposition: ist sie tatsächlich vorhanden beziehungsweise muss man mit ihrem Auftreten rechnen, oder ist sie ein logisch nicht auszuschließendes Ereignis, das mit großem Aufwand verhindert werden soll? Der erste Fall kennzeichnet die überwiegende Mehrzahl der umweltbedingten Gesundheitsrisiken; der zweite Fall ist typisch für die störfallbedingten Risiken einer Technik mit hohem und gut bekannten Gefahrenpotenzial, etwa der Kernenergie. Hier bleibt es auch – oder gerade – bei geringen Unsicherheiten in der Risikoabschätzung prinzipiell ungewiss, ob als sehr selten identifizierte Störfälle sich jemals ereignen.⁶⁰

⁵⁹ Vgl. dazu auch die Beitragsserie: Probabilistische Expositionsabschätzung in Umweltmedizin und Verbraucherschutz (Fehr & Mekel 2000).

⁶⁰ Zur Unterscheidung zwischen Unsicherheit und Ungewissheit siehe oben Kapitel 2.3.

5.1.1.1 Typen von Unsicherheit

Bei der Diskussion der Unsicherheit des Wissens über Risiken werden unterschiedliche Konzepte verwendet. Vielfach wird zwischen aleatorischer und epistemischer Unsicherheit unterschieden (z.B. Pate-Cornell 1996). Aleatorische Unsicherheit beschreibt einen Zufallsprozess (z. B. Würfeln, Zerfall radioaktiver Substanzen), epistemische Unsicherheit beschreibt begrenztes Wissen.

In vielen Fällen ist es aber eher eine Frage der Betrachtungsweise als eine Eigenschaft der betrachteten Größe, ob eine Unsicherheit aleatorisch oder epistemisch ist. So ist die unterschiedliche Sensitivität von Menschen gegenüber einer Noxe bei Betrachtung der ganzen Population eine aleatorische Unsicherheit. Betrachtet man jedoch eine einzelne Person, so ist diese Unsicherheit eher epistemisch, und die Sensitivität dieser Person kann unter Umständen sogar durch entsprechende Untersuchungen genauer ermittelt werden. Doch selbst wenn die Sensitivität jeder einzelnen Person exakt bekannt wäre, wäre es dennoch angemessen, für die Beschreibung der Sensitivität des Kollektivs eine Wahrscheinlichkeitsverteilung dieses Parameters zu verwenden, d. h. dies als einen Zufallsprozess (aleatorisch) zu beschreiben.

Eine andere, häufig getroffene Unterscheidung ist die zwischen *Variabilität* und (eigentlicher) *Unsicherheit* (z.B. Bogen 1990; NRC 1994). Erstere bezieht sich auf die Heterogenität innerhalb einer Population (dies entspricht der oben genannten aleatorischen Unsicherheit). Sie ist also ein objektives und im Prinzip exakt meßbares Merkmal der Natur. Dagegen bezieht sich Unsicherheit auf das unvollständige Wissen über einen bestimmten Sachverhalt. Unsicherheit ist kein Merkmal der Natur, sondern ein – sozusagen kognitives – Merkmal des jeweiligen Wissenschaftlers bzw. der jeweiligen *scientific community* (dies entspricht der epistemischen Unsicherheit).

Die Unterscheidung zwischen Variabilität und Unsicherheit ist nicht nur von theoretischem, sondern auch von praktischem Interesse. Denn während sich die Variabilität innerhalb einer Population auch durch zusätzliche Forschung nicht reduzieren, sondern bestenfalls genauer beschreiben lässt, kann die Unsicherheit durch weitere Forschung häufig verringert werden. In diesem Zusammenhang stellt sich dann die Frage, ob der Aufwand für eine weitere Reduzierung von Unsicherheit (für die jeweilige Risikobeurteilung) gerechtfertigt ist – eine Frage, die selbst wieder methodisch behandelt werden kann (*value of information analysis*; vgl. Clemen 1991; Morgan & Henrion 1990).

Für eine genauere Charakterisierung von Unsicherheit sind verschiedene Taxonomien vorgeschlagen worden (z.B. Bogen 1990; Morgan & Henrion 1990; NRC 1994). Eine häufig genutzte Unterscheidung ist die zwischen *Parameter-Unsicherheit* (z.B. Meßfehler, Validität von Indikatoren) und *Modell-Unsicherheit* (z.B. Angemessenheit von linearen Modellen ohne Schwellenwert bei der Bestimmung von Dosis-Wirkungsbeziehungen, Vollständigkeit von Modellen).

5.1.1.2 Unsicherheiten im Prozess der Risikoabschätzung

Die Vielschichtigkeit dieser Problematik wird deutlich, wenn man die Ursachen für Unsicherheiten und Variabilität für die verschiedenen Stufen des Prozesses der Risikoabschätzung betrachtet (siehe Tabelle 18).

Tabelle 18: Einige Ursachen für Unsicherheit und Variabilität in den Stufen der Risikoabschätzung (nach Bogen 1990, 30ff.)

	Unsicherheit	Variabilität
Hazard Identifikation	- Unterschiedliche Studienergebnisse - Unterschiedliche Studienqualität - Unterschiedliche Studientypen - Extrapolation der verfügbaren Evidenz auf die menschliche Zielpopulation	- Heterogene metabolische Kompetenz - Heterogene immunologische Kompetenz - Heterogene absolute Empfindlichkeit der Zielsysteme
Dosis-Wirkungs-Beziehung	- Extrapolation der im Experiment genutzten Dosis auf die für den Menschen relevante Dosis - Definition der „positiven Fälle“ in einer Untersuchung - Parameterschätzung - Modellauswahl für die Extrapolation in den Niedrigdosis-Bereich	- Heterogene metabolische Kapazität - Heterogene immunologische Kapazität - Heterogene quantitative Empfindlichkeit der Zielsysteme
Expositions-Abschätzung	- Charakterisierung des Kontaminations-szenarios (Produktion, Verteilung in den Umweltmedien, raum-zeitliche Eingrenzung etc.) - Charakterisierung des Expositionsszenarios (Expositionspfade, Expositionsdynamik) - Identifizierung der Zielpopulation	- Heterogenität in den maßgeblichen Expositionspfaden bezüglich Alter, Geschlecht, Beruf, Wohnort, Verhalten - Heterogenität in der Expositionsdynamik bezüglich Alter, Beruf, Wohnort, genetischer Disposition, Gesundheitsstatus, Zufallsprozessen
Risikocharakterisierung	Alle Unsicherheiten, die sich ergeben aus: - Hazard Identifikation - Dosis-Wirkungs-Beziehung - Expositionsabschätzung	Alle Variabilität, die sich ergibt aus: - Hazard Identifikation - Dosis-Wirkungs-Beziehung - Expositionsabschätzung

In der Vergangenheit wurde meist versucht, diese Unsicherheiten – vor allem für die beiden Schritte: Abschätzung der Dosis-Wirkungs-Beziehung und Expositionsabschätzung – durch „worst case“ Annahmen „aufzufangen“. Das heißt, die ungünstigsten Bedingungen gingen in Form von Punktschätzungen in die Risikoabschätzung ein. Ein solches Vorgehen bietet zwar auf der einen Seite weitgehende Sicherheit, dass Gesundheitsrisiken nicht unterschätzt werden, kann aber auf der anderen Seite zu einer deutlichen Überschätzung der Risiken und damit zu einer Fehlallokation von Ressourcen zum Gesundheitsschutz führen.

Neuere Ansätze, die auf probabilistischen Techniken beruhen, nutzen statt der „worst case“ Punktschätzungen für die zu berücksichtigenden Parameter Wahrscheinlichkeitsverteilungen, die mit Hilfe statistischer Verfahren („Monte Carlo“ Analyse u.ä.; vgl. Vose 1996) verrechnet werden (EPA 1996; Cullen & Frey 1999). Diese Vorgehensweise nutzt die verfügbare Information besser und kommt – so die Hoffnung – zu realistischeren Risikoabschätzungen. Allerdings müssen bei diesem Ansatz eine Reihe von Annahmen – etwa über Form der zugrunde liegenden Wahrscheinlichkeitsverteilungen – gemacht werden, die selbst wieder zu begründen sind und die Ergebnisse stark beeinflussen können. Eine umfassende Diskussion der Unterschie-

de beider Vorgehensweisen mit anschaulichen Beispielen für das Problem der Expositionsabschätzung bieten die Arbeiten von Fehr & Mekel (1999), Heinemann (2000) Mekel & Fehr (2000), Mosbach-Schulz (1999), Schümann (2000) und Wintermeyer (1999).

5.1.1.3 Unsicherheiten bei der Risikoabschätzung technischer Anlagen

Geht es bei der Risikoabschätzung um technische Anlagen (z.B. Chemieanlagen, Kernkraftwerke, Staudämme), so kommt ein weiterer Aspekt hinzu: die Unsicherheiten bei der Abschätzung von Störfallrisiken. Denn neben den Risiken, die mit dem bestimmungsgemäßen Betrieb technischer Anlagen verbundenen sind, müssen auch die Risiken von Störfällen betrachtet werden. Bezüglich der Risiken des bestimmungsgemäßen Betriebs technischer Anlagen ist die wesentliche Frage die nach den dabei gegebenen Umweltbelastungen und den möglicherweise resultierenden Gesundheits- bzw. Umweltschäden. Im Fall von Störfällen ist zu fragen, gegen welche Ereignisse aufgrund ihrer Plausibilität und ihrer Folgen Vorkehrungen getroffen werden müssen und ob ein dennoch verbleibendes Risiko aufgrund immer noch möglicher Störfälle akzeptabel ist. Dabei besteht die charakteristische Ungewissheit darin, ob ein als möglich – und auch als plausibel – erkannter Störfall tatsächlich in dem Zeitraum eintritt, der für eine Risikoabschätzung praktische Relevanz hat. Nicht mehr praktisch relevant in diesem Sinn ist beispielsweise ein Zeitraum, der wesentlich länger ist als die Betriebsdauer der betrachteten Anlagen beziehungsweise von Anlagen der betrachteten Art. Selbstverständlich überlässt man das Eintreten von Störfällen nicht dem Zufall. Zuverlässige Sicherheitseinrichtungen in technischen Anlagen bewirken, dass unakzeptable Störfälle – so die im Prinzip berechnete Denkweise von Genehmigungsverfahren – „praktisch ausgeschlossen werden können“. So lange aber die Fälle nicht logisch ausgeschlossen werden können, verbleibt ein Risiko aufgrund der genannten Ungewissheit.

Zur Ermittlung der Störfallhäufigkeiten werden (technische) Risikoanalysen durchgeführt. Sie fragen allgemein nach Ursachen für gefährliche Abweichungen von sicheren Prozessbedingungen und nach deren Folgen. Hierbei sind drei methodische Ansätze zu unterscheiden: Einmal der intuitiv empirische, auf dem beispielsweise ein Brainstorming, die Auswertung von Betriebserfahrungen in Unfallberichten oder auch die Abarbeitung von Checklisten beruhen. Der zweite methodische Ansatz ist das induktive Vorgehen. Hier wird beim auslösenden Ereignis begonnen und gefragt, welche Störfallabläufe sich daraus entwickeln können. Beispiele sind die Ausfalleffekt-Analyse und die Ereignisablauf-Analyse. Das dritte Verfahren, das deduktive Vorgehen, beginnt beim unerwünschten Anlagenzustand und fragt, wie es dazu kommen konnte. Dies entspricht dem Prinzip einer Fehlerbaumanalyse.

Ausgangspunkt eines Fehlerbaums ist ein unerwünschter Zustand eines Subsystems („Ausfall“), etwa eines Sicherheitssystems. Deduktiv werden die Ursachen ermittelt, die zu dem Zustand führen können. Dabei wird das System in seine Teilfunktionen und Komponenten zerlegt; der Funktionszusammenhang wird mit Hilfe der Booleschen Operatoren UND und ODER abgebildet. Dies wird bis zu einer Ebene fortgeführt, auf der komponentenbezogene Ausfallhäufigkeiten bekannt sind, aus denen dann die Häufigkeit des Systemausfalls berechnet wird. Ereignisbäume fragen dagegen induktiv nach den Folgen, die aus einem auslösenden Ereignis in Abhängigkeit von Reaktionen von Subsystemen (Aufgabenerfüllung oder Ausfall) entstehen. Durch

Kombination von Ereignis- und Fehlerbäumen werden Störfallszenarien entwickelt, an deren Ende Zustände des Gesamtsystems (der „Anlage“) stehen, deren bedingte *Wahrscheinlichkeiten* nach einem auslösenden Ereignis berechnet werden können. Für jeden dieser Anlagenzustände werden die *Folgen* ermittelt, indem die Vorgänge „Bildung und Transport von Noxen in der Anlage - Emission von Noxen - Ausbreitung - Exposition - Effekt“ modelliert werden unter der Annahme, der jeweilige Anlagenzustand sei gegeben. Auf diese Weise lässt sich das Risiko komplexer Anlagen ermitteln.

Ein prinzipielles Problem dieses Vorgehens besteht darin, dass nur Risiken gefunden werden können, nach denen gesucht wird (Betrachtungsrahmen), und dass nur nach Risiken gesucht werden kann, bei denen die schädliche Wirkung bekannt ist.

Zur Einschätzung und Bewertung von Unsicherheiten bei der Risikoabschätzung technischer Anlagen ist es daher erforderlich, dass die Dokumentation der Risikoanalysen aufzeigt,

- wo die Grenzen des Betrachtungsrahmens sind,
- an welchen Stellen subjektive Beurteilungen einfließen,
- an welchen Stellen Unsicherheiten bestehen und
- wie groß diese sind,
- wo es an Wissen mangelt und
- wie groß diese Wissenslücken sind.

5.1.2 Problem: Bewertungskriterien

Die Auswahl der Bewertungskriterien, nach denen eine vergleichende Risikobewertung vorgenommen werden soll, hängt ab von der Art der Zielstellung der CRA, speziell von den Schutzgütern, die bei der CRA berücksichtigt werden sollen. Ausgangspunkt für die Ableitung von Bewertungskriterien sind die Schutzgüter, in bezug auf welche die vergleichende Risikobewertung vorgenommen werden soll. Diese müssen in der Regel ausdifferenziert werden, um im Rahmen einer CRA berücksichtigt werden zu können. Dabei muss diese Ausdifferenzierung zu Kriterien führen, für die Meßverfahren (Metriken) verfügbar sind.

Betrachtet man zum Beispiel das allgemeine Schutzgut „menschliche Gesundheit“, so zeigt sich, dass der Begriff Gesundheit – und damit auch Gesundheitsrisiko – nicht so eindeutig ist, dass sich hieraus Bewertungskriterien von selbst ergeben würden. Schon die Definition von Gesundheit, die die Weltgesundheitsorganisation 1948 in der Präambel zu ihrer Gründung vorgelegt hat, macht deutlich, dass Gesundheit – wie ja auch Risiko – als multidimensionales Konzept verstanden werden kann:

"Gesundheit ist ein Zustand vollkommenen körperlichen, geistigen und sozialen Wohlbefindens und nicht allein das Fehlen von Krankheit und Gebrechen."⁶¹

⁶¹ Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

Für die vergleichende Risikobewertung ergibt sich damit das Problem, wie dieses vielschichtige Konzept „Gesundheit“ operationalisiert werden kann. Die in der WHO-Definition angesprochenen Aspekte des Wohlbefindens sind in bisherigen CRAs üblicherweise unter dem Aspekt „Lebensqualität“ gefasst und neben Gesundheit im engeren Sinne (d.h. dem Fehlen von Krankheit und Gebrechen) als gesonderter Bewertungsaspekt behandelt worden (vgl. EPA 1993). Ein Beispiel für das Spektrum von Bewertungskriterien für Lebensqualitätsrisiken ist die folgende Liste (EPA 1993, 2.4-6):

- das ästhetische Empfinden, z.B. durch Lärm- oder Geruchsbelästigungen;
- den ökonomischen Wohlstand, z.B. durch Arbeitsplatzverlust aufgrund von Umweltproblemen oder erhöhte Gesundheitsausgaben;
- Fairneß, beeinträchtigt z.B. durch ungleiche Verteilung von Nutzen und Risiken;
- zukünftige Generationen, z.B. indem die Beseitigung von Umweltschäden zukünftigen Generationen überlassen wird;
- „peace of mind“, z.B. durch Ängste vor Umweltgefährdungen;
- Freizeit- und Erholungsmöglichkeiten, z.B. durch den Verlust an Erholungsflächen;
- die soziale Gemeinschaft, z.B. durch Vereinsamung aufgrund veränderter Lebensbedingungen.

Bei der Ausdifferenzierung der Bewertungskriterien für Gesundheitsrisiken (im engeren Sinn) wird in der „Unfinished Business“ CRA der EPA (1987) beispielsweise zwischen kanzerogenen und nicht-kanzerogenen Effekten unterschieden (siehe oben Kapitel 3.2.1), wobei die nicht-kanzerogenen Effekte noch in drei Schadenskategorien unterteilt werden:

- *katastrophale Gesundheitseffekte*, z.B. Tod, Verringerung der Lebenserwartung, schwere körperliche oder geistige Behinderungen, Erbkrankheiten;
- *schwerwiegende Gesundheitseffekte*, z.B. Fehlfunktionen im Organ- und Nervensystem sowie Verhaltens- und Entwicklungsstörungen;
- *adverse Gesundheitseffekte*, z.B. Gewichtsverlust, Veränderungen bei Enzymen oder reversible organische Störungen.

Für die Kategorie der „adversen Gesundheitseffekte“ stellt sich die Frage, ab wann eine bestimmte Wirkung als „advers“ einzuschätzen ist. Im Sinne der WHO-Definition von „Adversität“ ist das der Fall, wenn es zu einer Beeinträchtigung der Funktionsfähigkeit eines Organismus kommt (siehe oben Kapitel 2.2). Ob dies für eine spezifische Wirkung angenommen werden kann, ist letztlich nur von Experten zu beurteilen. Dabei können diese zu durchaus unterschiedlichen Einschätzungen kommen. Sind beispielsweise Schlafstörungen, die von den Betroffenen auf den Straßenlärm oder auf nahestehende Mobilfunksendemasten zurückgeführt werden, als adverse Gesundheitseffekte oder „nur“ als Befindlichkeitsstörungen einzuordnen? Für eine CRA müssen hier klare Kriterien gesetzt werden.

Schließt man neben der menschlichen Gesundheit noch die Umwelt als Schutzgut mit ein, so ergibt sich ein noch komplexeres Bild, wie das folgende Beispiel für die Wirkung von Pestiziden zeigt (Levitan 1997). Hier werden als Kriterien genannt:

- Wirkungen auf die Gesundheit von Farmern, Konsumenten und der allgemeinen Öffentlichkeit, einschließlich besonders empfindlicher Subpopulationen;
- Letale und sub-letale Einwirkungen auf Flora und Fauna;
- Direkte und indirekte Einflüsse auf das natürliche und landwirtschaftliche Ökosystem, einschließlich der Effekte auf Habitate und Nahrungsmittelquellen;
- Verschmutzung von Boden, Luft und Wasser, die Konsequenzen für Menschen bzw. Flora und Fauna haben können; und
- Gesellschaftliche Kosten durch die Remediation von Umweltschäden.

Es ist offensichtlich, dass auch diese Liste von Bewertungskriterien weiter ausdifferenziert und vor allem operationalisiert werden muss, um im Rahmen einer vergleichenden Risikobewertung nutzbar zu sein. Interessant an dieser Liste ist auch, dass sie einzelne Wirkungsbereiche noch einmal unterteilt.

Ein weiteres, naheliegendes Bewertungskriterium ist der Nutzen, der mit der Anwendung von risikobehafteten Optionen (Technologien, Stoffen, Verhaltensweisen etc.) verbunden ist. Denn Risiken werden ja nicht um ihrer selbst Willen eingegangen, sondern weil man sich von den entsprechenden Technologien, Stoffen, Verhaltensweisen etc. einen (individuellen und/oder gesellschaftlichen) Nutzen verspricht. Und tatsächlich hat sich in verschiedenen Studien zur Risikowahrnehmung gezeigt, dass Risiko- und Nutzenurteile nicht unabhängig voneinander sind. Vielmehr besteht ein negativer Zusammenhang: hohe Nutzenurteile gehen meist mit geringen Risikoeinschätzungen einher und *vice versa* (vgl. Alhakami & Slovic 1994; Finucane et al. 2000; Fischhoff, Slovic, Lichtenstein, Read & Combs 1978; Gregory & Mendelsohn 1993; Harding & Eiser 1984; Vlek & Stallen 1981). Und natürlich spielt der Nutzenaspekt für die Frage der Akzeptanz bzw. Akzeptabilität von Risiken eine zentrale Rolle.

Im Rahmen einer CRA sollte die Bewertung von Risiken allerdings unabhängig von der des Nutzens sein, wie dies z.B. auch bei Risiko-Nutzen-Analysen vorausgesetzt wird (vgl. Merkhofer 1987). Denn es geht ja um die vergleichende Bewertung der Risiken. Und gerade, wenn risikobehaftete Optionen hinsichtlich ihrer Risiken *und* Nutzen bewertet werden sollen, muss die Bewertung beider Aspekte zunächst getrennt erfolgen, damit der Nutzen nicht doppelt in die Bewertung einfließt – einmal implizit über die Risikobewertung und ein zweites Mal explizit über die Nutzenbewertung.

5.1.3 Konsequenzen für die Durchführung einer CRA

Die Unsicherheiten, mit denen Risikoabschätzungen, die für eine CRA herangezogen werden, behaftet sind, können im Rahmen der CRA natürlich nicht aufgelöst werden. Umso wichtiger ist es, dass diese Unsicherheiten in der Risikocharakterisierung umfassend beschrieben werden. Gray (2000) nennt drei typische Probleme bzw. Unterlassungen bei der Risikocharakterisierung:

Falsche Präzision: Die Verwendung von Punktschätzungen zur Beschreibung eines Risikos kann – insbesondere bei Personen, die mit den Problemen der Risikoabschätzung und den daraus resultierenden Unsicherheiten nicht vertraut sind (Regulatoren, Journalisten, Öffentlichkeit) – zu einem irreführenden Eindruck von Genauigkeit und Eindeutigkeit in der Risikoabschätzung führen. Gerade für die Durchführung einer CRA unter Beteiligung von Stakeholdern, die keine Experten für Risikoabschätzung sind, ergibt sich damit die Notwendigkeit, in der Risikocharakterisierung solche Abschätzungsunsicherheiten nicht nur quantitativ zu beschreiben, sondern auch graphisch zu veranschaulichen und mit erläuternden Texten zu ergänzen.

Falsche Konsistenz: Im Rahmen von Risikoabschätzungen werden mitunter Standardannahmen (*default assumptions*) gemacht, die in unterschiedlichen Fällen unterschiedlich plausibel sind. Bei der Risikoabschätzung kanzerogener Stoffe wird beispielsweise häufig eine lineare Dosis-Wirkungs-Beziehung ohne Schwellenwert unterstellt, um zu einer Abschätzung der kanzerogenen Potenz im Niedrigdosisbereich zu kommen. Für bestimmte Agenzien (z.B. direkt mutagene Chemikalien) legen wissenschaftliche Daten einen solchen Zusammenhang in der Tat nahe, für andere Agenzien (z.B. nicht direkt mutagene Chemikalien) ist dies allerdings umstritten. Das heißt, Abschätzungen auf der Grundlage solcher Standardannahmen mögen in manchen Fällen zu präziseren Risikoabschätzungen führen, in anderen Fällen jedoch eine Überschätzung des Risikos zur Folge haben. Für die Risikocharakterisierung ist es erforderlich zu erläutern, inwieweit gleichartige Verfahren bzw. Annahmen bei der Risikoabschätzung auch zu gleichermaßen plausiblen Risikoabschätzungen führen.

Versteckte Entscheidungen: Im Rahmen von Risikoabschätzungen müssen häufig Annahmen gemacht und Entscheidungen getroffen werden, wie mit lückenhaften Daten und Defiziten im theoretischen Verständnis umgegangen werden soll. Die Anwendung der oben angesprochenen Standardannahmen bzw. „worst case“ Annahmen ist hierfür ein Beispiel. Ein anderes Beispiel sind die unterschiedlichen Erhebungsverfahren (*sampling methods*), die bei der Expositionsabschätzung eingesetzt werden (vgl. IPCS 2000). Je nach dem, welcher Expositionspfad betrachtet wird und an welcher Stelle des Expositionspfades die Erhebung durchgeführt wird (z.B. Menge von Pestiziden im Getreide auf dem Feld vs. Menge von Pestiziden in den entsprechenden Lebensmitteln), können sich völlig unterschiedliche Abschätzungswerte ergeben. Keines dieser Erhebungsverfahren ist *an sich* besser, vielmehr hängt es von den konkreten Randbedingungen (verfügbare Zeit und andere Ressourcen) ab, welches verwendet wird. Für die Risikocharakterisierung ist wichtig, offen zu legen, welche Annahmen getroffen oder welche Erhebungsverfahren gewählt wurden, damit dies bei der Bewertung der Ergebnisse berücksichtigt werden kann.

Betrachtet man die vielfältigen Unsicherheiten, die sich bei der Abschätzung von Risiken ergeben können – und die hier nur ansatzweise diskutiert werden können – so stellt sich die Frage, wie mit diesen Unsicherheiten im Rahmen einer CRA umgegangen werden soll.

Zunächst ist darauf hinzuweisen, dass eine vergleichende Risikobewertung immer nur so gut sein kann, wie es die zugrunde liegenden Daten aus den Risikoabschätzungen erlauben. Und da sich diese Unsicherheiten im Rahmen einer CRA nicht ausräumen lassen, sollte darauf geachtet werden, dass bei der vergleichenden Risikobewertung keine Risikoabschätzungen verwendet werden, deren Belastbarkeit stark divergiert. Das heißt, im Rahmen einer CRA sollten nur Risiken verglichen werden,

die auf Risikoabschätzungen beruhen, die konzeptionell äquivalent sind, d.h. auf Methoden, deren Anwendung zu ähnlich sicheren bzw. unsicheren Abschätzungen führen (Neus et al. 1995, 305).

Für den Umgang mit den verbleibenden Unsicherheiten sind zwei Problemaspekte zu unterscheiden: zum einen die Frage, wie die Unsicherheiten kommuniziert werden können, damit die Teilnehmer einer CRA verstehen, was diese Unsicherheiten bedeuten. Dies ist vor allem dann von großer Bedeutung, wenn bei der CRA auch Stakeholdergruppen beteiligt sind, die keine Experten für Risikoabschätzungen sind.⁶² Zum anderen muss entschieden werden, in welcher Weise die Unsicherheiten bei der vergleichenden Bewertung berücksichtigt werden sollen.

Die Möglichkeiten zur Kommunikation von Unsicherheiten hängen zunächst davon ab, ob diese sich quantitativ oder nur qualitativ beschreiben lassen. Liegen quantitative Charakterisierungen für die mit einer Abschätzung verbundenen Unsicherheiten vor, z.B. in Form von Konfidenzintervallen für Schätzwerte (Mittelwert, Media), Perzentilen oder Wahrscheinlichkeitsverteilungen, so können diese neben ihrer numerischen Darstellung auch graphisch illustriert werden.⁶³ Beispiele hierzu finden sich etwa in Morgan & Henrion (1990, Kap. 9), Thompson & Bloom (2000) und WHO (2000). Lassen sich Unsicherheiten nur qualitativ beschreiben (z.B. „Unsicherheit groß“, „Unsicherheit eher klein“), so kann auch hier neben der verbalen Charakterisierung versucht werden, eine graphische Veranschaulichung zu finden.⁶⁴ Bei der Nutzung verbaler Beschreibungen für das Ausmaß von Unsicherheit ist allerdings Vorsicht geboten, da – wie die in Kap. 4.3 dargestellten Untersuchungen zum Verständnis verbaler Charakterisierungen von Wahrscheinlichkeitsaussagen zeigen – die Bedeutung solcher Beschreibungen interindividuell stark variieren kann. In diesem Zusammenhang sind auch die (ebenfalls in Kap. 4.3 berichteten) Befunde von Erev & Cohen (1990) von Bedeutung, dass die meisten Menschen zwar verbale Beschreibungen vorziehen, wenn sie Wahrscheinlichkeiten kommunizieren sollen, selbst aber lieber in numerischer Form über Wahrscheinlichkeiten informiert werden möchten.

Die wesentliche Herausforderung für die CRA liegt aber wohl in der Frage, wie die Unsicherheiten, die bei der Abschätzung für die zu vergleichenden Risiken bestehen, im Rahmen der vergleichenden Risikobewertung berücksichtigt werden können. Grundsätzlich gibt es hier drei Möglichkeiten:

1. Aus dem durch die Unsicherheit charakterisierten Wertebereich wird ein bestimmter Schätzwert ausgewählt. Dies können zum Beispiel der Mittelwert oder Median sein oder auch ein oberer oder unterer Perzentilwert. Für den Vergleich bedeutet diese Auswahl, dass die Unsicherheit nicht mehr explizit berücksichtigt werden kann.

⁶² Aber auch für Experten mit verschiedenem professionellem Hintergrund können sich Verständnisprobleme ergeben, etwa zwischen Toxikologen und Epidemiologen (vgl. Bailar & Bailer 1999).

⁶³ Einen Überblick zu den verschiedenen Möglichkeiten graphischer Darstellung quantitativer Daten geben Cleveland (1993) und Tufte (1988).

⁶⁴ Siehe dazu auch die Diskussion der Darstellung von Unsicherheiten im WBGU Gutachten (1998) in Anhang 2.

2. Es werden verschiedene Bewertungsdurchgänge mit unterschiedlichen Schätzwerten durchgeführt, zum Beispiel zuerst für den Mittelwert, und dann jeweils für einen ausgewählten Perzentilwert. Es kann damit geprüft werden, inwieweit sich unterschiedliche Rankings für die verschiedenen Werte ergeben.⁶⁵
3. Unsicherheit wird als eigenständiger Bewertungsaspekt (Attribut) neben dem Mittelwert oder Median berücksichtigt. Im Rahmen eines multiattributen Bewertungsansatzes, wie er in Kapitel 5.2 erläutert wird, kann diesem Aspekt dann ein bestimmtes Gewicht zugeordnet und damit der Einfluss der Unsicherheit auf die Bewertung festgelegt werden.

Welche dieser drei Möglichkeiten für eine CRA angemessen ist, lässt sich nur im Hinblick auf die konkreten Rahmenbedingungen entscheiden. Aus theoretischer Sicht ist Alternative 2 die beste, da sie den Einfluss unterschiedlicher Annahmen über die tatsächliche Ausprägung eines Schätzwertes auf die vergleichende Risikobewertung transparent macht. Sie ist allerdings auch die aufwendigste.

5.2 Methodik vergleichender Risikobewertung

Den methodischen Rahmen für den Vergleich von Risiken bezüglich unterschiedlicher Bewertungskriterien liefert der Ansatz der multiattributen Entscheidungsfindung.⁶⁶ Dieser Ansatz – allgemein auch als entscheidungsanalytischer Ansatz bezeichnet – ist in der Vergangenheit schon explizit in einigen CRA-Verfahren genutzt worden, vor allem unter dem Gesichtspunkt der Einbeziehung unterschiedlicher Stakeholder-Perspektiven (z.B. Apostolakis & Pickett 1998; Bonano et al. 2000; McDaniels, Gregory & Fields 1999; Merkhofer, Conway & Anderson 1997). Implizit oder explizit liegt allen in Kapitel 3 dargestellten CRA-Verfahren eine solche multiattributen Struktur zugrunde.

Der Ansatz der multiattributen Entscheidungsfindung erlaubt es, verschiedene Entscheidungsoptionen in bezug auf unterschiedliche Bewertungsaspekte (Attribute) zu beurteilen und zu einer Gesamtbeurteilung zu integrieren, die dann als Basis für die Entscheidung dienen kann. Diese Konzeption lässt sich problemlos auf die CRA übertragen: Die Optionen sind entweder die zu bewertenden Risiken (Risikoquellen) oder bestimmte risikobehaftete Alternativen. Im ersten Fall werden verschiedene Risiken in bezug auf einen oder mehrere Bewertungsaspekt(e) beurteilt. Als Ergebnis erhält man eine Rangreihe von Risiken, die ihre Wichtigkeit / Schwere aus Sicht des Bewerter widerspiegelt. Im zweiten Fall geht es um die Auswahl der besten unter verschiedenen Alternativen, die ebenfalls im Hinblick auf ein oder mehrere Attribut(e) beurteilt werden.

Dabei basiert die Entscheidungsfindung auf zwei unterschiedlichen Arten von Information: Zum einen auf der Kenntnis bzw. den Erwartungen bezüglich der Folgen

⁶⁵ Dies entspricht der Durchführung einer Sensitivitätsanalyse, auf die unten in Abschnitt 5.2.2.7 eingegangen wird.

⁶⁶ Es gibt zahlreiche Texte, in denen die Theorie und Praxis multiattributer Entscheidungsfindung behandelt wird. Wir beziehen uns hier im wesentlichen auf Eisenführ & Weber (1993), Jungermann, Pfister & Fischer (1998); Keeney (1992), von Winterfeldt & Edwards (1986) und Yoon & Hwang (1995).

bzw. Effekte, die sich aus den jeweiligen Optionen ergeben können. Hier geht es also um mehr oder weniger gesichertes Faktenwissen. Zum anderen wird bei der Entscheidungsfindung auf Wertungen zurückgegriffen. Diese können sich zum Beispiel beziehen auf die Bewertung der Schwere, Wünschbarkeit, Akzeptabilität etc. bestimmter Folgen bzw. Effekte. Die Wertungen können sich aber auch auf die Bedeutung beziehen, die verschiedenen Beurteilungsdimensionen zugesprochen wird.

Pointiert könnte man also sagen: Entscheidungen basieren auf Fakten und Werten. Im Kontext vergleichender Risikobewertungen werden die Fakten durch die Verfahren der Risikoabschätzung geliefert. Natürlich – und darauf wurde und wird immer wieder hingewiesen (z.B. Shrader-Frechette 1995) – basieren die Verfahren der Risikoabschätzung selbst wiederum auf zahlreichen Wertungen. Beispielsweise müssen die Endpunkte ausgewählt werden, in bezug auf die eine Risikoabschätzung vorgenommen werden soll. Es müssen Messverfahren und Untersuchungsdesigns gewählt werden. Es müssen Vorgaben hinsichtlich der erforderlichen Datenqualität gemacht werden, usw. All dies sind Entscheidungen, die letztlich auf Wertungen der beteiligten Wissenschaftler beruhen. Allerdings – und dies ist ein wichtiger Unterschied zu anderen Wertungen – müssen für diese Entscheidungen in der Regel gute Gründe angeführt werden und sie müssen auch in der kritischen Auseinandersetzung zwischen Wissenschaftlern begründet werden. Diese Wertungen sind also rechtfertigungspflichtig (selbst wenn diese Rechtfertigungen in vielen Fällen im Laufe der Wissenschaftsgeschichte in den Hintergrund gerückt sind und für Nicht-Experten schon gar nicht nachvollziehbar sind).

Im folgenden soll zunächst auf den Unterschied zwischen *quantitativen* und *qualitativen Ansätzen* für die multiattribut Risikobewertung eingegangen werden. Danach wird der *Prozess* der multiattributen Bewertungen behandelt. Zunächst wird dabei von einem einzelnen Bewerter ausgegangen. Die besonderen Probleme von vergleichenden Risikobewertungen mit mehreren Stakeholdern werden danach in einem gesonderten Abschnitt diskutiert.

5.2.1 Multiattribut Risikobewertung: quantitative und qualitative Ansätze

Grundsätzlich ist die Struktur multiattributer Risikobewertung sehr einfach: Es gibt eine Reihe von Attributen ($x_{1...n}$), die aus Sicht des Bewerter für die Beurteilung u einer gegebenen Menge von m Risiken – oder allgemeiner: Optionen – von Bedeutung ist. Diese Attribute können für den Bewerter unterschiedlich wichtig sein und die Risiken können auf unterschiedliche Bewertungen den Attributen erhalten (siehe Tabelle 19).

Tabelle 19: Struktur multiattributer Entscheidungsfindung

	Attribut 1	Attribut 2	...	Attribut n	Gesamtbewertung
Risiko 1	$u(x_{11})$	$u(x_{12})$		$u(x_{1n})$	?
Risiko 2	$u(x_{21})$	$u(x_{22})$		$u(x_{2n})$	?
...	...	...		...	...
Risiko m	$u(x_{m1})$	$u(x_{m2})$		$u(x_{mn})$	?

Voraussetzung für eine solche multiattribut Bewertung ist, dass alle Risiken auf dem jeweiligen Attribut bewertet werden können. Die Bewertung kann dabei entweder *quantitativ* oder *qualitativ* erfolgen. Quantitative Bewertungen basieren auf quantitati-

ven Risikoabschätzungen und den dabei zugrunde gelegten quantitativen Risikometriken, z.B. „Zahl der Todesfälle“ oder „Relatives Risiko“. Auf dieser Basis lassen sich kardinal skalierte Risikowerte berechnen, die einen exakten quantitativen Vergleich ermöglichen (Risiko 1 ist drei mal so hoch wie Risiko 2).

Wie oben in Kapitel 5.1 gezeigt wurde, sind quantitative Risikoabschätzungen allerdings häufig mit großen Unsicherheiten behaftet, so dass es sinnvoll sein kann, Abschätzungen mit ähnlichen Werten und Unsicherheiten zu Kategorien zusammenzufassen. Beispielsweise könnten Häufigkeiten von Schadensereignissen in die Kategorien „selten“, „gelegentlich“, „häufig“ eingeordnet werden. Oder eine Kategorie, die sich auf die Schwere eines Risikos bezieht, könnte mit den Begriffen „gering“, „mittel“, „hoch“ bezeichnet werden. Da eine solche Kategorisierung quantitativer Daten immer mit einem Verlust an Information verbunden ist, erscheint dies nur angemessen, wenn die Unsicherheiten in den quantitativen Risikoabschätzungen so groß sind, dass sie sich nicht mehr sinnvoll quantitativ durch statistische Maßzahlen wie Mittelwert oder Varianz beschreiben lassen. Im Extremfall fehlen quantitative Daten völlig, so dass nur eine grobe, kategoriale Abschätzung der Risiken möglich ist.

Eine qualitative vergleichende Risikobewertung kann aber auch aus praktischen Gründen angezeigt erscheinen, z.B. um die Durchführung des Verfahrens für Laien zu vereinfachen. In jedem dieser Fälle stellt sich jedoch die Frage, ob auf der Basis dieser beschränkten Daten eine vergleichende Risikobewertung überhaupt noch einen Informationswert hat.

Eine spezielle Variante qualitativer Risikobewertungen sind diejenigen, bei denen die *Ergebnisdarstellung* in Form von Kategorien erfolgt, obwohl die zugrunde liegenden Risikoabschätzungen durchaus quantitativer Natur sein können. Vor allem im Kontext der Regulierung von Risiken werden häufig Risikoklassen gebildet, für die dann bestimmte Maßnahmen vorgesehen sind. Beispiele sind die Einordnung kanzerogener Chemikalien am Arbeitsplatz in das Klassifikationsschema der MAK-Kommission (Greim & Reuter 2001) oder die Einordnung krebserzeugender Luftschadstoffe in die Emissionsbegrenzungsklassen der TA-Luft (Schuhmacher-Wolz, Konietzka & Schneider 2002; siehe oben Kapitel 3.2.6). Grundlage für die Einordnung in die Risikoklassen ist hier eine Bewertung der Evidenz für die Kanzerogenität der Stoffe, in die sowohl quantitative Risikoabschätzungen wie auch die Beurteilung der Qualität der verfügbaren Daten eingehen. Auch im WBGU-Gutachten „Welt im Wandel. Strategien zur Bewältigung globaler Umweltrisiken“ (WBGU 1998) werden die Risikobewertungen in Kategorien zusammengefasst, die allerdings nicht mehr in einer ordinalen Relation zueinander stehen, sondern direkt auf bestimmte Risikomanagementstrategien verweisen.⁶⁷

⁶⁷ Siehe dazu das Kapitel 3.1.3 und Anhang 2.

5.2.2 Durchführung einer multiattributen Bewertung

Im multiattributen Bewertungsprozess können 7 Schritte unterschieden werden:⁶⁸

1. Problemstrukturierung
2. Auswahl und Kategorisierung der zu betrachtenden Risiken
3. Attributgenerierung und -strukturierung
4. Beurteilung der Risiken auf den Attributen
5. Attributgewichtung
6. Aggregation der Einzelbewertungen und (vorläufiges) Ranking
7. Sensitivitätsanalyse

Mittlerweile gibt es eine Reihe von Computerprogrammen, die die Durchführung multiattributer Bewertungsprozesse bzw. Entscheidungsfindung unterstützen. Vor allem bei der Beurteilung der Optionen (Risiken) auf den Attributen, bei der Attributgewichtung sowie bei der Aggregation der Einzelbewertungen und der Sensitivitätsanalyse sind diese Programme – jedenfalls, wenn eine quantitative multiattribute Bewertung vorgenommen werden soll – sehr hilfreich, da sie Verfahren etwa zur Elizitierung von Gewichtungen oder von Bewertungsfunktionen (Nutzenfunktionen) enthalten und Konsistenzprüfungen vornehmen können. Und in der Praxis ist eine Sensitivitätsanalyse ohne Computerunterstützung nur sehr umständlich durchzuführen. Viele Programme unterstützen aber auch die Auswahl von Optionen und die Generierung und Strukturierung von Attributen. Eine aktuelle Übersicht findet sich in Maxwell (2002). Keefer, Kirkwood & Corner (2003) geben einen Überblick zu Entwicklungsperspektiven von Entscheidungsunterstützungsverfahren und Kriterien für die Auswahl und Evaluation von *Decision Support Software* werden in Rohrmann & Schütz (1993) diskutiert.

5.2.2.1 Problemstrukturierung

Zu Beginn einer CRA müssen eine Reihe von Präzisierungen und Festlegungen getroffen werden, um die allgemeine Zielstellung der vergleichenden Risikobewertung auf eine operationale Ebene zu bringen:

- Welche Risiken sollen berücksichtigt werden?
- Für welchen geographischen Raum (national, regional, lokal) bzw. welche Population sollen Risiken verglichen werden?
- Welcher Zeitraum soll betrachtet werden: gegenwärtige Risiken oder zukünftige Risiken?
- Wer sind die Bewerter und welche Bewertungen sollen sie vornehmen (Gewichtung der Attribute, Beurteilungen auf den Attributen)?
- In welcher Weise soll die Bewertung der Risiken erfolgen: quantitativ oder qualitativ? Dies hängt vor allem von den folgenden Aspekten ab:

⁶⁸ In Anlehnung an das SMART-Verfahren von Edwards (1977) bzw. Edwards & Barron (1994).

- Ist eine quantitative Bewertung angesichts der Datenlage, d.h. wegen der evtl. bestehenden Unsicherheiten in den Risikoabschätzungen, möglich und sinnvoll?
- Wer sind die Adressaten der vergleichenden Risikobewertung, d.h. für wen sind die Ergebnisse der CRA bestimmt und welchen Zweck haben sie?
- Wie weit soll die Aggregation der Bewertungen auf den Attributen gehen?

In vielen Fällen wird dieser erste Schritt durch den Kontext, in dem eine CRA erstellt werden soll und die daraus resultierenden konkreten Anforderungen, schon weitgehend festgelegt.

5.2.2.2 Auswahl und Kategorisierung der zu betrachtenden Risiken

Auch dieser Schritt wird in vielen Fällen bereits durch den Kontext, in dem eine CRA durchgeführt wird, vorgegeben sein: wenn beispielsweise im Rahmen einer vergleichenden Bewertung chemischer Agenzien schon Stofflisten zusammengestellt sind.

Ist eine Liste der zu betrachtenden Risiken noch nicht vorhanden, so kann sie beispielsweise aus Umfragen gewonnen werden. Generell ist zu beachten, dass der Vergleich einer großen Zahl von Risiken schnell an praktische Grenzen stößt (vgl. Morgan et al. 2000). Es ist dann sinnvoll, die Risiken zu kategorisieren, indem zum Beispiel bestimmte Problemfelder ausgewählt werden, etwa Aktivitätsfelder (Ernährung etc.), die dann in einem weiteren Strukturierungsschritt in Problemfelder (z.B. Nahrungsmittel, Innenraumlufte) unterteilt werden. Für die unterste Ebene dieser Problemfelder müssen dann Leitindikatoren (z.B. Pestizidbelastung, Acrylamidbelastung oder Belastung durch Radon und Passivrauchen) festgelegt werden, die die zu vergleichenden Risikoquellen darstellen. Diese müssen so ausgewählt werden, dass damit die wesentliche Risikoproblematik im Problemfeld abgedeckt wird.

Schon in diesem frühen Schritt der Strukturierung sind zahlreiche Entscheidungen zu treffen, die das Ergebnis der CRA stark beeinflussen können. So muss etwa entschieden werden, ob für das Problemfeld Innenraumlufte neben Radon und Passivrauchen noch andere Indikatoren zu betrachten sind (z.B. Formaldehyd). Es ist leicht einzusehen, dass die Auswahl der Kategorien und Indikatoren das Ergebnis der CRA maßgeblich beeinflussen kann.

Bei der Strukturierung der Risiken ist auch zu beachten, dass nur solche Risiken für die CRA ausgewählt werden, für die konzeptionell äquivalente Risikoabschätzungen vorliegen, d.h. die auf Methoden beruhen, deren Anwendung zu ähnlich sicheren bzw. unsicheren Abschätzungen führt (Neus et al. 1995, 305).⁶⁹

5.2.2.3 Attributgenerierung und -strukturierung

Attribute bezeichnen die Aspekte, die bei der Risikobewertung aus Sicht des Bewerbers von Bedeutung sind. Hier kann es zum Beispiel um die betroffenen Schutzgüter

⁶⁹ Siehe oben Kapitel 5.1.3.

(z.B. menschliche Gesundheit, Umwelt, Lebensqualität) gehen oder um die Art des Schadens (Tod, chronische Schädigung etc.). Beispielsweise wurden für das *New Jersey Comparative Risk Project* (NJDEP 2003) die in Abbildung 13 wiedergegeben Attribute genutzt. Es werden dort drei Bereiche für die vergleichende Risikobewertung unterschieden: menschliche Gesundheit, ökologische Kriterien und sozio-ökonomische Kriterien, die mit jeweils drei Attributen charakterisiert werden. In dem Projekt werden diese drei Bereiche nicht zu einer Gesamtbewertung zusammengefaßt, sondern es bleiben als Ergebnis drei getrennte vergleichende Risikobewertungen (die auch zu unterschiedlichen Rangreihen führen).

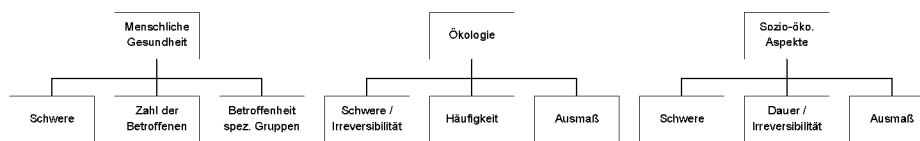


Abbildung 13: Beispiel für Attribute im Rahmen einer CRA (nach: NJDEP 2003)

Ob einzelne Bereiche, wie menschliche Gesundheit, Ökologie oder sozio-ökonomische Aspekte, im Rahmen einer CRA getrennt betrachtet werden oder – siehe das in Abbildung 14 dargestellte Beispiel – zu einer Gesamtbewertung zusammengefasst werden, hängt von der Zielsetzung der CRA ab. Bei einer Aggregation unterschiedlicher Teilbereiche zu einer Gesamtbewertung sollten diese aber in jedem Fall auch getrennt betrachtet werden, um zu verstehen, welche (möglicherweise) unterschiedlichen Rangfolgen der Risiken sich dort ergeben. Dies ist im Rahmen des multiattributen Ansatzes leicht möglich.

Die gewählten Attribute müssen – und dies gilt für quantitative und qualitative Bewertungsverfahren gleichermaßen – verschiedene Bedingungen erfüllen, um im Rahmen eines multiattributen Bewertungsverfahrens genutzt werden zu können. Kriterien dafür sind:

- (a) *Vollständigkeit*: Es muss sichergestellt werden, dass alle für die Bewertung wesentlichen Aspekte berücksichtigt werden. Diesem Aspekt kommt bei Multistakeholder-Bewertungen eine besondere Bedeutung zu (s.u.).
- (b) *Redundanzfreiheit*: Die Attribute sollten sich in ihrer Bedeutung nicht überschneiden, denn Bedeutungsüberschneidungen können dazu führen, dass ein Attribut einen größeren Einfluss bei der Entscheidungsfindung bekommt, als ihm eigentlich zukommt.
- (c) *Messbarkeit*: Die Attribute müssen so operationalisiert sein, dass es möglich ist, die Ausprägung der Risiken auf den Attributen zu messen. Solche Messungen sollen zum einen möglichst *treffend* sein, das heißt, es sollen diejenigen Aspekte gemessen werden, die dem Bewerter wirklich wichtig sind. Zum anderen sollen die Messungen möglichst *eindeutig* sein, das heißt, Unschärfen bei den Messungen sollen so gering wie möglich sein.
- (d) *Unabhängigkeit*: Die Beurteilung der Risiken auf einem Attribut soll unabhängig von der Beurteilung auf den anderen Attributen möglich sein. Eine Veränderung in der Beurteilung auf einem Attribut darf also nicht zu einer Veränderung auf einem (oder mehreren) anderen Attribut(en) führen. Diese Präferenzunabhängigkeit darf nicht mit der in einem konkreten Fall empirisch gegebenen Unabhängigkeit ver-

wechselt werden (statistische Unabhängigkeit). Für gegebene Alternativen können die Attribute faktisch sehr wohl korrelieren, d.h. die Ausprägung auf einem Attribut mit denen eines anderen Attributs zusammenhängen. Präferenzunabhängigkeit fordert nur, dass die Bewertungen im Prinzip unabhängig voneinander möglich sein müssen.

Häufig ist es sinnvoll, diese Attribute im Rahmen von Zielhierarchien oder Wertebäumen zu organisieren. Eine solche Hierarchisierung hilft dem Bewerter, sich über die verschiedenen Bewertungsaspekte und ihre Beziehung zueinander klar zu werden (siehe Abbildung 14). Wichtig ist, dass für die terminalen Attribute (d.h. die Attribute auf der untersten Ebene)⁷⁰ eine Metrik verfügbar ist, auf der die jeweiligen Risiken beurteilt werden können.

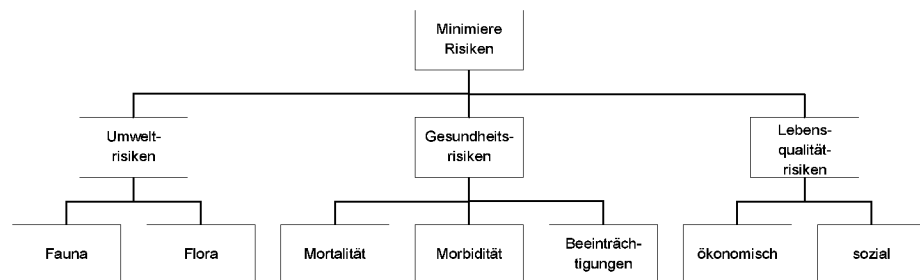


Abbildung 14: Beispiel für hierarchische Attributstruktur

Die Wahl der terminalen Attribute hängt deshalb auch von den zu bewertenden Risiken ab – genauer: vom Auflösungs-niveau, auf dem die Risiken betrachtet werden. Die in Anhang 1 dargestellten Risikometriken geben eine Übersicht zur Bandbreite der verfügbaren Meßverfahren. In Fällen, in denen keine Daten auf solchen Risikometriken verfügbar sind, können Indikatoren für Risiken, z.B. Emissionsmengen, genutzt werden. Für diese muss dann allerdings plausibilisiert werden, warum sie für die Bewertung des Risikos relevant sind.

Das Problem der Meßbarkeit der terminalen Attribute stellt sich natürlich vor allem bei quantitativen Risikobewertungen. In abgeschwächter Form gilt es aber auch für qualitative Bewertungen. Denn auch hier muss es möglich sein, eine Bewertung der Risiken durch Zuordnung in Kategorie für das jeweilige Attribut vorzunehmen.

Attributstrukturen können entweder *top-down* oder *bottom-up* entwickelt werden.⁷¹ Im ersten Fall versucht man zunächst, allgemeine Risikofelder, Ziele oder Werte zu benennen, die für die betrachteten Risiken relevant sind. Im nächsten Schritt werden diese allgemeinen Konzepte so weit ausdifferenziert, bis eine hierarchische Struktur gegeben ist, an deren unterem Ende die Attribute stehen, auf denen sie beurteilt werden können. Das *bottom-up* Verfahren beginnt mit einer Sammlung möglichst

⁷⁰ Die Verwendung des Begriffs Attribut ist in der Literatur nicht ganz eindeutig. Zum Teil werden nur die terminalen Attribute als Attribute bezeichnet und die darüberliegenden als Ziele. Wir sprechen hier unabhängig von der Ebene immer von Attributen.

⁷¹ Ausführlich wird die Entwicklung von Attributstrukturen in Keeney (1992) behandelt.

konkreter Beurteilungsaspekte für die Risiken. Diese werden unter Oberbegriffen zusammengefasst und so hierarchisch strukturiert.

Hinsichtlich der Anzahl der zu berücksichtigenden Attribute gilt die im Alltag bewährte Optimierungsregel: so viele Attribute wie nötig (um alle bewertungsrelevanten Aspekte zu erfassen), aber so wenige wie möglich (um die Übersicht nicht zu verlieren). Dabei ist auch zu berücksichtigen, dass bei qualitativen vergleichenden Risikobewertungen eine große Zahl von Attributen meist schwieriger zu handhaben sind, als bei quantitativen. Zu beachten ist auch, dass es wenig Sinn macht, Bewertungskriterien einzubeziehen, für die man keine belastbaren Daten hat. So mag man beispielsweise den Aspekt „Betroffenheit zukünftiger Generation“ für ein wesentliches Bewertungskriterium im Rahmen einer CRA halten. Wenn hierzu aber keinerlei wissenschaftliche Abschätzungen vorliegen oder die Abschätzungen sehr große Unsicherheiten enthalten, ist eine Einbeziehung dieses Kriteriums vermutlich nicht zielführend.

5.2.2.4 Beurteilung der Risiken auf den Attributen

Die Beurteilung der Risiken auf den Attributen ist die Aufgabe von Experten, denn hierbei geht es nicht um Bewertungsfragen, sondern um die Feststellung von Fakten, d.h. eine Beurteilung auf der Basis von Risikoabschätzungen – auch wenn diese mit mehr oder weniger großen Unsicherheiten behaftet sein mögen.

In der Art der Beurteilung bestehen zwischen qualitativen und quantitativen Bewertungsverfahren deutliche Unterschiede. Bei einem qualitativen Verfahren erfolgt diese Beurteilung durch Einordnung in Kategorien. Die Zahl der für die Charakterisierung eines Attributs verwendeten Beurteilungskategorien hängt im wesentlichen von praktischen Erfordernissen ab. Es gibt hierzu keine allgemeinen Richtlinien. Sinnvoll ist aber sicherlich, nicht zu viele Kategorien vorzusehen. Häufig werden drei Kategorien verwendet, wie in dem folgenden Beispiel aus dem *New Jersey Comparative Risk Project* (NJDEP 2003). Hier werden die Kategorien „gering“, „mittel“ und „hoch“ jeweils zur Beschreibung der Attribute „Schwere des Gesundheitseffekts“ („*Severity of specified health effects at current levels of exposure*“), „Zahl der Betroffenen“ („*Size of population at significant risk for each health effect*“) und „Gesamtrisikoeinschätzung“ („*Overall risk ranking (as a function of severity and population effected integrating across health effect)*“) verwendet. Außerdem wird noch eine dichotome Kategorie mit den Ausprägungen „ja“ und „nein“ für die Messung des Attributs „Betroffenheit spezifischer Gruppen“ („*Are there discrete communities at elevated risk*“) benutzt. Tabelle 20 zeigt einige Beispiele für Beurteilungen, die von Experten für den jeweiligen Bereich auf der Basis von Risikocharakterisierungen vorgenommen wurden.

Tabelle 20: Beispiel für eine qualitative multiattribute Risikobewertung⁷²

Risikoquelle	Schwere des Gesund- heitseffekts	Zahl der Betroffenen	Betroffenheit spezifischer Gruppen	Gesamt- risiko- einschätzung
Luftgängige Krankheitserreger	gering	unbekannt	ja	gering/mittel
Arsen	mittel	mittel	ja	mittel
Kadmium	gering	mittel	ja	gering
Chrom - insgesamt				groß/mittel
- Krebs	groß	groß/mittel	ja	(k. A.)
- Allergische Dermatitis	mittel	gering/mittel	ja	mittel
- Nicht-karzinogene Aufnahme	gering	groß/mittel	ja	gering/mittel
Dioxine / Furane - Krebs	groß	groß/mittel	ja	groß/mittel
Endokrine Disruptoren	gering/mittel	mittel	ja	mittel
Treibhausgase	gering	groß	ja	gering
Bodenozone	mittel	groß	ja	groß
Pestizide in Nahrungsmitteln	mittel	mittel	nein	mittel
Radon	groß	groß	ja	groß
Passivrauchen	groß	groß	nein	groß
UV Strahlung - insgesamt				mittel
- Haut	groß	mittel	ja	mittel
- Auge	groß	groß	nein	mittel
- Immunsystem	gering	gering	nein	gering

Eine Art semi-quantitativer Risikobeurteilung stellen Ratingskalen dar, bei denen das Ausmaß der Beurteilung auf dem Attribut numerisch ausgedrückt wird. Die in Tabelle 20 dargestellten kategorialen Beurteilungen wurden im *New Jersey Comparative Risk Project* zusätzlich auch noch in Form von Ratings ausgedrückt (siehe Tabelle 21).⁷³ Zwar lassen sich mit solchen numerischen Werten die Einzelbewertungen auf den Attributen rein rechnerisch zu Gesamtwerten aggregieren, solche Rechenoperationen sind aber nur dann sinnvoll, wenn vorausgesetzt werden kann, dass diese Beurteilungen zumindest Intervallskalenniveau (also äquidistante Skalenpunkte) haben. Ohne diese (starke) Annahme haben Ratingskalen nur ein ordinales Niveau und damit nicht mehr Aussagekraft als die verbalen Kategorien, die ja ebenfalls ordinale Relationen ausdrücken.

Anzumerken ist hier, dass verbale Charakterisierungen von Attributausprägungen in Form von Kategorien zwar auf den ersten Blick einfacher erscheinen mögen als Ratingskalen oder die im folgenden dargestellten quantitativen Beurteilungen. Sie erfassen die Beurteilung deswegen aber noch nicht unbedingt zuverlässiger. So zeigt ein Vergleich der kategorialen, verbalen Einschätzungen in Tabelle 20 und der numerischen Ratings in Tabelle 21, dass die Zuordnungen von verbalen Charakterisierungen und Zahlenwerten nicht immer konsistent sind – sowohl für eine Risikoquelle (die von jeweils einem Experten beurteilt wurde) wie zwischen Risikoquellen (die von

⁷² Auswahl aus *New Jersey Comparative Risk Project* (NJDEP 2003, Appendix 4).

⁷³ Die Anweisung für das Rating lautete hier (NJDEP 2003, Appendix 4):

- *Severity of specified health effects at current levels of exposure*: 1-5 with 1 being least severe.
- *Size of population at significant risk for each health effect*: 1-5 with 1 being smallest.
- *Are there discrete communities at elevated risk*: 1-5 with 1 being the lowest probability that there are discrete communities at elevated risk.
- *Overall risk ranking (as a function of severity and population effected integrating across health effect)*: 1-5 with 1 being the lowest overall risk.

verschiedenen Experten beurteilt wurden). Beispielsweise wird für „Bodenozon“ der Kategorie „groß“ sowohl der Wert 4 wie der Wert 5 zugeordnet. Und bei Arsen und Cadmium wird für die Kategorie „mittel“ einmal der Wert 3 und einmal der Wert 2 vergeben. Ruft man sich die in Kapitel 4.3 berichteten Diskrepanzen zwischen der Einschätzung von Wahrscheinlichkeiten in verbaler und numerischer Form in Erinnerung, so kommen diese Inkonsistenzen nicht überraschend.

Tabelle 21: Numerische Beurteilungen

Risikoquelle	Schwere des Gesundheitseffekts	Zahl der Betroffenen	Betroffenheit spezifischer Gruppen	Gesamtrisiko-einschätzung
Luftgängige Krankheitserreger	1	unbekannt	k.A. (ja)	2
Arsen	3	3	3 (ja)	3
Cadmium	1	2	3 (ja)	2
Chrom - insgesamt				4
- Krebs	5	4	4 (ja)	4
- Dermatitis (Allergie)	3	2	3 (ja)	3
- nicht-kanzerogene Aufnahme	2	3	3 (ja)	2
Dioxine / Furane - Krebs	5	4	3 (ja)	4
Endokrine Disruptoren	2,5	2,5	5 (ja)	2,5
Treibhausgase	1	4	3 (ja)	2
Lärm – Herz-Kreislauf-Wirkungen	2	2	5	2
Bodenozon	4	5	5 (ja)	4
Pestizide in Nahrungsmitteln	2	3	k.A. (nein)	3
Radon	5	5	5 (ja)	5
Passivrauchen	5	4	k.A. (nein)	5
UV Strahlung - insgesamt				3
- Haut	3	3	5 (ja)	3
- Auge	3	4	1 (nein)	3
- Immunsystem	2	2	1 (nein)	2

Die quantitative Beurteilung von Risiken auf den jeweiligen Attributen basiert auf den Abschätzungen, die in der jeweiligen Risikometrik vorliegen, mit der das Attribut gemessen wird. Um die Attributbeurteilungen auf den verschiedenen Metriken vergleichbar zu machen, werden sie über eine Bewertungsfunktion (die formal einer Wert- oder Nutzenfunktion⁷⁴ entspricht) auf eine gemeinsame Bewertungsskala reskaliert. Diese Bewertungsskala umfasst das Intervall [0;1], wobei 0 den geringsten und 1 den höchsten Wert in Bezug auf das jeweilige Attribut bezeichnet.

Zu unterscheiden sind monoton steigende bzw. monoton fallende Bewertungsfunktionen von nicht-monotonen Bewertungsfunktionen. Eine monoton steigende Bewertungsfunktion wird gewählt, wenn eine höhere Ausprägung auf dem Attribut auch einen höheren Wert bedeutet, z.B. wenn das Attribut „Ausmaß Reduktion von Schadstoffemissionen“ lautet. Monoton fallende Bewertungsfunktionen werden für Attribute verwendet, mit denen Schäden oder Kosten beurteilt werden, z.B. „Zahl der

⁷⁴ Wir unterscheiden hier nicht zwischen Nutzen- und Wertefunktionen. Vor allem in der entscheidungsanalytischen Literatur wird dieser Unterschied gemacht, wobei Nutzen (*utility*) für Entscheidungen unter Unsicherheit und Wert (*value*) für Entscheidungen unter Sicherheit verwendet wird. Bei der Konstruktion von *Utility*-Funktionen wird die Risikoeinstellung berücksichtigt, bei der Konstruktion von *Value*-Funktionen ist dies nicht der Fall. Zur Kritik dieser Unterscheidung siehe von Winterfeldt & Edwards 1986, Kap. 7.1.

Todesfälle“ für das Attribut „Mortalität“. Nicht-monotone Bewertungsfunktionen werden verwendet, wenn die beste Ausprägung des Attributs nicht am Ende der Attributskala liegt.

Die Konstruktion von Bewertungsfunktionen kann sehr aufwendig sein.⁷⁵ In vielen Fällen kann man aber davon ausgehen, dass lineare Bewertungsfunktion schon eine hinreichend gute Approximation der tatsächlichen Bewertungsfunktion darstellt.⁷⁶

Eine lineare Reskalierung der Attributbewertungen auf die Bewertungsskala kann etwa folgendermaßen vorgenommen werden: Der Wert 0 wird dem schlechtesten Wert der Risiken auf dem Attribut und der Wert 1 dem besten Wert der Risiken auf dem Attribut zugeordnet. Dann lässt sich der Wert $u_i(x_{ij})$ des Risikos i auf dem Attribut j berechnen nach:

$$u_i(x_{ij}) = \frac{x_{ij} - x_{ij}^-}{x_{ij}^+ - x_{ij}^-},$$

wobei Risiko x_{ij}^- der schlechteste Wert der Risiken auf dem Attribut und x_{ij}^+ der beste Wert der Risiken auf dem Attribut sind.

Im Zweifel muss aber geprüft werden, ob die Bewertungsfunktion des Entscheiders wirklich linear ist. Für Attribute beispielsweise, die eine monetäre Skala benutzen, ist eine Nutzenfunktion oftmals nicht linear – vor allem dann nicht, wenn die Skala einen großen Messbereich umfasst. Der Zuwachs des Nutzens von Geld ist im unteren Bereich größer als im oberen Bereich der Skala. Dies zeigt sich in einer konkaven Nutzenfunktion (siehe Abbildung 15).

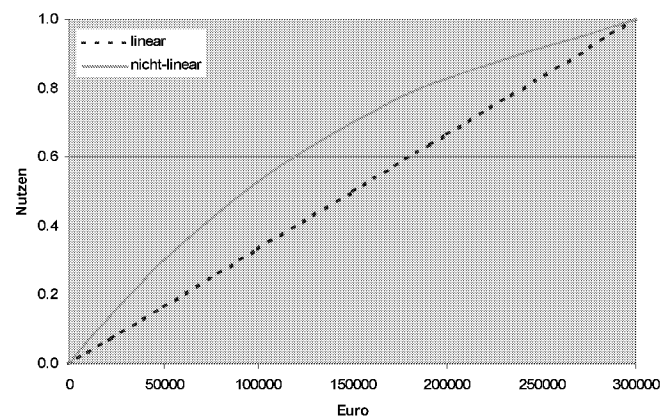


Abbildung 15: Beispiel für eine lineare und nicht-lineare Nutzenfunktion für Geld

⁷⁵ Verschiedene Verfahren zur Konstruktion von Nutzenfunktionen werden ausführlich in Eisenführ & Weber (1993), Jungermann, Pfister & Fischer (1998) sowie von Winterfeldt & Edwards (1986) beschrieben.

⁷⁶ Studien zeigen, dass Veränderungen der Gewichtungen die Entscheidung stärker beeinflussen als die Form (linear vs. nicht-linear) der Nutzenfunktion (vgl. Hobbs 1986; von Winterfeldt & Edward 1986).

Wenn im Rahmen einer CRA nur Risiken betrachtet werden, können die Attributskalen in der Regel so konstruiert werden, dass größere Risiken durch höhere Ausprägungen auf der Skala repräsentiert werden. Damit können monoton steigende Bewertungsfunktionen für die Bewertung des Risikos genutzt werden und die Bewertungsskala wird zu einer Risikoskala.

5.2.2.5 Attributgewichtung

Bei der Attributgewichtung geht es darum, die unterschiedliche Wichtigkeit, die verschiedene Attribute für den Bewerter haben können, explizit zu machen und sie bei der vergleichenden Bewertung zu berücksichtigen. Das Vorgehen ist dabei für qualitative und quantitative Risikobewertungen unterschiedlich.

Da bei qualitativen Risikobewertungen die Beurteilung der Risiken auf den Attributen lediglich auf einem ordinalen Skalenniveau vorliegen, ist es wenig sinnvoll, differenzierte numerische Gewichtungen vorzunehmen. Wenn bei den Attributen überhaupt unterschiedliche Wichtigkeiten berücksichtigt werden sollen, reicht es, die Attribute ebenfalls ordinal zu gewichten, also in eine Rangreihe zu bringen. Damit sind dann die Voraussetzungen für die Anwendung nicht-kompensatorischer Bewertungsverfahren gegeben.

Bei quantitativen Vergleichen erfolgt die Attributgewichtung durch die Abbildung der Wichtigkeit auf dem Intervall [0;1]. Dabei kann zwischen direkten und indirekten Gewichtsungsverfahren unterschieden werden.

Zu den direkten Verfahren gehören Gewichtsungsverfahren unter Nutzung von Rängen. Dabei werden die n Attribute vom Bewerter zunächst gemäß ihrer Wichtigkeit geordnet. Das wichtigste Attribut erhält die Rangordnungszahl $r_1 = 1$, das zweitwichtigste die Rangordnungszahl $r_2 = 2$, usw. Die Rangsummengewichte können dann nach der folgenden Formel berechnet werden:

$$w_j = \frac{(n - r_j + 1)}{\sum_{k=1}^n (n - r_k + 1)}$$

Die Rangsummengewichtung bildet die Ränge linear (d.h. mit gleichem Abstand) auf das Intervall [0;1] ab, wobei die Summe der Gewichte gleich 1 ist.

Ein anderes Berechnungsverfahren ist die reziproke Ranggewichtung:

$$w_j = \frac{\frac{1}{r_j}}{\sum_{k=1}^n \frac{1}{r_k}}$$

Hier werden die Ränge ebenfalls auf das Intervall [0;1] abgebildet und die Summe der Gewichte ist gleich 1. Allerdings nehmen die Gewichte der Ränge nicht linear ab, sondern im Verhältnis ihres Rangplatzes zum ersten Rang, d.h. das Gewicht des

zweiten Rangs ist nur halb so groß wie das des ersten; das des dritten Rangs nur ein Drittel des ersten usw. Mit diesem Berechnungsverfahren bekommen also die hohen Ränge ein besonderes Gewicht.

Rangsummengewichtungen sind zwar auch bei einer großen Zahl von Attributen noch relativ einfach durchzuführen, aber die Berechnungsvorschriften zur Transformation der Ordinalskaleninformation in das Intervallskalenniveau der Gewichte lässt dem Bewerter keine Möglichkeit, Unterschiede in der Wichtigkeit einzelner Attribute, die über eine bloße Rangbildung hinausgehen, auszudrücken.

Ein einfaches Verfahren, das dies erlaubt, ist die Punktegichtung. Hier werden von dem Bewerter eine bestimmte Zahl von Punkten (z.B. 100) so auf die Attribute verteilt, dass die Punkte die relative Wichtigkeit der Attribute zueinander widerspiegelt. Die Normierung auf das Intervall $[0;1]$ erfolgt durch die Division der Punkte durch die Gesamtzahl der zu vergebenden Punkte (z.B. 100). Die Punktegichtung ist einfach, solange nur einige wenige Attribute zu gewichten sind. Ist eine größere Zahl von Attributen zu gewichten, wird dieses Verfahren schnell unübersichtlich und eine konsistente Gewichtung ist schwer durchzuführen.

Ein zentraler Punkt bei der Attributgewichtung ist, dass die Einschätzung der Wichtigkeit von Attributen für den jeweiligen Entscheidungskontext (vor allem in Hinblick auf die anderen Attribute) vorgenommen werden muss. Es geht also nicht um die abstrakte oder generelle Wichtigkeit eines Attributs. Vor allem muss bei der Gewichtung auch die Bandbreite der Ausprägungsmöglichkeiten des Attributs berücksichtigt werden.

Ein Verfahren, das diesen Aspekt beim Erheben der Gewichte berücksichtigt, ist die *Swing-Gewichtung*.⁷⁷ Dieses Gewichtungsverfahren ist ein Beispiel für ein indirektes Verfahren zur Gewichtung von Attributen. Dabei wird zunächst für jedes der n Attribute die schlechteste und beste vorstellbare Ausprägung festgelegt. Dann wird ein hypothetisches Risiko angenommen, das auf allen Attributen die schlechtesten Ausprägungen hat: $R_0 = (x_{1-}, \dots, x_{j-}, \dots, x_{n-})$. Der Bewerter wählt dann das Attribut aus, für das es aus seiner Sicht am wichtigsten ist, von der schlechtesten zur besten Ausprägung verändert zu werden („Swing“). Es wird ein neues (hypothetisches) Risiko X_1 konstruiert, das dieses veränderte Attribut (x_{j+}) enthält: $R_1 = (\dots, x_{j+}, \dots)$. Der vorhergehende Schritt wird jetzt mit dem Attribut wiederholt, das aus Sicht des Bewerters als nächstes von der schlechtesten zur besten Ausprägung verändert werden soll. Damit erhält man ein zweites (hypothetisches) Risiko R_2 . Dieses Verfahren wird fortgesetzt, bis alle Attribute in die beste Ausprägung verändert worden sind. Man erhält dann das „am wenigsten schlimme“ (hypothetische) Risiko R_n . Diese n Risiken bilden eine Rangreihe. Der Bewerter gibt dann auf einer Skala von 0 bis 100 für jedes Risiko R_j an, wie groß die relative Verbesserung t_j ist, die durch den Swing des Attributs x_{j-} nach x_{j+} ist. Danach können die Gewichte für die n Attribute berechnet werden, indem die Werte für alle t_j auf das Intervall $[0;1]$ normiert werden, d.h. $w_j = t_j / \sum t_n$.

Wenn eine große Zahl von Attributen bei der Bewertung berücksichtigt werden muss, kann die Erhebung von Gewichten durch eine hierarchische Attributstruktur vereinfacht werden. Dann können zunächst nur die jeweils in einem Zweig und auf einer Ebene der Attributstruktur vorhandenen Attribute gegeneinander gewichtet werden.

⁷⁷ Siehe Jungermann, Pfister & Fischer (1998, 355).

Für die in Abbildung 14 dargestellte Attributstruktur heißt dies, dass auf der unteren Ebene jeweils getrennt die Attribute Flora und Fauna, die Attribute Mortalität, Morbidität und Beeinträchtigungen sowie soziale Schäden und ökonomische Schäden gegeneinander gewichtet werden müssen. Auf der oberen Ebene werden dann die drei Attribute Umwelt, Gesundheit und Lebensqualität gegeneinander gewichtet. Allerdings sind die so erhaltenen Gewichte zunächst *relative* Gewichte. Die für die multiattribut Bewertung erforderlichen absoluten Gewichte erhält man, indem man die Gewichte der terminalen Attribute mit den Gewichten der jeweils auf einer höheren Ebene ihres Zweiges liegenden Attributen multipliziert. Haben beispielsweise die Attribute Mortalität, Morbidität und Beeinträchtigungen die relativen Gewichte 0.5, 0.3 und 0.2 bekommen, und das Attribut Gesundheit das Gewicht 0.4, so ergeben sich als absolute Gewichte für Mortalität: $0.5 \times 0.4 = 0.20$, für Morbidität: $0.3 \times 0.4 = 0.12$ und für Beeinträchtigungen $0.2 \times 0.4 = 0.08$.

Die berechneten absoluten Gewichte der terminalen Attribute können also sehr stark von den erhobenen relativen Gewichten abweichen. Generell können die Gewichte der terminalen Attribute höchstens so groß wie das Gewicht des obersten Attributs des jeweiligen Zweiges sein. Die absoluten Gewichte werden im Regelfall deutlich geringer ausfallen als die erhobenen relativen Gewichte. In Abbildung 16 zeigt sich beispielsweise, dass die gesundheitlichen Beeinträchtigungen das mit Abstand geringste Gewicht aller Attribute aufweisen und Mortalität nur wenig mehr Gewicht hat als die sozialen Aspekte der Lebensqualität. Nach der Berechnung der absoluten Gewichte sollte deshalb immer geprüft werden, inwieweit diese tatsächlich die Einschätzung der Attributwichtigkeit der Bewerter wiedergeben.

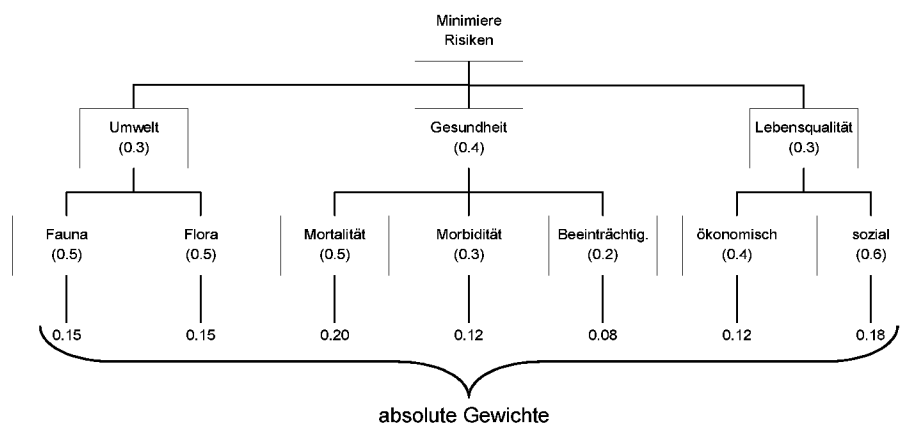


Abbildung 16: Beispiel für Attributgewichtung (relative Gewichte in Klammern)

Abbildung 16 macht auch deutlich, dass die Unterschiede in den Gewichten zwischen den Attributen gering sein können. Da sich die einzelnen Gewichte zum Wert 1 (oder 100%) aufsummieren müssen, wird – wenn es eine größere Zahl terminaler Attribute gibt – der „Spielraum“ für einzelne Attribute schnell eng. Rein rechnerisch lassen sich zwar beliebig differenzierte Attributgewichte bestimmen – deren praktische Aussagekraft ist dann allerdings fraglich. Der praktische Nutzen sehr differenzierter Attributgewichte bei multiattributen Risikobewertungen wird häufig auch dadurch begrenzt, dass für die Beurteilungen der Risiken auf den Attributen keine ge-

nauen wissenschaftlichen Abschätzungen vorliegen, sondern nur Abschätzungen mit einer Unsicherheit (d.h. einer gewissen Bandbreite) gegeben werden können. Je nach dem, welcher Wert aus dem Abschätzungsbereich für die Risikobeurteilung genommen wird, führt dies dann gerade bei in den Gewichtungen nur geringfügig unterschiedenen Attributen schnell zu Veränderungen in der Gesamtbewertung und u.U. auch in der Rangfolge der Risiken (siehe dazu auch unten Abschnitt 5.2.2.7 zur Sensitivitätsanalyse). Mit diesen Anmerkungen soll nicht die Bedeutung der Attributgewichtung in Frage gestellt werden – und schon gar nicht die Bedeutung der Transparentmachung von Gewichtungen. Man muss sich aber bewusst sein, dass man durch Präzision in der Attributgewichtung nicht die Unschärfen in der Risikoabschätzung ausgleichen kann. In Fällen, in denen große Unsicherheiten in den Risikoabschätzungen bestehen, ist es meist sinnvoller, sich bei der CRA auf einige wenige wichtige Attribute zu beschränken.

5.2.2.6 Aggregation der Einzelbewertungen und (vorläufiges) Ranking

Es gibt zwei Arten von Verfahren, nach denen die Einzelbewertungen der Risiken auf den Attributen zu einer Gesamtbewertung für jedes Risiko zusammengeführt werden können: *kompensatorische* und *nicht-kompensatorische* Verfahren. *Kompensatorische* Verfahren erlauben es, die Bewertungen der Risiken auf den Attributen gegeneinander aufzurechnen. Beispielsweise kann eine hohe Bewertung eines Risikos auf dem Attribut A durch eine niedrige Bewertung auf dem Attribut B kompensiert werden. Bei *nicht-kompensatorischen* Verfahren besteht diese Möglichkeit nicht. Wenn ein Risiko auf einem oder mehreren Attribut(en) einen festgelegten Schwellenwert nicht erreicht, wird es nicht gewählt – gleichgültig, wie es auf anderen Attributen abschneidet.

Die qualitative Bewertung von Risiken mit Hilfe von Kategorien erlaubt keine kompensatorischen Bewertungsverfahren, da die Kategorien nur in einer ordinalen Relation zueinander stehen und nicht gegeneinander „verrechnet“ werden können. In diesem Fall muss also auf nicht-kompensatorische Verfahren zurückgegriffen werden.

Nicht-kompensatorische Verfahren

Es gibt zahlreiche nicht-kompensatorische Bewertungsverfahren.⁷⁸ Für den vorliegenden Zusammenhang sind die folgenden von Interesse:

Dominanz: Ausgewählt wird das Risiko, das auf allen Attributen mindestens so hoch wie alle anderen Risiken eingeschätzt wird und das auf mindestens einem Attribut höher eingeschätzt wird.

Lexikographische Ordnung: Gewählt wird das Risiko, das auf dem wichtigsten Attribut die höchste Einschätzung bekommen hat. Haben mehrere Risiken gleich hohe Einschätzungen, so wird in der gleichen Weise das zweitwichtigste Attribut betrachtet, etc.

⁷⁸ Ausführlich dazu: Jungermann, Pfister & Fischer (1998), Kap. 4.4.

Konjunktion: Gewählt wird das Risiko, das auf allen Attributen einen festgesetzten Schwellenwert erreicht.

Disjunktion: Gewählt wird das Risiko, das auf mindestens einem Attribut einen festgesetzten Schwellenwert erreicht.

Entscheidungen nach nicht-kompensatorischen Regeln sind relativ leicht zu treffen. Im wesentlichen geht es darum, die Risiken gemäß der jeweiligen Verfahren zu ordnen. Es ist offensichtlich, dass die Einfachheit nicht-kompensatorischer Verfahren im Rahmen von Risikobewertungen ihre größte Schwäche ist. So bleibt offen, was geschehen soll, wenn mehrere Risiken bzw. Risikokategorien bei Anwendung einer Regel gleich gut abschneiden. Dann muss entweder eine Neubewertung unter Hinzunahme weiterer Attribute erfolgen oder es muss eine Zufallswahl getroffen werden. Letzteres dürfte im Rahmen von Risikobewertungen kaum akzeptabel sein. Wichtiger ist, dass die Reduktion einer zunächst durch viele Bewertungsaspekte gekennzeichneten Bewertungsaufgabe auf Ja/Nein-Entscheidungen bezüglich der Erfüllung eines Schwellenwertes (bei Konjunktion und Disjunktion) oder auf jeweils eindimensionale Vergleichssequenzen (bei der lexikographischen Ordnung) kaum dem Anspruch einer ausgewogenen Risikobewertung entspricht.

Kompensatorische Verfahren

Obwohl es auch bei den kompensatorischen Verfahren verschiedene Ansätze gibt, steht in der praktischen Anwendung eine Theorie im Mittelpunkt: die additive multiattribute Nutzentheorie (*Multi-Attribute Utility Theory*, MAUT). Nach der Entscheidungsregel der MAUT wählt ein Entscheider diejenige Option, die den größten Gesamtnutzen hat. Dieser Gesamtnutzen ist die Summe der gewichteten Einzelnutzen einer Option auf den Attributen.⁷⁹ Im Rahmen einer vergleichenden Risikobewertung bedeutet dies: Die Gesamtbewertung eines Risikos ist die Summe der gewichteten Einzelbeurteilungen des Risikos auf den Attributen. Dies kann in der folgenden Formel ausgedrückt werden:

$$u(R_i) = \sum_{j=1}^n w_j u(x_{ij})$$

wobei

$u(R_i)$ die Gesamtbewertung des Risikos R_i über die n Attribute ist,
 w_j das Gewicht des Attributs x_j ist (mit $w_1 + w_2 + \dots + w_n = 1$), und
 $u(x_{ij})$ die Beurteilung des Risikos R_i auf dem Attribut x_j ist.

Tabelle 22 veranschaulicht die dieser additiven, multiattributen Bewertung zugrunde liegenden Bewertungsmatrix. Als Ergebnis erhält man für jedes Risiko eine Gesamtbewertung, die mit denen der anderen Risiken verglichen werden kann und zu einem Ranking führt.

⁷⁹ Wird keine Gewichtung vorgenommen, so bedeutet dies eine Gleichgewichtung aller Attribute.

Tabelle 22: Generelle Struktur einer multiattributen Bewertung

	Attribut 1 Gewicht 1	Attribut 2 Gewicht 2	...	Attribut n Gewicht n	Σ
Risiko 1	$w_1 u(x_{11})$	$w_2 u(x_{12})$	...	$w_n u(x_{1n})$	$\sum_{j=1}^n w_j u(x_{1j})$
Risiko 2	$w_1 u(x_{21})$	$w_2 u(x_{22})$	...	$w_n u(x_{2n})$	$\sum_{j=1}^n w_j u(x_{2j})$
...	...	...	...	...	...
Risiko m	$w_1 u(x_{m1})$	$w_2 u(x_{m2})$	...	$w_n u(x_{mn})$	$\sum_{j=1}^n w_j u(x_{mj})$

5.2.2.7 Sensitivitätsanalyse

Zweck einer Sensitivitätsanalyse ist es festzustellen, wie robust das (vorläufige) Ranking gegen Veränderungen in den Beurteilungen auf den Attributen und/oder Veränderungen in den Attributgewichtungen ist.

Die Beurteilung von Risiken auf den Attributen ist selten so eindeutig möglich, dass keinerlei Abweichungen vorstellbar sind. Häufig bestehen Unsicherheiten hinsichtlich der Ausprägung der Risiken auf den Attributen. Meist stellen die tatsächlich verwendeten Beurteilungen nur eine – etwa die wahrscheinlichste oder plausibelste – der möglichen Beurteilungen dar. Im Rahmen von Sensitivitätsanalysen werden die Beurteilungen der Risiken auf den Attributen systematisch verändert und jeweils eine neue Gesamtbewertung vorgenommen. Dabei wird in der Regel eindimensional vorgegangen, das heißt, es wird immer die Beurteilung für ein Attribut variiert, damit der Zusammenhang zwischen Veränderung der Beurteilung und Gesamtbewertung nachvollziehbar bleibt. Im Prinzip sind aber natürlich auch mehrdimensionale Sensitivitätsanalysen, das heißt, die gleichzeitige Variation mehrerer Attributbeurteilungen, möglich. Damit liefern Sensitivitätsanalysen auch Informationen darüber, wie groß die Unsicherheiten bei der Abschätzung einzelner Effekte, die den Attributbeurteilungen zugrunde liegen, sein dürfen, ohne dass es zu Veränderungen im Ranking kommt.

Sensitivitätsanalysen können nicht nur hinsichtlich der Attributbeurteilung, sondern auch hinsichtlich der Attributgewichtungen vorgenommen werden. Die Wichtigkeit einzelner Attribute abzuwägen, ist für die Bewerter häufig schwierig, insbesondere, wenn es um Attribute geht, die sich auf elementare Werthaltungen beziehen. Ein Beispiel ist die Abwägung der Wichtigkeit von gesundheitlichen Folgen für Kinder, Erwachsene und alte Menschen. Sensitivitätsanalysen können hier zeigen, ob bzw. in welchem Maße Gewichtungsunterschiede zu Veränderungen im Ranking führen. Besondere Bedeutung kommt dieser Anwendung von Sensitivitätsanalysen im Rahmen von Multistakeholderbewertungen zu, da so geprüft werden kann, ob unterschiedliche Einschätzungen sich bezüglich der Wichtigkeit einzelner Attribute auf das Ranking auswirken (siehe Kap. 5.2.2.8).

Bei Sensitivitätsanalysen bezüglich der Gewichtungen ist zu beachten, dass nicht einfach die Gewichtung eines Attributs geändert werden darf, um zu sehen, welche Veränderungen sich ergeben. Denn wegen der Normierung der Gewichte auf den Wert 1 muss auch festgelegt werden, in welcher Weise die anderen Attributgewichte

mit verändert werden sollen, wenn für ein Attribut die Gewichtung variiert wird. Eine Möglichkeit ist, die Gewichte der anderen Attribute so zu verändern, dass die Gewichte weiterhin im gleichen Verhältnis zueinander stehen.

Die Durchführung von Sensitivitätsanalysen kann sehr zeitaufwendig sein. Dies hängt natürlich von der Zahl der zu prüfenden Attribute und der Feinheit der Variationen ab. Trotzdem – oder gerade deswegen – kommt diesem Schritt eine große Bedeutung bei der multiattributen Risikobewertung zu, denn hier zeigt sich, in welchem Maße die gemachten Annahmen bei der Beurteilung der Risiken auf den Attributen und auch die Wertungen, die die Gewichtung der Attribute bestimmen, die Risikobewertung insgesamt beeinflussen.

5.2.2.8 Multiattribute Risikobewertung mit mehreren Stakeholdern

Grundsätzlich ist die Struktur von Bewertungsproblemen mit mehreren Bewertern nicht anders als bei einzelnen Bewertern. Das heißt, es geht ebenfalls darum, Risiken auf verschiedenen Attributen zu bewerten und diese attributweisen Bewertungen dann zu einer Gesamtbewertung zu aggregieren. Der entscheidende Unterschied zum bzw. die notwendige Erweiterung des multiattributen Konzepts liegt darin, die (möglicherweise) unterschiedlichen Problemperspektiven, unterschiedlichen Risiken, unterschiedlichen Attributstrukturen und Attributgewichtungen und unterschiedlichen Risikobewertungen auf den Attributen zusammenzufügen. Abbildung 17 veranschaulicht diese Problematik.

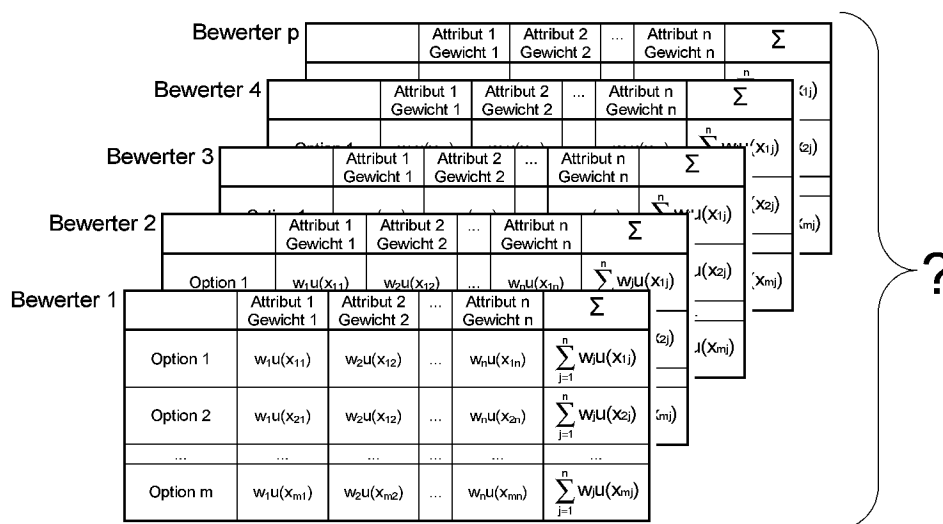


Abbildung 17: Multiattribute Multistakeholder Bewertung

Multiattribute Multistakeholder Bewertungen sind nur als kooperative Verfahren sinnvoll, d.h. alle Beteiligten sind sich einig, dass die oben behandelten Schritte (Aufgaben) der vergleichenden Risikobewertung in einem Gruppenprozess mit dem Ziel einer gemeinsamen Bearbeitung durchzuführen sind. Das bedeutet auch, dass Kriterien und Regeln dafür festgelegt werden müssen, wie mit Bewertungsdifferenzen

oder prozeduralen Dissensen umgegangen werden soll. Grundsätzlich sollten dabei die folgenden Kriterien beachtet werden (EPA 1993, 2.1-21):

- *Konsistenz*: Das Verfahren sollte einen konsistenten Vergleich von Risiken ermöglichen. Es sollte zum Beispiel sicherstellen, dass die ausgewählten Bewertungskriterien für alle Risiken zur Anwendung kommen und von allen Teilnehmern des Bewertungsprozesses beachtet werden. Dies kann etwa bei Verfahren, die geheime Abstimmungen zulassen, nicht ohne weiteres gewährleistet werden.
- *Fairneß*: Das Verfahren sollte offen und transparent sein, so dass etwaige Verzerrungen im Prozess der vergleichenden Risikobewertung von den Teilnehmern entdeckt und zur Sprache gebracht werden können.
- *Dokumentation*: Alle Schritte im Bewertungsprozess sollten dokumentiert werden, so dass Entscheidungen und Festlegungen auch später nachvollzogen und auch von Außenstehenden beurteilt werden können. Dies trägt wesentlich zu transparenten Risikovergleichen bei und erleichtert die Kommunikation der Ergebnisse.

In der konkreten Durchführung dieses Gruppenprozesses sind verschiedene Varianten möglich. Im einen Extrem wird der gesamte Bewertungsprozess gemeinsam durchlaufen, z.B. im Rahmen eines Workshops. Im anderen Extrem werden die einzelnen Bearbeitungsschritte von den Stakeholdern getrennt durchgeführt und die Ergebnisse dann jeweils bei gemeinsamen Treffen vorgestellt und diskutiert. Während Attributgewichtungen und Risikobeurteilungen auf den Attributen für die Stakeholder durchaus unterschiedlich sein können, ist eine Voraussetzung für eine Multistakeholder-Bewertung, dass Einigkeit bezüglich der Problemstruktur und der Attributstruktur erzielt wird.

Zunächst ist also eine gemeinsame Problemsicht zu entwickeln, d.h. es muss Einigung über die Problemstruktur erzielt werden. Dies betrifft zum einen die zu betrachtenden Risiken und den zu betrachtenden Raum-Zeit-Horizont. Geklärt werden muss auch, wer welche Bewertungen vornimmt. Zum anderen muss eine gemeinsame Attributstruktur entwickelt werden.

Ausgangspunkt für eine solche gemeinsame Attributstruktur können die jeweiligen Attributstrukturen der einzelnen Stakeholder sein. Oder die Attributstruktur wird in einem Gruppenprozess von den Beteiligten gemeinsam (*bottom-up* bzw. *top-down*) entwickelt. Wesentlich ist für beide Vorgehensweisen, dass am Ende eine Attributstruktur steht, die alle Beurteilungsaspekte enthält, die von den verschiedenen Beteiligten für wichtig gehalten werden. Darin können durchaus Attribute enthalten sein, die von einzelnen Stakeholdern für unwesentlich gehalten werden; diese können dann bei der Gewichtung von diesen Stakeholdern mit 0 gewichtet und damit effektiv aus ihrer Bewertung herausgenommen werden.

Die Festlegung der Attributgewichte durch die Stakeholder kann in dreierlei Weise erfolgen: (1) Im günstigsten Fall können sich alle Stakeholder – evtl. nach eingehenden Diskussionen – auf eine gemeinsame Gewichtung einigen. (2) Die verschiedenen Gewichtungen werden mathematisch, z.B. durch Mittelwertbildung, zu einer Gewichtung aggregiert. Dieses Vorgehen wird wohl nur dann die Zustimmung aller Beteiligten finden, wenn die Gewichtungen der Stakeholder nahe beieinander liegen. Andernfalls würde das „Ausmitteln“ der Gewichtungsunterschiede bestehende Bewertungsdifferenzen nur verdecken und zu nicht akzeptablen Entscheidungen bzw. Rankings führen. (3) Die unterschiedlichen Gewichtungen bleiben – zumindest zu-

nächst – nebeneinander stehen. Dies führt faktisch zu getrennten Bewertungen der jeweiligen Stakeholder.

Die Beurteilung der Risiken auf den Attributen durch die (ggf. von den Stakeholdern jeweils gewählten) Experten erfolgt in der gleichen Weise wie die Attributgewichtung: entweder durch Einigung auf eine gemeinsame Beurteilung oder durch mathematische Aggregation oder durch (vorläufiges) Bestehenlassen der Beurteilungsunterschiede.

Trotz bestehender Unterschiede in den Gewichtungen bzw. der Risikobeurteilungen kann sich natürlich das gleiche Ranking der Risiken ergeben. In diesem Fall sind die Unterschiede nicht so gravierend, dass sie sich auf das Ranking auswirken würden. Führen diese Unterschiede aber zu verschiedenen Rangfolgen der Risiko, so kann mit Hilfe von Sensitivitätsanalysen geprüft werden, bei welchen Veränderungen der Attributgewichte bzw. Risikobeurteilungen eine einheitliche Rangfolge erzielt werden kann. Zweck der Sensitivitätsanalyse ist in diesem Zusammenhang also nicht zu prüfen, wie robust das Ranking gegenüber Veränderungen in den Gewichtungen bzw. Risikobeurteilungen ist, sondern wie sich mit möglichst geringfügigen Änderungen eine einheitliche Rangfolge für alle Stakeholder erzielen lässt.

Ob die für eine einheitliche Rangfolge notwendigen Veränderungen in den Gewichtungen bzw. Risikobeurteilungen für die Stakeholder akzeptabel sind, kann durch die Sensitivitätsanalyse allerdings nicht geklärt werden. Hier müssen die Stakeholder entscheiden, ob diese Veränderungen noch mit ihren Einschätzungen verträglich sind.

6. Umsetzung der CRA in der Praxis

Für die praktische Umsetzung der CRA sind vier Aspekte zu bedenken. Zum ersten sind die Grenzen des Vergleichens zu bestimmen. Dabei geht es auch um sinnvolle Möglichkeiten für vergleichende Bewertungen von Risiken. Zum zweiten muss geklärt werden, wer die Risiken vergleichen soll und es sind Akteure zu bestimmen und deren Rollen festzulegen. Drittens müssen die Verfahrensabläufe für die CRA spezifiziert werden und viertens müssen Kommunikationsprobleme zwischen den Beteiligten minimiert werden.

6.1 Grenzen der Vergleichbarkeit

Einer der immer wieder geäußerten Einwände gegenüber Risikovergleichen betrifft die Frage, ob und welche Risiken sich überhaupt vergleichen lassen. Eine der Caveats der Deutschen Risikokommission bezüglich des Vergleichs von Risiken lautet, dass nur Vergleichbares verglichen werden sollte (Risikokommission 2003, S. 33).⁸⁰ Was aber ist vergleichbar?

Unstrittig ist, Risiken von Entscheidungsalternativen zu vergleichen. Dies hat im Feld der Technik eine lange Tradition. Beispielsweise werden Standorte oder technische Optionen für die Energiewirtschaft (Keeney 1980, siehe aber auch Kunreuther und Linnerooth 1983) oder die Abfallwirtschaft (Powell 1996), aber auch Varianten von Gefahrguttransporten (Schiene versus Straße) miteinander verglichen. Ebenso ist es sinnvoll, in der Medizin verschiedene Therapien bezüglich ihrer Risiken (aber auch ihres Nutzens) zu vergleichen. Ob sich ganz unterschiedliche Risiken – z.B. Radon, Kernkraftwerke, globale Klimaveränderungen und Innenraumluftverschmutzung – überhaupt vergleichen lassen, wird dagegen bezweifelt (siehe dazu Femers 1993). Diese Problematik soll im weiteren erörtert werden.

Dabei geht es um zwei Problemkreise: Zum einen um unterschiedliches Wissen über Risiken, zum anderen um den Gegensatz von faktischen versus hypothetischen Risiken.

Das Wissensproblem

Das Wissen um ein (tatsächliches oder vermeintliches) Risiko kann von ganz unterschiedlicher Qualität sein. Wir haben an anderer Stelle (Wiedemann et al. 2002) eine Klassifikation nach dem Risikowissen vorgeschlagen und – je nach Wissenslage – Typen unterschieden: neben bekannten Risiken sind das unbekannte Risiken und undeutliche Risiken.

Ein *bekanntes* Risiko liegt vor, wenn das Gefährdungspotenzial (der Hazard), die Dosis-Wirkungskurve sowie die Exposition der Bevölkerung hinreichend gut unter-

⁸⁰ „Bei vergleichender Risikoabschätzung muss besonders kritisch geprüft werden, dass auch nur Vergleichbares verglichen wird.“ (Risikokommission 2003, S. 33).

sucht sind und damit entsprechendes Wissen vorliegt. Hier kann eine vollständige Risikocharakterisierung vorgenommen werden. *Undeutliche* Risiken sind durch massive Informationsmängel geprägt. Es ist nicht einmal klar, ob überhaupt ein Gefährdungspotenzial vorliegt. Damit fehlen auch Informationen zur Dosis-Wirkungs-Beziehung; ebenso lassen sich kaum Angaben zu kritischen Expositionen machen. Eine Risikocharakterisierung lässt sich nur noch qualitativ vornehmen. *Unbekannte* Risiken schließlich sind solche, für die es keine konkreten Anhaltspunkte gibt.

Prinzipiell bieten sich zwei Möglichkeiten an, mit Abschätzungsunsicherheiten umzugehen: Entweder wird – wie es etwa der WBGU (1999) getan hat – die Abschätzungsunsicherheit als ein Attribut bei der vergleichenden Risikobewertung aufgenommen.⁸¹ Oder es werden Risikoklassen unterschieden und für diese dann getrennt Risikovergleiche vorgenommen.

Unserer Auffassung nach ist es jedoch nicht immer zielführend, die Abschätzungsunsicherheit als ein Attribut unter anderen in Risikovergleiche aufzunehmen und so zu versuchen, das Wissensproblem zu berücksichtigen. Denn es bleibt offen, wie dieses „Unsicherheits-Attribut“ skaliert und gewichtet werden soll. Zwar ist es möglich, hier Abstufungen anzugeben (wenig, mittel, viel Wissen, siehe auch Florig et al. 2001) und diese Wissensklassen zu gewichten. Es bleibt aber ein wesentliches Problem: Wie kann das Nichtwissen (geht von X eine Gefährdung aus?) mit anderen Attributen (z. B. Anzahl der Betroffenen) verrechnet werden? Beispielsweise ist der Vergleich von Radon und EMF so kaum möglich. Bei Radon ist das Gefährdungspotenzial bekannt, die Dosis-Wirkungs-Beziehung ist zwar nicht unumstritten, aber man geht von einer linearen Beziehung ohne Schwellenwert aus. Das kritische Wissensdefizit liegt hier allein bei der Exposition. Dagegen ist bei hochfrequenten EMF (bei Expositionen unterhalb der Grenzwerte) unklar, ob überhaupt ein Gefährdungspotenzial existiert (vgl. SSK 2001). Ein Risikovergleich scheitert daran, dass im Radonfall zwar abgeschätzt werden kann, mit wie vielen Krebsfällen pro Jahr zu rechnen ist, bei EMF bleibt dies völlig unklar. Dass bei dem Risikovergleich die unbestreitbare Ubiquität von EMF-Expositionen keine Bedeutung haben kann, liegt auf der Hand. Aus der Tatsache, dass EMF überall vorhanden sind, folgt solange nichts für die Abschätzung des Risikos, wie unklar ist, ob EMF überhaupt ein Gefährdungspotenzial besitzt.

Risikovergleiche über unterschiedliche Wissensklassen werfen erhebliche Schwierigkeiten auf. Nach unserer Auffassung sollten Risiken unterschiedlicher Wissensklassen (bekanntes versus undeutliches Gefährdungspotenzial) nicht verglichen werden. Anders als bei bekannten Risiken kommt es in der Bewertung bei undeutlichen Risiken (BUR) zunächst darauf an, eine systematische und transparente vergleichende Evidenzbewertung durchzuführen, das heißt festzustellen, wie die wissenschaftliche Evidenzlage ist und ggf. zu verdeutlichen, worin unterschiedliche wissenschaftliche Einschätzungen begründet sind. Die wesentliche Herausforderung für ein BUR liegt in der praktischen Organisation und Durchführung (siehe Anhang 3).

Faktizität und Hypothesizität

Technische Risiken werden oftmals über Schadens-Häufigkeits-Diagramme (auch als Farmer-Diagramme bekannt) angegeben, die aussagen, welche Schäden mit

⁸¹ Siehe hierzu das Beispiel in Anhang 2.

welcher statistischen Häufigkeit zu erwarten sind. Risikovergleiche zwischen verschiedenen technologischen Optionen (zum Beispiel Kernkraft versus fossile Brennstoffe) stehen dann vor dem Problem, wie sich Risiken vergleichen lassen, die einerseits durch eine Kombination von sehr kleinen Häufigkeiten mit sehr großen Schäden und andererseits durch relativ große Häufigkeiten bei kleineren Schäden entstehen. Anders formuliert stellt sich die Frage, wie Katastrophenpotenziale zu gewichten sind.⁸² Eine ganz andere Sache ist es, wenn Risiken von biologischen Endpunkten her gedacht werden.

Der Unterschied zwischen diesen beiden – den technischen und den biologischen – Risikokonzepten wird relevant, wenn Risiken aufgrund des bestimmungsgemäßen, aber mit geringen Emissionen verbundenen Betriebs von Anlagen („Normalbetriebs-Risiken“) mit Störfallrisiken verglichen werden.

Technikbezogene und biologische Risiken lassen sich nur schwer vergleichen, da sie von ganz verschiedenen Expositionsbegriffen ausgehen (vgl. Uth 1990). Bei technischen Risiken beziehen sich die Störfall-Wahrscheinlichkeiten auf die Möglichkeit der Exposition mit einer Noxe. Der Zusammenhang zwischen Exposition mit der Noxe und dem resultierenden Schaden kann dann entweder deterministisch oder stochastisch oder aber ungewiss sein. Dagegen beziehen sich biologische Risiken von vornherein auf eine vorhandene Exposition. Sie beschreiben die Wahrscheinlichkeit von adversen Wirkungen bei einer Exposition mit der Noxe. Im Unterschied zu den Störfallrisiken, bei denen lediglich eine potenzielle Exposition besteht, liegen hier reale Expositionen vor.

Es gibt also Risiken, die zwar nach logischen Maßstäben möglich sind, aber aller Wahrscheinlichkeit nach nie Realität werden oder dies mit nur einer extrem geringen Wahrscheinlichkeit. Typische Beispiele sind sehr gravierende Störfälle in einer technischen Anlage, die mit Hilfe probabilistischer Risikoanalysen identifiziert wurden und deren so ermittelte Eintrittswahrscheinlichkeit sehr klein ist. In Deutschland wurde ein derartiges Verfahren nur im Fall der Kernenergie praktiziert; hier hat sich für das entsprechende Risiko der Begriff „Restrisiko“ eingebürgert. Auch große Anlagen aus dem Bereich der Chemie und Staudämme gehören aufgrund ihres Gefahrenpotenzials zu den technischen Anlagen, bei denen entsprechende Störfälle mit großer Zuverlässigkeit verhindert werden müssen, und für die dann entsprechend geringe Störfallwahrscheinlichkeiten bestehen. Extrem geringe Wahrscheinlichkeiten ergeben sich auch für Risiken aufgrund extrem seltener Naturkatastrophen, etwa Erdbeben von einer Stärke, die in Deutschland geologisch gesehen denkbar, aber historisch nicht belegbar ist.

Bei technischen Anlagen, für die derartige Risiken ermittelt werden, sollen Maßnahmen (etwa konstruktive Gestaltung oder Sicherheitssysteme) gewährleisten, dass bestimmte Ereignisse gar nicht erst eintreten. Dann aber sind „alle Schadensüberlegungen, alle Berechnungen und alle faktisch-technischen Maßnahmen zur Schadensverhütung nicht mehr im Bereich der Faktizität, sondern im Bereich der Hypothetizität zu suchen“ (Häfele 1990, 33).

⁸² Dabei wäre konsequenterweise vorauszusetzen, dass bei fossilen Energieträgern ebenfalls eine Störfallanalyse über hypothetische *worst case* Fälle vorzunehmen ist.

Den Unterschied zwischen faktischen und hypothetischen Risiken kann man sich an folgendem Gedankenexperiment klar machen. Zwar ist es für ein Individuum gleichgültig, ob es während seines Lebens bei einem mit einer Wahrscheinlichkeit von 1:1 Million auftretenden Störfall an Krebs erkrankt oder ob es mit der gleichen Wahrscheinlichkeit durch eine ständig vorhandene Exposition eine Krebserkrankung bekommt. Auf der Ebene des kollektiven Risikos sieht die Sache aber ganz anders aus: Im Fall der gegebenen Exposition ist bei einem Individualrisiko von 1:1 Million in Deutschland mit 80 wirklichen Toten zu rechnen. Im Falle des Störfallrisikos kann keine derartige Aussage über real auftretende Folgen gemacht werden.⁸³

Wir schlagen daher vor, Störfall-Risiken *nicht* gegen Risiken zu verrechnen, die im Bereich des Faktischen liegen. Im ersten Fall wird nach der Realisierungsmöglichkeit eines schadenverursachenden Ereignisses, im zweiten Fall nach der Wirkung einer vorliegenden Noxe gefragt. Daraus folgt, dass bei Risikovergleichen die Vermengung von Expositionsumständen (potenziell versus real) vermieden werden sollte.

6.2 Ziele für eine vergleichende Bewertung von Risiken

Es gibt eine Reihe von Möglichkeiten für den Vergleich von Risiken. Dabei ist die klassische CRA nur ein Spezialfall von vergleichenden Analysen. Geht man von den Phasen der Risikoregulation (vgl. Risikokommission 2003) – vom Vorverfahren bis hin zum Risikomanagement – aus, so finden sich verschiedene sinnvolle Varianten. Tabelle 23 gibt dazu eine Übersicht.

6.2.1 Vorverfahren

In der Vorphase sind drei risikobezogene Vergleiche sinnvoll. Zum einen das Scanning von neuen Technikfeldern, um neue Risikopotenziale zu entdecken. Zum anderen geht es auch um vergleichende Studien zur Risikowahrnehmung und schließlich um die Priorisierung von Stoffen oder anderen Risikoquellen, die einer Risikobewertung zugeführt werden sollen.

Das *Scanning von neuen Technologiefeldern* für die Früherkennung von Risikopotenzialen basiert vorwiegend auf der Anwendung qualitativer Methoden, insbesondere der Szenariotechnik. Solche Szenarien müssen an dem Gerüst des *Risk Assessment* ausgerichtet werden. Leitfragen sind:

- Welche neuen Technologien, Werkstoffe und Produkte etc. kommen auf uns zu? Wie werden sich diese mengenmäßig entwickeln?
- Auf welche Aktivitätsfelder wirken sie sich aus?
- Welche potenziellen Stressoren/Risikofelder können damit verbunden sein?
- Welche Bevölkerungsgruppen werden diesen Stressoren ausgesetzt sein?

Das Kernproblem ist hier die Signalerkennung. Für diesen schwierigen Schritt stehen sowohl Kreativitätstechniken als auch Forecasting-Methoden zur Verfügung. Kreati-

⁸³ Dies mindert die Bedeutung des Störfallrisikos keinesfalls, zeigt aber, dass es getrennt behandelt werden muss.

vitätstechniken unterstützen das Denken außerhalb gewohnter Bahnen. Morphologische Methoden sind beispielsweise systematisch-analytische Verfahren, welche die Ideenfindung durch geordnetes Denken strukturieren. Das explorative Verfahren des Brainstorming hingegen lebt vom offenen Aussprechen der Gedanken und dem Aufgreifen und Weiterentwickeln der Ideen anderer Teilnehmer. Einen weiteren Zugang bieten Delphi-Verfahren. Der Kern dieses Verfahrens besteht darin, in mehreren Runden Expertenmeinungen, z.B. zur zukünftigen Technologieentwicklung, einzuholen. Anhand eines formalisierten Fragebogens beurteilt eine in der Regel große Anzahl von Experten bestimmte Thesen. In einer (oder mehreren) Feedback-Runden erhalten die Experten danach die Möglichkeit, vor dem Hintergrund der Einschätzungen der Fachkollegen die eigenen Urteile zu überdenken und ggf. zu modifizieren.

Zur vergleichenden Bewertung potenzieller zukünftiger Risiken braucht man Methoden, die eine möglichst transparente und nachvollziehbare Gewichtung der einzelnen Bewertungskriterien, wie z.B. der Neuheit oder der räumlichen Ausdehnung eines zukünftigen Risikofeldes, ermöglichen. Hierzu kann der in Kapitel 5.2 beschriebene Ansatz der multiattributen Bewertung herangezogen werden.

Um eine „dichte“ Beschreibung der als bedeutsam bewerteten möglichen Risikofelder zu erhalten, kann die Szenariomethode eingesetzt werden. Hier ist die Entwicklung mehrerer Zukunftsalternativen das Ziel. Ein Szenario ist ein Satz von Ausprägungen verschiedener Einflussfaktoren, der eine zukünftige Situation sowie die Ereignisse auf dem Weg dorthin und deren zeitlichen Ablauf beschreibt. Durch unterschiedliche Annahmen bezüglich der Richtung und Stärke, in der diese Faktoren wirken könnten, und durch konsistente Kombinationen solcher Annahmen für verschiedene Einflussfaktoren können dann unterschiedliche Szenarien entwickelt werden. Denn in der Einflussanalyse werden Einflussbereiche identifiziert, auf die Einflussbereiche wirkende Faktoren ermittelt und bewertet sowie die Vernetzung zwischen den Einflussbereichen erarbeitet. Im Ergebnis werden damit die treibenden Faktoren herausgearbeitet. Eine Beteiligung von Stakeholdern im Sinne von Information und Konsultation ist hier sinnvoll.

Für die *Abschätzung des Mobilisierungspotenzials von neuen Technologien* können vergleichende Studien zur *Risikowahrnehmung* hilfreich sein. Solche Untersuchungen gibt es seit den 70er Jahren des letzten Jahrhunderts. Solche Vergleiche sind mehrdimensional aufgebaut. Dabei beurteilen Personen zahlreiche Risikoquellen (z.B. in der Nähe einer Chemieanlage wohnen, Radfahren und Kernkraft) in Bezug auf ihre Riskantheit. Darüber hinaus werden diese Risikoquellen auf verschiedenen 'qualitativen' Dimensionen beurteilt, von denen man annimmt, dass sie für die Beurteilung von Risiken von Bedeutung sind. Zu diesen Dimensionen gehören beispielsweise die Schwere oder Ernsthaftigkeit des Risikos, die Freiwilligkeit des Eingehens eines Risikos, die Kontrollierbarkeit und die Bekanntheit des Risikos oder das Katastrophenpotenzial, das mit einer Risikoquelle verbunden ist. In der statistischen Analyse werden die so gewonnenen Daten für jede Risikoquelle und Beurteilungsdimension über die Befragten aggregiert, so dass für jede Risikoquelle und Beurteilungsdimension jeweils ein Mittelwert resultiert. Diese Daten werden dann mit Hilfe multivariater Analyseverfahren wie Faktoren- oder Regressionsanalyse ausgewertet. Solche Studien haben gezeigt, dass die Risikowahrnehmung u.a. von folgenden Faktoren abhängt (vgl. National Research Council, 1996, 147, Wiedemann et. al. 2000):

- Das Risiko ist auf den ersten Blick verständlich und anschlussfähig an bestehende Überzeugungen.

- Es weist dramatische Qualitäten auf: Die Noxe ist charakterisiert durch einen hohen „Dread“-Faktor (sehr stark angst- und empörungsauslösend) sowie durch eine unfreiwillige Exposition einer großen Anzahl von Menschen.
- Die Betroffenen sind identifizierbar und „Täter“ können rasch gefunden werden.
- Die gesellschaftliche Risiko-Nutzen-Verteilung wird als ungerecht erlebt.
- Strukturähnliche Risiken haben in der Vergangenheit zu Kontroversen geführt.
- Den verantwortlichen Institutionen wird nicht vertraut bzw. ihnen fehlt die gesellschaftliche Legitimation.

Bei vergleichenden Risikowahrnehmungsstudien stehen definitionsgemäß die Sichtweisen der Betroffenen (Stakeholder, Öffentlichkeit) im Mittelpunkt.

Zuweilen sind aufgrund einer Vielzahl von regelungsbedürftigen Stoffen besondere Dringlichkeitsfälle zu bestimmen, d.h. diejenigen Fälle, für die prioritär eine Risikoabschätzung durchzuführen ist (siehe Risikokommission 2003). Solche *Dringlichkeitsbewertungen* basieren ebenfalls auf Vergleichen. Hier sind in der Regel mehrere Merkmale zu bewerten: Emissionsdaten, Expositionscharakteristika, Hinweise auf Toxizität. Ein Beispiel ist der DYNAMEC Ansatz der OSPAR Kommission (2002). Hier werden hauptsächlich auf der Basis von PBT-Kriterien⁸⁴ Einstufungen in die Liste der „*Substances of Possible Concern*“ vorgenommen. Welche Merkmale auch immer in die Abschätzung einbezogen werden – methodisch bietet auch für diese Art der vergleichenden Risikobewertung der in Kapitel 5.2 beschriebene Ansatz der multiattributen Bewertung eine Grundlage. Eine Beteiligung der Öffentlichkeit erfolgt hier im Regelfall durch Information und Konsultation.

⁸⁴ PBT = Persistence, Bioaccumulation, Toxicity

Tabelle 23: Möglichkeiten vergleichender Analysen

	Ziel	Fokus des Vergleichs	Modell	Beteiligung
Voruntersuchung	Abschätzung von neuen Risikopotenzialen	Vergleich neuer Technologiefelder	Szenario-Technik und MAUT	Information, Konsultation
	Abschätzung des Mobilisierungspotenzials von neuen Technologien	Vergleich von Risikowahrnehmungen	Psychometrischer Ansatz	Teilhabe
	Bewertung der Dringlichkeit einer Risikobewertung	Vergleich von Stoffen	MAUT	Information, Konsultation
Risikoabschätzung	Abschätzung der Evidenzen für ein Risiko	Vergleich unterschiedlicher Evidenzbewertungen	Evidence based medicine	Information
	Bewertung der Evidenz für ein Risiko	Einstufung der Evidenzen in eine Evidenzbewertungsskala	Konsensus Modell	Information
	Bewertung der Potenz eines Schädigungspotenzials	Vergleich unterschiedlicher Stoffe in Bezug auf ihre Potenz, z.B. nach endokrin modulierender Wirkung	Ranking	Information
Risikobewertung	Bewertung der Exposition mit einer Noxe	Vergleich von Expositionsquellen	Ranking	Information, Konsultation
	Bewertung der Vulnerabilität von Populationen	Kinder versus Erwachsene	Vergleich von Risikoabschätzungen	Information, Konsultation
	Bewertung von Risiken	Programmatische Risikovergleiche bis hin zu weitgehend homogenen Risikolisten	MAUT	Teilhabe
Risikomanagement	Vergleich von Technologien	Kosten-Nutzen Vergleiche zur Auswahl von Managementoptionen, Risk Tradeoff Analysen Vergleiche des Mobilisierungspotenzials, Standorte	MAUT, Kosten-Nutzen-Analyse	Information, Konsultation

6.2.2 Risikoabschätzung

In der Phase der Risikoabschätzung geht es um die Bewertung der Evidenzen, die für eine adverse Wirkung eines Agens sprechen. Hier kann sich das Problem ergeben, dass verschiedene Experten(gruppen) zu unterschiedlichen Risikoabschätzungen kommen. Dies ist in der Wissenschaft zwar nichts ungewöhnliches, sorgt jedoch in der Öffentlichkeit häufig für Verunsicherung. Da sich in der Praxis an einem solchen Expertendissens häufig Risikokonflikte⁸⁵ entzünden, ist der Vergleich von unterschiedlichen Evidenzbewertungen eine essenzielle Variante der CRA.

Grundsätzlich können unterschiedliche Experteneinschätzungen bei allen Schritten des Prozesses der Risikoabschätzung vorliegen: bei Bestimmung des Gefährdungspotenzials, bei der Abschätzung der Dosis-Wirkungs-Beziehung und der Exposition sowie bei der Risikocharakterisierung. Besondere Resonanz in der Öffentlichkeit findet aber die Frage, ob überhaupt ein Gefährdungspotenzial gegeben ist. Es erscheint deshalb ratsam, Unterschiede und Gemeinsamkeiten in den Experteneinschätzungen mit Hilfe eines transparenten Verfahrens in einem systematischen Vergleich solcher Risikobewertungen zu verdeutlichen. Erste Vorschläge für die Durchführung und auch praktische Umsetzungen eines solchen Vergleichsverfahrens sind vom *California EMF Project* (Neutra, DelPizzo & Lee 2002) sowie von Wiedemann, Schütz & Thalmann (2002) vorgelegt worden.

Es kommt hier darauf an, die divergierenden Experten an einem Tisch zu versammeln und in nachvollziehbarer Weise die Pro- und Kontra-Argumente für das Vorliegen eines Schädigungspotenzials herauszuarbeiten (siehe dazu das Beispiel in Anhang 3). Diese argumentative Grundstruktur kann als Ausgangspunkt für ein Verfahren zur Verdeutlichung der Unterschiede in der Evidenzbewertung zwischen verschiedenen Experten bzw. Expertengruppen dienen. Ein solches Verfahren umfasst drei wesentliche Schritte, welche die teilnehmenden Experten bzw. Expertengruppen bearbeiten müssen:

- (1) die Beurteilung der Befundlage in Bezug auf Qualität der vorliegenden wissenschaftlichen Studien;
- (2) darauf aufbauend, die Formulierung von Pro- bzw. Kontra-Argumenten in Bezug auf das Vorliegen eines Kausalzusammenhangs;
- (3) die abschließende (qualitative oder quantitative) Evidenzbewertung.

Für die Auswahl und Bewertung wissenschaftlicher Evidenz gibt es eine Reihe von Leitlinien.⁸⁶ So liefert beispielsweise das *Cochrane Reviewers Handbook* für den medizinischen Bereich – speziell für randomisierte, kontrollierte Studien – sehr detaillierte Informationen zu allen Facetten einer systematischen Literatursuche und

⁸⁵ Tatsächlich wird man sagen können, dass eine wesentliche Ursache aller gesellschaftlichen Konflikte über Risiken die unterschiedlichen Risikoeinschätzungen von Experten sind (vgl. Mazur 1981; Peters 1993).

⁸⁶ Für einen allgemeinen Überblick siehe: Cooper, H. & Hedges, L.V. (eds.) (1994): *The handbook of research synthesis*. New York: Russell Sage Foundation, Part III.

-bewertung.⁸⁷ Ähnlich detailliert ist das Handbuch des englischen NHS Centre for Reviews and Dissemination.

Die Möglichkeit zur Beteiligung von Stakeholdern und Öffentlichkeit liegt bei dieser Evidenzbewertung in der Bewertung der Überzeugungskraft von Pro- und Kontra-Argumenten.

Eine andere Variante des Evidenzvergleichs sind Evidenzbewertungsskalen wie sie beispielsweise von der IARC vorliegen. Dazu zählt auch das Klassifikationsschema der deutschen Strahlenschutzkommission, die bei ihrer Beurteilung der Evidenzlage zu gesundheitlichen Effekten hoch- und niederfrequenter elektromagnetischer Felder drei Evidenzkategorien unterscheidet: *wissenschaftlicher Nachweis*, *wissenschaftlich begründeter Verdacht* und *wissenschaftlicher Hinweis* (SSK 2001, 7). Diese sind wie folgt definiert:

- Wissenschaftlich nachgewiesen ist ein Zusammenhang zwischen einer Gesundheitsbeeinträchtigung und elektromagnetischen Feldern, wenn wissenschaftliche Studien voneinander unabhängiger Forschungsgruppen diesen Zusammenhang reproduzierbar zeigen und das wissenschaftliche Gesamtbild das Vorliegen eines kausalen Zusammenhangs stützt.
- Ein wissenschaftlich begründeter Verdacht auf einen Zusammenhang zwischen einer Gesundheitsbeeinträchtigung und elektromagnetischen Feldern liegt vor, wenn die Ergebnisse bestätigter wissenschaftlicher Untersuchungen einen Zusammenhang zeigen, aber die Gesamtheit der wissenschaftlichen Untersuchungen das Vorliegen eines kausalen Zusammenhangs nicht ausreichend stützt. Das Ausmaß des wissenschaftlichen Verdachts richtet sich nach der Anzahl und der Konsistenz der vorliegenden wissenschaftlichen Arbeiten.
- Wissenschaftliche Hinweise liegen vor, wenn einzelne Untersuchungen, die auf einen Zusammenhang zwischen einer Gesundheitsbeeinträchtigung und elektromagnetischen Feldern hinweisen, nicht durch voneinander unabhängige Untersuchungen bestätigt sind und durch das wissenschaftliche Gesamtbild nicht gestützt werden.

Für eine vergleichende Evidenzbewertung sind solche Kategorien hilfreich, da sie schon eine ordinale Metrik bieten, anhand derer die Evidenz für verschiedene Noxen und/oder Endpunkte verglichen werden kann. Für eine große Zahl von Noxen liegt die Evidenzklassifikation bezüglich bestimmter Endpunkte (vor allem Kanzerogenität) mittlerweile schon in Form öffentlich (über das Internet) zugänglicher Datenbanken vor, so dass solche Rankings ohne großen Rechercheaufwand abgerufen bzw. vorgenommen werden können. Hier sind vor allem die Datenbank der IARC⁸⁸, das IRIS-System⁸⁹ der EPA und das Scorecard-System⁹⁰ der Environmental Defense Organisation zu nennen. Auch für diesen Fall der Evidenzskalierung ist die Beteiligung der Öffentlichkeit wohl am besten über Information und Konsultation vorzunehmen.

⁸⁷ Clarke, M. & Oxman, A.D. (eds.) (2002): Cochrane Reviewers Handbook 4.1.5 [updated April 2002]. In: The Cochrane Library, Issue 2, 2002. Oxford: Update Software. Updated quarterly. Im Internet unter: Cochrane Collaboration: <http://www.cochrane.org/cochrane/resource.htm>

⁸⁸ <http://www.iarc.fr/>

⁸⁹ <http://www.epa.gov/iris/index.html>

⁹⁰ <http://www.scorecard.org/>

6.2.3 Risikobewertung

In der Phase der Risikobewertung sind vier verschiedene Möglichkeiten für Vergleiche gegeben: (1) die Bewertung der Potenz einer Noxe, (2) die Bewertung der Exposition mit einer Noxe, (3) die Bewertung der Vulnerabilität von Populationen und (4) die vergleichende Bewertung verschiedener Risiken.

Die vergleichende Bewertung der Potenz einer Noxe stellt sich u.a. bei der Bewertung der endokrin modulierenden Wirkung (vgl. Gaido et al. 1997). Der Vergleich von Expositionsquellen ist beispielsweise für die Auswahl von Minimierungsmaßnahmen wichtig; etwa wenn es darum geht, diejenigen Quellen auszuschalten oder zu begrenzen, die den größten Beitrag zur Exposition leisten. Schließlich ist der Vergleich der Vulnerabilität von Teilpopulationen bei der Festlegung von Vorsorgemaßnahmen von Bedeutung. Hier interessiert die Frage, ob es solche Gruppen gibt und wieviel sensibler als die „Normalpopulation“ sie auf eine Noxe reagieren. Diese Varianten vergleichender Risikobewertung erfordern spezielle Expertisen und bieten nur eingeschränkt Platz für eine Öffentlichkeitsbeteiligung.

Die zentrale Variante der vergleichenden Risikobewertung, die unter Beteiligung von Laien erfolgt, stellt jedoch die bereits in Kapitel 3.1.1 bzw. 3.2.1 beschriebene programmatische vergleichende Risikobewertung dar. Zwar gibt es in Deutschland dazu noch keine Praxis, es ist aber durchaus vorstellbar, eine solche CRA im Rahmen des Aktionsprogramms „Umwelt und Gesundheit“ oder ähnlicher Programme durchzuführen. In Anhang 2 finden sich dazu weitere Hinweise.

6.2.4 Risikomanagement

Vergleichende Risikobewertungen beziehen sich hier auf die Handlungsoptionen, die für die Bewältigung eines Risikos bestehen. Typischerweise geht es dabei entweder um die Frage, welche Option im Hinblick auf einen gewünschten Nutzen die beste ist, oder darum, welche Option zur Risikoreduzierung unter dem Gesichtspunkt der Kosteneffizienz die günstigste ist. Beispiele für den ersten Fall sind etwa die Auswahl von geeigneten Standorten für eine Mülldeponie oder die vergleichende Bewertung von Energieversorgungssystemen (siehe oben Kapitel 3.2.4 und 3.2.5). Beispiele für den zweiten Fall sind die Bewertungen von Maßnahmen zur Reduzierung von Gesundheitsrisiken im Hinblick auf die geschätzten Kosten, die sie pro gerettetes Leben verursachen (z.B. Ramsberg 1997; Tengs et al. 1995).⁹¹ Wie bereits oben in Kapitel 2.7 ausgeführt, muss in beiden Fällen beachtet werden, dass die Vermeidung bzw. Verminderung von Risiken selbst wieder Risiken zur Folge haben kann. Etwaige Risikoverschiebungen sollten deshalb bei der Beurteilung und Auswahl von Risikomanagement-Optionen im Rahmen einer Risk-Tradeoff-Analyse berücksichtigt werden (Graham & Wiener 1997; Presidential Commission 1997a, 35).

Diese Form vergleichender Risikobewertungen muss neben den Risikoabschätzungen auch Nutzenabschätzungen miteinbeziehen. Hinzu kommt, dass für die Entscheidung über Risikomanagementoptionen unter Umständen auch weitere Kriterien

⁹¹ Der Katalog der untersuchten Optionen reicht dabei von Maßnahmen zur Erhöhung der Verkehrssicherheit über Maßnahmen zur Kontrolle von Schadstoffemissionen bis hin zu medizinischen Therapien und Vorsorgemaßnahmen.

herangezogen werden, die jenseits quantifizierbarer Risiko- und Nutzenabschätzungen liegen. Neus et al. (1995, 261) nennen in diesem Zusammenhang:

- die Berücksichtigung zeitlicher Trends (vor allem im Hinblick auf eine mögliche Verschärfung von Problemlagen);
- die Einbeziehung von Leitideen, wie zum Beispiel der Verwirklichung einer nachhaltigen Entwicklung;⁹²
- die Überprüfung der sozialen Ausgewogenheit von Handlungsoptionen;⁹²
- die Prüfung rechtlicher bzw. institutioneller Handlungsspielräume.

Dies zeigt, dass die Anwendung vergleichender Risikobewertungen im Rahmen des Risikomanagements weitergehende methodische (z.B. Einbeziehung von Kosten-Nutzen-Analysen) und sogar gesundheits- bzw. umweltpolitische Probleme aufwirft. Sie geht damit deutlich über den für diesen Bericht (und das Projekt) gesetzten Rahmen hinaus.

6.3 Beteiligungsmodelle

Die Partizipation der Öffentlichkeit kann durch verschiedene Modelle umgesetzt werden.⁹³ Zwei Grundvarianten sind zu unterscheiden: die Partizipation organisierter Interessengruppen oder die von Bürgern und Bürgerinnen, die keinen organisierten Gruppen angehören.⁹⁴ Vorstellbar ist auch ein Verfahren, das beide Varianten miteinander kombiniert. In der nachstehenden Abbildung 18 werden fünf Stufen der Beteiligung unterschieden.

Die Stufen ergeben sich aus dem Umfang der Informationsbereitschaft und der Bereitschaft, auf Anliegen der Öffentlichkeit einzugehen sowie dem Ausmaß, in dem die Bürger und Bürgerinnen in die CRA einbezogen werden.

Die **erste Stufe** der Beteiligung umfasst den Zugang zu Risiko-Informationen der CRA. Hier wird nicht nur das Ergebnis der CRA veröffentlicht, sondern auch alle Daten und Gutachten, auf denen die Prozesse der Risikoabschätzung und -bewertung aufsetzen.

Die **zweite Stufe** der Beteiligung wird erreicht, wenn alle Informationen der CRA für die Bürger und Bürgerinnen entsprechend aufbereitet sind, so dass die Kriterien Verständlichkeit, Klarheit und Prägnanz erfüllt werden. Insbesondere geht es um transparente und nachvollziehbare Information (z.B. Zusammenfassungen der Risikoinformationen im Frage & Antwort-Format), um das Bereitstellen von Bewertungshilfen und Hintergrundinformationen sowie um Hinweise auf weiterführende Informationen.

⁹² Für die Einbeziehung dieses Aspekts gibt das in Kapitel 3.2.5 dargestellte GaBE-Projekt ein Beispiel.

⁹³ Siehe dazu CMA (1988) Community Awareness Workbook, Edison Electric Institute (1994); Public Participation Manual, Canada Standards Association (1996) A Guide to Public Involvement.

⁹⁴ Ein Beispiel dafür ist der Planungszellen-Ansatz von Dienel (2002), siehe auch <http://www.planungszelle.de/>

Auf der **dritten Stufe** der Beteiligung geht es um eine aktive Verbreitung von Informationen über die CRA. Das CRA-Projektmanagement wartet also nicht auf die Nachfrage der Bürgerinnen und Bürger, sondern informiert aus eigenem Anliegen heraus. Im Mittelpunkt stehen: (1) Frühzeitige Information über Risikoabschätzungen und -bewertungen, (2) Definieren der Zielgruppen der Risikokommunikation, (3) Auswahl geeigneter Kommunikationskanäle und -mittel für die Zielgruppen, (4) Bekanntmachen der Kommunikationskanäle in der Öffentlichkeit, (5) Test, ob die Kommunikation auch „ankommt“.

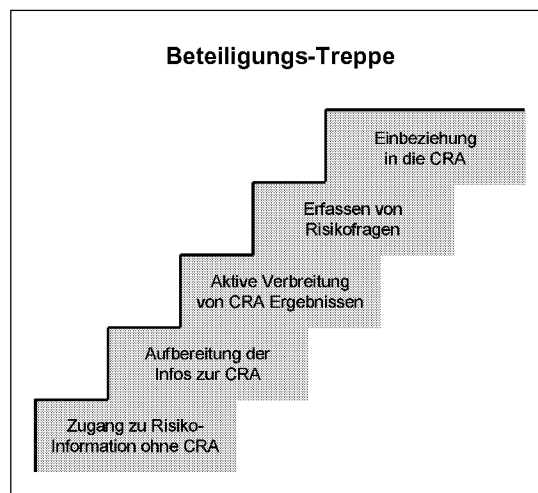


Abbildung 18: Stufen der Beteiligung

Die **vierte Stufe** der Beteiligung ist erreicht, wenn das CRA-Projektmanagement – z.B. via Fokusgruppen oder Umfragen – die Anliegen der Bürgerinnen und Bürger erfasst und diese bei der Planung und Durchführung der CRA berücksichtigt.

Für die **fünfte Stufe** der Beteiligung sind Konsultationen mit Betroffenen charakteristisch. Das CRA-Projektmanagement bezieht Bürgerinnen und Bürger in die vergleichende Risikobewertung ein.

6.4 Akteure der CRA unter Beteiligung der Öffentlichkeit

Der Planung der Beteiligung geht eine systematische Stakeholder-Analyse voraus. Dazu ist zu ermitteln:

- Wer wird von der CRA betroffen sein?
- Welche Bereiche / Sektoren (z.B. Wirtschaft, Umweltverbände, Verbraucherschutz), aber auch welche geographischen Regionen werden von wem vertreten?
- Gibt es dominante Risikokonflikte?
- Gibt es in bestimmten Regionen und/oder zu bestimmten Problemfeldern Bürgerinitiativen?

- Welche Interessen vertreten die verschiedenen Stakeholder? Welche Positionen haben sie zu verschiedenen Risikoproblemen, und wie stehen sie zu einer CRA?
- Wie gut sind die Stakeholder organisiert? Welche Unterstützung brauchen sie?
- Welche Kosten entstehen?

CRA-Projektmanagement

Das CRA-Projektmanagement ist für die Planung und Durchführung der CRA verantwortlich. Zu seinen Aufgaben gehören die Präzisierung der Zielstellung der CRA, die Entwicklung eines ersten Rahmengerüsts für die CRA sowie die Einrichtung einer Lenkungsgruppe.

Das Projektmanagement soll das Verfahren in Gang bringen und erfolgreich abschließen. Alle Planungstätigkeiten zur erfolgreichen Durchführung der CRA werden von ihm koordiniert. Es hat die Aufgabe, Informations-, Wissens- und andere Ungleichgewichte auszubalancieren. Dabei muss es in Rechnung stellen, dass neben den Wissensvoraussetzungen auch die Interessen der Beteiligten verschieden sind und dass unterschiedliche Erfahrungszusammenhänge und Erwartungen die CRA beeinflussen können. Weiterhin obliegt ihm die Sicherung der Qualität des Experteninputs für die Risikoabschätzungen. Es ist auch für die projektbezogene Öffentlichkeitsarbeit verantwortlich.

Lenkungsgruppe

Die Lenkungsgruppe besteht aus Mitgliedern der mit der CRA beauftragten Organisation (z.B. der Bundesoberbehörden) und anderen Vertretern staatlicher und nicht-staatlicher Einrichtungen. Es empfiehlt sich, bereits hier auf eine Repräsentanz, zumindest aber die Informierung entsprechender gesellschaftlicher Gruppen zu achten. Daneben können auch Fachleute herangezogen werden, die Spezialkenntnisse in Risikokommunikation und bezüglich der Organisation von Beteiligungsverfahren besitzen. Die Gruppe dient der Unterstützung des Projektmanagements. In der Regel ist sie während der Vorbereitungsphase besonders aktiv und berät bei der Entwicklung des Rahmenkonzepts der CRA.

Die Lenkungsgruppe bietet eine praktikable Möglichkeit, den Sachverstand verschiedener Institutionen zu bündeln und in die Planung sowie Durchführung der CRA einfließen zu lassen. Solche Lenkungsgruppen haben sich schon in anderen Projekten bewährt.⁹⁵

Bürgerbeirat

Es ist Aufgabe des Bürgerbeirats, auf der Basis der Expertenbewertungen der Risiken das Risiko-Ranking zu erstellen bzw. dazu Empfehlungen zu geben. In Fällen, in denen die Teilnahme der Öffentlichkeit besonders kritisch ist, wäre zu überlegen, wie

⁹⁵ Siehe Wiedemann et al. (2002) Machbarkeitsstudie zur Risikofrüherkennung im Bereich Umwelt und Gesundheit; sowie Renn et al. (2003).

durch öffentliche Anhörungen zusätzliche Inputs für die CRA gegeben werden können, beispielsweise bei der Auswahl der Problemlisten für das Ranking.

Wer am Bürgerbeirat beteiligt werden sollte, muss jeweils immer wieder neu überlegt werden. Dafür gibt es kein Rezept. Es ist Verhandlungssache, die von den Gegebenheiten vor Ort und den Zielstellungen des Verfahrens abhängt. Neben gesellschaftlichen Interessengruppen (Umweltverbände, Patientengruppen, Verbraucherschutz, Industrie) ist jedoch auch an betroffene Bürger zu denken. Bei der Auswahl der Teilnehmer ist darauf zu achten:

- Sind alle wichtigen Gruppierungen berücksichtigt?
- Welche Bedingungen stellen die Parteien an das Verfahren?
- Lässt sich eine Gesprächsbereitschaft erkennen?
- Gibt es für die einzelnen Parteien Gründe, die CRA zu blockieren?
- Durch welche Anreize kann die Motivation zur Mitarbeit an der CRA verbessert werden?

Expertenteams

Die Expertenteams stellen die Daten für die Risikoabschätzung zusammen, bewerten die Risiken und können auch ein erstes Ranking vornehmen, das später vom Bürgerbeirat erörtert und modifiziert wird. Sie haben in Anhörungen die ihren Bewertungen zugrundeliegenden Annahmen und die verbleibenden Unsicherheiten zu erläutern. Die Experten kommen aus den Behörden, aus privaten Instituten oder von Universitäten und anderen Forschungseinrichtungen.

Je nachdem, welche Risiken in welcher Hinsicht bewertet werden sollten, sind unterschiedliche Kenntnisse gefragt.

Adressaten

Adressaten der CRA sind alle interessierten Parteien. Es empfiehlt sich deshalb, eine Zusammenstellung aller Gruppen zu machen, die über das Verfahren zu informieren sind. Dazu gehören potenzielle Teilnehmer, interessierte Experten, Nichtregierungsorganisationen, die Politik und alle Gruppen, die von den Ergebnissen der CRA unmittelbar betroffen sind.

Zur Sicherstellung eines erfolgreichen Starts der CRA sind diese Gruppen rechtzeitig über die Ziele und Vorgehensweisen der CRA zu informieren. Dazu empfiehlt es sich, ein Faltblatt oder eine Kurzbroschüre zu erstellen sowie eine Präsentation zu entwickeln, die bei den verschiedenen, vorbereitenden Veranstaltungen verteilt werden kann.

Von besonderer Bedeutung ist die Klärung der Zusammenarbeit des Bürgerbeirats mit den Expertenteams. Das beginnt bei der Auswahl der Experten. Die CRA-Lenkungsgruppe mag berechtigterweise dazu neigen, die Expertengruppen mit Behördenvertretern zu besetzen, um Kosten zu sparen. Dies kann jedoch auf Widerstand beim Bürgerbeirat stoßen, der eine möglichst breit angelegte Expertise präfe-

riert und auch Experten seiner Wahl berücksichtigt wissen will. Hier ist eine Konsensorientierung anzuraten. Es empfiehlt sich, dem Bürgerbeirat ein Vorschlagsrecht einzuräumen.

Ein zweiter, klärungsbedürftiger Punkt betrifft die Verteilung der Rechte und Entscheidungskompetenzen zwischen den Experten, dem Planungsteam und dem Bürgerbeirat. Bewährt hat sich eine Teilung zwischen Abschätzungs- und Bewertungsaufgaben. Die Experten übernehmen – nachdem die Liste der zu bewertenden Risikoprobleme abgestimmt ist – die notwendigen Arbeiten zur Risikoabschätzung und Risikocharakterisierung, auf deren Grundlage dann der Bürgerbeirat das Ranking vornimmt.

6.5 Ablauf der CRA

Um die Mitglieder des CRA-Teams miteinander bekannt zu machen und um sie auf die anstehenden Aufgaben vorzubereiten, ist ein Workshop empfehlenswert. In den meisten Fällen werden dazu zwei Tage reichen. Hier geht es darum, die Teilnehmer in die CRA einzuführen. Zu Beginn sollten ihnen die Zielstellungen verdeutlicht werden. Weiterhin sind die Teilnehmer mit den Verfahren der CRA vertraut zu machen. Dazu kann ein Übungsranking durchgeführt werden. Im Workshop sind weiterhin Kommunikationsregeln zu verabschieden und der Umgang mit möglichen Konflikten zu erläutern.

Bei der tatsächlichen Bearbeitung der CRA muss zunächst Einvernehmen darüber hergestellt werden, mit welchen Problembereichen man sich in der CRA befassen will. Beispielsweise ist zu klären, für welche Regionen und Zeitskalen die CRA durchgeführt werden soll. Im Weiteren ist zu klären, welche Risikofelder (Verkehrsbezogene Risiken, Risiken der Lebensführung, umweltbezogene Risiken usw.) betrachtet werden sollen. Das erfordert die Bereitschaft, Problemhorizonte zu öffnen. Denn eng gefasste Problemsichten blockieren das Verfahren. Andererseits ist für die CRA auch die Ausweitung ins Unendliche problematisch.

Es kann mit einem Brainstorming begonnen werden, das nach Risiken in Aktivitätsfeldern sucht (Wohnen und Bauen, Kommunikation, Ernährung, Mobilität usw.). Anzustreben ist eine praktikable Liste der zu behandelnden Problemfelder. Innerhalb dieser Felder werden dann Risikokategorien (z.B. Luftverschmutzung, Abfälle, Innenraumluftverschmutzung, Kontaminanten in Nahrungsmitteln) gesucht und hier werden wiederum Zielgrößen, d.h. Leitstoffe ausgewählt (z.B. Acrylamid für Nahrungsmittel oder Passivrauchen für Innenraumluftverschmutzung sowie Feinstäube, O₃, NO_x, SO₂ für Luftverschmutzung im Außenbereich), die als Indikatoren für die Risikokategorie dienen können. Für diese sind ausgehend von Schutzgütern (Gesundheit, Umwelt, Lebensqualität) die entsprechenden Attribute zu bestimmen. Das können zum einen Endpunkte wie kanzerogene Erkrankungen, zum anderen Attribute wie Beeinträchtigungen durch Lärm oder Gerüche sein. Für diese sind dann Metriken bzw. Risikometriken, beispielsweise NOAELs, DALYs oder QALYs, auszuwählen, auf die hin eine (Risiko)beurteilung vorgenommen werden kann. Welche Metriken und welche Vergleichsmethode gewählt werden, hängt von der Datenlage und dem Ziel der Vergleiche ab. Schließlich sind Vergleiche vorzunehmen, die ein Ranking der Risiken ermöglichen.

In Abhängigkeit davon, welche Problembereiche und welche Schutzgüter betrachtet werden, sind von den Expertenteams unterschiedliche Aufgabengebiete zu bearbeiten, um den Prozess der Risikoabschätzung nicht über Gebühr in die Länge zu ziehen. Hierbei kann ganz unterschiedlich verfahren werden. Beispielsweise können Arbeitsgruppen nach Umweltmedien (Boden, Luft, Wasser), nach Schutzgütern oder nach Aktivitätsfeldern gegliedert werden.⁹⁶

Für das Risiko-Ranking können verschiedene Vergleichsverfahren genutzt werden. Grob unterschieden stehen drei Varianten zur Verfügung. Neben dem von uns präferierten multiattributen Bewertungsverfahren gibt es Konsensbildungsverfahren sowie einfache Abstimmungsverfahren. Im multiattributen Verfahren gewichten die Teilnehmer des Bürgerbeirats die Bewertungen der Risiken auf den verschiedenen Dimensionen. Daraus wird das Ranking errechnet. Im Unterschied dazu werden die Risiken bei den beiden übrigen Vorgehensweisen direkt auf der Basis konsensualer Entscheidungen bzw. von Abstimmungsergebnissen in eine Rangfolge gebracht.

Die jeweils eingesetzten Instrumente bzw. Verfahren wirken sich auf das Resultat der CRA aus. Sie lenken die Aufmerksamkeit, strukturieren das Problemfeld und beeinflussen das Ranking. Über die hier zu lösenden Probleme haben wir in Kapitel 4 berichtet.

Schließlich muss festgelegt werden, wie der Prozess der CRA und seine Ergebnisse dokumentiert werden sollen, und welche Teilergebnisse vor Abschluss des Prozesses in welcher Form der Öffentlichkeit zugänglich gemacht werden. Weitgehend unkritisch ist hier die Publikation der erfolgreich abgeschlossenen Verfahrensschritte und der im Konsens abgestimmten Zwischenergebnisse.

Zur Erfolgssicherung der CRA sind vorab eine Reihe von Vorkehrungen zu treffen. Dabei spielen die übergeordneten Ziele eine bedeutende Rolle. Diese Ziele, die vor allem die Umsetzung der CRA-Ergebnisse und deren weitere Wirkungen im Politikbereich „Umwelt und Gesundheit“ betreffen, sollten von vornherein bedacht werden. Mögliche Barrieren, wie institutionelle Dogmatik, Denken in Referatsschablonen, Ideologie-Konflikte und versteckte Interessen, sind zu erkennen und zu überwinden. Es wäre naiv anzunehmen, dass dies im Selbstlauf geschieht.

6.6 Organisation der Kommunikation

Bei der Vorbereitung der CRA sollte die Problemlage erfasst werden. Neben der sachlichen Ebene (Welches Verfahren sollte mit welcher Zielstellung durchgeführt werden?) spielt die personale Seite eine wesentliche Rolle. Hier muss bedacht werden:

- Welche besonderen Empfindlichkeiten haben die Betroffenen?
- Stimmt das Timing für eine CRA derzeit?

⁹⁶ Erfahrungen aus den USA demonstrieren, dass insbesondere die Gliederung nach Umweltmedien, die Gefahr beinhaltet, dass die jeweiligen Experten (die in ihren Behörden diese Bereiche auch bearbeiten) dazu neigen, „ihre“ Risiken besonders hoch einzuschätzen, um ihre Tätigkeitsgebiete „hoch“ zu halten. Solche interessenbedingten Verzerrungen sind zwar nie völlig auszuschalten, sie lassen sich aber bei einer anderen Aufteilung – z.B. nach Aktivitätsfeldern – minimieren.

- Inwieweit sind die „Issues“ (Konfliktthemen/Risikofragen) bereits ausgeüfert?
- Haben alle Parteien ein Interesse an einer CRA? Und wenn ja, welche?

Für die Beteiligten spielt die gemeinsame Klärung des zukünftigen Umgangs miteinander (z.B. durch eine Geschäftsordnung), sowie die Festlegung von zu behandelnden Themen eine wichtige Rolle. Wesentlicher Punkt dabei ist, einen Konsens über das Ziel der CRA und die Rolle der Beteiligten zu erlangen. Das Verfahren kann darauf abzielen, Empfehlungen vorzubereiten, Optionen zu generieren und zu bewerten oder gemeinsame Entscheidungen zu treffen. Je nachdem, welche Ziele die Teilnehmer verfolgen, haben sie unterschiedliche Rollen: Als Gutachter, Berater, oder als (Mit)entscheider. Klare Kommunikation kann helfen, Rollenkonfusionen zu vermeiden.

Jede CRA setzt voraus, dass die Beteiligten ausreichend informiert sind, um kompetent an einer Entscheidung mitwirken zu können. Daher ist es erforderlich, die vorhandenen Informationsdefizite auszugleichen. Alle Fakten, die für sie relevant sind, müssen auf den Tisch. Es kann durchaus erforderlich sein, Hearings mit Experten zu veranstalten, damit sich alle Teilnehmer eine kompetente Meinung bilden können.

Spannungen entstehen, wenn unterschiedliche Auffassungen aufeinander stoßen. In einer solchen Situation ist es notwendig, den Beteiligten gleichen Raum zur Darlegung ihrer Standpunkte zu geben, die Interessen, die hinter den Standpunkten stehen, zu explorieren, und sich schon in dieser Phase zu bemühen, Gemeinsamkeiten der Konfliktparteien zu erkunden und zu verdeutlichen.

Die CRA setzt weiterhin voraus, dass die zu vergleichenden Risiken bestmöglich charakterisiert werden. Die EPA erörtert vier Kriterien (siehe Tabelle 24), nach denen die Qualität einer Risikocharakterisierung beurteilt werden kann: Transparenz, Klarheit, Konsistenz und Angemessenheit / Vernünftigkeit (vgl. EPA 2000, S. 14ff.).

Tabelle 24: Kriterien zur Risikocharakterisierung

Risiko-Charakterisierung		
Prinzip	Definition	Kriterien
Transparenz	Offenlegung des Risikoabschätzungsprozesses	Beschreibung der Vorgehensweise, Annahmen, Modelle und Extrapolationen Identifikation von Datenlücken Darstellung der Unsicherheiten
Klarheit	Verständlichkeit der Darstellungen	Kürze Logischer Aufbau Umgangssprache
Konsistenz	Lege artis Vorgehen	Peer Reviewing Übereinstimmung mit Leitfäden und „Gute Praxis“-Vorgehen
Angemessenheit	Die Bewertung erfolgt nach bestem wissenschaftlichen Wissen	Peer Reviewing Neueste Daten „Good judgement“

Transparenz meint das Offenlegen der Vorgehensweise bei der Risikoabschätzung. D.h. Vorgehensweise, Annahmen und Modelle, und Extrapolationen müssen beschrieben werden, Datenlücken müssen identifiziert und Unsicherheiten deutlich gemacht werden. Klarheit bezieht sich im wesentlichen auf die Verständlichkeit der

Darstellungen. Kriterien sind hier Prägnanz und logischer Aufbau der Darstellung sowie eine verständliche Sprache. Mit Konsistenz ist die Übereinstimmung mit etablierten Verfahrensweisen oder Leitlinien für die Risikoabschätzung gemeint. Hier muss deutlich gemacht werden, ob die Schlussfolgerungen, die aus der Befundlage gezogen werden, in Übereinstimmung mit den jeweiligen wissenschaftlichen Standards bzw. Verfahrensweisen stehen. Ist dies nicht der Fall, so muss das begründet werden. Angemessenheit schließlich ist ein Gesamt-Kriterium, das noch einmal verdeutlicht, wie „vernünftig“ die Bewertung des Risikos ist. Kriterien hierfür sind, dass die Risikocharakterisierung die Zustimmung der relevanten „wissenschaftlichen Gemeinschaft“ findet, und dass sie auf den besten verfügbaren Daten und allgemein akzeptierten wissenschaftlichen Erkenntnissen basiert.

Allerdings ist hier auch mit einer Reihe von Fehlern und Fallen zu rechnen. Gerade, wenn arbeitsteilig – in verschiedenen Gruppen – Risiken bewertet werden, können verschiedene Annahmen, die nicht explizit angegeben werden (z.B. die Annahme von ‚worst case‘-Expositionen versus ‚reasonable case‘-Expositionen) zu Konfusionen führen (fehlende Transparenz). Ein Übermaß an Risikojargon geht zu Lasten der Klarheit. Ein besonderes Problem stellt die Konsistenz dar, die durch unterschiedliche Annahmen und Vorgehensweisen bei den Risikoabschätzungen beeinträchtigt werden kann.

Neben solchen Verfahrens- und konzeptionellen Problemen sind aber auch prozessuale Aspekte bei der Kommunikation zu beachten. Im Prinzip geht es dabei um die Umsetzung von Fairness, Kompetenz und Vertrauen (vgl. Renn, Webler & Wiedemann 1995). Dies stellt hohe Anforderungen an die Kommunikation.

Parallel zu der internen Kommunikation ist auf eine ausreichende externe Kommunikation zu achten, um die Akzeptanz des CRA-Projekts zu sichern. Tabelle 25 bietet einen Überblick zu den Kommunikationsproblemen bei einer CRA.

Tabelle 25: Ursachen von Kommunikationsproblemen

Phase der CRA	Kommunikationsproblem	Schwerpunkte der Kommunikation
Vorphase	Misstrauen gegenüber den Trägern der CRA	Darlegung der Zielvorstellungen (Schutzgüter und Schutzziele). Information über Vorgehensweisen und Beteiligung bei der CRA.
Risikoabschätzung	Expertenkontroversen	Information und Erörterung über das Vorgehen und die Ergebnisse bei der Identifikation von Gefahrenquellen (Hazard Identifikation), der Dosis-Wirkungs-Abschätzung und der Expositionsabschätzung.
Risikocharakterisierung	Intransparenz, Selektion von bestätigenden Risikoeinschätzungen	Transparente Darstellung der eingesetzten Verfahren Nachvollziehbare Zusammenfassung und Vermittlung der wichtigsten Ergebnisse der Risikoabschätzung Auskunft über die Abschätzungsunsicherheiten.
Risiko-Ranking	Nichtberücksichtigung des gesellschaftlichen Wertspektrums	Information über das Verfahren und die beteiligten Gruppen. Offener Zugang zu den Ergebnissen. Möglichkeit zur Kommentierung.

In sachlicher Hinsicht ist Kommunikation dann besonders schwierig, wenn sich andeutet, dass bei der CRA von Risikokonzepten und -perspektiven ausgegangen wird, die grundsätzlich von der Betrachtungsweise der Öffentlichkeit abweichen. Dieses

Problem kann nie vollständig gelöst werden. Ein Spannungsfeld bleibt immer bestehen. Wenn die CRA vollständig den Risikowahrnehmungen der Öffentlichkeit folgt, verfehlt sie ihr Ziel und gerät zu einer Risikowahrnehmungsstudie – von der sie sich ja gerade unterscheiden will. Blendet die CRA jedoch alle Aspekte öffentlicher Risikowahrnehmung aus und pocht allein auf wissenschaftliche Fakten, so kann sie ins gesellschaftliche Abseits geraten: Die Akzeptanz für ihre Resultate bleibt aus. Daraus erwächst die Aufgabe, die Vorgehensweise bei der CRA – insbesondere die verwendeten Attribute – genau darzulegen.

Weitere Risikokommunikationsprobleme mit der Öffentlichkeit können aus Ressourcenmängeln (Zeit, Personal und Finanzen) sowie aus mangelnden Absprachen zwischen dem CRA Projekt-Team und dem CRA-Projektmanagement erwachsen.

Ob eine CRA erfolgreich ist, hängt auch davon ab, wie mit den Entscheidern kommuniziert wird. Neben situativen Faktoren (Zeit, Umstände) spielen auch Präsentationsformate eine Rolle. Dabei geht es vor allem um ein gutes „executive summary“ und um graphische Hilfsmittel, die die wesentlichen Informationen zusammenfassen und anschaulich darstellen. Bislang gibt es jedoch zur Frage der Eignung von Präsentationsformaten für eine CRA so gut wie keine empirischen Untersuchungen.

Insgesamt bieten die Erfahrungen mit CRAs in den USA zwar reiches Anschauungsmaterial zur Vielfalt der Umsetzungsmöglichkeiten, sind aber wegen der administrativen, politischen und nicht zuletzt auch kulturellen Unterschiede zu Deutschland bzw. Europa nur begrenzt übertragbar. Will man, dass die vergleichende Risikobewertung einen festen Platz im Prozess der Risikoregulierung hat, so sind praktische Erprobungen von Verfahren, wie sie in diesem Bericht ausgearbeitet worden sind, unverzichtbar.

Anhang 1: Risikometriken

Die Auswahl einer angemessenen Risikometrik ist eine Aufgabe der Experten für die jeweiligen Risikofelder. Die Verfügbarkeit einer solchen Metrik ist Voraussetzung für die Beurteilung der Risiken hinsichtlich ihrer gesundheitlichen Folgen. In der Literatur werden verschiedene Metriken vorgeschlagen, von denen die wichtigsten im folgenden kurz dargestellt werden sollen.

Metriken zur Charakterisierung der Dosis-Wirkungs-Beziehungen: NOAEL, Benchmark Dose, Reference Dose

Die Wirkung einer Noxe in einem Organismus ist abhängig vom zeitlichen Verlauf der Dosis⁹⁷ und vom Weg, über den die Dosis vom Organismus aufgenommen wird. Wenn von einer Dosis-Wirkungs-Beziehung gesprochen wird, wird daher immer (implizit oder explizit) ein bestimmter Aufnahmeweg und ein bestimmter Zeitverlauf vorausgesetzt.

Die Beziehung zwischen Dosis und Wirkung wird als stetige (meist auch als monoton steigende) Funktion angenommen. Sie variiert z.B. mit dem Alter und dem allgemeinen Gesundheitszustand des Individuums, von Individuum zu Individuum (Intraspezies-Variabilität) und zwischen verschiedenen Arten (insbesondere zwischen Mensch und Tier, Interspezies-Variabilität). In epidemiologischen Untersuchungen und Laborversuchen können immer nur bestimmte Punkte dieser Funktion ermittelt werden, und diese auch nur mit begrenzter Genauigkeit.

NOEL, NOAEL, LOEL, LOAEL und Benchmark-Dosis sind repräsentative Punkte der Dosis-Wirkungs-Funktion:

NOAEL: *No Observed Adverse Effect Level*: Höchste Dosis, bei der kein adverser (krankmachender) Effekt festgestellt werden konnte oder bei der kein adverser Effekt zu erwarten ist.

Nicht immer ist die Angabe eines NOAEL möglich, ersatzweise kommt dann einer der folgenden Werte zur Anwendung:

LOAEL: *Lowest Observed Adverse Effect Level*: Niedrigste Dosis, bei der ein adverser (krankmachender) Effekt festgestellt werden konnte oder bei der ein adverser Effekt zu erwarten ist.

NOEL: *No Observed Effect Level*: Höchste Dosis, bei der kein Effekt festgestellt werden konnte oder bei der kein Effekt zu erwarten ist.

LOEL: *Lowest Observed Effect Level*: Niedrigste Dosis, bei der ein Effekt festgestellt werden konnte oder bei der ein Effekt zu erwarten ist.

Bei der Nutzung des NOEL bzw. LOEL ist zu berücksichtigen, dass sich diese auf biologische Effekte allgemein beziehen. Da es im Rahmen einer CRA aber um mög-

⁹⁷ z. B. einmalige Aufnahme der gesamten Dosis, Aufnahme in mehreren Einzeldosen, kontinuierliche Aufnahme einer konstanten Dosisrate während einer begrenzten Zeit, oder kontinuierliche Aufnahme einer konstanten Dosisrate über die gesamte Lebensdauer.

liche Gesundheitsschäden geht, ist eine Anwendung von NOEL bzw. LOEL nur dann sinnvoll, wenn es gute Gründe für die Annahme gibt, dass die dabei betrachteten Effekte als Indikatoren für adverse Effekte herangezogen werden können.

Für die Anwendung des NOAEL ist zu beachten, dass dieser *nicht* besagt, dass bei der jeweiligen Dosis *sicher* keine adverse Wirkung auftritt, sondern dass bei dieser Dosis *mit der durch die Versuchsbedingungen gegebenen Aussagesicherheit* keine adverse Wirkung auftritt. Diese Wahrscheinlichkeit wird aber nicht explizit angegeben und kann für unterschiedliche Substanzen (auf Grund unterschiedlicher Versuchsbedingungen) verschieden sein. U. a. deshalb wird als präziser definierter Wert die *Benchmark Dose* (BMD) vorgeschlagen:

»Seit Mitte der 80er Jahre wird für Stoffe mit Wirkungsschwelle bei der Ableitung von Referenzdosen der EPA (RfD-Werte) als Ersatz für den NOAEL die Verwendung einer sog. „Benchmark Dose“ diskutiert. In statistischer Sprechweise entspricht sie dem unteren Konfidenzbereich der Effekt-Dosis ED_x , bei der in x% der untersuchten Testpopulation eine Wirkung aufgetreten ist.« (Neus et al. 1997, S. 173-174)

Ein Effekt, der bei weniger als 10% oder 5% der Testpopulation auftritt, sei (je nach Versuchsbedingungen) an der Grenze dessen, was noch identifizierbar ist, so dass *de facto* die Benchmark Dose in der Nähe des LOAEL oder NOAEL liege (Barnes et al. 1995; zitiert in Neus et al. 1997, S. 174).

Bei allen diesen Schwellwerten muss unterschieden werden zwischen dem NOAEL (oder der Benchmark Dose), der das Ergebnis von Laborversuchen mit *Versuchstieren* darstellt als höchste Dosis, bei der kein adverser Effekt *festgestellt werden konnte*, und dem NOAEL (oder BMD), der auf den *Menschen* übertragen die höchste Dosis repräsentiert, bei der kein adverser Effekt *zu erwarten ist*.

Das Konzept der *Reference Dose* wurde von der US-EPA entwickelt, um verschiedene Probleme mit den Konzepten des NOAEL und des ADI (acceptable daily intake) zu beheben.

»The RfD [Reference Dose] is a benchmark dose operationally derived from the NOAEL by consistent application of generally order-of-magnitude uncertainty factors (UFs) that reflect various types of data sets used to estimate RfDs. [...] In addition, a modifying factor (MF), is sometimes used which is based on a professional judgment of the entire data base of the chemical. [...]

The RfD is determined by use of the following equation: $RfD = NOAEL / (UF \times MF)$

[...] In general, the RfD is an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. The RfD is generally expressed in units of milligrams per kilogram of bodyweight per day (mg/kg/day).

The RfD is useful as a reference point from which to gauge the potential effects of the chemical at other doses. Usually, doses less than the RfD are not likely to be associated with adverse health risks, and are therefore less likely to be of regulatory concern. As the frequency and/or magnitude of the exposures exceeding the RfD increase, the probability of adverse effects in a human population increases. However, it should not be categorically concluded that all doses below the

RfD are "acceptable" (or will be risk-free) and that all doses in excess of the RfD are "unacceptable" (or will result in adverse effects).« (US-EPA 1993)

All diese aus der Dosis-Wirkungs-Beziehung abgeleiteten Metriken sind Schwellenwerte und als solche nicht kumulierbar, d. h. sie sind nicht geeignet zur Beurteilung der kumulierten Wirkung mehrerer Noxen.

Unit Risk

Unit Risk bezeichnet das geschätzte zusätzliche Krebsrisiko eines Menschen nach konstanter Exposition über 70 Jahre gegenüber einer Konzentration von 1 µg Schadstoff pro m³ Luft (oder entsprechenden Konzentrationen in anderen Medien).

Becher et al. (1995, S. 88) definieren Unit Risk (UR) als Differenz $UR = P_1 - P_0$. Dabei sind P_1 und P_0 über die gesamte Lebenszeit integrierte Sterbewahrscheinlichkeiten infolge der durch die Noxe induzierten Krankheit, P_0 ohne Belastung mit der Noxe und P_1 bei einer Belastung mit der Noxe in der Konzentration 1 µg/m³ Luft (oder entsprechende Konzentration in anderen Medien). Weiter wird dargelegt, dass es zur Ermittlung von P_0 und P_1 verschiedene Schätzverfahren (mittels Approximation des Integrals durch die Summe) gibt, die eine (jeweils konsistente) Über- oder Unterschätzung dieser Werte bedingen. Ein weiteres Schätzverfahren, bei dem das bei einer Konzentration x beobachtete zusätzliche Risiko R_x durch die Konzentration x dividiert wird ($UR = R_x/x$), setzt eine lineare Dosis-Wirkungs-Beziehung voraus. Es sollten daher nur solche Unit Risk-Werte miteinander verglichen werden, die mit dem gleichen Schätzverfahren ermittelt wurden.

Hier ist anzumerken, dass Wahrscheinlichkeiten und damit auch Unit Risk-Werte im Prinzip nicht additiv sind. Die arithmetische Addition und Subtraktion von Wahrscheinlichkeiten ist eine Näherung, die voraussetzt, dass $P_0 < P_1 \ll 1$ ist.⁹⁸ Werden Unit Risk-Werte mehrerer Noxen linear addiert, können sich dadurch Risikowerte >1 ergeben.

Aus dem gleichen Grund ergeben sich Probleme, wenn die Krankheit oder der Endpunkt, für die das Unit Risk bestimmt werden soll, mehr als nur einen geringen Teil aller Krankheiten ausmacht. Im Grenzfall, für alle Krankheiten (Endpunkte) zusammen, ist $P_1 = P_0 = 1$. Für eine (hypothetische) Noxe, die die Mortalitätsraten aller Krankheiten erhöht, wäre also $UR = 0$ (Becher et al. 1995, S. 108).

Die US-EPA definiert Unit Risk wie folgt (Kincaid et al. 1995):

Unit Risk: The upper-bound excess lifetime cancer risk estimated to result from continuous exposure to an agent at a concentration of 1 µg/L in water or 1 µg/m³ in air (with units of risk per µg/m³ air or risk per µg/L water). (p.5-38)
[...]

⁹⁸ Die exakte Formel für die Wahrscheinlichkeit des kombinierten Risikos ist $P_1 = P_0 + UR - P_0 \cdot UR$; dies berücksichtigt die Tatsache, dass jemand nur entweder an den Folgen der Exposition oder in Folge von etwas anderem, aber nicht zuerst an dem einen und dann noch einmal an dem anderen sterben kann. Damit ist $UR = (P_1 - P_0) / (1 - P_0)$.

The slope factor, or q_1^* , can also be used to determine the incremental cancer risk that would occur if the chemical was present in an environmental medium such as drinking water at a unit concentration (i.e., 1 μg of chemical per liter of drinking water). The calculation for drinking water usually assumes the person weighs 70 kg and drinks 2 liters of water per day:

$$\text{Drinking Water Unit Risk} = q_1^* \times 1/70 \text{ kg} \times 2 \text{ L/day} \times 10^{-3}$$

Air unit risk (risk per $\mu\text{g}/\text{m}^3$) is derived from the linearized multistage procedure and calculated using the GLOBAL program. (p.5-46)

Der hier angesprochene Slope Factor ist wie folgt definiert:

Slope Factor (q_1^*): A measure of an individual's excess risk or increased likelihood of developing cancer if exposed to a chemical. It is determined from the upperbound of the slope of the dose-response curve in the low-dose region of the curve. More specifically, q_1^* is an approximation of the upper bound of the slope when using the linearized multistage procedure at low doses. The units of the slope factor are usually expressed as $1/(\text{mg}/\text{kg}\cdot\text{day})$ or $(\text{mg}/\text{kg}\cdot\text{day})^{-1}$. (p.5-37)

[...]

The slope factor is a measure of the incremental risk or increased likelihood of an individual developing cancer if exposed to a unit dose of the chemical for a lifetime. The risk is expressed as a probability (i.e., one chance in ten or one chance in one million), and the unit dose is normally expressed as 1 mg of the chemical per unit body weight (kg) per day:

$$\text{Slope Factor} = \text{Risk per unit dose, or Risk per mg/kg-day}$$

When based on animal data, the slope factor is derived by extrapolating from the incidences of tumors occurring in animals receiving high doses of the chemical to low exposure levels expected for human contact in the environment. The EPA uses q_1^* for its risk assessments (see definition of slope factor). The q_1^* for a chemical, in units of $(\text{mg}/\text{kg}\cdot\text{day})^{-1}$, is based on the linearized multistage procedure for carcinogenesis and can be calculated by computer program (e.g., GLOBAL). (p.5-45)

Auch die Definitionen der US-EPA setzen voraus, dass das zusätzliche Risiko $\ll 1$ ist. Das „Unit Risk“ wird von der US-EPA auch zur Ermittlung des kumulierten Risikos bei Belastung mit mehreren Kanzerogenen benutzt (US-EPA, *Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures*, 2000). Auch dies gilt nur so lange, wie das resultierende Risiko $\ll 1$ ist.

Relatives Risiko (RR)

Das Relative Risiko ist definiert als das Verhältnis des Risikos für einen adversen Gesundheitseffekt bei den Exponierten zum Risiko des adversen Gesundheitseffekts bei den Nicht-Exponierten (Kreienbrock & Schach 1995, 47f.). Es gibt an, um wieviel sich die Wahrscheinlichkeit für den adversen Gesundheitseffekt erhöht, wenn man einer definierten Exposition ausgesetzt ist. Das Relative Risiko kann also als Maß für die Stärke des Zusammenhangs zwischen Exposition und adversen Effekt verstanden werden.

Individual Exposure Ratio (IER)

Für nicht-kanzerogene Effekte nutzt die EPA das *Individual Exposure Ratio* (IER), das als Quotient aus einer Expositions-dosis und der *Reference Dose* (RfD) berechnet wird. Das IER gibt an, in welchem Ausmaß die aufgenommene Dosis einer Substanz den Wert einer sicheren Dosis überschreitet. Je höher das IER, desto größer die Wahrscheinlichkeit für adverse Gesundheitswirkungen. Allerdings weist die EPA auf einer Reihe konzeptioneller Probleme bei der Verwendung des IER als Risikometrik für CRAs hin (EPA 1993, 2.2-22). Zum Beispiel setzt ein Risikovergleich auf der Basis der IER voraus, dass für alle betrachteten Substanzen die gleiche Dosis-Wirkungs-Beziehung gilt – was nicht der Fall ist.

Margin of Exposure (MOE) bzw. Margin of Safety (MOS)

Nach Neus et al. (1997, S. 173) ist „Margin of Exposure“ wie folgt definiert:

»Insbesondere im Rahmen der Regulation durch Zulassung oder Genehmigung - so z.B. in dem Technical Guidance Document der EU (CEC 1996) ist es üblich, die Gesundheitsverträglichkeit einer Exposition danach zu beurteilen, ob der „margin of safety“ bzw. der „margin of exposure“ ausreichend ist. Diese Größe ist nach der Zusammenstellung von OECD und IPCS international übereinstimmend als Quotient aus NOAEL (oder einem anderen Schätzwert für die Wirkungsschwelle) und der abgeschätzten Exposition (Zufuhr) definiert; sie kann definitionsgemäß nur für Stoffe mit Wirkungsschwelle angegeben werden.«

„Margin of Exposure“ und „Margin of Safety“ werden bedeutungsgleich genutzt (vgl. auch KEMI 2003, 35). Zu beachten ist, dass „Margin of Safety“ mitunter auch in einer anderen Bedeutung, nämlich im Sinne von „Sicherheitszuschlag“ oder „Sicherheitsfaktor“ (*safety factor*), benutzt wird (z.B. Goldstein 1990).

Das BfR (vormals BgVV) verwendet den Begriff „Margin of Exposure“ wie folgt: „Die MOE errechnet sich, indem man die Dosis, die bei Tieren zu Tumoren führt, durch die Aufnahmemenge des Menschen teilt.“ (BfR 2002) Beim Vergleich von MOE ist also darauf zu achten, ob die Wirkungsschwelle, auf die Bezug genommen wird, die im Tierversuch beobachtete Dosis oder die auf den Menschen extrapolierte Dosis ist. Dieses Maß ist nicht unmittelbar kumulierbar (zur Beurteilung der gleichzeitigen Wirkung mehrerer Noxen) sondern allenfalls sein Kehrwert, und dieser auch nur für Noxen, die über den selben Toxizitäts-Mechanismus zum selben Endpunkt führen (Sielken 2000).

Intake Fraction

Die „Intake Fraction“ ist ein Maß zur Abschätzung des Transports einer Noxe zwischen Freisetzung und Exponierten (Bennett, Margni, et al. 2002):

»*iF* is the integrated incremental intake of a pollutant released from a source or source category (such as mobile sources, power plants, or refineries) and summed over all exposed individuals during a given exposure time, per unit of emitted pollutant.

$$iF = \frac{\sum_{\text{people, time}} \text{mass intake of pollutant by an individual (mass)}}{\text{mass released into the environment (mass)}}$$

Although the pollutant intake is summed over population and time, in actuality, when a pollutant is released, there is a distribution of individual exposures within the exposed population. Individual exposure can be quantified in terms of the *individual intake fraction*, iF_i . The total intake fraction comprises iF s summed over all members of a potentially exposed population.«

Dieses Maß wäre nur dann kumulierbar, wenn alle Noxen in gleicher Menge emittiert würden, was in der Regel nicht gegeben ist.

Mortalität, Morbidität und Years of Life Lost (YLL)

Mortalität bezeichnet die Sterblichkeit in einer Population. Im Rahmen von Risikobetrachtungen interessieren vor allem die einer bestimmten Risikoquelle zurechenbaren Todesfälle, die als Maß für die Größe des Risikos interpretiert werden können. In diesem Sinne wird Mortalität zum Beispiel bei der vergleichenden Analyse von Alltagsrisiken verwendet (z.B. Cohen 1991; Wilson 1979).

Morbidität meint das populationsbezogene Auftreten von Krankheiten. Im Rahmen von Risikobetrachtungen interessiert vor allem die Inzidenz einer Krankheit, das heißt die Zahl der „Neuzugänge“ für eine Krankheit. Auch hier ist wieder vor allem die einer bestimmten Risikoquelle zurechenbare Inzidenz von Interesse, zum Beispiel die Anzahl von Krebsfällen, die der Exposition mit einem Karzinogen zugeschrieben werden können.

YLL (*Years of Life Lost, Expected Loss of Lifetime* (ELL), erwarteter Lebenszeitverlust) gibt an, um wie viel die durchschnittliche Lebenserwartung exponierter Personen geringer ist als die durchschnittliche Lebenserwartung nicht-exponierter Personen. Im Gegensatz zur reinen Zählung der Todesfälle bewertet YLL einen Todesfall um so höher, in je geringerem Lebensalter er eintritt. Die YLL-Metrik kann (im Gegensatz zu Unit Risk und Margin of Exposure) auch endpunktübergreifend, d.h. nicht-krankheitsspezifisch angewendet werden.

Disability Adjusted Life Years (DALYs), Quality Adjusted Life Years (QALYs)

Diese Metriken wurden entwickelt, um neben verschiedenen Arten von Todesfällen auch nicht unmittelbar zum Tod führende Erkrankungen in den Vergleich einbeziehen zu können.

QALYs sind das Integral des Gesundheitszustands über die Lebenszeit. Dabei wird ein Lebensjahr in perfekter Gesundheit mit dem Wert 1 gezählt. Krankheitszeiten werden mit einer Gewichtung <1 gewertet. Dem Zustand „tot“ wird die Gewichtung 0 zugeordnet. Die Gewichtungen werden durch Befragung von Personen ermittelt. Daraus können für Krankheiten, die schlimmer als der Tod empfunden werden, Gewichtungen <0 resultieren.

DALYs sind die Summe der verlorenen Lebensjahre und der mit Behinderung verbrachten Lebensjahre: $DALYs = YLL + YLD$. Bei der Ermittlung der DALYs können die Lebensjahre wahlweise diskontiert (Bewertung umso geringer je weiter in der Zukunft) und mit einem vom Lebensalter abhängigen Faktor gewichtet werden.

Nach Hofstetter & Hammitt (2002) gibt es nur wenige wesentliche Unterschiede zwischen QALYs und DALYs:

1. Sie sind komplementär, und Änderungen ihrer Werte sind invers zu einander ($\Delta QALY \sim -\Delta DALY$).
2. Es werden unterschiedliche Methoden zur Ermittlung der Gewichtungen für Lebensqualität und Behinderung angewandt.
3. DALYs nehmen als Referenz vollkommen gesunde Personen an, die nach der mittleren erwarteten Lebenszeit sterben. QALYs berücksichtigen typische altersbedingte Behinderungen und tatsächliche Lebenserwartungen.
4. DALYs können eine vom Lebensalter abhängige Bewertung vornehmen, die bei QALYs nicht benutzt wird.

Hierzu ist zweierlei anzumerken: Erstens, es mag zwar üblich sein, zur Ermittlung der Gewichtungen für Lebensqualität und Behinderung bei DALYs andere Methoden anzuwenden als bei QALYs, die Definition von DALYs und QALYs erzwingt dies aber nicht. Ebenso könnte man auch bei QALYs wahlweise eine vom Lebensalter abhängige Bewertung vornehmen. Zweitens, wenn (laut Punkt 2) unterschiedliche Methoden zur Ermittlung der Gewichtungen für Lebensqualität und Behinderung angewandt werden oder (laut Punkt 4) bei DALYs eine vom Lebensalter abhängige Bewertung vorgenommen wird, bei QALYs aber nicht, gilt Punkt 1 (Metriken sind komplementär, Änderungen der Werte invers zu einander) nur näherungsweise.

DALYs und QALYs vergleichen verschiedene Gesundheitszustände oder Lebensverläufe. Diese können sich zum einen auf eine einzelne Person oder eine Gruppe von Personen beziehen, für die die Vor- und Nachteile z.B. einer medizinischen Behandlung abzuwägen sind; auf dieser Ebene werden DALYs und QALYs in der „evidence based medicine“ verwendet. Zum anderen können sie sich auch auf ein oder mehrere Kollektive von Personen beziehen, die einer oder mehreren Noxen ausgesetzt sind. Auf dieser Ebene können DALYs oder QALYs über die Wirkungen verschiedener Noxen und über Kollektive exponierter Personen aggregiert werden (z.B. in US-EPA *Comparative Risk Framework Methodology and Case Study*, SAB Review Draft 1998). Bei der Aggregation über Kollektive wird allerdings das Risiko um so geringer gewichtet, je weniger Personen zum betrachteten Kollektiv, z.B. zu einer Gruppe besonders sensibler Personen, gehören. Ob in Anbetracht der Verpflichtung zum Schutz auch sensibler Gruppen diese Gewichtung „richtig“ ist, ist umstritten (Hubbell 2002).

Probleme von QALYs und DALYs:

- Bei QALYs gehen sowohl Schwere als auch Dauer einer Krankheit linear ein. Dies entspricht nicht immer der tatsächlichen Einschätzung durch die betroffenen Personen.
- Es ist unklar, was „perfekte Gesundheit“ ist. (Der bestmögliche während eines Lebens erreichbare Gesundheitszustand? Altersabhängig der bestmögliche im

jeweiligen Alter erreichbare Gesundheitszustand? Welche Personengruppe wird als Referenz für „perfekte Gesundheit“ zu Grunde gelegt?)

- Es ist umstritten, ob diskontiert werden soll oder nicht, und ob eine Wertung abhängig vom Lebensalter vorgenommen werden soll.

Willingness To Pay (WTP)

Willingness To Pay (WTP) ist ein Ansatz, Gesundheits- und Todesrisiken auf Geldwerte umzurechnen. WTP ist der Geldbetrag, den der Betroffene zu zahlen bereit ist, um eine Verringerung des Risikos zu erzielen. WTA (Willingness To Accept) ist der Geldbetrag, für den der Betroffene bereit ist, auf eine Verringerung des Risikos zu verzichten. In der Regel ist $WTA > WTP$, manchmal ist der Unterschied aber nur gering. In der Praxis werden WTP und WTA durch Befragungen ermittelt. (Näheres zum Unterschied zwischen WTP und WTA findet sich bei Hammitt, 2002).

WTP vermeidet die Probleme, die bei QALY durch die Linearitäts-Annahme entstehen. Andererseits muss bei WTP berücksichtigt werden, dass die Fähigkeit zu zahlen die Bereitschaft zu zahlen beschränkt. Hubbell (2002) zeigt die Unterschiede in der Bewertung durch QALY und WTP in der folgenden Tabelle:

Tabelle 26: Vergleich des QALY- und WTP-Ansatzes (nach Hubbell 2002, S.33)

Parameter	QALY	WTP
Risk Aversion	Risk neutral	Empirically determined
Relation of duration and quality	Independent	Empirically determined
Proportionality of duration/quality tradeoff	Constant	Variable
Treatment of time/age in utility function	Utility linear in time	Empirically determined
Preferences	Community	Individual
Source of preference data	Stated	Revealed and stated
Treatment of Income and Prices	Not explicitly considered	Constrains choices

Eignung der Risikometriken für die CRA

Für die Nutzung dieser Risikometriken im Rahmen einer CRA ist zu beachten, dass sie sich auf unterschiedliche Dimensionen beziehen. Einige Metriken betrachten die Dosis-Wirkungs-Beziehung, sie charakterisieren die Toxizität von Stoffen und damit das Gefährdungspotenzial. Andere Metriken beziehen sich auf die Exposition und wieder andere auf die Ausmaß der feststellbaren bzw. zu erwartenden Gesundheitseffekte. Es ist klar, dass Metriken, die sich z.B. auf Wirkungsschwellen beziehen (NOAEL etc.), nicht ohne weiteres mit hochaggregierten Metriken wie etwa DALYs oder QALYs verglichen werden können.

Zu beachten ist auch, dass Abschätzungen für die gleiche Risikometrik (z.B. Lebenszeitrisiko), aber unterschiedliche Endpunkt (z.B. Herzinfarkt vs. Krebs) auf unterschiedlichen Methoden beruhen können, die eine unterschiedliche „Empfindlichkeit“ im Aufweisen von Risiken haben (vgl. Neus et al. 1995, 305f.). So liegt etwa die statistisch gesicherte Nachweisgrenze für die auf die Lebenszeit bezogenen Mortalität bei umweltepidemiologischen Untersuchungen oberhalb von 10^{-2} bis 10^{-3} . Dagegen können für krebserzeugende Umweltbelastungen Abschätzungen vorgenommen werden, die um Größenordnungen darunter (im Bereich von 10^{-5}) liegen.

Schließlich muss man feststellen, dass die meisten Risikometriken sich auf schwerwiegende oder gar tödliche Gesundheitseffekte beziehen. Bisher fehlen Metriken, mit denen schwächere Formen adverser Effekte (z.B. Befindlichkeitsstörungen durch Lärm oder Gerüche) gemessen und verglichen können.

Welche Risikometriken sich unter welchen Bedingungen (und mit welchen Zusatzannahmen) für eine vergleichende Risikobewertung nutzen lassen, kann nur durch die jeweiligen Experten entschieden werden. Bei Beteiligung anderer Stakeholdergruppen ist es wichtig, die Aussagekraft der einzelnen Metriken für die CRA zu verdeutlichen.

Anhang 2: Multiattribute vergleichende Risikobewertung

Bei der Durchführung einer multiattributen vergleichenden Risikobewertung (MRB) kommt der Vorbereitung und Strukturierung des Bewertungsprozesses eine große Bedeutung zu – vor allem dann, wenn neben den (verschiedenen) Experten auch noch Stakeholdergruppen beteiligt werden sollen. Für die multiattribute vergleichende Risikobewertung ergibt sich dann das in Tabelle 27 dargestellte Profil.

Tabelle 27: Profil der MRB

Zielstellung	Vergleich verschiedener Risiken auf einem oder mehreren Attribut(en) durch einen oder mehrere Bewerter
Anwendungsbereich	Bekannte Risiken
Fokus	Abschätzung von Gesundheitsrisiken
Beteiligung	Experten und ggf. Stakeholder
Output	Risiko-Ranking, Risiko-Sortierung

Im folgenden soll an zwei Beispielen erläutert werden, wie eine multiattribute vergleichende Risikobewertung durchgeführt werden kann.

Beispiel 1: Vorgehensweise bei einer MRB für Umwelt-, Gesundheits- und Lebensqualitätsrisiken

Das bislang wohl umfassendste Konzept für eine vergleichende Bewertung von Umwelt-, Gesundheits- und Lebensqualitätsrisiken hat die EPA (1993) mit ihrem „*Guidebook to Comparing Risks and Setting Environmental Priorities*“ vorgelegt. Es basiert auf den vielfältigen Erfahrungen, die die EPA aus den zahlreichen nationalen, regionalen und lokalen CRAs gezogen hat. Diese wurden in Kapitel 3 ausführlich dargestellt.

Der Ansatz der EPA (1993) ist zwar grundsätzlich multidimensional, sie führt diesen Gedanken aber nicht konsequent zu Ende, sondern belässt es schließlich bei getrennten Rankings für verschiedene Effektgruppen. Die Gründe hierfür liegen zum einen in der zum Teil mangelhaften Datenlage, zum anderen und vor allem aber in den Schwierigkeiten, eine gemeinsame Risikometrik für die vergleichende Risikobewertung zu finden (EPA 1993, 2.2-24).

Unseres Erachtens bietet der in Kapitel 5.2 vorgestellte Ansatz zur multiattributen Bewertung die Möglichkeit zu einer abschließenden, gemeinsamen Risikobewertung und damit zu einem einheitlichen Ranking. Das Problem mangelhafter Daten kann damit allerdings nicht gelöst werden. Wie oben erläutert, basieren multiattribute vergleichende Risikobewertungen auf den Daten der Risikoabschätzungen. Wo diese fehlen, hilft auch der multiattribute Ansatz nicht weiter. Es kann dann allenfalls versucht werden, mit Hilfe qualitativer Beurteilungen zu einer vergleichenden Risikobewertung zu kommen. Ein solches Vorgehen wird weiter unten am Beispiel des WBGU Gutachtens demonstriert.

Übersicht zu den Schritten bei der MRB

Das Vorgehen bei einer MRB kann in drei Phasen mit einer Reihe von Schritten unterteilt werden. In der ersten, vorbereitenden Phase sind zunächst die Teilnehmer des Bewertungsprozesses zu bestimmen, d.h. es müssen Expertenteams und ein Bürgerbeirat ausgewählt werden. Des Weiteren müssen die Ziele der MRB festgelegt werden. Die folgende Analysephase ist Arbeitsgebiet der Experten. Sie umfasst die Risikoabschätzung und -charakterisierung. Letztere ist die Grundlage für die abschließende Phase, das Ranking der Risiken durch den Bürgerbeirat. Diese Vorgehensweise ist aber keinesfalls linear – vor allem, wenn Bürger einbezogen werden. Dann sind im Verfahren Rückkoppelungen vorzunehmen, die sich aus der Zusammenarbeit zwischen Experten und Bürgerbeirat ergeben.

Schritt 1: Zusammensetzung des MRB-Teams

Nach einer ersten Sichtung der Problemliste sind die Expertenteams und der Bürgerbeirat zusammenzustellen. Beide Gruppen sollten an den Entscheidungen über die Einengung des Problembereichs und der Auswahl der Kategorien beteiligt werden. Aufgabe der Experten ist es hierbei vor allem, die Vorstellungen des Bürgerbeirats auf ihre Realisierbarkeit hin zu überprüfen, da hierbei die verfügbare Datenlage eine wesentliche Rolle spielt und die Auswahl der Risikometriken entscheidend von der Festlegung der Kategorien abhängt. Ergebnis dieses Prozessschrittes ist eine Liste von Risiken, die in der MRB berücksichtigt werden sollen.

Im weiteren haben die Experten die Aufgabe, sowohl die zu betrachtenden Schutzgüter wie auch die Attribute und die zu ihrer Messung nutzbaren Risikometriken genauer festzulegen. Sie liefern auch die Risikoabschätzungen und – darauf aufbauend – die Risikocharakterisierung in Form von Datenblättern, die dem Bürgerbeirat für die vergleichende Risikobewertung zur Verfügung gestellt werden.

Schritt 2: Festlegung der Eckpunkte der CRA

Bei der Durchführung der MRB müssen eine Reihe von Entscheidungen getroffen werden. Es geht um die Erstellung einer Problemliste, um die Festlegung der Attribute sowie um die Wahl des Vergleichsverfahrens.

Die Problemliste kann z.B. aus Umfragen gewonnen werden, sie ist aber auch über ein top-down Verfahren zu gewinnen. Hierbei werden für eine Region Problemfelder ausgewählt, z.B. nach Aktivitätsfeldern (Ernährung etc.) und im nächsten Schritt die Problemfelder (z.B. Nahrungsmittel, Trinkwasser), um schließlich dann die Leitindikatoren (z.B. Pestizidbelastung, Acrylamidbelastung) festzulegen. Diese müssen so ausgewählt werden, dass damit die wesentliche Risikoproblematik im Problemfeld abgedeckt wird. Dabei sind schwierige Entscheidungen zu treffen. So wäre zu bestimmen, ob im Bereich der Innenraumluft neben Radon und Passivrauchen noch andere Indikatoren zu betrachten sind (z.B. Formaldehyd). Es ist leicht einzusehen, dass die Auswahl der Kategorien und Indikatoren das Ergebnis der CRA maßgeblich beeinflussen kann.

Um die MRB übersichtlich zu halten, wird man in der Regel die Zahl der zu vergleichenden Risiken auf max. 20 begrenzen. Dies gilt vor allem dann, wenn eine größere Zahl von Attributen für den Vergleich herangezogen werden sollen.

Die Wahl der Attribute richtet sich nach den Schutzgütern, in unserem Fall sind das Gesundheit und eventuell auch die Lebensqualität. Die EPA (1993) hatte hier für die Beurteilung der Gesundheitsrisiken zwischen kanzerogenen und nicht-kanzerogenen Effekten unterschieden. Für beide Effekttypen steht – zumindest im Prinzip – eine Risikometrik zur Verfügung. Für kanzerogene Effekte ist dies die Krebs-Inzidenz, die auf Abschätzungen der Dosis-Wirkungs-Beziehung und der Exposition basiert. Für nicht-kanzerogene Effekte nutzt die EPA das IER (*Individual Exposure Ratio*), das als Quotient aus einer Expositionsdosis und der *Reference Dose* (RfD) berechnet wird. Das IER gibt an, in welchem Ausmaß die aufgenommene Dosis einer Substanz den Wert einer sicheren Dosis überschreitet. Je höher das IER, desto größer die Wahrscheinlichkeit für adverse Gesundheitswirkungen. Allerdings weist die EPA auf eine Reihe konzeptioneller Probleme bei der Verwendung des IER als Risikometrik für CRAs hin (EPA 1993, 2.2-22). Zum Beispiel setzt ein Risikovergleich auf der Basis der IER voraus, dass für alle betrachteten Substanzen die gleiche Dosis-Wirkungs-Beziehung gilt – was nicht der Fall ist.

Ein zentrales Problem dieses Schrittes besteht also in der Auswahl von geeigneten Risikometriken (siehe dazu Anhang 1). Die Entscheidung darüber, welche Risikometriken geeignet sind, können letztlich nur Experten treffen.

Als Alternative schlägt die EPA vor, Risikometriken zu verwenden, die sich direkt auf Schadenskategorien beziehen. Für kanzerogene Risiken ist hier die Krebs-Inzidenz geeignet. Für nicht-kanzerogene Risiken werden drei Schadenskategorien vorgeschlagen:

- *katastrophale Gesundheitseffekte*, z.B. Tod, Verringerung der Lebenserwartung, schwere körperliche oder geistige Behinderungen, Erbkrankheiten;
- *schwerwiegende Gesundheitseffekte*, z.B. Fehlfunktionen im Organ- und Nervensystem sowie Verhaltens- und Entwicklungsstörungen;
- *adverse Gesundheitseffekte*, z.B. Gewichtsverlust, Veränderungen bei Enzymen oder reversible organische Störungen.

In ihrem „*Guidebook*“ diskutiert die EPA noch zwei weitere Aspekte, die für die Bewertung der Gesundheitsrisiken relevant sind: Unsicherheiten in der Risikoabschätzung und sensitive Subpopulationen. Unsicherheiten in der Risikoabschätzung ergeben sich vor allem aus lückenhaften Analysen, fragwürdiger Datenqualität und Unsicherheiten bei der Expositionsabschätzung. Eine Abschätzung dieser Unsicherheiten ist derzeit in den meisten Fällen wohl nur qualitativ möglich, kann aber als eigenständiges Attribut in der MRB berücksichtigt werden.

Sensitive Subpopulationen, das heißt Personengruppen, die auf bestimmte Noxen besonders empfindlich reagieren, stellen ein besonderes Problem für eine MRB dar. Einerseits ist offensichtlich, dass in einer vergleichenden Risikobewertung sensitive Subpopulationen berücksichtigt werden müssen, andererseits ist die (nicht immer nur) wissenschaftliche Diskussion darüber, welche Gruppen in bezug auf welche Stoffe besonders sensitiv sind, äußerst kontrovers. Spezielles Interesse in der aktuellen Diskussion findet die Frage, ob Kinder in besonderem Maße sensitiv für um-

weltbedingte Gesundheitsgefährdungen sind (z.B. Heinemeyer & Dürkop 2002; Landrigan 1999; Tamburlini, von Ehrenstein & Bertollini 2002). Sensitive Gruppen können ebenfalls als eigenständiges Attribut bzw. als Attributgruppe – gegebenenfalls mit einer besonderen Gewichtung – berücksichtigt werden. Falls quantitative Abschätzungen verfügbar sind, können diese, wie oben für die Normalpopulation beschrieben, genutzt werden. Andernfalls muss auch hier auf Einschätzungen mit Hilfe von Ratingskalen zurückgegriffen werden.

Als zweiter Aspekt neben den Gesundheitsrisiken sollen hier noch Risiken für die Lebensqualität betrachtet werden. Üblicherweise werden hierbei ökonomische Effekte betrachtet, zum Beispiel Kosten für das Gesundheitswesen, Verluste in der Landwirtschaft oder Schäden an Bauwerken. Dies sind Schadenskategorien, für die in der Regel quantitative Daten vorliegen, und die in Form von Währungseinheiten (Euro, Dollar) mit einer einheitlichen Metrik gemessen werden können. Neuere MRB-Verfahren berücksichtigen darüber hinaus auch intangible Aspekte der Lebensqualität, wie beispielsweise den Verlust an Erholungsflächen oder „ästhetische“ Belästigungen durch Lärm, Smog oder Gerüche (EPA 1993, Kap. 2.4).

Für diese intangiblen Aspekte ergibt sich das Problem der Messung. Grundsätzlich können auch hier wieder Ratingskalen genutzt werden. Dabei stellt sich die Frage, wer die Ratings abgeben soll – wer also als „Experte“ für die Einschätzung dieser Lebensqualitätsfragen ist. Anders als bei den Gesundheitsrisiken, die nur von Experten abgeschätzt werden können, ist in Fragen der Lebensqualitätsverluste jeder (betroffene) Bürger ein potenzieller Experte. Das Problem ist hier, Beurteilungen zu bekommen, die für die Betroffenen repräsentativ sind. Im einfachsten Fall kann dabei auf ein Bürgerkomitee zurückgegriffen werden – wenn dieses als repräsentativ für die von den Lebensqualitätsverlusten betroffene Bevölkerung gelten kann. Andernfalls müssen repräsentative Surveys durchgeführt werden, um an die notwendigen Daten zu gelangen. Das Problem der repräsentativen Beurteilungen stellt sich auch für ein anderes Verfahren zur Erhebung von Beurteilungen intangibler Aspekte: die Erfassung von Zahlungsbereitschaften (*willingness to pay*).

Die letzte Aufgabe in diesem Schritt ist die Auswahl eines Vergleichsverfahrens. Hier stehen im Prinzip die in Kapitel 5.2 erläuterten Verfahren zur Verfügung. Welchem Verfahren der Vorzug gegeben wird, hängt sowohl von der Komplexität des Problems und den verfügbaren zeitlichen Ressourcen wie auch von dem Anspruch an die Qualität der Entscheidungsfindung ab. Das in Kapitel 5.2 erläuterte, multiattribute Bewertungsverfahren ist zwar sicherlich das aufwendigste, bietet aber auch das höchste Maß an Transparenz und Rationalität bei der Bewertung.

Schritt 3: Risikoabschätzung

Die Durchführung der Risikoabschätzungen ist Sache der Experten. Zu ihrer Methodik gibt es mittlerweile zahllose Abhandlungen, deren Darstellung den Rahmen dieser Arbeit bei weitem sprengen würde (siehe dazu auch Kap. 2.4). Besonders interessant für die Durchführung von CRAs sind neuere Ansätze, die auf eine Harmonisierung der in der Praxis sehr vielfältigen Verfahrensweisen abzielen (EC 2000).

Aus den im Schritt 2 festgelegten Eckpunkten ergeben sich eine Reihe von Anforderungen an die Risikoabschätzung. So muss etwa geprüft werden, für welche der

ausgewählten Problemfelder bzw. Indikatoren bereits Risikoabschätzungen existieren und auf welchen Risikometriken diese basieren.

Liegen keine Abschätzungen für den ausgewählten Raum-Zeit-Bereich vor, so kann man versuchen, diese Abschätzungen aus ähnlich gelagerten Fällen bzw. Situationen zu übertragen. Das gleiche gilt für Attribute, die sich auf sensitive Subpopulationen beziehen. Die dabei unvermeidlichen Abschätzungsunsicherheiten müssen dann bei vergleichender Bewertung beachtet werden. Die explizite Darlegung der mit den jeweiligen Risikoabschätzungen verbundenen Unsicherheiten ist ein elementarer Bestandteil der Kommunikation von Abschätzungsergebnissen.⁹⁹

Schritt 4: Risikocharakterisierung

Zweck der Risikocharakterisierung im Rahmen der MRB ist eine übersichtliche und aussagekräftige Darstellung der Ergebnisse der Risikoabschätzungen. Eine Möglichkeit ist, für jedes Risiko auf einem separaten Datenblatt die Abschätzungsergebnisse für die verschiedenen Attribute zusammenzufassen. Abbildung 19 zeigt ein Beispiel für das Risiko „Schulbusunfall“ aus einer vergleichenden Risikobewertung zu Risiken für Schüler einer amerikanischen Schule (Florig et al. 2000). Das Datenblatt enthält eine kurze Textzusammenfassung der Risikoabschätzung und eine Tabelle mit den quantitativen Abschätzungen für die einzelnen Attribute. In dieser MRB wurden für die Attribute Mortalität und Morbidität verschiedene Metriken angeboten (z.B. Häufigkeiten und Wahrscheinlichkeiten). Damit sollte ein besseres Verständnis des Risikos erzielt werden. Im Rahmen einer multiattributen Bewertung müsste dann allerdings eine dieser Risikometriken für jedes Attribut (Mortalität und Morbidität) ausgewählt werden. Dies muss so geschehen, dass diese Metrik mit den Risikocharakterisierungen für die anderen Risiken kompatibel ist.

Das Datenblatt in Abbildung 19 gibt auch ein Beispiel dafür, wie die Unsicherheiten bei Risikoabschätzungen dargestellt werden können. Angegeben werden nicht nur die beste Schätzung, sondern auch die Werte für die höchsten und niedrigsten Abschätzungen. Andere Möglichkeiten der Kommunikation von Unsicherheiten sind die Angabe von Konfidenzintervallen oder Perzentilen. Natürlich muss bei der Nutzung solcher Darstellungsweisen von Unsicherheit darauf geachtet werden, dass die an der MRB Beteiligten solche Angaben auch verstehen.

In der Praxis werden Risikocharakterisierungen meist umfangreicher und auch komplexer sein, als es in diesem einfachen Beispiel von Florig und Mitarbeitern der Fall ist. Aber auch dann sollte es möglich sein, Risikocharakterisierungen für die einzelnen Risiken in Form zusammenfassender Darstellungen auf Datenblättern anzufertigen, die den Teilnehmern eine übersichtliche Bewertung gestatten. Beispiele für solche umfassenden Datenblätter finden sich im *New Jersey Comparative Risk Project* (NJDEP 2003, Appendix 4).

⁹⁹ Die Bedeutung der Berücksichtigung von Unsicherheiten bei der Risikoabschätzung wird in zahlreichen Arbeiten zu Problemen von Risikoabschätzungen betont. Siehe z.B. Bailer & Bailer (1999); EPA (1996); NRC (1994).

School Bus Accidents			
Summary: Most school bus-related deaths occur among students who are outside the bus either getting on or getting off. Half of school bus injuries occur among students on the bus. At Centerville Middle School half of the 430 students ride the school, almost identical to the national average. Accidents involving more than one death are very rare. Because CMS buses use the Alvaroz Expressway and cross the C&LL rail line, the risk of a catastrophic bus accident in Centerville is estimated to be between four and six times higher than the national average.			
School bus accident risk for Centerville Middle School*			
Student deaths	<i>Low estim.</i>	<i>Best estimate</i>	<i>High estim.</i>
Number of deaths per year	.0001	.0002	.0004
Chance in a million of death per year for the average student	25	0.5	1
Chance in a million of death per year for the student at highest risk	0.5	1	2
Greatest number of deaths in a single episode		20 - 50	
Student illness or injury			
More serious long-term cases per year	.0002	.0006	.002
Less serious long-term cases per year	.0004	.0015	.004
More serious short-term cases per year	.001	.002	.006
Less serious short-term cases per year	.002	.005	.015
Other Factors			
Time between exposure and health effects		immediate	
Quality of scientific understanding		high	
Combined uncertainty in death, illness, injury		1.6 (low)	
Ability of student/parent to control exposure		moderate	
*See "Notes on the Numbers" for definitions and explanations of assumptions.			

Abbildung 19: Beispiel für Datenblatt zur Risikocharakterisierung bei einer MRB (aus: Florig et al. 2000, 919)

Schritt 5: Ranking

Der abschließende Schritt des Ranking kann sich – je nach gewähltem Vergleichsverfahren – unterschiedlich gestalten. Im Prinzip sind hier drei Varianten praktikabel:

- (1) Die Teilnehmer bewerten auf der Basis der Datenblätter jedes Risiko insgesamt und bringen die Risiken dann aufgrund ihres holistischen Urteils in eine Rangreihe. Verfahrenstechnisch ist dies vermutlich der einfachste Weg. Darüber hinaus entspricht er der intuitiven Urteilsweise im Alltag. Dieses Vorgehen liegt deshalb nahe, wenn die vergleichende Bewertung durch einen Bürgerbeirat erfolgen soll. Allerdings stößt dieses holistische Verfahren schnell an die Grenzen der kognitiven Informationsverarbeitung von Menschen, sobald eine größere Zahl von Risiken und Attributen bewertet werden müssen.
- (2) Die Bewertung und das Ranking erfolgen nach einer nicht-kompensatorischen Ordnungsregel, zum Beispiel der lexikographischen Regel.¹⁰⁰ Danach wird von jedem Teilnehmer (bzw. einer Teilnehmergruppe) zunächst das wichtigste Attribut gewählt und für jedes Risiko die Risikoabschätzung auf diesem Attribut aus dem Datenblatt entnommen. Aus diesen Abschätzungen ergibt sich dann die Rangfolge der Risiken. Finden sich für zwei oder mehr Risiken auf einem Attribut gleiche

¹⁰⁰ Siehe dazu oben Kapitel 5.2.2.

Abschätzungen und ergeben sich damit gleiche Rangplätze, kann die Bewertung anhand des zweitwichtigsten Attributs (usw.) fortgeführt werden. Dieses Vorgehen hält zwar den kognitiven Aufwand in Grenzen, ist aber insbesondere dann, wenn es zu unterschiedlichen Rankings kommt, wenig transparent.

- (3) Die vergleichende Risikobewertung erfolgt mit Hilfe eines multiattributen Bewertungsverfahrens, wie es in Kapitel 5.2 erläutert wurde. Die Durchführung eines solchen Verfahrens ermöglicht eine transparente Problemstrukturierung, in der Bewertungsunterschiede und ihr Einfluss auf das resultierende Risikoranking identifiziert und damit auch im Bewertungsprozess thematisiert werden können. Dies führt natürlich nicht automatisch zur Beseitigung solcher Unterschiede, es ist aber eine wesentliche Voraussetzung für eine rationale Diskussion.

Im folgenden soll der Aspekt der Attributstrukturierung und -gewichtung und die sich dabei ergebenden Probleme an dem oben begonnenen EPA-Beispiel weiter illustriert werden.

Die EPA (1993) unterscheidet zwischen kanzerogenen und nicht-kanzerogenen Risiken. Dies legt eine Attributstruktur nahe, wie sie in Abbildung 20a dargestellt ist. Bei der Gewichtung der Attribute ist hier zu beachten, dass die drei terminalen Attribute zusammen maximal nur so wichtig sein können wie ihr übergeordnetes Attribut „nicht-kanzerogen“. Gewichtet man die beiden oberen Attribute gleich stark, so bedeutet dies, dass das terminale Attribut „Krebs-Inzidenz“ schon 50 Prozent des Gesamtgewichts aller Attribute für Gesundheitsrisiken erhält und die drei Attribute des Aspekts „nicht-kanzerogen“ sich die verbleibenden 50 Prozent teilen müssen. In der Abbildung sind die sich ergebenden Gewichte für jede Attributebene eingetragen.

Noch stärker fällt dieser Effekt aus, wenn man kanzerogene Risiken höher gewichtet als nicht-kanzerogene – was vermutlich eine häufig anzutreffende intuitive Bewertung ist. Unmittelbarer und einsichtiger kann die Gewichtung erfolgen, wenn die Attribute für Gesundheitsrisiken nicht-hierarchisch strukturiert werden, wie es in Abbildung 20b dargestellt ist. Hier können alle vier Attribute direkt miteinander verglichen und gegeneinander gewichtet werden. Die sich hier durch eine einfache Ranggewichtung ergebenden Werte zeigen weniger starke Unterschiede.

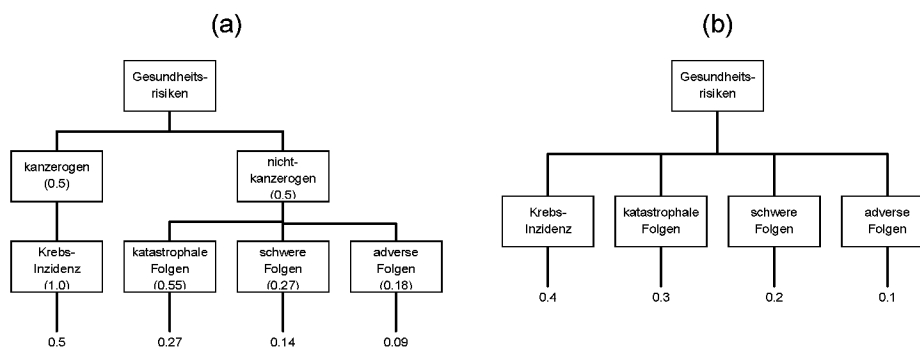


Abbildung 20: Zwei Möglichkeiten zur Strukturierung der Attribute

Eine MRB für diese Attribute lässt sich beispielhaft für drei Risiken durchspielen, zu denen das „Guidebook“ der EPA Daten enthält (EPA 1993, 2.2-31). Der obere Block von Tabelle 28 enthält Abschätzungen der Personenzahlen, die durch Luftschadstoffe, Trinkwasserverunreinigungen und Sondermüll in Bezug auf die vier Schadenskategorien gefährdet sind. Die Risiko-Rangfolge, die sich schon auf der Basis dieser Rohdaten ergibt, ist eindeutig: Die Gesundheitsrisiken durch Luftschadstoffe liegen bei weitem über denen durch Trinkwasser und Sondermüll. An dieser Rangfolge ändert sich auch nichts, wenn die Daten nach dem in Kapitel 5.2.2 erläuterten Verfahren normiert (mittlerer Block der Tabelle) und dann gewichtet werden (unterer Block der Tabelle). Dies ist auch ein Beispiel für eine ausgeprägte Dominanzstruktur in den Risiken. Luftschadstoffe haben auf allen Attributen höhere Werte als Trinkwasser und dieses hat wiederum auf allen Attributen höhere Werte als Sondermüll. In einem solchen Fall sind Normierung und Gewichtung bedeutungslos, denn die Rangordnung ergibt sich schon zwingend aus den Rohdaten.

Tabelle 28: Beispiel für multiattribut Risikobewertung auf der Basis der EPA Daten

Risiken	Krebs-Inzidenz	Katastrophale Folgen	Schwere Folgen	Adverse Folgen	Σ
Luftschadstoffe	12.5	5000	250000	800000	
Trinkwasser	4.0	2000	50000	250000	
Sondermüll	1.2	100	5000	10000	
		↓			
Luftschadstoffe	1.000	1.000	1.000	1.000	
Trinkwasser	0.248	0.388	0.184	0.304	
Sondermüll	0.000	0.000	0.000	0.000	
		↓			
Gewichte	0.400	0.300	0.200	0.100	
Luftschadstoffe	0.400	0.300	0.200	0.100	1.000
Trinkwasser	0.099	0.116	0.037	0.030	0.283
Sondermüll	0.000	0.000	0.000	0.000	0.000

Mit zunehmender Komplexität der Attributstruktur wächst auch die Bedeutung der für die Erhebung der Attributgewichte verwendeten Verfahren. In der Regel wird man hier auf Verfahren zurückgreifen, in denen die einzelnen Teilgruppen der Attribute getrennt gewichtet werden und die absoluten Gewichte durch Ausmultiplizieren der Äste berechnet werden (siehe dazu Kap. 5.2.2).

Beispiel 2: Vorgehensweise bei einer MRB für sehr heterogene Risiken

Insbesondere bei dem Vergleich sehr heterogener Risiken mit unterschiedlichen Schadensfolgen und unterschiedlichem Raum-Zeit-Horizont ist eine quantitative Abschätzung von Schadensfolgen nicht mehr möglich. Hier muss auf eine qualitative Beurteilung der Schadensfolgen zurückgegriffen werden. Diese kann allerdings mit Hilfe von Ratingskalen quantifiziert werden.

Dieses Vorgehen soll am Beispiel des Jahresgutachtens des WBGU (1998) „Welt im Wandel. Strategien zur Bewältigung globaler Umweltrisiken“¹⁰¹ illustriert werden. Ziel des WBGU Gutachtens war zwar nicht, ein Risikoranking zu erstellen, sondern durch diese Charakterisierung unterschiedliche Risikomanagementstrategien zu begründen. Trotzdem bietet die Analyse des WBGU ein gutes Beispiel für die Probleme vergleichender Risikobewertung an, wenn die Beurteilung der Risiken auf den Attributen mit hohen Unsicherheiten verbunden sind bzw. wenn quantitative Abschätzungen nicht möglich sind.

In dem Gutachten des WBGU wird eine sehr heterogene Liste von (potenziellen) Umweltrisiken auf verschiedenen Attributen beurteilt. Diese Attribute beziehen sich zum einen auf den möglichen Schaden (unterschieden in Schadensausmaß und Abschätzungssicherheit für das Schadensausmaß) und die Wahrscheinlichkeit (Eintrittswahrscheinlichkeit und Abschätzungssicherheit für die Eintrittswahrscheinlichkeit). Zum anderen werden Bewertungsaspekte genutzt, die die Risiken hinsichtlich raum-zeitlicher Aspekte als „Proxies“ charakterisieren. Dies sind: Ubiquität, Persistenz, Irreversibilität, Verzögerungswirkung. Ein weiteres Beurteilungskriterium: das Mobilisierungspotenzial, bezieht sich auf die Abschätzung des gesellschaftlichen Konfliktpotenzials eines Risikos. Der WBGU begründet diese Auswahl mit Ergebnissen der Risikowahrnehmungsforschung, wo sich diese Aspekte als bedeutsam für die intuitive Risikobeurteilung erwiesen haben. Abbildung 21 zeigt die Strukturierung der WBGU Kriterien in einer hierarchischen Attributstruktur.

Auf diesen Attributen bewertet der WBGU eine Liste von mehr als zwanzig (potenziellen) Risiken, die von technischen Risiken (z.B. Kernenergie, großchemischen Anlagen) über Gesundheitsrisiken (z.B. AIDS-Infektion, BSE-nv-CJD-Infektion), Umweltrisiken (z.B. Freisetzung transgener Pflanzen) und Stoffrisiken (z.B. persistente organische Verbindungen, endokrin wirksame Stoffe) bis hin zu Klimarisiken (z.B. anthropogen schleichender Klimawandel, Instabilität der westantarktischen Eisschilde) und Risiken aus Naturereignissen (z.B. Überschwemmungen, Erdbeben) reichen.

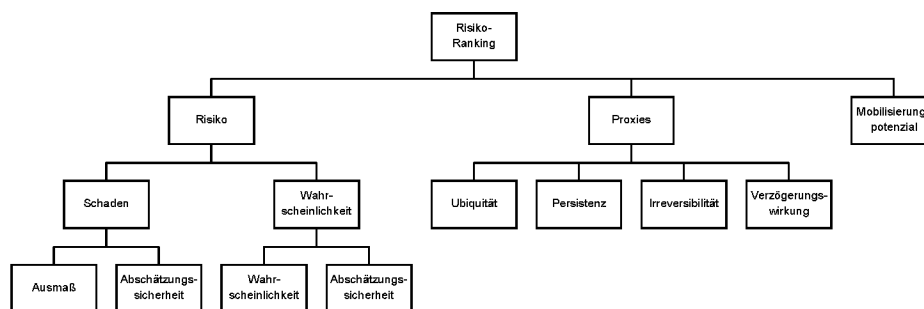


Abbildung 21: Multiattribute Struktur der WBGU Risikobewertung

Die Bewertung dieser Risiken auf den Attributen erfolgte mit Hilfe von vierstufigen Ratingskalen mit verbalen Labeln („gering“, „eher gering“, „eher hoch“, „hoch“ sowie als gesonderte Kategorie: „ungewiss“), wobei die Bewertung auch mehrere Stufen

¹⁰¹ Siehe auch oben Kapitel 3.

umfassen konnte (was durch schraffierte Flächen auf der Skala dargestellt wurde). Die Bewertungen der Risiken selbst basieren auf umfassenden Problemanalysen, die aber eben nicht zu Ergebnissen geführt haben, die es erlauben, eine quantitative Abschätzung auf den Attributen vorzunehmen. Das hohe Maß an Unsicherheit bei praktisch allen betrachteten Risiken erklärt sich dadurch, dass der WBGU sich in diesem Gutachten ausdrücklich nur mit Risiken beschäftigt, die nicht in die Kategorie „Normalrisiken“ fallen und deshalb durch hohe Unsicherheit gekennzeichnet sind.¹⁰²

Dem Unsicherheitsaspekt trägt der WBGU insbesondere durch die beiden Aspekte *Abschätzungssicherheit für Schadensfolgen* und *Abschätzungssicherheit für Eintrittswahrscheinlichkeit* Rechnung. Eine solche Konzeptualisierung von Unsicherheiten bei der Risikoabschätzung mag für den Zweck der Klassifizierung von Risiken zur Begründung von Risikomanagementstrategien – und darum geht es in dem WBGU-Gutachten ja – hilfreich sein. Im Rahmen einer multiattributen Bewertung ergibt sich jedoch das grundsätzliche Problem, dass nicht klar ist, welche Wertigkeit eine hohe bzw. geringe Abschätzungssicherheit hat. Soll eine geringe Abschätzungssicherheit als Hinweis für oder gegen das Vorliegen eines Risikos gewertet werden?

In Bezug auf die Eintrittswahrscheinlichkeit lässt sich die Abschätzungssicherheit als Wahrscheinlichkeitsabschätzung zweiter Ordnung verstehen. Je nach Form, die dieser Wahrscheinlichkeitsverteilung zugewiesen wird (z.B. Rechteck- oder Normalverteilung; siehe Abbildung 22), ergeben sich unterschiedliche Interpretationsmöglichkeiten: Bei der Rechteckverteilung beispielsweise sind alle Werte in dem betrachteten Intervall gleichermaßen wahrscheinlich; hier bietet es sich an, jeweils getrennte Analysen mit der oberen und der unteren Grenze der Verteilung vorzunehmen. Unterstellt man eine (angenäherte) Normalverteilung (was in diesem Beispiel durch die Art der Schraffierung nahegelegt wird),¹⁰³ so kann man den Skalenwert für die Analyse nutzen, der dem Mittelwert entspricht.

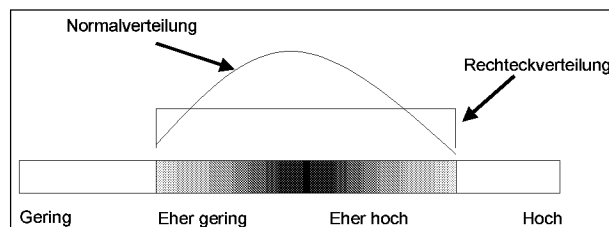


Abbildung 22: Beispiel für mögliche Interpretationen der unscharfen Bewertung auf der Ratingskala des WBGU (1998)

¹⁰² Normalrisiken zeichnen sich folgendermaßen aus: Geringe Ungewissheiten in Bezug auf die Wahrscheinlichkeitsverteilung von Schäden, insgesamt eher geringes Schadenspotential, insgesamt geringe bis mittlere Eintrittswahrscheinlichkeit, geringe Persistenz und Ubiquität (zeitliche und örtliche Ausdehnung), weitgehende Reversibilität des potenziellen Schadens, geringe Schwankungsbreiten von Schadenspotenzial und Eintrittswahrscheinlichkeiten und geringes soziales Konflikt- und Mobilisierungspotenzial (v. a. keine deutlichen Bewertungsdiskrepanzen zwischen der Gruppe der Risikoträger und der Gruppe der Chancen- bzw. Nutzengewinner). (WBGU 1998, 6f.)

¹⁰³ Legt die Schraffierung eine schiefe Verteilung nahe, so kann der Skalenwert, der dem Median entspricht, genutzt werden.

Für eine multiattribut Strukturierung mit Hilfe der WBGU Kriterien kann also das Kriterium Abschätzungssicherheit durch eine getrennte Analyse der oberen und unteren Intervallgrenzen bzw. durch den Mittelwert für die Eintrittswahrscheinlichkeit ersetzt werden. Es ergibt sich damit die in Abbildung 23 dargestellte Attributstruktur.

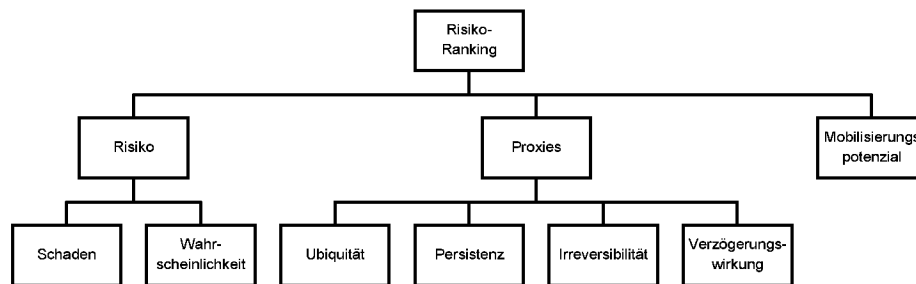


Abbildung 23: Neue Attributstruktur der WBGU Risikobewertung

In dieser Strukturierung bilden Schaden und Wahrscheinlichkeit die beiden zu beurteilenden Attribute für das übergeordnete Attribut Risiko. Zu beachten ist, dass diese Konzeptualisierung von Risiko sich grundsätzlich von der in der Risikoabschätzung üblichen unterscheidet (siehe oben Kapitel 2). Hier werden Schaden und Wahrscheinlichkeit additiv verknüpft. Damit kann ein hoher Schaden nicht durch eine entsprechend kleine Wahrscheinlichkeit kompensiert werden (und *vice versa*). Diese Konzeptualisierung erscheint vor dem Hintergrund der qualitativ gehaltenen Beurteilungen von Schaden und Wahrscheinlichkeit dennoch angemessen. Grundsätzlich wäre es natürlich auch hier möglich, Risiko als Produktterm zu nutzen, zumal der WBGU selbst das übliche Wahrscheinlichkeitsintervall $[0;1]$ als die seinen Eintrittswahrscheinlichkeitsabschätzungen zugrunde liegende Skala definiert.¹⁰⁴ Allerdings fällt es schwer, die extrem geringen Wahrscheinlichkeiten, die für einige der betrachteten Risiken (z.B. Kernenergie; Meteoriten)¹⁰⁵ angenommen werden müssen, zusammen mit den eher hohen Wahrscheinlichkeiten für andere Risiken (Destabilisierung der Ökosysteme, schleichender Klimawandel) auf dieser Skala abzutragen (siehe Tabelle 29).

Im folgenden soll beispielhaft ein auf einer multiattributen Bewertung beruhender Vergleich von einigen der mehr als zwanzig vom WBGU analysierten Risiken vorgenommen werden. Abbildung 24 zeigt noch einmal die Attributstruktur – jetzt mit beispielhaften Gewichtungen. Diese Gewichtungen wurden (der Einfachheit halber)

¹⁰⁴ Der WBGU (1998, 60) erläutert seine Skala zur Eintrittswahrscheinlichkeit folgendermaßen: „gering heißt: „sehr unwahrscheinlich“ (gegen 0)“; „hoch heißt „sehr wahrscheinlich“ (gegen 1)“.

¹⁰⁵ Die Wahrscheinlichkeit für katastrophale Kernkraftunfälle kann mit 10^{-5} bis 10^{-6} pro Betriebsjahr angesetzt werden (NEA 2000, 38). Eine (sehr grobe) Schätzung für die Wahrscheinlichkeit einer globalen Katastrophe durch einen Meteoreinschlag mit bis zu einer Milliarde Toten beläuft sich auf 4 Einschlüsse pro eine Millionen Jahre (Gerrard 2000, 901). Auf einem Workshop der OECD (2003) wurde die Wahrscheinlichkeit, dass sich im 21. Jahrhundert ein katastrophaler Asteroideneinschlag mit dem Potenzial, einen ganzen Kontinent zu verwüsten, ereignet, auf 1:5000 geschätzt. Die Wahrscheinlichkeit für einen global katastrophalen Asteroideneinschlag im 21. Jahrhundert wird auf 10^{-6} geschätzt.

nach dem Ranggewichtungsverfahren vorgenommen (siehe Kapitel 5.2.2). Dem Attribut „Risiko“ wurde Rang 1, „Proxies“ Rang 2 und „Mobilisierungspotenzial“ Rang 3 zugewiesen. Schaden und Wahrscheinlichkeit wurden gleichgewichtet und den vier Proxy-Attributen wurde die Rangreihe Irreversibilität, Ubiquität, Persistenz und Verzögerungswirkung zugewiesen. In der unteren Zeile der Abbildung finden sich die absoluten Gewichte, die später in die Berechnung eingehen. Interessant ist, dass bei dieser spezifischen Strukturierung und dem genutzten Gewichtungungsverfahren das Attribut Mobilisierungspotenzial das dritthöchste Gewicht bekommt, obwohl (auf der obersten Attributebene) das Attribut „Proxies“ ein deutlich höheres Gewicht hat. In Kapitel 5.2.2 wurde darauf hingewiesen, dass geprüft werden muss, ob solche Effekte hierarchischer Attributstrukturierung mit den tatsächlichen Einschätzungen der Attributwichtigkeit durch den Bewerter übereinstimmen.

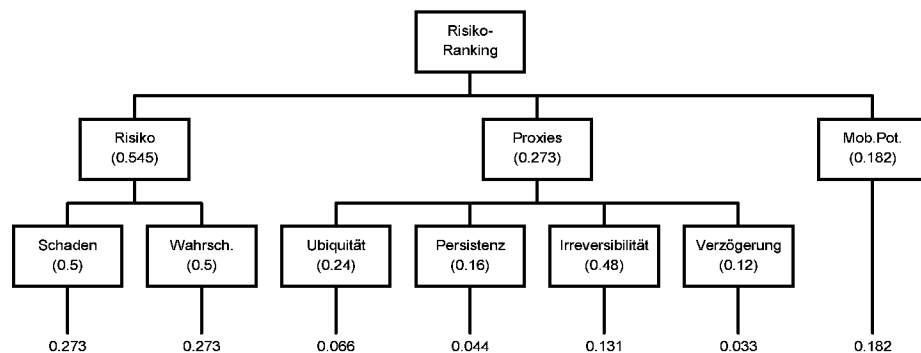


Abbildung 24: Attributstruktur der WBGU Risikobewertung mit absoluten Gewichten (relative Gewichte in Klammern)

Für den Beispiel-Risikovergleich wurden folgende Risiken aus der Liste des WBGU ausgewählt: Destabilisierung der Ökosysteme, EMF, Kernenergie, Meteoriten, natürlicher Klimawandel, schleichender Klimawandel und Überschwemmungen. Im WBGU-Gutachten finden sich für jedes Risiko Beurteilungen auf den Attributen. Diese Beurteilungen liegen in Form von (mehr oder weniger eindeutig) markierten Flächen auf der vierstufigen Ratingskala vor. Für die multiattribut Bewertung wurden dem Label der Ratingskala Zahlenwerte zugeordnet: „gering“ = 1; „eher gering“ = 2; „eher hoch“ = 3; „hoch“ = 4. Bei unscharfen Markierungen, die sich über mehrere Skalenstufen erstrecken, wurde ein Mittelwert geschätzt (siehe oben das Beispiel in Abbildung 22). Der obere Block in Tabelle 29 zeigt die so extrahierten Beurteilungen; die Spaltenspalte ganz rechts enthält die jeweilige Zeilensumme (und gibt damit einen ersten Hinweis auf die Rangfolge der Risiken).

Im nächsten Schritt wurden – wie in Kapitel 5.2.2 erläutert – diese Rohwerte linear auf einen Risikoindex mit dem Intervall [0;1] normiert. Der mittlere Block in Tabelle 29 zeigt in der obersten Zeile die Gewichte, wie sie in Abbildung 24 wiedergegeben sind, und darunter die (noch ungewichteten) normierten Werte des Risikoindex. Im letzten Schritt schließlich werden diese normierten Werte mit den Gewichten multipliziert. Das Ergebnis ist im unteren Block von Tabelle 29 wiedergegeben. Die letzte Spalte enthält die für den Vergleich und das Ranking wesentliche Information. Sie enthält die Summe der gewichteten Einzelbeurteilungen auf den Attributen.

Tabelle 29: Beispiel für multiattribut Risikobewertung auf der Basis der WBGU Daten

Risikoquelle	Eintrittswahrscheinlichkeit	Schadensausmaß	Ubiquität	Persistenz	Irreversibilität	Verzögerungswirkung	Mobilisierungspotenzial	Σ
Dest. Ökosysteme	3.500	4.000	4.000	3.700	3.700	4.000	3.000	25.900
EMF	2.000	2.000	3.500	1.500	3.000	2.000	4.000	18.000
Kernenergie	1.000	4.000	3.700	3.700	3.700	1.700	3.700	21.500
Meteoriten	1.200	3.600	3.600	3.300	3.700	1.000	1.000	17.400
Nat. Klimawandel	2.000	3.000	2.500	1.300	2.000	1.300	2.000	14.100
Schl. Klimawandel	3.000	3.700	3.600	3.700	3.600	3.700	2.000	23.300
Überschwemmungen	1.300	3.600	1.700	3.500	1.300	1.300	1.200	13.900
				↓				
Gewichte	0.273	0.273	0.066	0.044	0.131	0.033	0.182	1.000
Dest. Ökosysteme	1.000	1.000	1.000	1.000	1.000	1.000	0.667	6.667
EMF	0.400	0.000	0.783	0.083	0.708	0.333	1.000	3.308
Kernenergie	0.000	1.000	0.870	1.000	1.000	0.233	0.900	5.003
Meteoriten	0.080	0.800	0.826	0.833	1.000	0.000	0.000	3.539
Nat. Klimawandel	0.400	0.500	0.348	0.000	0.292	0.100	0.333	1.973
Schl. Klimawandel	0.800	0.850	0.826	1.000	0.958	0.900	0.333	5.668
Überschwemmungen	0.120	0.800	0.000	0.917	0.000	0.100	0.067	2.003
				↓				
Dest. Ökosysteme	0.273	0.273	0.066	0.044	0.131	0.033	0.121	0.941
EMF	0.109	0.000	0.052	0.004	0.093	0.011	0.182	0.450
Kernenergie	0.000	0.273	0.057	0.044	0.131	0.008	0.164	0.677
Meteoriten	0.022	0.218	0.055	0.037	0.131	0.000	0.000	0.462
Nat. Klimawandel	0.109	0.137	0.023	0.000	0.038	0.003	0.061	0.371
Schl. Klimawandel	0.218	0.232	0.055	0.044	0.126	0.030	0.061	0.765
Überschwemmungen	0.033	0.218	0.000	0.040	0.000	0.003	0.012	0.307

Das Ergebnis dieses multiattributen Bewertungsprozesses ist noch einmal in Abbildung 25 wiedergegeben. Sie zeigt für diese Auswahl von Risiken, dass auf der Basis der WBGU-Beurteilungen die Destabilisierung von Ökosystemen, der schleichende Klimawandel und die Kernenergie die deutlich größten Risiken sind. Interessant ist, dass Meteoriten und EMF ein nahezu gleich großes Risiko haben. An diesem Fall lässt sich auch die Bedeutung von Sensitivitätsanalysen illustrieren. Ein Blick in Tabelle 29 zeigt, dass Meteoriten und EMF sich am deutlichsten auf dem Attribut „Mobilisierungspotenzial“ unterscheiden. Oben war schon darauf hingewiesen worden, dass durch die Art der Strukturierung und Gewichtung dieses Attribut ein relativ hohes Gewicht erhalten hat. Eine Sensitivitätsanalyse, in der das Gewicht für „Mobilisierungspotenzial“ systematisch verringert und das der beiden anderen Attribute „Risiko“ und „Proxies“ proportional erhöht wird, ergibt, dass tatsächlich EMF und Meteoriten ihre Ränge tauschen, sobald das Gewicht von „Mobilisierungspotenzial“ ungefähr halbiert wird. Eine noch stärkere Verringerung des Gewichts von „Mobilisierungspotenzial“ führt allerdings zu keinen weiteren Veränderungen in der Rangfolge.

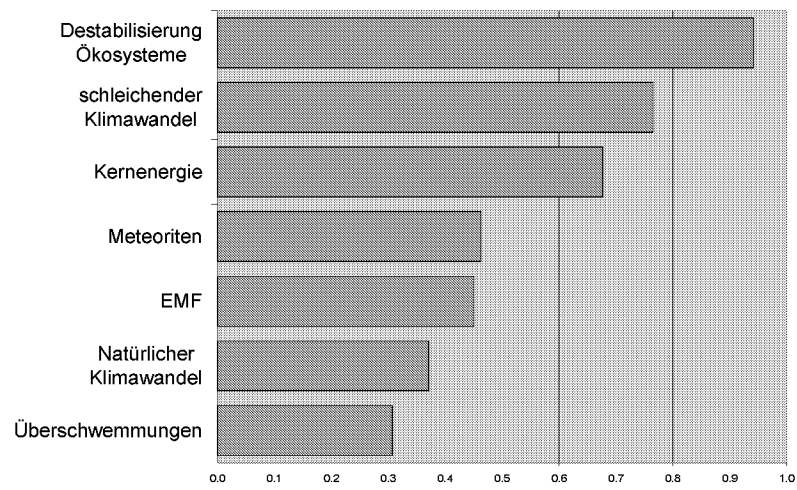


Abbildung 25: Rangfolge der Umweltrisiken auf der Basis des multiattributen Risikoindex

Anhang 3: Vergleichende Bewertung bei undeutlichen Risiken

Die wesentliche Problematik bei der Bewertung undeutlicher Risiken (BUR) besteht darin, dass verschiedene Bewerter bereits bei der Beurteilung der wissenschaftlichen Befunde zu unterschiedlichen Ergebnissen kommen können. Aus diesem Grund soll bei der Darstellung der BUR dieser Fall in den Mittelpunkt gerückt werden. Gegenstand der BUR ist der Vergleich von Bewertungen eines undeutlichen Risikos durch verschiedene Bewerter mit dem Ziel der Konsensbildung hinsichtlich der Beweislage für das Vorliegen eines Gefährdungspotenzials (siehe Tabelle 30). Der Übertrag auf die vergleichende Bewertung mehrerer undeutlicher Risiken ergibt sich relativ einfach.

Tabelle 30: Profil der BUR

Zielstellung	Bestimmung des Gefährdungspotenzials (Hazard); Ermittlung der Dissense/ Konsense zwischen verschiedenen Bewertern bei der Identifikation eines Hazards
Anwendungsbereich	Undeutliche Risiken
Fokus	Evidenzen für ein Gefährdungspotenzial
Beteiligung	Experten mit unterschiedlichen Hazardbewertungen, Bürgerbeirat
Output	Evidenzbewertungen und Evidenz-Rankings

Die amerikanische Umweltbehörde EPA, der australische National Health and Medical Research Council NHMRC, die Cochrane Collaboration sowie die Internationale Agentur für Krebsforschung IARC, die OECD und die EU haben eine Reihe von Standards entwickelt, auf die aufgebaut werden kann. Dabei handelt es sich um Leitlinien zur Literatursuche, zur Analyse und zur Darstellung von wissenschaftlichen Untersuchungen, zur Hazardidentifikation und zur Hazardcharakterisierung.¹⁰⁶ Diesen Leitlinien folgen wir bei unseren Vorschlägen zur BUR.

Im folgenden wird die BUR am Beispiel der Bewertung der Gefährdungspotenziale hochfrequenter elektromagnetischer Felder dargestellt, da es sich hierbei um einen prototypischen Fall handelt, der seit den letzten Jahren kontrovers diskutiert wird und immer wieder zu Konflikten führt. Einzelheiten dazu finden sich im DIFU- und im WIK-Gutachten (Drücke et al. 2003; Büllingen et al. 2002, siehe aber auch Wiedemann et al. 2001).

Übersicht zu den Schritten der BUR

Wie Abbildung 26 zeigt, sind hierbei eine Reihe von Aufgaben zu lösen. Zuerst ist die Zielstellung der BUR zu präzisieren. Weiterhin ist das BUR-Team zu bilden, das aus Experten und Bürgerbeirat besteht. Danach nehmen die Experten die Arbeit auf. Die erste Aufgabe besteht in der Bestimmung der Datenbasis für ihre Hazardabschätzung. Des weiteren sind die Studien zu den verschiedenen Endpunkten zusammen-

¹⁰⁶ Siehe hier u.a.: ICNIRP Statement 2002, National Health and Medical Research Council, 1999, 2000, The Presidential/Congressional Commission on Risk Assessment and Risk Management Final Report Volume 2, 1997, National Research Council: Understanding Risk 1996. EPA Risk Assessment Guidelines (<http://www.epa.gov/ncea/raf/rafguid.htm>), Technical guidance document

zufassen und zu bewerten. Danach folgt die Hazardcharakterisierung in Form von Pro- und Kontra-Argumenten, auf der die Bewertung durch den Bürgerbeirat aufbaut.

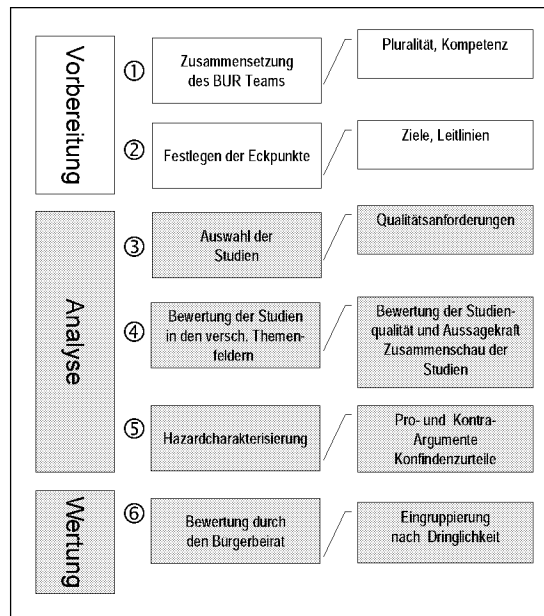


Abbildung 26: Arbeitsschritte bei der BUR

Schritt 1: Bestimmung des BUR-Teams

Ausgangspunkt einer BUR ist in der Regel ein Dissens über das Vorliegen eines Risikos, d.h. im Mittelpunkt stehen undeutliche Risiken.

Wie in Kapitel 6 bereits ausgeführt, sind Expertenteams und ein Bürgerbeirat zu etablieren. Aufgabe des Expertenteams ist die Abschätzung der Evidenzbasis für eine Risikoannahme. Aufgabe des Bürgerbeirats ist die Gewichtung von Pro- und Kontra-Hazard-Argumenten oder die Einordnung in eine Handlungsklasse (Prioritätensetzung).

Da der Bürgerbeirat erst nach den Arbeiten der Experten tätig werden kann, soll die Expertenauswahl zuerst behandelt werden.

Expertenteam: Die Zusammenstellung des Expertenteams ist bei der BUR von essenzieller Bedeutung. Zum einen geht es ja gerade um unterschiedliche Einschätzungen des Gefährdungspotenzials seitens der Experten. Zum anderen sind komplexe Detailfragen zu klären, die das Hinzuziehen von sachkundigen Experten für die einzelnen Themenbereiche erforderlich machen.

Wir schlagen deshalb vor Experten einzubeziehen, die die Pole des Dissens repräsentieren und zusätzlich solche Experten zu beteiligen, die in jeweils einem Themenfeld / Untersuchungsbereich über spezielle Fachkompetenz verfügen, aber nach

Möglichkeit sich in der Vergangenheit nicht an der Kontroverse über das Vorliegen eines Gefährdungspotenzials beteiligt haben (siehe Abbildung 27).

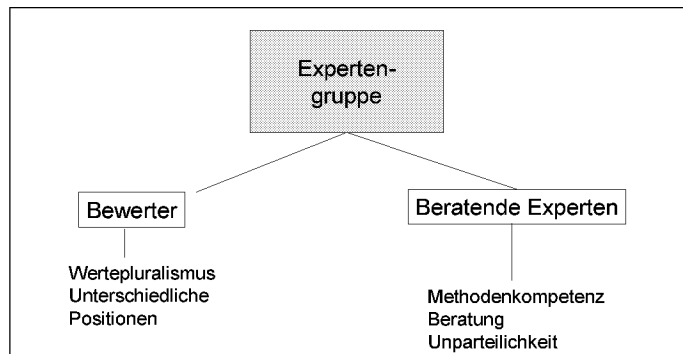


Abbildung 27: Zusammensetzung der Expertengruppe

Die Gründe dafür sind im Einzelnen:

- Für die Hazardbewertung müssen Forschungsergebnisse aus zahlreichen, zum Teil sehr unterschiedlichen Themenfeldern bewertet werden. Da die Bewertung ganz unterschiedliche Kompetenzen verlangt, ist Sorge zu tragen, dass möglichst alle wesentlichen Bereiche vertreten sind.
- Neben dem selbstverständlichen Kriterium der Sachkompetenz für ein Fachgebiet sind noch zwei weitere Kriterien für die Auswahl der beratenden Experten zu beachten: Sie sollten unabhängig von wirtschaftlichen oder politischen Interessen in der jeweiligen Risikodebatte sein. Damit soll sichergestellt werden, dass die beratenden Experten nicht unter den Vorwurf der Parteistellung geraten, der in der öffentlichen Diskussion hochpolitisierter Themen schnell geäußert wird, etwa um unliebsame Einschätzungen zu diskreditieren.

Zur besseren Transparenz des in der Gruppe vorhandenen Expertenwissens sollten Kompetenzprofile erstellt und öffentlich zugänglich gemacht werden. Das heißt, jeder Experte hat anzugeben, über welche wissenschaftlichen Kernkompetenzen er oder sie verfügt, welche Arbeitsgebiete er oder sie vertritt und welchen wissenschaftlichen Gremien er oder sie angehört. In ähnlicher Weise argumentiert auch die Cochrane Collaboration (2003).

Vorschläge für die Zusammensetzung der Expertengruppe sollten auch unter den beteiligten Experten konsensfähig sein, so dass die Arbeitsfähigkeit gewahrt bleibt.

Der Bürgerbeirat sollte verschiedene Teile der Gesellschaft repräsentieren, z.B. Kommunen, Verbände, Industrie und Bürger/Innen. Es kann aber auch nach dem Modell der Planungszelle verfahren werden und eine nach dem Zufallsprinzip ausgewählte Gruppe von nicht-organisierten Bürgern und Bürgerinnen einbezogen werden.

Der Bürgerbeirat sollte in die Evidenzbewertung einbezogen werden. Ihm kommt die Aufgabe zu, die Pro- und Kontra-Argumente zu bewerten (siehe dazu Schritt 6).

Schritt 2: Festlegung der Eckpunkte der BUR

Wesentlich ist, dass eine präzise definierte Zielstellung vorliegt. Zwar ist klar, dass es um die Einschätzung der Evidenzen für ein undeutliches Risiko geht. Was spricht für das Vorliegen eines Risikos, was dagegen? Die Schwierigkeiten liegen aber – wie immer – in den Details. Die erste notwendige Klärung betrifft die Datenbasis. Zur Auswahl stehen wissenschaftliche Untersuchungen, professionelles Wissen von Ärzten über Einzelfälle sowie Laienwissen. Welches Wissen soll ein- und welches ausgeschlossen werden? Und wenn alle drei Wissensklassen betrachtet werden, wie soll im Streitfall entschieden werden?

Bei der Hazardanalyse sollte dem wissenschaftlichen Wissen die entscheidende Rolle eingeräumt werden. Dies bedeutet jedoch nicht, dass den anderen Wissensklassen keinerlei Bedeutung zukommt. Professionelles und Laienwissen sollten bei der Erörterung eines Risikoverdachts hinzugezogen werden; wegen der bekannten Fehlermöglichkeiten intuitiver Urteilsprozesse (siehe Nisbett & Ross 1980) sind sie jedoch für die Prüfung einer Risikohypothese nicht geeignet.

Eine zweite Klärung muss bezüglich der Ziele der Abschätzung des Gefährdungspotenzials (Hazards) durchgeführt werden. Dabei spielt die Diskussion um das Vorsorgeprinzip eine Rolle. Geklärt werden müssen die zentralen Begriffe, mit denen die Befundlage charakterisiert werden soll. Für die Verwendung von Begriffen wie Nachweis, Verdacht, Hinweis, Adversität etc. muss Einigkeit über operationale Definitionen erzielt werden. Wo immer möglich, sollte hier auf eingeführte Definitionen anerkannter Gremien (z.B. WHO) zurückgegriffen werden.

Eine weitere Frage ist, ob neben den gesundheitsbezogenen Endpunkten weitere Attribute für die Hazardanalyse ausgewählt werden sollten. Insbesondere betrifft dies das sogenannte Mobilisierungspotenzial (siehe WBGU 1999), das über Risikowahrnehmung und Mechanismen der sozialen Verstärkung bestimmt wird. Der Einbeziehung solcher Attribute stehen wir skeptisch gegenüber, da es zweifelhaft erscheint, dass diese zur Hazardanalyse einen Beitrag leisten können. Gleiches gilt für die WBGU Kriterien Ubiquität, Irreversibilität, Verzögerungswirkung und Persistenz. Diese können allenfalls, wenn eine Hazardidentifikation nicht gelingt, als zusätzliche Kriterien für ein vorsorgendes Management herangezogen werden.

Schritt 3: Auswahl relevanter Studien

Für die BUR ist zwischen den Bewertern eine Einigung darüber zu erzielen, welche Studien in die Bewertung der gesundheitlichen Auswirkungen einbezogen werden sollen. Vorklärungen sind zwar bereits in Schritt 1 (Zielstellungen) getroffen. Es muss aber noch eine einheitliche Konzeption für die Auswahl und Bewertung der wissenschaftlichen Literatur sowie der Erstellung des wissenschaftlichen Gesamtbildes entwickelt werden.

Relevant ist auch das Verfahren zur Literatursuche. Hier kann auf die Suchstrategie zur Identifikation von Meta-Analysen in Cinahl und Medline¹⁰⁷ zurückgegriffen werden; weiteres findet sich im Cochrane Reviewers Handbook. Denn es zeigt sich,

¹⁰⁷ <http://www.york.ac.uk/inst/crd/search.htm>

dass verschiedene Bewerter – trotz gleicher Auftragslage – bei der Zusammenstellung der Literatur beträchtlich differieren können (vgl. Wiedemann, Schütz & Thalmann 2002).

Für die Auswahl und Bewertung wissenschaftlicher Literatur liefert das *Cochrane Reviewers Handbook* – speziell zu randomisierten, kontrollierten Studien – sehr detaillierte Informationen zu allen Facetten einer systematischen Literatursuche und -bewertung.¹⁰⁸ Ähnlich ausführlich ist das Handbuch des englischen NHS Centre for Reviews and Dissemination.¹⁰⁹ Bei der Auswahl wissenschaftlicher Studien zu den gesundheitlichen Wirkungen sind vor allem Qualitätsanforderungen festzulegen. Im Mittelpunkt stehen die folgenden Fragen:

- Wie zuverlässig ist die Expositions-Abschätzung?
- Welche Anforderungen werden an die Spezifität und Sensibilität der verwendeten Verfahren gestellt?
- Was wird als gesundheitlich relevanter Effekt betrachtet (Adversitätsfrage)?
- Wie wird mit Heterogenität umgegangen?

Darüber hinaus ist eine Einigung darüber zu erzielen, welches Protokoll zur Erfassung der relevanten Studien verwendet werden soll, und nach welchen Kriterien die Bewertung der Qualität der Studien zu erfolgen hat (siehe Tabelle 31).

Tabelle 31: Beispiele für Kriterien für die Qualitätsbewertung zur Entscheidung über die Einbeziehung einer Studie

	Gewichtung hoch	Gewichtung niedrig
Tierversuche	Exposition zuverlässig abgeschätzt	Expositionsmessung eher fehlerhaft
	Niedrige Rate	Hohe Spontan-Tumor-Inzidenz
	Hohe Spezifität des Untersuchungsverfahrens	Geringe Spezifität des Untersuchungsverfahrens
	Dosis-Wirkungs-Beziehung	Keine Daten
Untersuchungen am Menschen	Exposition zuverlässig abgeschätzt	Expositionsmessung eher fehlerhaft
	Hohe Spezifität des Untersuchungsverfahrens	Geringe Spezifität des Untersuchungsverfahrens
	Kontrolle von Biases und Confoundern	Geringe oder keine Kontrolle

Im Protokoll muss ausgewiesen werden, welche Studien aus welchen Gründen ausgeschlossen wurden.

¹⁰⁸ Clarke, M. & Oxman, A.D. (eds.) (2002): *Cochrane Reviewers Handbook* 4.1.6 [updated January 2003]. In: *The Cochrane Library*, Issue 2, 2002. Oxford: Update Software. Updated quarterly. Im Internet unter: [Cochrane Collaboration: http://www.cochrane.org/cochrane/resource.htm](http://www.cochrane.org/cochrane/resource.htm)

¹⁰⁹ NHS Centre for Reviews and Dissemination (2001): *Undertaking Systematic Reviews of Research on Effectiveness: CRD's Guidance for Carrying Out or Commissioning Reviews* (2nd Edition 2001). Im Internet unter: <http://www.york.ac.uk/inst/crd/report4.htm>

Schritt 4: Prüfung der wissenschaftlichen Untersuchungen in den verschiedenen Themenfeldern

Nach der Einigung auf die Datenbasis steht die Prüfung der Aussagekraft der Untersuchungen an. Als Referenzmethode sei hier auf die evidence based medicine verwiesen, die Regeln vorgibt, wie bei der Prüfung von wissenschaftlichen Untersuchungen zur Erstellung eines wissenschaftlichen Gesamtbildes vorzugehen ist.¹¹⁰ Für den Bereich elektromagnetischer Felder heißt das, die vorhandenen Untersuchungen sind im Hinblick auf die in Tabelle 32 aufgeführten Kriterien zu bewerten.

Tabelle 32: Beispiele für Kriterien zur Bewertung von Studien

Kriterium	Erläuterung
Aussagekraft des Untersuchungsansatzes (Generelle Eignung zur Prüfung eines Kausalzusammenhangs)	Untersuchungen können zu falsch positiven oder zu falsch negativen Resultaten führen. Sie sind deshalb kritisch zu prüfen. Zuerst ist die prinzipielle Eignung der Untersuchungen zu beurteilen. Ist die Fehleranfälligkeit der in den Studien angewandten Tests bzw. Messinstrumente tolerierbar?
Absicherung der Befunde gegen den Zufall, Störvariablen und Messfehler	Auch Resultate von im Prinzip geeigneten Untersuchungsansätzen sind kritisch zu prüfen: So ist zu entscheiden, ob die Befunde nicht auch durch Zufall produziert sind. Sind Stichproben- und andere Fehler minimiert? Sind Störvariablen ausgeschaltet oder zumindest kontrolliert?
Zuverlässigkeit der Expositionsabschätzung (Dosimetrie)	Die Expositionen müssen zuverlässig gemessen sein.
Relevanz des Effekts für die Bewertung gesundheitlicher Auswirkungen auf den Menschen	Die gesundheitliche Relevanz der Studien für den Menschen ist von zwei Faktoren abhängig: Zum einen, ob die Studie, wenn sie nicht am Menschen durchgeführt wurde, auf den Menschen übertragbar ist, und zum anderen davon, ob der gewählte Messpunkt gesundheitlich relevant ist. Dabei sind vier Abstufungen denkbar: - Die Effekte sind gesundheitsbeeinflussend - Die Effekte weisen plausibel auf eine Schädigung hin - Für die Effekte kann weder ausgeschlossen noch angenommen werden, dass sie gesundheitsschädlich sind - Die Effekte liegen in einem auch für Langzeitexpositionen tolerierbaren Bereich

Für jedes Themenfeld ist eine Zusammenfassung zu erstellen, die Grundlage für die spätere Ausformulierung der Pro- und Kontra-Argumente bildet. Die Gutachter haben sich dabei auf eine Darstellungsform zu einigen, die den Kriterien in Tabelle 32 folgen sollte.

Schritt 5: Hazard-Charakterisierung in den Themenfeldern

Auf dieser Zusammenfassung aufbauend, muss das wissenschaftliche Gesamtbild in den verschiedenen Untersuchungsbereichen charakterisiert werden. Dabei sind Unstimmigkeiten und Widersprüche zwischen den Untersuchungen – so weit es geht – zu verdeutlichen.

¹¹⁰ Vgl. National Health and Medical Research Council (2000): „How to use the evidence: assessment and application of scientific evidence“ (online verfügbar unter: <http://www.nhmrc.health.gov.au/publications/pdf/cp69.pdf>)

Zentral sind die Pro- und Kontra-Argumente, die für oder gegen das Vorhandensein eines Hazards sprechen. Sie müssen deutlich herausgearbeitet werden.¹¹¹ Als Pro-Argument fassen wir alle Argumente auf, die auf ein Gefährdungspotenzial beim Menschen hinweisen. Als Kontra-Argument, betrachten wir alle Argumente, die gegen einen solchen Verdacht sprechen. Diese qualitative Hazard-Charakterisierung soll dem Bürgerbeirat als Bewertungsgrundlage dienen. Deshalb ist hier insbesondere auf Transparenz und Verständlichkeit zu achten.

Notwendige Voraussetzung für das weitere Vorgehen bei der BUR ist ein Konsens über die Pro- und Kontra-Argumente. Denn jeder Bewerter hat diese Argumente zu gewichten. Dabei muss jedoch keine Einigung erzielt werden. Vielmehr wird damit deutlich, wie sich die Bewerter in der Beurteilung des Hazards unterscheiden.

Eine quantitative Bewertung der Pro- und Kontra-Argumente lässt sich auf verschiedene Weise durchführen. Zunächst muss eine Abstufung von Konfidenzgraden vorgenommen werden. Ein Beispiel hierfür ist das California EMF Projekt (Neutra, DelPizzo & Lee 2002). Dort werden sieben Abstufungen für das Ausmaß von Konfidenz vorgenommen, denen Zahlenwerte zugeordnet sind. Diese drücken das Ausmaß der Konfidenz auf einer Skala von 0 bis 100 Prozent Konfidenz aus (siehe Tabelle 33).

Tabelle 33: Definitionen des Konfidenzgrades (übersetzt nach Neutra, DelPizzo & Lee 2002, 61)

Qualitative Beschreibung	% Konfidenz
Bin praktisch sicher, dass EMF Exposition das Risiko in gewissem Maße erhöht	> 99.5
Glaube stark, dass EMF Exposition das Risiko in gewissem Maße erhöht	90 – 99.5
Bin geneigt zu glauben, dass EMF Exposition das Risiko in gewissem Maße erhöht	60 – 90
An der Grenzlinie zwischen Glauben und Nicht-Glauben, dass EMF Exposition das Risiko in gewissem Maße erhöht	40 – 60
Bin geneigt zu glauben, dass EMF Exposition das Risiko nicht in irgendeinem Maße erhöht	10 – 40
Glaube stark, dass EMF Exposition das Risiko nicht in irgendeinem Maße erhöht	0.5 – 10
Bin praktisch sicher, dass EMF Exposition das Risiko nicht in irgendeinem Maße erhöht	< 0.5

Die Konfidenzurteile geben an, wie gut die Evidenz dafür eingeschätzt wird, dass für den jeweiligen Endpunkt Evidenz für einen Kausalzusammenhang zwischen Exposition und adversen Effekt und damit ein Hazard vorliegt. Die Konfidenzurteile ermöglichen auch einen direkten Vergleich der Experten hinsichtlich ihrer Evidenzeinschätzung. Als zusammenfassender Kennwert kann hier das arithmetische Mittel herangezogen werden.

Schritt 6: Bewertung durch den Bürgerbeirat

Auf der Basis der Pro- und Kontra-Argumente, der Konfidenztabelle sowie der Kompetenzprofile der Experten können dann die Vertreter des Bürgerbeirats ein Ranking vornehmen und das bewertete Hazardpotenzial einer Dringlichkeitsklasse zuordnen. Dazu können Kategorien wie „keine Dringlichkeit“, „mittlere Dringlichkeit“ und „hohe Dringlichkeit“ verwendet werden. Eine Alternative dazu ist die Zuordnung zu verschiedenen Vorsorgeoptionen.

¹¹¹ Siehe dazu das Handbuch der EPA zur Risikocharakterisierung (EPA 2000)

Um einem möglichen Missverständnis vorzubeugen: Der Bürgerrat soll keine Evidenz beurteilen, er soll jedoch auf der Basis der Bewertung der Evidenz durch die Experten und in Kenntnis der Expertendifferenzen und ihrer Gründe (d.h. der jeweiligen Gewichtungen der Pro- und Kontra-Argumente) zu einer reflektierten Einschätzung der Notwendigkeit des Handlungsbedarfs kommen.

Anhang 4: Vergleichende Bewertung von Störfallrisiken

Bei Störfallrisiken liegt die risikotypische Ungewissheit weniger darin, ob eine Exposition zu einem Schaden führt, sondern hauptsächlich darin, ob überhaupt eine Emission stattfindet. Für die Risikocharakterisierung ist im Unterschied zu den Normalrisiken als weiteres Attribut die Wahrscheinlichkeit der Realisierung der Noxe (z. B. der Emission) zu berücksichtigen. Deren Skala reicht von zwar nicht alltäglichen, aber doch im menschlichen Erfahrungsbereich liegenden Ereignissen (z. B. Störfall in einer Chemieranlage, Flugzeugabsturz) bis zu Ereignissen, die so unwahrscheinlich sind, dass man praktisch nicht damit rechnet (z. B. Flugzeugabsturz auf eine Innenstadt, Bruch eines Staudamms) und die man deshalb als hypothetisch bezeichnen kann.

Die Durchführung einer vergleichenden Bewertung von Störfallrisiken (VSR) unterscheidet sich von der für die vergleichende multiattribut Risikobewertung (MRB) für Normalrisiken nur geringfügig. Wie Tabelle 34 zeigt, sind Zielstellung, Fokus, Beteiligung und Output die gleichen wie für Normalrisiken. Der Unterschied liegt darin, dass die Wahrscheinlichkeit der Realisierung der Noxe nicht (zumindest nicht generell) linear gegen das Schadensausmaß verrechnet werden kann, und dass daher die unterschiedlichen Wahrscheinlichkeitsbereiche getrennt bewertet werden müssen. Auch die Organisation und Durchführung einer VSR entspricht den für die MRB (in Anhang 2) geschilderten Schritten. Das soll deshalb hier nicht noch einmal wiederholt werden.

Tabelle 34: Profil der VSR

Zielstellung	Vergleich verschiedener Störfallrisiken auf einem oder mehreren Attribut(en) durch einen oder mehrere Bewerter
Anwendungsbereich	Störfallrisiken, d.h. Risiken, bei denen die risikotypische Ungewissheit darin liegt, ob überhaupt eine Exposition mit der Noxe stattfindet
Fokus	Abschätzung von Gesundheitsrisiken
Beteiligung	Experten und ggf. Bürger
Output	Risiko-Ranking

Im folgenden sollen deshalb nur noch einige für Störfallrisiken spezifische Probleme am Beispiel der vergleichenden Bewertung von Technikrisiken dargestellt werden.

Zunächst ist die Klärung und Eingrenzung des Problembereichs sowie Auswahl der zu betrachtenden technischen Systeme erforderlich. Betrachtet werden im folgenden beispielhaft Technikrisiken für die Bereiche Energie, Mobilität und Chemie. Das heißt, es geht um Risiken, die mit der Energieerzeugung, mit Transportsystemen oder chemischen Produktionsanlagen verbunden sind. Diese betrachteten Technikrisiken sind Risiken, die sowohl eine Normalrisiko- wie auch eine Störfallrisikokomponente haben.

Eine anschauliche Möglichkeit der gleichzeitigen Darstellung von Schadenshäufigkeit und Schadensumfang ist ein Schadens-Häufigkeits-Diagramm. Dies kann ein kontinuierliches Diagramm sein, wie dies z. B. in DRS-A (1979) angegeben wurde. Darin sind die Daten für Flugzeugkatastrophen, Grubenkatastrophen, Zugkatastrophen sowie Brände und Explosionen wiedergegeben. Diese Daten beruhen zum Teil auf

statistischen Daten für Großbritannien, zum Teil auf Extrapolation aufgrund weltweiter Daten¹¹² (Abbildung 28).

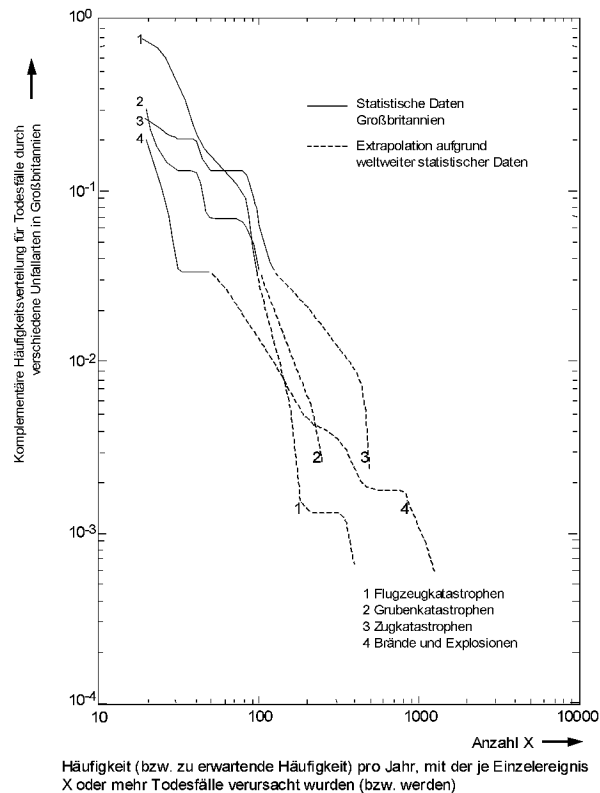


Abbildung 28: Schadens-Häufigkeits-Diagramm (nach DRS-A, 1979)

Aus Abbildung 28 lässt sich ablesen, dass es keine für alle Häufigkeitsbereiche gültige Rangfolge gibt. Vielmehr hängt das Ranking davon ab, welcher Abschnitt der Schadens-Häufigkeits-Verteilung betrachtet wird. Beispielsweise ergibt sich für Brand- und Explosionsunfälle bei Ereignissen, die häufiger als 1 mal in 100 Jahren sind, ein geringerer Schadensumfang als für die anderen dargestellten Risiken. Bei Ereignissen, die seltener sind als etwa 1 mal in 500 Jahren, haben Brand- und Explosionsunfälle dagegen den größeren Schadensumfang.

Dies ist ein Beispiel für eine Vergleichssituation, in der quantitative Abschätzungen von Schadenshäufigkeiten und Schadensumfängen gegeben sind. Das dürfte gerade für den Bereich der Störfallrisiken nicht immer der Fall sein.

Liegen nicht genügend oder nicht genügend genaue Daten vor, dann kann man Kategorien jeweils vergleichbarer Risikobereiche bilden. Die Bildung von Kategorien ist

¹¹² Die Daten, auf denen die Abbildung basiert, sind wahrscheinlich veraltet, das hat aber keinen Einfluss auf die prinzipielle Vorgehensweise.

zum Zweck der sicherheitstechnischen Auslegung von Anlagen gängige Praxis. So wird z. B. in den *Sicherheitskriterien für Kernkraftwerke* (1977) unterschieden zwischen

- anomalen Betriebszuständen – diese können während der Lebensdauer der Anlage auftreten und müssen so beherrscht werden, dass Störfälle mit ausreichender Zuverlässigkeit vermieden werden,
- Störfällen – es sind Systeme zu ihrer Beherrschung vorzusehen (die StrlSchV fordert außerdem den Nachweis, dass dabei die Werte des §28 Abs.3 StrlSchV eingehalten werden), und
- Unfällen – hierfür sind Maßnahmen zur Feststellung und Eindämmung von Unfallfolgen vorzusehen.

Auch in der chemischen Industrie werden bei der sicherheitstechnischen Auslegung von Anlagen und Prozessen Kategorien betrachtet. Madjar & von Rohr (1995) unterscheiden dabei folgende Vorgehensweisen:

qualitativer Ansatz: Ein Ereignis wird relativ zu anderen Ereignissen verbal beschrieben. Es fehlen absolute Bezugspunkte und Angaben zur Bandbreite (welcher Bereich wird dem Begriff "hoch" zugeordnet?), so dass kein Vergleich verschiedener Prozesse / Anlagen möglich ist.

Semi-quantitativer Ansatz: Einem qualitativen Begriff wird eine Größenklasse zugeordnet. Die Klassenbreite soll derart gewählt werden, dass die Ereignisse auch bei ungenauem Datenmaterial noch innerhalb der von der Klasse gebildeten Bandbreite liegen.

quantitativer Ansatz: Einem qualitativen Begriff wird ein Zahlenwert zugeordnet. Die Verifikation der Ergebnisse wird durch den Nachvollzug der einzelnen Erkenntnisschritte ermöglicht.

Ein Beispiel für eine semi-quantitative Risikoanalyse für chemische Anlagen geben die folgenden Tabellen. Zunächst wird eine Kategorisierung von Wahrscheinlichkeiten (im Sinne relativer Häufigkeiten) vorgenommen. Tabelle 35 zeigt die Häufigkeitsintervalle, deren verbale Beschreibung sowie die jeweilige Wahrscheinlichkeitskategorie.

Tabelle 35: »Beispiel für eine semi-quantitative Bewertung der Wahrscheinlichkeit auf Prozessebene (Quelle: Hoffmann-La Roche)« (Madjar & von Rohr, 1995, Tabelle 4.6)

Wahrscheinlichkeit	Beschreibung
A	häufig: mehr als 1 mal pro Jahr
B	oft: 1 mal pro Jahr
C	gelegentlich: 1 mal in 5 Jahren
D	selten: 1 mal pro 30 Jahre
E	unwahrscheinlich: 1 mal in 100 Jahren
F	sehr unwahrscheinlich: 1 mal in 1000 Jahren

Tabelle 36 enthält die Kategorisierung der Schadensfolgen in Tragweitenklassen, die aus Kombinationen von Schadensfolgen für Mensch, Umwelt und Sachwerte gebildet worden sind. Dies ist zugleich ein Beispiel, wie eine multidimensionale Bewertungsstruktur (Mensch, Umwelt, Sachwerte) unter einer spezifischen Zielstellung aggregiert und kategorisiert werden kann. Die inhaltliche Festlegung dieser Tragweiten-

klassen resultiert aus den spezifischen Zielstellungen dieser Analysen im Rahmen der chemischen Sicherheitsanalysen.

Tabelle 36: »Beispiel für eine semi-quantitative Tragweitenabschätzung (Quelle Hoffmann-La Roche)« (Madjar & von Rohr, 1995, Tabelle 4.2)

Tragweiteklasse	Beschreibung
I	Mensch: Tote, außerhalb des Areals Evakuierte Umwelt: arealüberschreitende Langzeitschäden Sachwerte: >10 Millionen sFr., Ausfall der Anlage > 1 Jahr
II	Mensch: Verletzte, außerhalb des Areals Reizungen Umwelt: reversible Schäden in der Nachbarschaft Sachwerte: <10 Millionen sFr., Ausfall der Anlage: Monate
III	Mensch: Verletzte nur im Areal, Belästigungen in der Nachbarschaft Umwelt: Areal, allenfalls Abwasserreinigungsanlage betroffen Sachwerte: <2 Millionen sFr., Ausfall der Anlage: Wochen
IV	Mensch: Leichtverletzte nur im Areal Umwelt: nur Anlage bis Areal betroffen Sachwerte: <1 Millionen sFr., Ausfall der Anlage: Tage

»Eine semi-quantitative Einteilung (Tab. 4.2) reicht häufig aus, um die Größenordnung der Tragweite einzelner Szenarien auf der Prozess-/ Anlagenebene abzuschätzen. Sollen jedoch Auswirkungen großer Tragweite bestimmt werden, so sind hierzu Modelle zur Freisetzungs- und Ausbreitungsrechnung erforderlich.«

Nach den Kategorien aus den Tabellen 35 und 36 können verschiedene Störfallszenarien nach ihrer Wahrscheinlichkeit und Tragweite in eine Risikomatrix eingetragen werden. Tabelle 37 zeigt ein Beispiel für eine solche Matrix mit vier fiktiven Szenarien.

Tabelle 37: »Beispiel einer Risikomatrix« (modifiziert nach Madjar & von Rohr, 1995, Abb. 4.6)

↑ Wahrscheinlichkeit					
A					
B			●4		
C					
D	●2				
E		●1	●3		
F					
	IV	III	II	I	→ Tragweite

●1 = Szenario Nr. 1

Aus Tabelle 37 lässt sich ablesen, dass Szenario 4 in Tragweite-Kategorie II mit Abstand die höchste Wahrscheinlichkeit hat. Zwei Szenarien (Szenario 1 und Szenario 3) fallen in den Wahrscheinlichkeitsbereich E. Damit ergibt sich für die vergleichende Risikobewertung dieser beiden Risiken, dass Szenario 3 im Risikoranking vor Szenario 1 liegt, da es die größeren zu erwartenden Schadensfolgen hat.

Die Risikomatrix in Tabelle 37 verdeutlicht allerdings auch ein grundsätzliches Problem, dass mit solchen qualitativen Kategorisierungen für die vergleichende Risikobewertung verbunden ist. Betrachtet man alle vier Szenarien, so zeigt sich, dass sich nicht für alle Szenarien ein eindeutiger Rangplatz ergibt. Nur Szenario 4 liegt eindeutig auf Rang 1, da es die höchste Wahrscheinlichkeit hat und es kein Szenario gibt, das einen höheren Tragweitenrang hat. Für die anderen drei Szenarien ist dagegen das Risikoranking von der Methode der Rangbildung abhängig. Bei einer lexikogra-

phischen Ordnung¹¹³, die hier naheliegt, hängt der Rang davon ab, welches der beiden Bewertungskriterien: Wahrscheinlichkeit oder Tragweite, als das bedeutsamere angesehen wird. Gewichtet man die Wahrscheinlichkeit höher als die Tragweite, so ergibt sich die Rangfolge Szenario 4, Szenario 2, Szenario 3, Szenario 1. Gewichtet man die Tragweite höher als die Wahrscheinlichkeit, so ergibt sich die Rangfolge Szenario 4, Szenario 3, Szenario 1, Szenario 2.

Nicht-kompensatorische Bewertungsverfahren geben hier keine Möglichkeit zu einer weitergehenden Analyse – die Rangfolge ist schlicht Folge der relativen Gewichtung von Wahrscheinlichkeit und Tragweite (und natürlich der Einschätzungen der Szenarien auf den beiden Attributen).

Eine Möglichkeit zur differenzierteren Bewertung auf der Basis der gegebenen Kategorien bietet der in Kapitel 5.2 erläuterte und in Anhang 2 illustrierte Ansatz der multiattributen Risikobewertung. Bei diesem Vorgehen sind Wahrscheinlichkeit und Tragweite die Attribute, auf denen die vier Szenarien beurteilt werden. Beide Attribute könnten, entsprechend der Einschätzungen der Bewerter, unterschiedlich oder auch gleich gewichtet werden. Die Szenarien können dann in bezug auf ihre Wahrscheinlichkeit und Tragweite von den Bewertern auf Ratingskalen beurteilt werden und sind so einer differenzierteren Beurteilung zugänglich. Im Prinzip ist es auch möglich, für die beiden Attribute nicht-lineare Nutzenfunktionen zu konstruieren, die eine Berücksichtigung der großen Ausprägungsunterschiede, die sich in der Wahrscheinlichkeits- bzw. Tragweite-Kategorie finden, bei der Beurteilung gestatten.

¹¹³ Die Rangbildung erfolgt danach, welches Szenario auf dem wichtigsten Attribut die beste höchste Einschätzung bekommen hat. Haben mehrere Szenarien gleich hohe Einschätzungen, so wird in der gleichen Weise das zweitwichtigste Attribut betrachtet (lexikographische Regel; siehe oben Kapitel 5.2.2.)

Literatur

- Ale, B.J.M. (2002): "Risk assessment practices in The Netherlands." *Safety Science*, 40: 105-126.
- Alhakami, A.S. & Slovic, P. (1994): A psychological study of the inverse relationship between perceived risk and perceived benefit. *Risk Analysis*, 14, 1085-1096.
- Alhakami, A.S. (1991): *A psychological study of the inverse relationship between perceived risk and perceived benefit of technological hazards*. Dissertation Abstracts, Vol. 52/09-B.
- Apostolakis, G.E. & Pickett, S.E. (1998): Deliberation: Integrating analytical results into environmental decisions involving multiple stakeholders. *Risk Analysis*, 18(5), 621-634.
- Ayyub, B.M. & Bender, W.J. (1999): *Assessment of the Construction Feasibility of the Mobile Offshore Base. Part I – Risk-Informed Assessment Methodology*. College Park, MD: Center for Technology and Systems Management. University of Maryland.
- Bailar III, J.C. & Bailer, A.J. (1999): Risk assessment – The mother of all uncertainties. Disciplinary perspectives on uncertainty in risk assessment. In A.J. Bailer & C. Maltoni & J.C. Bailar III & F. Belpoggi & J.V. Brazier & M. Soffritti (Eds.), *Uncertainty in the risk assessment of environmental and occupational hazards: An international workshop* (Vol. 895, pp. 273-285): New York: The New York Academy of Sciences.
- Balderjahn, I. & Wiedemann, P.M. (1999): *Bedeutung von Risikokriterien bei der Bewertung von Umweltproblemen*. Arbeiten zur Risiko-Kommunikation Heft 74. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.
- Baron, J. & Spranca, M. (1997): Protected values. *Organizational Behavior and Human Decision Processes*, 70, 1-16.
- Baron, J. (1988): *Thinking and deciding*. Cambridge: University Press.
- Bazerman, M.H. (1990): *Judgment in Managerial Decision Making*. New York: John Wiley & Sons.
- Bazerman, M.H. & Lewicki, R.J. (1983): *Negotiation in organizations*. Beverly Hills, London, New Delhi: Sage Publications
- Bazerman, M.H., Moore, D.A., Tenbrunsel, K.A., Wade-Benzoni, K.A. & Blount, S. (1999): Explaining How Preferences Change across Joint versus Separate Evaluation. *Journal of Economic Behavior and Organization*, 39, 41-58.
- Beck, U. (1986). *Risikogesellschaft. Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp.

- Bennett, D. H., Margni, M. D., McKone, T. E., et al. (2002): Defining Intake Fraction. *Environmental Science & Technology*, 36 (9): 206 A - 211 A. URL (2003-05-14): <http://pubs.acs.org/cgi-bin/article.cgi/esthag-a/0000/36/i09/html/09benn.html>.
- Berg, M., Erdmann, G., Leist, A., Renn, O., Schaber, P., Scheringer, M., Seiler, H. & Wiedenmann, R. (1995): *Risikobewertung im Energiebereich* Zürich: vdf Hochschulverlag AG an der ETH Zürich.
- Beyth-Marom, R. (1982): How probable is probable? A numerical translation of verbal probability expression. *Journal of Forecasting*, 1, 257 – 269.
- Böhm, G. & Pfister, H.-R. (2000): Action tendencies and characteristics of environmental risks. *Acta Psychologica*, 104, 317-337.
- Bonano, E.J., Apostolakis, G.E., Salter, P.F., Ghassemi, A. & Jennings, S. (2000): Application of risk assessment and decision analysis to the evaluation, ranking and selection of environmental remediation alternatives. *Journal of Hazardous Materials*, 71(1-3), 35-57.
- Breyer, S. (1993): *Breaking the Vicious Circle: Toward Effective Risk Regulation*. Cambridge, MA: Harvard University Press.
- Brun, W. & Teigen, K. H. (1988): Verbal probabilities: ambiguous, context-dependent, or both? *Organizational Behavior and Human Decision Processes*, 41, 390 – 404.
- Budescu, D. V. & Wallsten, T. S. (1985): Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes*, 36, 391 – 405.
- Budescu, D. V. & Wallsten, T. S. (1995): Processing linguistic probabilities: General principles and empirical evidence. *The Psychology of Learning and Motivation*, 32, 275 – 318.
- Büllingen, F., Hillebrand, A. & Wörter, M. (2002): *Elektromagnetische Verträglichkeit zur Umwelt (EMVU) in der öffentlichen Diskussion. Situationsanalyse, Erarbeitung und Bewertung von Strategien unter Berücksichtigung der UMTS-Technologien im Dialog mit dem Bürger*. Studie im Auftrag des Bundesministeriums für Wirtschaft und Technologie (BMWi): WIK Consult, Bad Honnef.
- Bundesinstitut für Risikobewertung (BfR)* (2002): *Acrylamid in Lebensmitteln - ernstes Problem oder überschätzte Gefahr?* Informationsveranstaltung des BgVV vom 29.08.2002. (* vormals Bundesinstitut für gesundheitlichen Verbraucherschutz und Veterinärmedizin (BgVV)) URL (2003-05-21): http://www.bgvv.de/cms/media.php/70/acrylamid_in_lebensmitteln_ernstes_problem_oder_ueberschaetzte_gefahr.pdf.
- Canadian Standards Association (1996): *A Guide to Public Involvement*. Z764-96. CSA: Canada.
- CEC (1996): *Technical Guidance Document in support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No. 1488/94 on Risk Assessment for Existing Substances*. 1996. Luxembourg: Office for Official Publications of the European Communities.

Champaud, C. & Bassano, D. (1987): Argumentative and Informative Functions of French Intensity Modifiers „presque“ (almost), „à peine“ (just, barely) and „à peu près“ (about): an Experimental Study of Children and Adults. *European Bulletin of Cognitive Psychology*, 7, 6, 605 – 631.

Chapman, G.B. & Johnson, E.J. (1995): Preference Reversals in Monetary and Life Expectancy Evaluations. *Organizational Behavior and Human Decision Processes*, 62, 300-317.

Charlton Research Company (2001): *Poll on comparative risk*. The Congressional Institute.

Chemical Manufacturers Association (CMA) (1988): *Title III Community Awareness Workbook*. Washington, DC.

Clarke, M. & Oxman, A.D. (eds.) (2002): Cochrane Reviewers Handbook 4.1.6 (updated January 2003): In: The Cochrane Library, Issue 2, 2002. Oxford: Update Software. Updated quarterly. In Internet unter: Cochrane Collaboration: <http://www.cochrane.org/cochrane/resource.htm>.

Clarke, V. A., Ruffin, C. L., Hill, D. J. & Beamen, A. L. (1992): Ratings of orally presented verbal expressions of probability by a heterogeneous sample. *Journal of Applied Social Psychology*, 22, 638 – 656.

Clemen, R.T. (1991). *Making hard decisions*. Boston, MA: PWS-Kent Publishing Company.

Cleveland, W.S. (1993). *Visualizing data*. Murray Hill, N.J: Hobart Press.

Cohen, J.T., Beck, B.D. & Rudel, R. (1997): Life years lost at hazardous waste sites: remediation worker fatalities vs. cancer deaths to nearby residents. *Risk Analysis*, 17(4), 419-425.

Commoner, B. (1992): Pollution Prevention: Putting Comparative Risk Assessment in its Place. Presentation to Resources for the Future Conference: Setting National Environmental Priorities, November 2002.

Covello, V.T., McCallum, D.B. & Pavlova, M.T. (1989): Principles and guidelines for improving risk communication. In: V.T. Covello, D.B. McCallum and M.T. Pavlova (eds.), *Effective risk communication*. New York: Plenum Press, 3-16.

Covello, V.T., Sandman, P.M. & Slovic, P. (1988): *Risk Communication, Risk Statistics, and Risk Comparisons: A Manual for Plant Managers*. Washington, DC: Chemical Manufacturers Association.

Cropper, M.L. & Subramanian, U. (1995): *Public Choices between Lifesaving Programs*. Policy Research Working Paper 1497. New York: Policy Research Dissemination Center.

Crouch, E.A.C. & Wilson, R. (1982): *Risk/benefit analysis*. Cambridge, MA: Ballinger Publishing.

Cullen, A.C. & Frey, H.C. (1999). Probabilistic techniques in exposure assessment : a handbook for dealing with variability and uncertainty in models and inputs. New York, NY: Plenum Press.

Dienel, P. (2002): *Die Planungszelle*. Der Bürger als Chance. Opladen: Westdeutscher Verlag.

DRS-A (1979): *Deutsche Risikostudie Kernkraftwerke*. Eine Studie der Gesellschaft für Reaktorsicherheit im Auftrag des Bundesministeriums für Forschung und Technologie. Hauptband. Köln: Verlag TÜV Rheinland.

Drücke, H., Henckel, D., Reidenbach, M. & Seidel-Schulze, A. (2003): *Monitoring zur Selbstverpflichtung der Netzbetreiber gegenüber der Bundesregierung Verbesserung der Kooperation mit den Kommunen beim Aufbau von Mobilfunknetzen. Ergebnisse einer Befragung von Kommunen und Netzbetreibern Gutachten im Auftrag des Informationszentrums Mobilfunk e.V. (IZMF)*: Berlin: Deutsches Institut für Urbanistik (DIFU):

EC (2000a): *First Report on the Harmonisation of Risk Assessment Procedures - Part 1* (report): Brussels: European Commission; Scientific Committee's Working Group; Directorate C.

EC (2000b): *First Report on the Harmonisation of Risk Assessment Procedures - Part 2* (report): Brussels: European Commission; Scientific Committee's Working Group; Directorate C.

Edision Electric Institute (1994): *Public Participation Manual*.

Edwards, W. & Barron, F.H. (1994): SMARTS and SMARTER - Improved simple methods for multiattribute utility measurement. *Organizational Behavior and Human Decision Processes*, 60, 306-325.

Edwards, W. (1977): How to use multiattribute utility measurement for social decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-7, 326-340.

Einhorn, H. J. & Hogarth, R. M. (1985): Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, 92, 433-461.

Einhorn, H.J. & Hogarth, R.M. (1978): Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, 85, 395-416.

Eisenführ, F. & Weber, M. (1993): *Rationales Entscheiden*. Berlin: Springer-Verlag.

EPA – United States Environmental Protection Agency (2000): *Risk characterization handbook*. U.S. Environmental Protection Agency. Washington, DC.

EPA (1987): *Unfinished business: A comparative assessment of environmental problems*. Washington, DC: U.S. Environmental Protection Agency.

EPA (1993): *A guidebook to comparing risks and setting environmental priorities*. Washington, DC: US Environmental Protection Agency (Online: <http://www.epa.gov/opperspd/guidbook/lb08007t.pdf>):

EPA (1996): *Characterization of Uncertainties in Risk Assessment with Special Reference to Probabilistic Uncertainty Analysis* (U.S. Environmental Protection Agency Information Brief EH-413-068/0496): Washington, DC: U.S. Environmental Protection Agency.

EPA (1998): *Guidelines for Ecological Risk Assessment* (EPA/630/R-95/002F): Washington: EPA - U.S. Environmental Protection Agency.

EPA (2003): *Draft Final Guidelines for Carcinogen Risk Assessment (external review draft, February 2003)* (Other NCEA-F-0644A): Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum.

Erev, I & Cohen, V. L. (1990): Verbal versus Numerical Probabilities: Efficiency, Biases, and the Preference Paradox. *Organizational Behavior and Human Decision Processes*, 45, 1 – 18.

European Environment Agency (2001): *Late lessons from early warnings: the precautionary principle 1896-2000*. Environmental Issue report No. 22. Luxembourg: Office for Official Publications of the European Communities.

Evans, J.St.B.T. (1990): *Bias In Human Reasoning. Causes and Consequences. Essays in Cognitive Psychology*. Lawrence Erlbaum Associates Publishers, Hove and London.

ExternE - Methodology (1995): ExternE - Externalities of Energy, Vol. 2: Methodology, Luxembourg: Office for Official Publications of the European Communities (EUR 16521 EN).

ExternE - Summary (1995): ExternE - Externalities of Energy, Vol. 1: Summary, Luxembourg: Office for Official Publications of the European Communities (EUR 16520 EN).

ExternE Methodology Annexes (1997): DGXII 'ExternE' Project. Brussels: European Commission. URL (2003-03-12): <http://externe.jrc.es/append.pdf>.

ExternE - Vol. 10: National Implementation (1999). ExternE - Externalities of Energy, Luxembourg: Office for Official Publications of the European Communities (EUR 18528).

ExternE Overview (o. D.; Web-Seite): Externalities of Energy. A research project of the European Commission. URL (2003-03-18): <http://externe.jrc.es/overview.html>.

- Ezzati, M., Lopez, A.D., Rodgers, A., Vander Hoorn, S. & Murray, C.J. (2002): Selected major risk factors and global and regional burden of disease. *Lancet*, 360(9343), 1347-1360.
- Fehr, R. & Mekel, O.C.L. (1999): Probabilistische Expositionsabschätzung in Umweltmedizin und Verbraucherschutz (Workshop) - Präambel. *UWSF - Z. Umweltchem. Ökotox.*, 12(3), 169.
- Feldman, D.L., Hanahan, R.A. & Perhac, J., Ralph M. (1999): Environmental Priority-Setting Through Comparative Risk Assessment. *Environmental Management*, 23(4), 0483-0493.
- Femers, S. (1993): *Information über technische Risiken: Zur Rolle fehlenden direkten Erfahrbarkeit von Risiken und den Effekten abstrakter und konkreter Informationen*. Frankfurt: Peter Lang Verlag.
- Fillenbaum, S., Wallsten, T.S. Cohen, B., and Cox, J. (1991): Some effects of vocabulary and communication task on the understanding and use of vague probability expression. *American Journal of Psychology*, 140, 36-60.
- Finkel, A.M. & Golding, D. (1994): *Worst things first? The debate over risk-based national environmental priorities*. Washington, DC: Resources for the Future.
- Finucane, M.L., Alhakami, A., Slovic, P. & Johnson, S.M. (2000): The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, 13, 1-17.
- Fischer, K. Jungermann, H. (1996): Rarely Occuring Headaches and Rarely Occuring Blindness: Is Rarely = Rarely? *Journal of Behavioral Decision Making*, 9, 153-172.
- Fischhoff, B. (1997): Ranking Risks. In: M. Bazerman, D. Messick, A. Tenbrunsel & K. Wade-Benzoni (eds.), *Environment and ethics: Psychological contributions*. San Francisco: Jossey-Bass, 342-371.
- Fischhoff, B., Slovic, P. & Lichtenstein, S. (1978): Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 330-344.
- Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S. & Combs, B. (1978): How safe is safe enough? A psychometric study of attitudes toward technological risks and benefits. *Policy Science*, 9, 127-152.
- Fischhoff, B., Watson, S. & Hope, C. (1984): Defining risk. *Policy Science*, 17, 123-139.
- Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S.L. & Keeney, R.L. (1981): *Acceptable Risk*. Cambridge University Press.
- Fisher, R. & Ury, W. (1981): *Getting to Yes: Negotiating Agreement Without Giving*. Boston: Houghton Mifflin.

Fiske, A.P. & Tetlock, P.E. (1997): Taboo tradeoffs: Reactions to Transactions that Transgress Spheres of Exchange. *Journal of Political Psychology*, 18, 255-297.

Florida Center for Public Management (1995). *Comparing Florida's environmental risks*. Tallahassee, FL: Florida Department of Environmental Protection.

Florig, H.K., Morgan, M.G., Morgan, K.M., Jenni, K.E., Fischhoff, B., Fischbeck, P.S. & DeKay, M.L. (2001): A deliberative method for ranking risks (I): Overview and test bed development. *Risk Analysis*, 21(5), 913-921.

Freudenburg, W.R. & Rursch, J.A. (1994): The risks of "Putting the Numbers in Context". A cautionary tale. *Risk Analysis*, 14, 6, 949-958.

Frischknecht, R. (Ed) et al. (1996): *Ökoinventare von Energiesystemen - Grundlagen für den ökologischen Vergleich von Energiesystemen und den Einbezug von Energiesystemen in Ökobilanzen für die Schweiz*, 3rd Ed., ETHZ/PSI, Zürich.

Gaido, K.W., Leonard, L.S., Lovell, S., Gould, J.C., Babai, D., Portier, C.J. & McDonnell, D.P. (1997): Evaluation of chemicals with endocrine modulating activity in a yeast-based steroid hormone receptor gene transcription assay. *Toxicol Appl Pharmacol*, 143(1), 205-212.

GBD 1996: Murray, C.J.L. & Lopez, A.D. (1996): *The Global Burden of Disease*. World Health Organization, Harvard School of Public Health, World Bank.

GBD 2000: Kay, D, Prüss, A. & Corvalan, C. (2000): *Methodology of assessment of environmental burden of disease*, World Health Organization, Geneva.

GBD 2001: Mathers, C.D., Vos, T., Lopez, A.D., Salomon, J. and Ezzati, M. (ed): *National Burden of Disease Studies: A Practical Guide. Edition 2.0*. Global Program on Evidence for Health Policy. Geneva: World Health Organization.

Gerrard, M.B. (2000): Risks of hazardous waste sites versus asteroid and comet impacts: accounting for the discrepancies in U.S. resource allocation. *Risk Analysis*, 20(6), 895-904.

Gethmann, C.F. (2002): *Ethische Aspekte der technischen Sicherheit*. Europäische Akademie zur Erforschung von Folgen wissenschaftlicher-technischer Entwicklungen, Bad-Neuenahr-Ahrweiler.

Goklany, I.M. (2001): *The Precautionary Principle. A Critical Appraisal of Environmental Risk Assessment*. Washington, DC: CATO Institute.

Gold, L.S., Slone, T.H., Manley, N.M. & Ames, B.N. (2002): *Misconceptions About the Causes of Cancer* (Risk Controversy Series 3): Vancouver, Canada: Fraser Institute.

Goldstein, B.D. (1990): The problem with the margin of safety: Toward the concept of protection. *Risk Analysis*, 10(1), 7-10.

- Gonzales, M. & Frenck-Mestre, C. (1993): Determinants of numerical versus verbal probabilities *Acta Psychologica*, 83, 33-51.
- Gonzales-Vallejo, C. C. & Wallsten, T. S. (1992): Effects of probability modes on preference reversal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 855 – 864.
- Gonzales-Vallejo, C. C., Erev, I. & Wallsten, T. S. (1994): Do decision quality and preference order depend on whether probabilities are verbal or numerical? *American Journal of Psychology*, 107, 57 – 172.
- Graham, J.D. & Wiener, J.B. (1997): Confronting risk tradeoffs. In J.D. Graham & J.B. Wiener (Eds.), *Risk versus risk. Tradeoffs in protecting health and the environment* (pp. 1-41): Cambridge, MA: Harvard University Press.
- Gray, G.M. (2000). *Characterizing Risks for Sound Risk Management*. Paper presented at the 3rd International Workshop on Risk Evaluation and Management of Chemicals, Yokohama National University, Yokohama, Japan (27 -28 January 2000). (URL: <http://risk.kan.ynu.ac.jp/rmg/Gray.PDF>)
- Gray, P.C.R. (1996): *Risk Indicators: Types, Criteria, Effects. A framework for analysing the use of indicators and comparisons in risk communication*. Arbeiten zur Risiko-Kommunikation Heft 56. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.
- Gregory, R. & Mendelsohn, R. (1993): Perceived risk, dread and benefits. *Risk Analysis*, 13, 259-264.
- Gregory, R., Flynn, J. & Slovic, P. (1995): Technological stigma. *American Scientist*, 83, 220-223.
- Greim, H. & Reuter, U. (2001): Classification of carcinogenic chemicals in the work area by the German MAK Commission: current examples for the new categories. *Toxicology*, 166(1-2), 11-23.
- Grossi, P., Kleindorfer, P. & Kunreuther, H. (1999): *The impact of uncertainty in managing seismic risk: the case of earthquake frequency and structural vulnerability*. Philadelphia, PA: The Wharton School, University of Pennsylvania.
- Häfele, W. (1990): Energiesysteme: Eine einführende Problematisierung. In: W. Häfele (Hrsg), *Energiesysteme im Übergang: unter den Bedingungen der Zukunft; Ergebnisse einer Studie des Forschungszentrums Jülich GmbH* (pp. 1-48). Landsberg/Lech: mi-Poller.
- Haimes, Yakov Y. (1998): *Risk Modeling, Assessment and Management*. Wiley Series in Systems Engineering. New York: John Wiley & Sons (A Wiley-Interscience Publication).
- Halpern, D. F., Blackman, S. & Salzman, B. (1989): Using statistical risk information to assess oral contraceptive safety. *Applied Cognitive Psychology*, 3, 251-260.

Hamm, R. M. (1991): Selection of verbal probabilities: A solution for some problems of verbal probability expression. *Organization Behavior and Human Decision Processes*, 45, 1 – 18.

Hammit, J. K. (2002): "QALYs versus WTP." *Risk Analysis*, 22 (5): 985-1001.

Harding, C.M. & Eiser, J.R. (1984): Characterizing the perceived risk of some health issues. *Risk Analysis*, 4, 131-141.

Heinemeyer, G. (2000): Probabilistische Expositionsabschätzung in Umweltmedizin und Verbraucherschutz - Teil IV: Probabilistische Schätzung der inhalativen Exposition vs. Punktschätzung, dargestellt am Beispiel der Freisetzung von Xylol aus Farben. UWSF - Z. Umweltchem. Ökotox., 12(2), 105-112.

Heinemeyer, G. & Dürkop, J. (2002): Müssen Kinder bei der Risikobewertung besonders berücksichtigt werden? *Umweltmedizinischer Informationsdienst*, 1/2002.

Hill, A.B. (1965): The environment and disease: Association and causation. *Proceedings of the Royal Society of Medicine*, 295-300.

Hirschberg, S., Spiekernan, G. & Dones, R. (1998): *Project GaBE: Comprehensive Assessment of Energy Systems — Severe Accidents in the Energy Sector* (PSI Bericht Nr. 98-16): Villigen, Schweiz: Paul Scherrer Institut.

Hirschberg, S., Voss, A.: *Nachhaltigkeit und Energie: Anforderungen der Umwelt*. In: PSI-Proceedings 99-01, Mai 1999.

Hobbs, B.F. (1986): What can we learn from experiments in multiobjective decision analysis. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-16, 384-394.

Hofstetter, Patrick; Hammit, James K. (2002): "Selecting Human Health Metrics for Environmental Decision-Support Tools." *Risk Analysis*, 22 (5): 965-983.

Hohenemser, C., Kasperson, R.E. & Kates, R.W. (1985): Causal structure. In: R.W. Kates, C. Hohenemser and J.X. Kasperson (eds.), *Perilous progress. Managing the hazards of technology*. Boulder, CO: Westview Press, 25-66.

Hood, C., Rothstein, H. & Baldwin, R. (2001): *The Government of Risk. Understanding Risk Regulation Regimes*. Oxford, University Press.

Hornstein, D. T. (1992): Reclaiming environmental law: a normative critique of comparative risk analysis. *Columbia Law Review*, 92(3), 562-633.

Hsee, C. (1996): The Evaluability Hypothesis: An Explanation of Preference Reversals Between Joint and Separate Evaluations of Alternatives. *Organizational Behavior and Human Decision Processes*, 46, 247-257.

Hubbell, B. J. (2002): *Implementing QALYs in the Analysis of Air Pollution Regulations*. Draft paper. Research Triangle Park, North Carolina 27711: US-EPA, Office of Air Quality Planning and Standards, Innovative Strategies and Economics Group. URL (2003-04-17): <http://www.epa.gov/ttn/ecas/workingpapers/ereqaly.pdf>.

Hutter, B.M. (2001): *Regulation and Risk. Occupational Health and Safety on the Railways*. Oxford, University Press.

ICNIRP (2002): ICNIRP Statement: General Approach to Protection against Non-ionizing Radiation. *Health Physics*, 82(4), 540-548.

IPCS - International Programme on Chemical Safety (2000). *Human exposure assessment*. Geneva, Switzerland: World Health Organization.

Irwin, J. & Baron, J. (2001): Values and Decisions. In: S.J. Hoch, H.C. Kunreuther, R.E. Gunther (Eds.), *Wharton on Making Decisions*, John Wiley & Sons Inc., New York, pp. 131-157.

Irwin, J.R. (1994): Elicitation Rules and Incompatible Goals. *Behavioral and Brain Sciences*, 17, 20-21.

Jablonowski, M. (1994): Communicating Risk. Words or Numbers? *Risk Management*, Dezember, 47-50.

Jaffe-Katz, A., Budescu, D. V. & Wallsten, T. S. (1989): Timed magnitude comparisons of numerical and nonnumerical expressions of uncertainty. *Memory & Cognition*, 17, 249 – 264.

Johnson, B.B. (2003): Are some risk comparisons more effective under conflict?: A replication and extension of Roth et.al., *Risk Analysis*, 23 (4), 767-780

Johnson, B.B. (2003): Further notes on public response to uncertainty in risk and science. *Risk Analysis*, 23 (4), 771 –789

Johnson, B.B. & C. Chess (2003): How reassuring are risk comparisons to pollution standards and emission limits? *Risk Analysis*, 23 (5), 999 -1007.

Johnson, B.B. & Slovic, P. (1995): Presenting Uncertainty in Health Risk Assessment: Initial Studies of its Effects on Risk Perception and Trust. *Risk Analysis*, 15, 4, 485 – 494.

Johnson, B.B. (2002): Stability and Inoculation of Risk Comparisons' Effects Under Conflict: Replicating and Extending the „Asbestos Jury“ Study by Slovic et al. *Risk Analysis*, 22, 4, 777-788.

Jungermann, H. & Slovic, P. (1993): Die Psychologie der Kognition und Evaluation von Risiko. In G. Bechmann (Ed.), *Risiko und Gesellschaft* (pp. 167-207). Opladen: Westdeutscher Verlag.

Jungermann, H., Pfister, H.R. & Fischer, K. (1998): *Die Psychologie des Entscheidens*. Heidelberg: Spektrum Akademischer Verlag.

Kahneman, D. & Tversky, A. (1979): Prospect Theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.

Kaplan, S. & Garrick, B.J. (1981): On the quantitative definition of risk. *Risk Analysis*, 1, 11-27.

Kaplan, S. (1997): The words of risk analysis. *Risk Analysis*, 17, 407-417.

Karger, C.R. & Wiedemann, P.M. (1998): *Kognitive und affektive Determinanten der intuitiven Bewertung von Umweltrisiken*. Arbeiten zur Risiko-Kommunikation, Heft 64. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Kasperson, R.E. & Kasperson, J.X. (1988): *Guides for the presentation of quantitative risk information*. Worcester, MA: Center for Technology, Environment, and Development (CENTED), Clark University.

Keefer, D.L., Kirkwood, C.W. & Corner, J.L. (2003): Perspective on decision analysis applications, 1990-2001. *Decision Analysis*, 1(1).

Keeney, R.L. & Raiffa, H. (1976): *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Wiley.

Keeney, R.L. (1980): *Siting Energy Facilities*. New York: Academic Press.

Keeney, R.L. (1992): *Value-focused thinking: A path to creative decision making*. Cambridge, MA: Harvard University Press.

KEMI (2003). Human health risk assessment (1/03). Solna, Sweden: The Swedish National Chemicals Inspectorate.
(Online: http://www.kemi.se/publikationer/pdf/rapport1_03.pdf)

Kincaid, L. E., Meline, J. D. & Davis, G. A. (1995): Chapter 5: Chemical & Process Information: Human Health Hazards Summary. In: *Cleaner Technologies Substitutes Assessment - A Methodology and Resource Guide*. (EPA Grant X821-543) Washington: US-EPA, Office of Pollution Prevention and Toxics. Project homepage (2003-05-23): <http://www.epa.gov/oppt/dfe/pubs/tools/ctsa/index.htm>, document source: <http://www.epa.gov/oppt/dfe/pubs/tools/ctsa/ch5/mod5-4.pdf>.

Kommission der Europäischen Gemeinschaften (2001): Weissbuch. Strategie für die zukünftige Chemikalienpolitik. Brüssel, 27.3.2001.

Krewitt, W., Mayerhofer, P., Friedrich, R., et al. (1997): *ExternE National Implementation Germany*. Final Report, prepared by IER. Stuttgart: Institut für Energiewirtschaft und Rationelle Energieanwendung (IER): URL (2003-03-12): <http://externe.jrc.es/ger.pdf>.

Kröger, W. (1998): *Nachhaltigkeit und Energie – Weniger Beliebtheit*. In: *Nachhaltigkeit – ein Konzept ohne Inhalt?* BATS-Report 5/98, Basel: Fachstelle BATS. http://www.bats.ch/publications/report5-98/bericht5-98_4.html.

Kröger, W. (2000): *Nachhaltigkeit ist messbar*. Energie-Spiegel Nr. 3 / Sept. 2000, Paul Scherrer Institut, Projekt GaBE, Villigen, Schweiz.

Kröger, W., Hirschberg, S., Foskolos, K. (2001): *The Attributes of Sustainability of Energy Options*. MIT CANES-Symposium, April 2001.

Kunreuther, H. & Linnerooth, J. (Eds.) (1983): *Risk analysis and decision processes: the siting of liquefied energy gas facilities in four countries*. Berlin/ New York: Springer Verlag.

Kunreuther, H. (2002): Risk analysis and risk management in an uncertain world. *Risk Analysis*, 22(4), 655-664.

Landrigan, P.J. (1999): Risk assessment for children and other sensitive populations. In A.J. Bailer & C. Maltoni & J.C. Bailar III & F. Belpoggi & J.V. Brazier & M. Soffritti (Eds.), *Uncertainty in the risk assessment of environmental and occupational hazards: An international workshop* (Vol. 895, pp. 1-9.): New York: The New York Academy of Sciences.

Leviton, L. (1997). *An Overview of Pesticide Impact Assessment Systems a.k.a. Pesticide Risk Indicators* (Background Paper Prepared for the Organisation of Economic Cooperation and Development (OECD) Workshop on Pesticide Risk Indicators 21-23 April, 1997, Copenhagen, Denmark). Ithaca, New York: Cornell University. (URL: <http://environmentalrisk.cornell.edu/PRI/Publications/LCL-PestRiskInd7-97.pdf>)

Lion, R., Meertens, R.M. & Bot, I. (2002): Priorities in information desire about unknown risks. *Risk Analysis*, 22 (4), 765-776.

Lindell, M.K. & Earle, T.C. (1980): *Public Attitudes toward Risk Tradeoffs in Energy Policy Choices*. Abschlussbericht, Seattle, Battelle Human Affairs Research Centers.

Madjar, M. & von Rohr, P.R. (1995): *Risikoanalyse verfahrenstechnischer Anlagen*. Polyprojekt Risiko und Sicherheit, Leitfäden, Nr. 5. Zürich: vdf Hochschulverlag AG an der ETH Zürich.

Magat, W.A., Viscusi, W. K. & Huber, J. (1987): Risk-dollar tradeoffs, risk perceptions, and consumer behavior. In W.K. Viscusi & Magat (Eds.), *Learning about risk* (pp. 83 – 97): Cambridge, MA: Harvard University Press.

Mathers, C. D., et al (2002): *Global Burden of Disease 2000: Version 2 methods and results*. Global Programme on Evidence for Health Policy, Discussion Paper No. 50. Geneva: WHO.

URL (2003-11-18): http://www3.who.int/whosis/discussion_papers/zip/paper50.zip

Maxwell, D.T. (2002): Decision analysis survey: Adding insight VI. *OR/MS Today*, June.

Mayerhofer, P., Krewitt, W. & Friedrich, R. (eds.) (1997): *Extension of the Accounting Framework*. ExternE Core Project (Maintenance, Improvement, Extension and Application of the ExternE Accounting Framework): Final Report. Stuttgart: Institut für Energiewirtschaft und Rationelle Energieanwendung (IER): URL (2003-03-12): <http://externe.jrc.es/meth2.pdf>.

Mazur, A. (2003): *True warnings and false alarms about technology: 1948 - 1971*. Washington, DC: Resources for the Future.

McCloskey, M. (1994): Problems with Inappropriately Broad Use of Risk Terminology. *The Comparative Risk Bulletin*, January 1994.

Mekel, O.C.L. & Fehr, R. (2000): Probabilistische Expositionsabschätzung in Umweltmedizin und Verbraucherschutz - Teil III: Berücksichtigung von Variabilität und Unsicherheit in quantitativen Risikoabschätzungen (QRA). *UWSF - Z. Umweltchem. Ökotox.*, 12(1), 42-50.

Merkhofer, M.W. (1987). *Decision science and social risk management*. Dordrecht: Reidel.

Mertens, J. (1993): Risikoanalyse. In H. Schütz & P.M. Wiedemann (Eds.), *Schlüsselbegriffe der Technikbewertung* (pp. 187-192): Frankfurt/Main: IMK.

Minard, R.A. (1993): Critical Values at Risk: Lessons from Vermont's Quality of Life Analysis. Northeast Center for Comparative Risk, Vermont Law School, South Royalton, Vermont.

Morgan, K.M., DeKay, M.L., Fischbeck, P.S. et al. (2001): A deliberative method for ranking risks (II): Evaluation of validity and agreement among risk managers. *Risk Analysis*, 21, 923-937.

Morgan, M.G. & Henrion, M. (1990): *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy research*. Cambridge: Cambridge University Press.

Morgan, M.G., Florig, H.K., DeKay, M.L. & Fischbeck, P. (2000): Categorizing risks for risk ranking. *Risk Analysis*, 20(1), 49-58.

Morgenstern, R. D. & Sessions, S. L. (1988): "Weighing environmental risks: EPA's unfinished business." *Environment*, 30 (6): 14-30.

Morgenstern, R.D., Shih, J.-S. & Sessions, S.L. (2000): Comparative risk assessment: an international comparison of methodologies and results. *Journal of Hazardous Materials*, 78(1-3), 19-39.

Mosbach-Schulz, O. (1999): Probabilistische Expositionsabschätzung in Umweltmedizin und Verbraucherschutz - Teil II: Methodische Aspekte probabilistischer Modellierung. *UWSF - Z. Umweltchem. Ökotox.*, 11(5), 292-298.

Mosteller, F. & Youtz, C. (1990): Quantifying probabilistic expressions. *Statistical Science*, 5, 2 -16.

Moxey, L. M. & Sanford, A. J. (2000): Communicating Quantities: A Review of Psycholinguistic Evidence of How Expressions Determine Perspective. *Applied Cognitive Psychology*, 14, 237 - 255.

Moxey, L.M. & Sanford, A.J. (1993): *Communicating Quantities. A Psychological Perspective. Essays in Cognitive Psychology*. Hove, UK: Erlbaum.

Müller-Herzog, U. & Scheringer, M. (1999): *Zur Umweltgefährdungsbewertung von Schadstoffen und Schadstoffkombinationen durch Reichweiten- und Persistenzanalysen* (Graue Reihe Nr. 18): Bad Neuenahr-Ahrweiler: Europäische Akademie zur Erforschung der Folgen wissenschaftlich-technischer Entwicklungen.

Murphy (2001): Framing the nicotine debate: A cultural approach to risk. *Health Communication*, 13, 2, 119-140.

Murray C.J.L. and Lopez, A.D. (eds) (1998): *Health dimensions of sex and reproduction: The global burden of sexually transmitted diseases, HIV, maternal conditions, perinatal disorders and congenital anomalies*. (GBD Series Vol. III. Harvard School of Public Health on behalf of the World Health Organization and the World Bank, Cambridge, Massachusetts):

Murray, C.J.L. and Lopez, A.D. (eds) (1996): *The Global Burden of Disease. A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. (GBD Series Vol. I. Harvard School of Public Health on behalf of the World Health Organization and the World Bank, Cambridge, Massachusetts, 1996):

Murray, C. J. L. & Lopez, A. D. (1996): *The Global Burden of Disease and Injury Series, Executive Summary*. Burden of Disease Unit, Harvard School of Public Health. URL (2003-11-18): <http://www.hsph.harvard.edu/organizations/bdu/GBDseries.html>

Neale, M.A. & Bazerman, M.H. (1985): The effects of framing and negotiator overconfidence on bargaining behaviors and outcomes. *Academy of Management Journal*, 28 (1), 34-49.

National Research Council. (1989): *Improving risk communication*. Washington, DC: National Academy Press.

National Research Council (1994): *Science and judgment in risk assessment*. Washington, DC: National Academy Press.

National Health and Medical Research Council (2000): *How to use the evidence: assessment and application of scientific evidence* (online verfügbar unter: <http://www.nhmrc.health.gov.au/publications/pdf/cp69pdf>).

National Health and Medical Research Council (NHMRC) (1999): *A Guide to the Development, Implementation and Evaluation of Clinical Practice Guidelines*. Canberra, NHMRC.

National Research Council (NRC) (1996): *Understanding Risk. Informing Decisions in a Democratic Society*. National Academic Press, Washington, DC.

NEA – Nuclear Energy Agency (2000): *Nuclear energy in a sustainable development perspective*. Paris, France: Nuclear Energy Agency.

Neubert, D. (1994): Möglichkeiten und Methoden der quantitativen Risikoabschätzung. In H. Marquardt & S.G. Schäfer (Eds.), *Lehrbuch der Toxikologie* (pp. 840-913): Mannheim: BI Wissenschaftsverlag.

Neus, H., Ollroge, I., Schmid-Höpfner, S. & Kappos, A. (1997): *Zur Harmonisierung gesundheitsbezogener Umweltstandards - Probleme und Lösungsansätze*. Bericht zum F+E-Vorhaben 116 01 001, im Auftrag des Umweltbundesamts. Hamburg: Behörde für Arbeit, Gesundheit und Soziales, Amt für Gesundheit. Veröffentlicht als: Neus, H., Ollroge, I., Schmid-Höpfner, S. & Kappos, A. (1998): *Aktionsprogramm Umwelt und Gesundheit, Teilvorhaben: Zur Harmonisierung gesundheitsbezogener Umweltstandards - Probleme und Lösungsansätze*. Forschungsbericht 116 01 001, UBA-Projekt Nr. 98/01, UBA-FB Nr. 97-095. Behörde für Arbeit, Gesundheit und Soziales, Amt für Gesundheit, Hamburg, im Auftrag des Umweltbundesamts. Berlin: Erich Schmidt Verlag.

Neus, H., Sagunski, H., Kappos, A. & Schümann, M. (1995): Zur administrativen Umsetzung von Risikoabschätzungen. (Teil I). *Bundesgesundheitsblatt*, 7, 258-264.

Neutra, R., DelPizzo, V. & Lee, G.M. (2002). *An evaluation of the possible risks from electric and magnetic fields (EMFs) from power lines, internal wiring, electrical occupations and appliances* (Final Report). Oakland, CA: California EMF Project. (Online unter: <http://www.dhs.ca.gov/ehib/emf/RiskEvaluation/riskeval.html>).

Newstead, S.E. & Collis, J.M. (1987): Context and the interpretation of quantifiers of frequency. *Ergonomics*, 30, 1447-1462.

NHS Centre for Reviews and Dissemination (2001): *Undertaking Systematic Reviews of Research on Effectiveness: CRD's Guidance for Carrying Out or Commissioning Reviews*. Im Internet unter: <http://www.york.ac.uk/inst/crd/report4.htm>.

Nisbett, R.E. & Ross, L. (1980): *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.

NJDEP (2003). *Final report of the New Jersey Comparative Risk Project*. Trenton, NJ: New Jersey Department of Environmental Protection. (Online: <http://www.state.nj.us/dep/dsr/njcrp/>)

OECD (2003): *Workshop on near earth objects: Risks, policies and actions*. Frascati, Italy: Organisation for Economic Co-operation and Development.

OSPAR Commission (2000): *Briefing Document on the Work of DYNAMEC and the DYNAMEC Mechanism for the Selection and Prioritisation of Hazardous Substances*.

OSPAR Commission (2002): *Dynamic Selection and Prioritisation Mechanism for Hazardous Substances*.

Pate-Cornell, M.E. (1996): Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety*, 54(2-3), 95-111.

Peters, H.P. (1993): Risikokonflikt / Risikokontroverse. In H. Schütz & P.M. Wiedemann (Eds.), *Schlüsselbegriffe der Technikbewertung* (pp. 203-208): Frankfurt/Main: IMK.

Powell JC, 1996. The Evaluation of Waste Management Options, *Waste Management and Research*, 14, 515-526.

Presidential/Congressional Commission on Risk Assessment and Risk Management (1997): *Framework for Environmental Health Management. Volumes 1 & 2*. Washington, DC (Im Internet unter: <http://cfpub.epa.gov/ncea/cfm/pcrarm.cfm?ActType=default>).

Pruitt, D. & Rubin, J.Z. (1986): *Social conflict: escalation, impasse, and resolution*. Reding, MA: Addison-Wesley.

Raiffa, H. (1982): *The art and science of negotiation*. Cambridge, MA: Harvard University Press.

Ramsberg, J.A. & Sjöberg, L. (1997): The cost-effectiveness of lifesaving interventions in Sweden. *Risk Analysis*, 17(4), 467-478.

Reagan, R. T., Mosteller, F. & Youtz, C. (1989): Quantifying meanings of verbal probability expression. *Journal of Applied Psychology*, 74, 433 – 442

Rehbinder, E. (1988): Vorsorgeprinzip im Umweltrecht und präventive Umweltpolitik. In E.U. Simonis (Ed.), *Präventive Umweltpolitik* (pp. 129-143): Frankfurt am Main: Campus.

Renn, O., Carius, R., Kastenholz, H. & Schulze, M. (2003). *Entwicklung eines mehrstufigen Verfahrens der Risikokommunikation (ERiK). Ein Leitfaden für die Oberen Bundesbehörden*. Stuttgart: Akademie für Technikfolgenabschätzung in Baden-Württemberg (im Auftrag des Umweltbundesamtes).

Renn, O., Webler, Th. & Wiedemann, P.M. (Eds.), (1995): *Fairness and Competence in Citizen Participation. Evaluating Models for Environmental Discourse*. Dordrecht, Kluwer Academic Publishers.

Reyna, V. F. & Brainerd, C. J. (1991): Fuzzy-trace and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making*, 4, 249-262.

Risikokommission (2002): *1. Bericht der Risikokommission*. Salzgitter: Geschäftsstelle der Risikokommission, Bundesamt für Strahlenschutz (Im Internet unter: <http://www.apug.de>).

Risikokommission (2003). *Abschlussbericht der Risikokommission*. Salzgitter: Geschäftsstelle der Risikokommission, Bundesamt für Strahlenschutz (Im Internet unter: <http://www.apug.de>).

Ritov, I. & Baron, J (1992): Status-Quo and omission biases. *Journal of Risk and Uncertainty*, 5, 49-62.

Ritov, I. & Baron, J. (1999): Protected Values and Omission Bias. *Organizational Behavior and Human Decision Processes*, 79, 79-94.

Rodricks, J.V. (1992): *Calculated risks*. Cambridge: Cambridge University Press.

Rohrmann, B. & Schütz, H. (1993). The evaluation of decision aiding programs. In S. Nagel (Ed.), *Computer-aided decision analysis* (pp. 5-29). Westport: Greenwood.

Roth, E., Morgan, M.G., Fischhoff, B., Lave, L. & Bostrom, A. (1990): What Do We Know About Making Risk Comparisons? *Risk Analysis*, 10, 3, 375-387.

Rothman, K.J. & Greenland, S. (1998). Causation and causal inference. In K.J. Rothman & S. Greenland (Eds.), *Modern epidemiology* (2 ed.). Philadelphia: Lippincott Wilkins & Wilkins.

Russo, J.E. & Schoemaker, P.J.H. (1990): *Decision Traps*. New York: Simon & Schuster, 1990, 75-76.

Salomon, J. A. et al. (2001): *Methods for life expectancy and healthy life expectancy uncertainty analysis*. Global Programme on Evidence for Health Policy, Working Paper No. 10. Geneva: WHO.

URL (2003-11-18): http://www3.who.int/whosis/discussion_papers/zip/paper10.zip

Sanford, A. J., Fay, N., Stewart, A. & Moxy, L. (2002): Perspective in Statements of Quantity, with Implications for Consumer Psychology. *Psychological Science*, 13, 2, 130 – 134.

Schneider, K. (2000): *Anwendungsbeispiele für quantitative Krebsrisikoabschätzungen in der regulatorischen Toxikologie*. Freiburg: Forschungs- und Beratungsinstitut Gefahrstoffe GmbH (FoBiG).

URL (2003-11-12): <http://www.fobig.de/Downloads/FoBiG-Risk.PDF>

Schneider, K., Konietzka, R. & Schuhmacher-Wolz, U. (2001): "Klassierung krebserzeugender Luftschadstoffe für die TA-Luft-Novelle." *Gefahrstoffe - Reinhaltung der Luft*, 61 (10): 454-458.

Schoemaker, P.J.H. & Russo, J.E. (2001): Managing Frames to Make Better Decisions. In: S.J. Hoch, H.C. Kunreuther, R.E. Gunther (Eds.), *Wharton on Making Decisions*, John Wiley & Sons Inc., New York, pp. 131-157.

Schümann, M. (2000): Probabilistische Expositionsabschätzung in Umweltmedizin und Verbraucherschutz - Teil V: Probabilistische Modelle der Expositionsabschätzung - Möglichkeiten der Validierung und des Vergleichs am Beispiel einer Altlast mit Bodenbelastungen. *UWSF - Z. Umweltchem. Ökotox.*, 12(3), 169-179.

Schütz, H. & Wiedemann, P.M. (2003): Risikowahrnehmung in der Gesellschaft. *Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz*, 46(7), 549-554.

Schütz, H., Wiedemann, P.M. & Gray, P.C.R. (2000): *Risk perception beyond the psychometric paradigm*. Arbeiten zur Risiko-Kommunikation, Heft 78. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich GmbH.

Schuhmacher-Wolz, U., Konietzka, R. & Schneider, K. (2002): Using Carcinogenic Potency Ranking to Assign Air Contaminants to Emission Classes. *Regulatory Toxicology and Pharmacology*, 36 (3): 221-233.

Shrader-Frechette, K.S. (1995): Comparative risk assessment and the naturalistic fallacy. *Tree*, 10(1), 50.

Sicherheitskriterien für Kernkraftwerke (1977): Der Bundesminister des Innern: *Bekanntmachung von Sicherheitskriterien für Kernkraftwerke*. Bonn: Bundesanzeiger Nr.206 vom 3.11.1977

Sielken, R.L. (2000): Risk metrics and cumulative risk assessment methodology for the FQPA. *Regulatory Toxicology and Pharmacology*, 31(3), 300-307.

Silbergeld, E. K. (1995): The hazards of synthetic (anthropogenic) chemicals. *Toxicology Letters*, 82/83, 835-841.

Sjöberg, L. (1996): A discussion of the limitations of the psychometric and Cultural Theory approaches to risk perception. *Radiation Protection Dosimetry*, 68, 219-225.

Slovic, P. (1987): Perceptions of Risk. *Science*, 236, 280-285.

Slovic, P. ; Kraus, N. & Covello, V.T. (1990): What Should We Know About Making Risk Comparisons. Comment. *Risk Analysis*, 10, 3, 389-392.

SRU - Der Rat von Sachverständigen für Umweltfragen SRU (1996): *Umweltgutachten 1996. Zur Umsetzung einer dauerhaft umweltgerechten Entwicklung*. Stuttgart: Metzler-Poeschel.

SSK (1994): Strahlenschutzgrundsätze zur Begrenzung der Strahlenexposition durch Radon und seine Zerfallsprodukte in Gebäuden, veröffentlicht im Bundesanzeiger Nr. 155 vom 18. August 1994

SSK – Strahlenschutzkommission (2001): Grenzwerte und Vorsorgemaßnahmen zum Schutz der Bevölkerung vor elektromagnetischen Feldern – Empfehlungen der Strahlenschutzkommission. Bonn, SSK.

Stone, E.R., Yates, J.F. & Parker, A.M. (1994): Risk communication:- Absolute versus relative expressions of low-probability risks. *Organizational Behavior and Human Decision Processes*, 60, 387-408.

Strahlenschutzkommission (2001): *Grenzwerte und Vorsorgemaßnahmen zum Schutz der Bevölkerung vor elektromagnetischen Feldern - Empfehlungen der Strahlenschutzkommission*. Bonn: SSK.

Streffer, C., Bückner, J., Cansier, A., Cansier, D., Gethmann, C.F., Guderian, R., Hanekamp, G., Henschler, D., Pösch, G., Rehbinder, E., Renn, O., Slesina, M. & Wuttke, K. (2000). *Umweltstandards: kombinierte Expositionen und ihre Auswirkungen auf den Menschen und seine Umwelt*. Berlin: Springer.

Swartjes, F.A. (2002): *Variation in calculated human exposure. Comparison of calculations with seven European human exposure models* (RIVM report 711701030 /2002): Bilthoven, The Netherlands: RIVM - National Institute for Public Health and the Environment.

Tamburini, G., von Ehrenstein, O.S. & Bertollini, R. (2002): *Children's health and environment: A review of evidence*. Luxembourg: World Health Organization and European Environment Agency.

Taubes, G. (1995): Epidemiology faces its limits. *Science*, 269, 164-169.

Teigen, K. H. (1988): The language of uncertainty. *Acta Psychologica*, 68, 27 – 38.

Teigen, K. H. & Brun, W. (2000): Ambiguous Probabilities: When Does $p = 0.3$ Reflect a Possibility, and When Does It Express a Doubt? *Journal of Behavioral Decision Making*, 13, 345 – 362.

Teigen, K. H. & Brun, W. (1999): The Directionality of Verbal Probability Expressions: Effects on Decisions, Predictions, and Probabilistic Reasoning. *Organizational Behavior and Human Decision Processes*, 80, 2, 155 – 190.

Ten Berge, W.F. & Stallen, J.P.M. (1995): How to compare the risk assessments for accidental and chronic exposure. *Risk Analysis*, 15, 2, 111-113.

Tengs, T.O., Adams, M.E., Pliskin, J.S., Safran, D.G., Siegel, J.E., Weinstein, M.C. & Graham, J.D. (1995): Five-hundred life-saving interventions and their cost-effectiveness. *Risk Analysis*, 15(3), 369-390.

Thalmann, A. T. (in Vorbereitung): Bewertung von Strategien zur Vermittlung undeutlicher Risiken und Identifizierung von Kriterien zur Optimierung Laien-gerechter Kommunikation bei unterschiedlicher Evidenzstärken. Dissertation.

Thompson, L. (1990): Negotiation behavior and outcomes: empirical evidence and theoretical issues. *Psychological Bulletin*, 108 (3), 515-532.

Thompson, L. & Hastie, R. (1990): Social perception in negotiation. *Organizational Behavior and Human Decision Processes*, 47, 98-123.

Thompson, K.M. & Bloom, D.L. (2000): Communication of risk assessment information to risk managers. *Journal of Risk Research*, 3, 333-352.

Tufte, E.R. (1983). *The visual display of quantitative information*. Cheshire, Conn.: Graphics Press.

Tversky, A. & Kahnemann, D. (1974): Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.

Tversky, A. & Kahneman, D. (1981): The framing of decisions and the psychology of choice. *Science*, 211, 453-458.

U.S. Environmental Protection Agency (U.S. EPA), Office of Policy Analysis 1987: Unfinished Business: A Comparative Assessment of Environmental Problems, Overview Report. Washington, D.C.: U.S. EPA.

U.S. EPA Science Advisory Board (2000): Toward Integrated Environmental Decision-Making. EPA-SAB-EC-00-011.

US-EPA (1993): *Reference Dose (RfD): Description and Use in Health Risk Assessments*. Background Document 1A. Washington: United States Environmental Protection Agency (EPA): URL (2003-05-20): <http://www.epa.gov/iris/rfd.htm>.

US-EPA National Center for Environmental Assessment (1998): *Comparative Risk Framework Methodology and Case Study*. SAB Review Draft. Cincinnati: United States Environmental Protection Agency (EPA): URL (2003-04-16): <http://www.epa.gov/ncea/frame.htm>.

US-EPA Risk Assessment Forum (Oct. 2000): *Benchmark Dose Technical Guidance Document*. Washington: United States Environmental Protection Agency (EPA): URL (2003-05-19): http://www.epa.gov/ncea/pdfs/bmds/BMD-External_10_13_2000.pdf, homepage: http://www.epa.gov/ncea/bnchrk/bmds_peer.htm.

US-EPA Risk Assessment Forum Technical Panel (2000): *Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures* (EPA/630/R-00/002): Washington: United States Environmental Protection Agency (EPA): URL (2003-05-23): http://www.epa.gov/nceawww1/pdfs/chem_mix/chem_mix_08_2001.pdf.

US-EPA Science Advisory Board (1999): *An SAB Report on the National Center for Environmental Assessment's Comparative Risk Framework Methodology. A Review by the Drinking Water Committee*. Washington: United States Environmental Protection Agency (EPA) Science Advisory Board.

Uth, H.-J. (1991): Expertise: Risiko-Kommunikation in der Chemie. In: H. Jungermann, B. Rohrmann & P.M. Wiedemann (eds.), *Risiko-Kontroversen – Konzepte, Konflikte, Kommunikation*. Berlin: Springer Verlag.

Viscusi, W. K. & Magat, W. A. (1987): *Learning about risk*. Cambridge, MA: Harvard University Press.

Vlek, C. & Stallen, P.J. (1981): Judging risks and benefits in the small and in the large. *Organizational Behavior and Human Decision Processes*, 28(2), 235-271.

Von Winterfeldt, D. & Edwards, W. (1986): *Decision analysis and behavioral research*. Cambridge, MA: Cambridge University Press.

Vose, D. (1996). *Quantitative risk analysis*. New York: Wiley.

Wallsten, T., Budescu, D. V., Zwick, R. & Kemp, S. M. (1993): Preference and reasons for communicating probabilistic information in numerical or verbal terms. *Bulletin of the Psychonomic Society*, 31, 135 – 138.

Wallsten, T.S., Fillenbaum, S. & Cox, J. A., (1986): Base rate effects on the interpretation of probability and frequency expressions. *Journal of Memory and Language*, 25, 571 – 587.

WBGU Wissenschaftlicher Beirat "Globale Umweltveränderungen" (1998): *Welt im Wandel - Strategien zur Bewältigung globaler Umweltrisiken. Jahresgutachten 1998*. Berlin: Springer.

Weber, E., U. & Hilton, D., J. (1990): Contextual Effects in the Interpretation of Probability Words: Perceived Base Rate and Severity of Events. *Journal of Experimental Psychology: Human Perception and Performance* Vol. 16. No. 4, 781-789.

Weinstein, N.D., Sandman, P.M. & Robert, N.E. (1989): Communicating effectively about risk magnitudes. Washington; DC. United States Environmental Protection Agency, Office of Policy, Planning and Evaluation (EPA 230/08-89-064):

WHO (2002) Establishing a dialogue on risks from electromagnetic fields. Geneva, Switzerland: Radiation and Environmental Health, Department of Protection of the Human Environment, World Health Organization.

Wiedemann, P.M. & Balderjahn, I. (1998): *Komparative Risikobewertung und akteursbezogene Konsensancen*. Arbeiten zur Risiko-Kommunikation, Heft 67. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Wiedemann, P.M. & Balderjahn, I. (1999): *Akteursspezifische Urteilsmodelle zur Bewertung von Risiken*. Arbeiten zur Risiko-Kommunikation, Heft 72. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Wiedemann, P.M. & Brüggemann, A. (2001): *Vorsorge aus der Perspektive der Sozialwissenschaft: Probleme, Sachstand und Lösungsansätze*. Arbeiten zur Risiko-Kommunikation Heft 82. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Wiedemann, P.M. & Kresser, R. (1997): *Intuitive Risikobewertung – Strategien der Bewertung von Umweltrisiken*. Arbeiten zur Risiko-Kommunikation Heft 62. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Wiedemann, P.M. & Schütz, H. (2002): *Wer fürchtet den Mobilfunk? Gruppenspezifische Differenzen bei der Risikowahrnehmung*. Arbeiten zur Risiko-Kommunikation Heft 84. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Wiedemann, P.M. (1993a): Risiko-Kommunikation. In H. Schütz & P.M. Wiedemann (Eds.), *Schlüsselbegriffe der Technikbewertung* (pp. 197-202): Frankfurt/Main: IMK.

Wiedemann, P.M. (1993b): Risikomanagement. In H. Schütz & P.M. Wiedemann (Eds.), *Schlüsselbegriffe der Technikbewertung* (pp. 209-213): Frankfurt/Main: IMK.

Wiedemann, P.M. (1998): *Die Wahrnehmung von Umweltproblemen und die Beurteilung von Handlungspräferenzen*. Arbeiten zur Risiko-Kommunikation, Heft 66. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Wiedemann, P.M., Karger, C.R. & Clauberg, M. (2002): Machbarkeitsstudie Risikofrüherkennung im Bereich Umwelt und Gesundheit: Teilvorhaben 9; Aktionsprogramm Umwelt und Gesundheit (im Auftrag des Umweltbundesamtes Berlin) Forschungszentrum Jülich, Programmgruppe Mensch, Umwelt, Technik, 2002. Umweltforschungsplan des Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit, F+E-Vorhaben , 200 61 218/09.

Wiedemann, P.M., Karger, C.R. & Homburg, A. (1996): *Subjektive Taxonomien gefährdeter Umweltgüter*. Arbeiten zur Risiko-Kommunikation Heft 58. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Wiedemann, P.M., Schütz, H. & Thalmann, A.T. (2002): *Mobilfunk und Gesundheit. Risikobewertung im wissenschaftlichen Dialog*. Programmgruppe Mensch, Umwelt, Technik. Forschungszentrum Jülich.

Williams, P.R. & Paustenbach, D.J. (2002): Risk characterization: principles and practice. *J Toxicol Environ Health B Crit Rev*, 5(4), 337-406.

Wilson, R. & Crouch, E.A.C. (1987): Risk assessment and comparisons: An introduction. *Science*, 236, 268-270.

Wintermeyer, D. (1999): Probabilistische Expositionsabschätzung in Umweltmedizin und Verbraucherschutz - Teil I: Probabilistische Expositionsabschätzung zur Beurteilung der gesundheitlichen Auswirkung von Umweltbelastung. *UWSF - Z. Umweltchem. Ökotox.*, 11(4), 228-233.

Yoon, K.P. & Hwang, C.-L. (1995): *Multiple attribute decision making: An introduction*. Thousand Oaks, CA: Sage.

Zadeh, L. A. (1978): Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and System*, 1, 3 – 28

Zimmer, A. C. (1983): Verbal versus numerical processing of subjective probabilities. in R. W. Scholz (Ed.), *Decision making under uncertainty*. Amsterdam: Elsevier.

Glossar

ACP: Advisory Committee on Pesticides

ADI-Werte (*Acceptable Daily Intake*): akzeptable tägliche Aufnahmemenge für Substanzen. Der ADI-Wert stellt die maximale Menge einer Substanz dar, der ein Subjekt täglich über die ganze Lebensspanne ausgesetzt sein kann, ohne das ein Gesundheitsrisiko besteht.

Adverse Effekte: unerwünschte Wirkung. Veränderung in der Morphologie, Physiologie, Wachstum, Entwicklung oder Lebensspanne eines Organismus, einer Population oder eines ökologischen Systems mit dem Resultat einer Beeinträchtigung der funktionalen Kapazität oder der Kapazität zusätzlichen Stress oder schädliche Effekte zu kompensieren.

Attributal Fraction (AF): Attribuierbarer Anteil ist definiert als die prozentuale Reduzierung des Auftretens einer Behinderung oder der Sterblichkeit aufgrund eines Risikofaktors, wenn dieser Faktor eliminiert wird.

ALARA (*As Low As Reasonably Achievable*): so niedrig wie vernünftigerweise zu erreichen.

APUG: Aktionsprogramm "Umwelt und Gesundheit". Gemeinsame Aktionsprogramm des Ministeriums für Gesundheit und des Ministeriums für Umwelt, Naturschutz und Reaktorsicherheit (www.apug.de).

Attribuierbares Risiko: Das attribuierbare Risiko (*attributable risk*) misst den spezifischen Beitrag einer Exposition zur Gesamtinzidenz bemessen.

Basisrate: Die Häufigkeit, mit der ein bestimmtes Ereignis auftritt. Liefert die Basis für die Bestimmung der a-priori Wahrscheinlichkeit eines Ereignisses.

Benchmark Dose (BMD): entspricht dem unteren Konfidenzbereich der Effekt-Dosis, bei der in x% der untersuchten Testpopulation eine Wirkung aufgetreten ist.

BfG: Bundesanstalt für Gewässerkunde (<http://www.bafg.de/>).

BfR: Bundesinstitut für Risikobewertung (www.bgvv.de).

BfS: Bundesamt für Strahlenschutz (<http://www.bfs.de/>).

BGVV: ehemals: Bundesinstitut für gesundheitlichen Verbraucherschutz und Veterinärmedizin; jetzt: Bundesinstitut für Risikobewertung (BfR).

BMU: Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (<http://www.bmu.de>).

BVEL: Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft (<http://www.verbraucherministerium.de/>).

CERA (*Comparative Extreme Risk Assessment*): Vergleich mehrerer Extremrisiken auf mehreren Beurteilungsdimensionen durch verschiedene Bewerter.

CHA (*Comparative Hazard Assessment*): vergleichende Gefährdungsbewertung. Vergleich der Einschätzung mehrerer Bewerter auf mehreren Beurteilungsdimensionen in Bezug auf ein undeutliches Risiko.

Conjoint Analyse: Ein ursprünglich in der mathematischen Psychologie entwickeltes, der Multidimensionalen Skalierung (MDS) ähnliches Analyseverfahren der Interdependenzanalyse zur Messung von subjektiven Werturteilen.

Counterfactual distribution: Vergleich der Konsequenzen aus der vorliegenden Expositionsverteilung mit denjenigen Konsequenzen aus einer fiktiven, postulierten Verteilung, die einem bestimmten Szenario entspricht zur Ermittlung des Beitrags eines Risikofaktors.

CPSC: Consumer Product Safety Commission.

DALY (*Disability Adjusted Life Years*): Anzahl der Behinderungs-justierten Lebensjahre. Summe der verlorenen Lebensjahre und der mit Behinderung verbrachten Lebensjahre.

DFLE (*Disability-free life expectancy*): Behinderungsfreie Lebenserwartung. Lebenserwartung, die sich für den Gesundheitsstatus ergibt, in dem ein bestimmter Grad an Behinderung nicht überschritten wird.

DIFU: Deutsches Institut für Urbanistik, Berlin.

Dobris-Assessment: Projekt der EEA von 1995 zur Bewertung von 56 verschiedenen Umweltproblemen.

EEA (*European Environmental Agency*): Europäische Umweltbehörde, Kopenhagen.

Endpunkte: Die in einer Untersuchung interessierende Veränderungen oder Krankheiten.

EPA (*Environmental Protection Agency*): Umweltschutzbehörde der USA.

Expositionssituation: Umstand, bei dem man bestimmten schädlichen Noxen ausgesetzt ist.

GaBE: „Ganzheitliche Betrachtung von Energiesystemen“ – Projekt des Paul Scherrer Instituts (PSI) mit dem Ziel, eine umfassende Methodik zur konsistenten und detaillierten Bewertung von Energiesystemen zu entwickeln.

GBD / Global Burden of Disease Study: von der Weltbank, der Harvard School of Public Health und der WHO initiiertes Projekt mit dem Ziel einer umfassenden Analyse der Belastung der Bevölkerung durch Krankheit und Verletzungen.

Gefahr: das Potenzial einer Risikoquelle (Technologie, Produkt, Substanz, Situation, Aktivität etc.), adverse Effekte zu verursachen.

HALE (*Health-adjusted life expectancy*): gesundheits-justierte Lebenserwartung. Diejenigen summarischen Messgrößen, bei denen explizite Gewichtungen verwendet werden, um Gesundheitserwartungen einzelner Gesundheitszustände zusammenzufassen.

HE (*Health expectancy*): Gesundheitserwartung. Darunter werden alle summarische Messgrößen für den Gesundheitsstatus der Bevölkerung zusammengefasst, die abschätzen, wie viele Jahre an Lebenszeiten in einem bestimmten Gesundheitszustand – etwa als gesunder Jugendlicher – zu erwarten sind.

Human – Biomonitoring: Mit Monitoring ist die kontinuierliche oder wiederholte Beobachtung, Messung, und Evaluation der Gesundheit und/oder Umwelt oder technischer Daten für bestimmte Zwecke gemeint. Das Human-Biomonitoring (HBM) spielt bei Bewertungsfragen in der Umweltmedizin und für die Beurteilung der korporalen Schadstoffbelastung der Bevölkerung sowie von Personengruppen und Einzelpersonen eine wesentliche Rolle.

IARC (*International Agency for Research on Cancer*): Internationale Krebsforschungsagentur, Lyon.

ICD (*International Classification of Diseases*): eine internationale Klassifikation für Krankheiten und Todesursachen. Aktuell ist ICD 10 = Internationale statistische Klassifikation der Krankheiten und gesundheitsbezogenen Probleme, 10. Revision.

Individual Exposure Ratio (IER): Quotient aus einer Expositions-dosis und der Reference Dose (RfD). Das IER gibt an, welchem Ausmaß die aufgenommene Dosis einer Substanz den Wert einer sicheren Dosis überschreitet.

Individual Risiken: Wahrscheinlichkeit, dass eine individuelle Person adversen Effekte ausgesetzt sein wird.

Intake Fraction: Maß zur Abschätzung des Transports einer Noxe zwischen Freisetzung und Exponierten.

Inzidenzrate: Häufigkeit des Neuauftretens einer Krankheit innerhalb eines bestimmten Zeitraumes unabhängig davon, ob die Erkrankung am Ende der Zeitperiode noch besteht oder nicht.

IRIS: Integrated Risk Information System, US-EPA, USA.

Kollektivrisiko: Risiko, welches die Gesellschaft als Ganzes betrifft.

Konfidenzintervalle: ist der auch Vertrauensbereich genannte Bereich für eine Parameterschätzung im Rahmen eines statistischen Modells. In diesem Intervall liegt unter Berücksichtigung der Zufälligkeit des Ergebnisses der "wahre" Parameter. Je sicherer diese Aussage sein soll, d.h. je niedriger die Irrtumswahrscheinlichkeit, umso größer wird dieser Bereich. Zwei geschätzte Parameter können als signifikant ver-

schieden angesehen werden, wenn sich ihre Vertrauensbereiche nicht überschneiden, d.h. kein gemeinsamer Wert möglich ist.

LAI: Länderausschuss für Immissionsschutz (<http://www.lai-immissionsschutz.de/>).

LAUG: Länderarbeitsgemeinschaft umweltbezogener Gesundheitsschutz.

LCA (*Life Cycle Assessment*): Lebenszyklusanalyse.

Life-style Risiken: Aus der Lebensweise bzw. aus dem eigenen Handeln resultierende Risiken.

LOAEL (*Lowest Observed Adverse Effect Level*): Niedrigste Dosis, bei der ein adverser (krankmachender) Effekt festgestellt werden konnte oder bei der ein adverser Effekt zu erwarten ist.

LOEL (*Lowest Observed Effect Level*): Niedrigste Dosis, bei der ein Effekt festgestellt werden konnte oder bei der ein Effekt zu erwarten ist.

MAK-Werte: Maximaler Arbeitsplatzkonzentration: Maximale Konzentration eines Arbeitsstoffes in der Luft am Arbeitsplatz, die bei einem 8h-Arbeitstag über die Arbeitslebenszeit (ca. 40 Jahre) für gesunde Personen im erwerbstätigen Alter, die zudem einer kontinuierlichen betriebsärztlichen Überwachung unterliegen, noch gesundheitlich zulässig ist.

Margin of Exposure (MOE): Der Quotient aus NOAEL (oder einem anderen Wert für die Wirkungsschwelle) und der abgeschätzten Exposition.

MAUT (*Multi-Attribute-Utility Theory*): multiattribute Nutzentheorie.

MCRA (*Multi Attribute Risk Assessment*): Multiattribute vergleichende Risikobewertung. Vergleich mehrerer Risiken auf mehreren Beurteilungsdimensionen durch verschiedene Bewerter.

Mortalität – Morbidität: Zahl der Todes- bzw. Krankheitsfälle.

Mutagene: Stoffe, Strahlen oder Mikroorganismen, die eine erbgut-schädigende Wirkung haben.

NAS: National Academy of Science, USA.

NEA: (Nuclear Energy Agency) Internationale Kernenergie Behörde der OECD.

NOAEL (*No Observed Adverse Effect Level*): Höchste Dosis, bei der kein adverser (krankmachender) Effekt festgestellt werden konnte oder bei der kein adverser Effekt zu erwarten ist.

NOEL (*No Observed Effect Level*): Höchste Dosis, bei der kein Effekt festgestellt werden konnte oder bei der kein Effekt zu erwarten ist.

Odds Ratio: Mit Odds beschreibt man das Verhältnis der Wahrscheinlichkeit zweier komplementärer, einander ausschliessender Ereignisse. Mit Odds Ratio wird das Verhältnis zweier Odds beschrieben. In der Gruppe der Exponierten sind a die Erkrankten und b die Nichterkrankten. In der Gruppe der Nichtexponierten sind c die Erkrankten und d die Nichterkrankten. Odds für die Exponierten= a/b . Odds für die Nichtexponierten= c/d . Das Verhältnis der zwei Odds zueinander lautet dann $a/b/c/d$ oder $a*d/c*b$.

OECD: (*Organization for Economical Cooperation and Development*): Organisation für wirtschaftliche Zusammenarbeit und Entwicklung der UN, Paris.

OSHA: Occupational Safety and Health Administration USA.

PAK: polycyclische aromatische Kohlenwasserstoffe.

PCB: Polychlorierte Biphenyle, fallen unter die Gruppe der POP.

Perzentile: Perzentil, Quantil: Ein Perzentil, zu dem noch die n %-Angabe gehört, (z.B. das 95%-Perzentil) bezeichnet die Stelle in einer nach Größe geordneten Reihe von Beobachtungswerten, auf die bezogen n % aller Werte kleiner/gleich diesem Wert sind. Perzentile können auch für ordinale Variable angegeben werden. Spezielle Perzentile sind Median (50%), Quartile (25%, 50%, 75%) oder Quintile (20%, 40%, 60%, 80%).

POP: persistente organische Schadstoffe.

Proxyvariable: Ein Indikator, der zur Charakterisierung einer möglichen Gefahrenquelle herangezogen werden kann, wenn das Schädigungspotenzial selbst nicht abgeschätzt werden kann.

PSI: Paul Scherrer Institut (Schweiz).

QALY (*Quality Adjusted Life Years*): Integral des Gesundheitszustand über die Lebenszeit.

Relatives Risiko: Verhältnis der Krankheitshäufigkeit in einer Bevölkerungsgruppe mit einem Risikofaktor im Vergleich zu einer Bevölkerung ohne diesen Risikofaktor: Es gibt somit an, wieviel Mal häufiger bzw. seltener eine bestimmte Erkrankung bei Exponierten im Vergleich zu nicht exponierten auftritt.

Residual-Risiko: Risiko, welches trotz Regulierung (z.B. durch Grenzwerte) noch besteht.

RfD (*Reference Dose*): Referenzdosis – eine Abschätzung der lebenslangen, täglichen oralen Exposition, bei der für Menschen (einschließlich sensibler Menschen) keine Gesundheitsschäden zu erwarten sind.

Risiko: Eine Funktion der Wahrscheinlichkeit einer gesundheitsschädlichen Wirkung sowie des Schweregrads dieser Wirkung als Folge der Exposition mit einer oder mehreren Gefahrenquelle(n).

Risikoabschätzung: Prozess von der Identifizierung des Gefährdungspotentials bis zur quantitativen Charakterisierung von Risiken.

Risiko-Akzeptabilität: annehmbares Risiko. Forderung, die jeweiligen Risiken zu billigen. Als normatives Konzept nicht mit Akzeptanz von Risiken (= akzeptiertes Risiko) gleichzusetzen, was sich als deskriptiver Begriff auf die empirisch feststellbare Billigung von Risiken durch Personen, Gruppen oder einer Gesellschaft bezieht.

Risikobewertung: die Erörterung und Bewertung eines Risikos bezug auf seine Bedeutsamkeit, Schwere oder Ernsthaftigkeit aus naturwissenschaftlicher und gesellschaftlicher Sicht in Konsultation mit allen Betroffenen.

Risikomanagement: Risikomanagement ist der von der Risikoabschätzung und Risikobewertung getrennte Prozess der Abwägung von Handlungsoptionen, unter Berücksichtigung der Risikoabschätzung und anderer für den Gesundheitsschutz relevanter Faktoren und erforderlichenfalls unter Auswahl geeigneter Vorbeugungs- und Kontrollmöglichkeiten in Konsultation mit allen Betroffenen.

Risikoranking: Ordnen von Risiken nach Merkmalen, z.B. des Gefahrenpotenzials, der Exposition. Ergebnis der vergleichenden Risikobewertung.

Risikoregulation: Risikoregulation ist der gesamte Prozess der Identifikation, Abschätzung, Bewertung und des Managements von Risiken.

RKI: Robert Koch Institut (<http://www.rki.de/>).

SAB (Scientific Advisory Board): wissenschaftliches Beratergremium der US-EPA.

Schädigungspotential (*hazard*): Das Potenzial einer Risikoquelle, adverse Effekte zu verursachen.

Schwellenwert: Die kleinste Dosis einer chemischen Substanz, bei welcher ein bestimmter messbarer Effekt beobachtet werden kann und unterhalb derer kein solcher Effekt auftritt.

Sensitive Subpopulationen: Personengruppen, die auf bestimmte Noxen besonders empfindlich reagieren.

Sensitivitätsanalyse: Feststellen der Empfindlichkeit der Analyseergebnisse gegenüber Änderung der Annahmen und Eingangsdaten.

Slope Factor: Risiko pro Dosis-Einheit oder Risiko pro mg/kg-Tag.

SRU (Rat von Sachverständigen für Umweltfragen): wissenschaftliches Beratungsgremium der Deutschen Bundesregierung (<http://www.umweltrat.de/>).

SSK: Deutsche Strahlenschutzkommission (<http://www.ssk.de/>).

StFV: Schweizer Störfallverordnung.

Stressoren: Stressverursachende Ereignisse, Gegebenheiten, Substanzen etc.

TRK-Werte: technische Richtkonzentrationen.

UBA: Umweltbundesamt (<http://www.umweltbundesamt.de/>).

UN/IAEA: Vereinte Nationen / International Atomic Energy Agency, Wien.

Unit Risk: gibt das geschätzte zusätzliche Krebsrisiko eines Menschen nach konstanter inhalativer Exposition über 70 Jahre gegenüber einer Konzentration von 1 µg Schadstoff pro m³ Luft.

Vorsorgeprinzip: Prinzip, dass Umweltpolitik über die Beseitigung von eingetretener Schäden und die Abwehr konkreter Gefahren hinaus schon im Vorfeld von Gefahren das Entstehen von Umweltbelastungen verhindern oder einschränken und dadurch Risiken vermindern soll. Verfassungsrechtlich verankertes Vorsorgeprinzip (Art. 20a GG): Der Staat ist schon dann zum Handeln aufgerufen, wenn Schadensmöglichkeiten gegeben sind, die sich nur deshalb nicht ausschliessen lassen, weil nach dem derzeitigen Wissenstand bestimmte Ursachenzusammenhänge weder bejaht noch verneint werden können und daher insoweit noch keine Gefahr, sondern nur ein Gefahrenverdacht oder ein Besorgnispotential besteht.

VSL (*value of a statistical life*): Wert eines statistischen Menschenlebens, Bewertungsansatz, der versucht, die Präferenz des Individuums für eine marginale Verringerung des eigenen Todesfallrisikos zu messen.

Vulnerabilität: Anfälligkeit. Eine Vulnerable Gruppe stellt eine Subpopulation dar, welche möglicherweise ein größeres Risiko Schaden zu nehmen haben als die Mehrheit der Gesamtbevölkerung.

WBGU: Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen (www.wbgu.de).

WHO: Weltgesundheitsorganisation.

WTP (*willingness to pay*): Ansatz, Gesundheits- und Todesrisiken auf Geldwerte umzurechnen.

YLD (*years of life lived with disability*): mit einer Beeinträchtigung oder Behinderung lebenden Jahre.

YLL (*years of life lost*): aufgrund vorzeitiger Sterblichkeit verlorenen Jahre.

1. **Energiemodelle in der Bundesrepublik Deutschland. Stand der Entwicklung**
IKARUS-Workshop vom 24. bis 25. Januar 1996
herausgegeben von S. Molt, U. Fahl (1997), 292 Seiten
ISBN: 3-89336-205-3
2. **Ausbau erneuerbarer Energiequellen in der Stromwirtschaft**
Ein Beitrag zum Klimaschutz
Workshop am 19. Februar 1997, veranstaltet von der Forschungszentrum Jülich GmbH und der Deutschen Physikalischen Gesellschaft
herausgegeben von J.-Fr. Hake, K. Schultze (1997), 138 Seiten
ISBN: 3-89336-206-1
3. **Modellinstrumente für CO₂-Minderungsstrategien**
IKARUS-Workshop vom 14. bis 15. April 1997
herausgegeben von J.-Fr. Hake, P. Markewitz (1997), 284 Seiten
ISBN: 3-89336-207-X
4. **IKARUS-Datenbank - Ein Informationssystem zur technischen, wirtschaftlichen und umweltrelevanten Bewertung von Energietechniken**
IKARUS. Instrumente für Klimagas-Reduktionsstrategien
Abschlußbericht Teilprojekt 2 „Datenbank“
H.-J. Laue, K.-H. Weber, J. W. Tepel (1997), 90 Seiten
ISBN: 3-89336-214-2
5. **Politiksznarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 1. Szenarien und Maßnahmen zur Minderung von CO₂-Emissionen in Deutschland bis zum Jahre 2005
herausgegeben von G. Stein, B. Strobel (1997), 410 Seiten
ISBN: 3-89336-215-0
6. **Politiksznarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 2. Emissionsminderungsmaßnahmen für Treibhausgase, ausgenommen energiebedingtes CO₂
herausgegeben von G. Stein, B. Strobel (1997), 110 Seiten
ISBN: 3-89336-216-9

7. **Modelle für die Analyse energiebedingter Klimagasreduktionsstrategien**
IKARUS. Instrumente für Klimagas-Reduktionsstrategien
Abschlußbericht Teilprojekt 1 „Modelle“
P. Markewitz, R. Heckler, Ch. Holzapfel, W. Kuckshinrichs, D. Martinsen,
M. Walbeck, J.-Fr. Hake (1998), VI, 276 Seiten
ISBN: 3-89336-220-7
8. **Politikszzenarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 3. Methodik-Leitfaden für die Wirkungsabschätzung von Maßnahmen zur Emissionsminderung
herausgegeben von G. Stein, B. Strobel (1998), VIII, 95 Seiten
ISBN: 3-89336-222-3
9. **Horizonte 2000**
6. Wolfgang-Ostwald-Kolloquium der Kolloid-Gesellschaft
3. Nachwuchstage der Kolloid- und Grenzflächenforschung
Kurzfassungen der Vorträge und Poster
zusammengestellt von F.-H. Haegel, H. Lewandowski, B. Krah-Urban (1998),
150 Seiten
ISBN: 3-89336-223-1
10. **Windenergieanlagen - Nutzung, Akzeptanz und Entsorgung**
von M. Kleemann, F. van Erp, R. Kehrbaum (1998), 59 Seiten
ISBN: 3-89336-224-X
11. **Policy Scenarios for Climate Protection**
Study on Behalf of the Federal Environmental Agency
Volume 4. Methodological Guideline for Assessing the Impact of Measures for Emission Mitigation
edited by G. Stein, B. Strobel (1998), 103 pages
ISBN: 3-89336-232-0
12. **Der Landschaftswasserhaushalt im Flußeinzugsgebiet der Elbe**
Verfahren, Datengrundlagen und Bilanzgrößen
Analyse von Wasserhaushalt, Verweilzeiten und Grundwassermilieu im
Flußeinzugsgebiet der Elbe (Deutscher Teil). Abschlußbericht Teil 1.
von R. Kunkel, F. Wendland (1998), 110 Seiten
ISBN: 3-89336-233-9

13. **Das Nitratabbauvermögen im Grundwasser des Elbeeinzugsgebietes**
Analyse von Wasserhaushalt, Verweilzeiten und Grundwassermilieu im
Flußeinzugsgebiet der Elbe (Deutscher Teil). Abschlußbericht Teil 2.
von F. Wendland, R. Kunkel (1999), 166 Seiten
ISBN: 3-89336-236-3

14. **Treibhausgasminderung in Deutschland zwischen nationalen Zielen und internationalen Verpflichtungen**
IKARUS-Workshop am 27.05.1998, Wissenschaftszentrum Bonn-Bad
Godesberg. Proceedings
herausgegeben von E. Läge, P. Schaumann, U. Fahl (1999), ii, VI, 146 Seiten
ISBN: 3-89336-237-1

15. **Satellitenbildauswertung mit künstlichen Neuronen Netzen zur Umweltüberwachung**
Vergleichende Bewertung konventioneller und Neuronaler Netzwerkalgorithmen
und Entwicklung eines integrierten Verfahrens
von D. Klaus, M. J. Canty, A. Poth, M. Voß, I. Niemeyer und G. Stein (1999),
VI, 160 Seiten
ISBN: 3-89336-242-8

16. **Volatile Organic Compounds in the Troposphere**
Proceedings of the Workshop on Volatile Organic Compounds in the
Troposphere held in Jülich (Germany) from 27 – 31 October 1997
edited by R. Koppmann, D. H. Ehhalt (1999), 208 pages
ISBN: 3-89336-243-6

17. **CO₂-Reduktion und Beschäftigungseffekte im Wohnungssektor durch das CO₂-Minderungsprogramm der KfW**
Eine modellgestützte Wirkungsanalyse
von M. Kleemann, W. Kuckshinrichs, R. Heckler (1999), 29 Seiten
ISBN: 3-89336-244-4

18. **Symposium über die Nutzung der erneuerbaren Energiequellen Sonne und Wind auf Fischereischiffen und in Aquakulturbetrieben**
Symposium und Podiumsdiskussion, Izmir, Türkei, 28.-30.05.1998.
Konferenzbericht
herausgegeben von A. Özdamar, H.-G. Groehn, K. Ülgen (1999), IX, 245 Seiten
ISBN: 3-89336-245-9

19. **Das Weg-, Zeitverhalten des grundwasserbürtigen Abflusses im Elbeeinzugsgebiet**
Analyse von Wasserhaushalt, Verweilzeiten und Grundwassermilieu im Flußeinzugsgebiet der Elbe (Deutscher Teil). Abschlußbericht Teil 3.
von R. Kunkel, F. Wendland (1999), 122 Seiten
ISBN: 3-89336-249-5

20. **Politiksznarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 5. Szenarien und Maßnahmen zur Minderung von CO₂-Emissionen in Deutschland bis 2020
herausgegeben von G. Stein, B. Strobel (1999), XII, 201 Seiten
ISBN: 3-89336-251-7

21. **Klimaschutz durch energetische Sanierung von Gebäuden. Band 1**
von J.-F. Hake, M. Kleemann, G. Kolb (1999), 216 Seiten
ISBN: 3-89336-252-2

22. **Electroanalysis**
Abstracts of the 8th International Conference held from 11 to 15 June 2000 at the University of Bonn, Germany
edited by H. Emons, P. Ostapczuk (2000), ca. 300 pages
ISBN: 3-89336-261-4

23. **Die Entwicklung des Wärmemarktes für den Gebäudesektor bis 2050**
von M. Kleemann, R. Heckler, G. Kolb, M. Hille (2000), II, 94 Seiten
ISBN: 3-89336-262-2

24. **Grundlegende Entwicklungstendenzen im weltweiten Stoffstrom des Primäraluminiums**
von H.-G. Schwarz (2000), XIV, 127 Seiten
ISBN: 3-89336-264-9

25. **Klimawirkungsforschung auf dem Prüfstand**
Beiträge zur Formulierung eines Förderprogramms des BMBF
Tagungsband des Workshop „Klimaforschung“, Jülich, vom 02. bis 03.12.1999
von J.-Fr. Hake, W. Fischer (2000), 150 Seiten
ISBN: 3-89336-270-3

26. **Energiezukunft 2030**
Schlüsseltechnologien und Techniklinien
Beiträge zum IKARUS-Workshop 2000 am 2./3. Mai 2000
herausgegeben von U. Wagner, G. Stein (2000), 201 Seiten
ISBN: 3-89336-271-1

27. **Der globale Wasserkreislauf und seine Beeinflussung durch den Menschen**
Möglichkeiten zur Fernerkundungs-Detektion und -Verifikation
von D. Klaus und G. Stein (2000), 183 Seiten
ISBN: 3-89336-274-6

28. **Satelliten und nukleare Kontrolle**
Änderungsdetektion und objektorientierte, wissensbasierte Klassifikation von
Multispektralaufnahmen zur Unterstützung der nuklearen Verifikation
von I. Niemeyer (2001), XIV, 206 Seiten
ISBN: 3-89336-281-9

29. **Das hydrologische Modellsystem J2000**
Beschreibung und Anwendung in großen Flußgebieten
von P. Krause (2001), XIV, 247 Seiten
ISBN: 3-89336-283-5

30. **Aufwands- und ergebnisrelevante Probleme der Sachbilanzierung**
von G. Fleischer, J.-Fr. Hake (2002), IV, 64 Blatt
ISBN: 3-89336-293-2

31. **Nachhaltiges Management metallischer Stoffströme**
Indikatoren und deren Anwendung
Workshop, 27.-28.06.2001 im Congresscentrum Rolduc, Kerkrade (NL)
herausgegeben von W. Kuckshinrichs, K.-L. Hüttner (2001), 216 Seiten
ISBN: 3-89336-296-7

32. **Ansätze zur Kopplung von Energie- und Wirtschaftsmodellen zur
Bewertung zukünftiger Strategien**
IKARUS-Workshop am 28. Februar 2002, BMWi, Bonn. Proceedings
herausgegeben von S. Briem, U. Fahl (2003), IV, 184 Seiten
ISBN: 3-89336-321-1

33. **TRACE. Tree Rings in Archaeology, Climatology and Ecology**
Volume 1: Proceedings of the Dendrosymposium 2002,
April 11th – 13th 2002, Bonn/Jülich, Germany
edited by G. Schleser, M. Winiger, A. Bräuning et al., (2003), 135 pages, many
partly coloured illustrations
ISBN: 3-89336-323-8

34. **Klimaschutz und Beschäftigung durch das KfW-Programm zur CO₂-
Minderung und das KfW-CO₂-Gebäudesanierungsprogramm**
von M. Kleemann, R. Heckler, A. Kraft u. a., (2003), 53 Seiten
ISBN: 3-89336-326-2

35. **Klimaschutz und Klimapolitik: Herausforderungen und Chancen**
Beiträge aus der Forschung
herausgegeben von J.-Fr. Hake, K. L. Hüttner (2003), III, 231 Seiten
ISBN: 3-89336-327-0

36. **Umweltschutz und Arbeitsplätze, angestoßen durch die Tätigkeiten des
Schornsteinfegerhandwerks**
Auswertung von Schornsteinfeger-Daten
von M. Kleemann, R. Heckler, B. Krüger (2003), VII, 66 Seiten
ISBN: 3-89336-328-9

37. **Die Grundwasserneubildung in Nordrhein-Westfalen**
von H. Bogen, R. Kunkel, T. Schöbel, H. P. Schrey, F. Wendland (2003), 148
Seiten
ISBN: 3-89336-329-7

38. **Dendro-Isotope und Jahrringbreiten als Klimaproxy der letzten 1200 Jahre
im Karakorumgebirge/Pakistan**
von K. S. Treydte (2003), XII, 167 Seiten
ISBN: 3-89336-330-0

39. **Das IKARUS-Projekt: Energietechnische Perspektiven für Deutschland**
herausgegeben von P. Markewitz, G. Stein (2003), IV, 274 Seiten
ISBN: 3-89336-333-5

40. **Umweltverhalten von MTBE nach Grundwasserkontamination**
von V. Linnemann (2003), XIV, 179 Seiten
ISBN: 3-89336-339-4

41. **Climate Change Mitigation and Adaptation: Identifying Options for Developing Countries**
Proceedings of the Summer School on Climate Change, 7-17 September 2003,
Bad Münstereifel, Germany
edited by K. L. Hüttner, J.-Fr. Hake, W. Fischer (2003), XVI, 341 pages
ISBN: 3-89336-341-6

42. **Mobilfunk und Gesundheit: Risikobewertung im wissenschaftlichen Dialog**
von P. M. Wiedemann, H. Schütz, A. T. Thalmann (2003), 111 Seiten
ISBN: 3-89336-343-2

43. **Chemical Ozone Loss in the Arctic Polar Stratosphere: An Analysis of Twelve Years of Satellite Observations**
by Simone Tilmes (2004), V, 162 pages
ISBN: 3-89336-347-5

44. **TRACE. Tree Rings in Archaeology, Climatology and Ecology**
Volume 2: Proceedings of the Dendrosymposium 2003,
May 1st – 3rd 2003, Utrecht, The Netherlands
edited by E. Jansma, A. Bräuning, H. Gärtner, G. Schleser (2004), 174 pages
ISBN: 3-89336-349-1

45. **Vergleichende Risikobewertung: Konzepte, Probleme und Anwendungsmöglichkeiten**
von H. Schütz, P.M. Wiedemann, W. Hennings et al. (2004), 231 Seiten
ISBN: 3-89336-350-5

1. **Energiemodelle in der Bundesrepublik Deutschland. Stand der Entwicklung**
IKARUS-Workshop vom 24. bis 25. Januar 1996
herausgegeben von S. Molt, U. Fahl (1997), 292 Seiten
ISBN: 3-89336-205-3
2. **Ausbau erneuerbarer Energiequellen in der Stromwirtschaft**
Ein Beitrag zum Klimaschutz
Workshop am 19. Februar 1997, veranstaltet von der Forschungszentrum Jülich GmbH und der Deutschen Physikalischen Gesellschaft
herausgegeben von J.-Fr. Hake, K. Schultze (1997), 138 Seiten
ISBN: 3-89336-206-1
3. **Modellinstrumente für CO₂-Minderungsstrategien**
IKARUS-Workshop vom 14. bis 15. April 1997
herausgegeben von J.-Fr. Hake, P. Markewitz (1997), 284 Seiten
ISBN: 3-89336-207-X
4. **IKARUS-Datenbank - Ein Informationssystem zur technischen, wirtschaftlichen und umweltrelevanten Bewertung von Energietechniken**
IKARUS. Instrumente für Klimagas-Reduktionsstrategien
Abschlußbericht Teilprojekt 2 „Datenbank“
H.-J. Laue, K.-H. Weber, J. W. Tepel (1997), 90 Seiten
ISBN: 3-89336-214-2
5. **Politiksznarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 1. Szenarien und Maßnahmen zur Minderung von CO₂-Emissionen in Deutschland bis zum Jahre 2005
herausgegeben von G. Stein, B. Strobel (1997), 410 Seiten
ISBN: 3-89336-215-0
6. **Politiksznarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 2. Emissionsminderungsmaßnahmen für Treibhausgase, ausgenommen energiebedingtes CO₂
herausgegeben von G. Stein, B. Strobel (1997), 110 Seiten
ISBN: 3-89336-216-9

7. **Modelle für die Analyse energiebedingter Klimagasreduktionsstrategien**
IKARUS. Instrumente für Klimagas-Reduktionsstrategien
Abschlußbericht Teilprojekt 1 „Modelle“
P. Markewitz, R. Heckler, Ch. Holzapfel, W. Kuckshinrichs, D. Martinsen,
M. Walbeck, J.-Fr. Hake (1998), VI, 276 Seiten
ISBN: 3-89336-220-7
8. **Politiksznarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 3. Methodik-Leitfaden für die Wirkungsabschätzung von Maßnahmen zur Emissionsminderung
herausgegeben von G. Stein, B. Strobel (1998), VIII, 95 Seiten
ISBN: 3-89336-222-3
9. **Horizonte 2000**
6. Wolfgang-Ostwald-Kolloquium der Kolloid-Gesellschaft
3. Nachwuchstage der Kolloid- und Grenzflächenforschung
Kurzfassungen der Vorträge und Poster
zusammengestellt von F.-H. Haegel, H. Lewandowski, B. KrahI-Urban (1998),
150 Seiten
ISBN: 3-89336-223-1
10. **Windenergieanlagen - Nutzung, Akzeptanz und Entsorgung**
von M. Kleemann, F. van Erp, R. Kehrbaum (1998), 59 Seiten
ISBN: 3-89336-224-X
11. **Policy Scenarios for Climate Protection**
Study on Behalf of the Federal Environmental Agency
Volume 4. Methodological Guideline for Assessing the Impact of Measures for Emission Mitigation
edited by G. Stein, B. Strobel (1998), 103 pages
ISBN: 3-89336-232-0
12. **Der Landschaftswasserhaushalt im Flußeinzugsgebiet der Elbe**
Verfahren, Datengrundlagen und Bilanzgrößen
Analyse von Wasserhaushalt, Verweilzeiten und Grundwassermilieu im
Flußeinzugsgebiet der Elbe (Deutscher Teil). Abschlußbericht Teil 1.
von R. Kunkel, F. Wendland (1998), 110 Seiten
ISBN: 3-89336-233-9

13. **Das Nitratabbauvermögen im Grundwasser des Elbeeinzugsgebietes**
Analyse von Wasserhaushalt, Verweilzeiten und Grundwassermilieu im
Flußeinzugsgebiet der Elbe (Deutscher Teil). Abschlußbericht Teil 2.
von F. Wendland, R. Kunkel (1999), 166 Seiten
ISBN: 3-89336-236-3

14. **Treibhausgasminderung in Deutschland zwischen nationalen Zielen und
internationalen Verpflichtungen**
IKARUS-Workshop am 27.05.1998, Wissenschaftszentrum Bonn-Bad
Godesberg. Proceedings
herausgegeben von E. Läge, P. Schaumann, U. Fahl (1999), ii, VI, 146 Seiten
ISBN: 3-89336-237-1

15. **Satellitenbilddauswertung mit künstlichen Neuronalen Netzen zur
Umweltüberwachung**
Vergleichende Bewertung konventioneller und Neuronaler Netzwerkalgorithmen
und Entwicklung eines integrierten Verfahrens
von D. Klaus, M. J. Canty, A. Poth, M. Voß, I. Niemeyer und G. Stein (1999),
VI, 160 Seiten
ISBN: 3-89336-242-8

16. **Volatile Organic Compounds in the Troposphere**
Proceedings of the Workshop on Volatile Organic Compounds in the
Troposphere held in Jülich (Germany) from 27 – 31 October 1997
edited by R. Koppmann, D. H. Ehhalt (1999), 208 pages
ISBN: 3-89336-243-6

17. **CO₂-Reduktion und Beschäftigungseffekte im Wohnungssektor durch das
CO₂-Minderungsprogramm der KfW**
Eine modellgestützte Wirkungsanalyse
von M. Kleemann, W. Kuckshinrichs, R. Heckler (1999), 29 Seiten
ISBN: 3-89336-244-4

18. **Symposium über die Nutzung der erneuerbaren Energiequellen Sonne und
Wind auf Fischereischiffen und in Aquakulturbetrieben**
Symposium und Podiumsdiskussion, Izmir, Türkiye, 28.-30.05.1998.
Konferenzbericht
herausgegeben von A. Özdamar, H.-G. Groehn, K. Ülgen (1999), IX, 245 Seiten
ISBN: 3-89336-247-9

19. **Das Weg-, Zeitverhalten des grundwasserbürtigen Abflusses im Elbeeinzugsgebiet**
Analyse von Wasserhaushalt, Verweilzeiten und Grundwassermilieu im Flußeinzugsgebiet der Elbe (Deutscher Teil). Abschlußbericht Teil 3.
von R. Kunkel, F. Wendland (1999), 122 Seiten
ISBN: 3-89336-249-5

20. **Politiksznarien für den Klimaschutz**
Untersuchungen im Auftrag des Umweltbundesamtes
Band 5. Szenarien und Maßnahmen zur Minderung von CO₂-Emissionen in Deutschland bis 2020
herausgegeben von G. Stein, B. Strobel (1999), XII, 201 Seiten
ISBN: 3-89336-251-7

21. **Klimaschutz durch energetische Sanierung von Gebäuden. Band 1**
von J.-F. Hake, M. Kleemann, G. Kolb (1999), 216 Seiten
ISBN: 3-89336-252-2

22. **Electroanalysis**
Abstracts of the 8th International Conference held from 11 to 15 June 2000 at the University of Bonn, Germany
edited by H. Emons, P. Ostapczuk (2000), ca. 300 pages
ISBN: 3-89336-261-4

23. **Die Entwicklung des Wärmemarktes für den Gebäudesektor bis 2050**
von M. Kleemann, R. Heckler, G. Kolb, M. Hille (2000), II, 94 Seiten
ISBN: 3-89336-262-2

24. **Grundlegende Entwicklungstendenzen im weltweiten Stoffstrom des Primäraluminiums**
von H.-G. Schwarz (2000), XIV, 127 Seiten
ISBN: 3-89336-264-9

25. **Klimawirkungsforschung auf dem Prüfstand**
Beiträge zur Formulierung eines Förderprogramms des BMBF
Tagungsband des Workshop „Klimaforschung“, Jülich, vom 02. bis 03.12.1999
von J.-Fr. Hake, W. Fischer (2000), 150 Seiten
ISBN: 3-89336-270-3

26. **Energiezukunft 2030**
Schlüsseltechnologien und Techniklinien
Beiträge zum IKARUS-Workshop 2000 am 2./3. Mai 2000
herausgegeben von U. Wagner, G. Stein (2000), 201 Seiten
ISBN: 3-89336-271-1

27. **Der globale Wasserkreislauf und seine Beeinflussung durch den Menschen**
Möglichkeiten zur Fernerkundungs-Detektion und -Verifikation
von D. Klaus und G. Stein (2000), 183 Seiten
ISBN: 3-89336-274-6

28. **Satelliten und nukleare Kontrolle**
Änderungsdetektion und objektorientierte, wissensbasierte Klassifikation von
Multispektralaufnahmen zur Unterstützung der nuklearen Verifikation
von I. Niemeyer (2001), XIV, 206 Seiten
ISBN: 3-89336-281-9

29. **Das hydrologische Modellsystem J2000**
Beschreibung und Anwendung in großen Flußgebieten
von P. Krause (2001), XIV, 247 Seiten
ISBN: 3-89336-283-5

30. **Aufwands- und ergebnisrelevante Probleme der Sachbilanzierung**
von G. Fleischer, J.-Fr. Hake (2002), IV, 64 Blatt
ISBN: 3-89336-293-2

31. **Nachhaltiges Management metallischer Stoffströme**
Indikatoren und deren Anwendung
Workshop, 27.-28.06.2001 im Congresszentrum Rolduc, Kerkrade (NL)
herausgegeben von W. Kuckshinrichs, K.-L. Hüttner (2001), 216 Seiten
ISBN: 3-89336-296-7

32. **Ansätze zur Kopplung von Energie- und Wirtschaftsmodellen zur
Bewertung zukünftiger Strategien**
IKARUS-Workshop am 28. Februar 2002, BMWi, Bonn. Proceedings
herausgegeben von S. Briem, U. Fahl (2003), IV, 184 Seiten
ISBN: 3-89336-321-1

33. **TRACE. Tree Rings in Archaeology, Climatology and Ecology**
Volume 1: Proceedings of the Dendrosymposium 2002,
April 11th – 13th 2002, Bonn/Jülich, Germany
edited by G. Schleser, M. Winiger, A. Bräuning et al., (2003), 135 pages, many
partly coloured illustrations
ISBN: 3-89336-323-8

34. **Klimaschutz und Beschäftigung durch das KfW-Programm zur CO₂-
Minderung und das KfW-CO₂-Gebäudesanierungsprogramm**
von M. Kleemann, R. Heckler, A. Kraft u. a., (2003), 53 Seiten
ISBN: 3-89336-326-2

35. **Klimaschutz und Klimapolitik: Herausforderungen und Chancen**
Beiträge aus der Forschung
herausgegeben von J.-Fr. Hake, K. L. Hüttner (2003), III, 231 Seiten
ISBN: 3-89336-327-0

36. **Umweltschutz und Arbeitsplätze, angestoßen durch die Tätigkeiten des
Schornsteinfegerhandwerks**
Auswertung von Schornsteinfeger-Daten
von M. Kleemann, R. Heckler, B. Krüger (2003), VII, 66 Seiten
ISBN: 3-89336-328-9

37. **Die Grundwasserneubildung in Nordrhein-Westfalen**
von H. Bogen, R. Kunkel, T. Schöbel, H. P. Schrey, F. Wendland (2003), 148
Seiten
ISBN: 3-89336-329-7

38. **Dendro-Isotope und Jahrringbreiten als Klimaproxy der letzten 1200 Jahre
im Karakorumgebirge/Pakistan**
von K. S. Treydte (2003), XII, 167 Seiten
ISBN: 3-89336-330-0

39. **Das IKARUS-Projekt: Energietechnische Perspektiven für Deutschland**
herausgegeben von P. Markewitz, G. Stein (2003), IV, 274 Seiten
ISBN: 3-89336-333-5

40. **Umweltverhalten von MTBE nach Grundwasserkontamination**
von V. Linnemann (2003), XIV, 179 Seiten
ISBN: 3-89336-339-4

41. **Climate Change Mitigation and Adaptation: Identifying Options for Developing Countries**
Proceedings of the Summer School on Climate Change, 7-17 September 2003, Bad Münstereifel, Germany
edited by K. L. Hüttner, J.-Fr. Hake, W. Fischer (2003), XVI, 341 pages
ISBN: 3-89336-341-6

42. **Mobilfunk und Gesundheit: Risikobewertung im wissenschaftlichen Dialog**
von P. M. Wiedemann, H. Schütz, A. T. Thalmann (2003), 111 Seiten
ISBN: 3-89336-343-2

43. **Chemical Ozone Loss in the Arctic Polar Stratosphere: An Analysis of Twelve Years of Satellite Observations**
by Simone Tilmes (2004), V, 162 pages
ISBN: 3-89336-347-5

44. **TRACE. Tree Rings in Archaeology, Climatology and Ecology**
Volume 2: Proceedings of the Dendrosymposium 2003,
May 1st – 3rd 2003, Utrecht, The Netherlands
edited by E. Jansma, A. Bräuning, H. Gärtner, G. Schleser (2004), 174 pages
ISBN: 3-89336-349-1

45. **Vergleichende Risikobewertung: Konzepte, Probleme und Anwendungsmöglichkeiten**
von H. Schütz, P.M. Wiedemann, W. Hennings et al. (2004), 231 Seiten
ISBN: 3-89336-350-5

Politische Entscheidungen über den Umgang mit umweltbezogenen Gesundheitsrisiken verlangen eine Bewertung, deren Grundlage die wissenschaftliche Risikoabschätzung bietet. In vielen Fällen sind dabei Prioritäten zu setzen, d.h. Risiken in ihrer Bedeutsamkeit gegeneinander abzuwägen. Hierzu bedarf es eines Vergleichs von Risiken, der auf klaren Maßstäben und einer nachvollziehbaren Methodik beruht. Darüber hinaus müssen solche vergleichenden Risikobewertungen auch für die Öffentlichkeit transparent sein.

In diesem Band, der die Ergebnisse des BfS-Projekts „Risikobewertung und -management: Ausarbeitung von Konzepten eines integrierten und vergleichenden Risikoansatzes“ darstellt, werden vorliegende Ansätze zur vergleichenden Risikobewertung aufgearbeitet und Möglichkeiten erkundet, solche Vergleiche zu verbessern, Partizipationschancen zu erörtern und Verfahren der vergleichenden Risikobewertung so zu gestalten, dass sie bei politischen Entscheidungen über umweltbezogene Gesundheitsrisiken mehr Akzeptanz finden.

Vorgeschlagen wird ein multiattributes Bewertungsverfahren, das ein theoretisch fundiertes und transparentes Vorgehen in überschaubaren Einzelschritten ermöglicht. Für die praktische Umsetzung der vergleichenden Risikobewertung wird ein Prozessvorschlag erarbeitet, welcher der besonderen Problematik der gemeinsamen Risikobewertung durch Experten und gesellschaftliche Stakeholder Rechnung trägt.

Autoren

Holger Schütz ist wissenschaftlicher Mitarbeiter der Programmgruppe Mensch, Umwelt, Technik (MUT), die von **Peter M. Wiedemann** geleitet wird. Beide beschäftigen sich aus sozialwissenschaftlicher Sicht mit Fragen der Risikowahrnehmung, Risikokommunikation und Risikobewertung. **Wilfried Hennings** hat bei MUT speziell die Methoden zur Risikoermittlung – vorwiegend für technische Anlagen – als Arbeitsschwerpunkt; **Johannes Mertens** konzentriert sich in der Programmgruppe auf prinzipielle Aspekte der Ermittlung und Bewertung technikinduzierter Risiken. **Martin Clauberg** war im Projekt „Risikobewertung und -management“ bei MUT wissenschaftlicher Mitarbeiter und arbeitet jetzt als international tätiger Consultant zu Fragen der Toxikologie und des Wissensmanagements im Prozess der Risikoregulierung.

Programmgruppe Mensch, Umwelt, Technik (MUT)

Die Programmgruppe Mensch, Umwelt, Technik (MUT) des Forschungszentrums Jülich untersucht den **gesellschaftlichen Umgang mit Chancen und Risiken wissenschaftlicher und technischer Entwicklungen**. Dabei konzentriert sie sich auf drei Schwerpunkte: auf „**Management von Innovationen**“, „**Umgang mit Unsicherheit**“ und „**Öffentlichkeit, Politik und Massenmedien**“. MUT führt hierzu sozialwissenschaftliche, naturwissenschaftlich-technische und interdisziplinäre Forschung durch. Die Ergebnisse werden umgesetzt in praktische Ansätze zur Verbesserung des Umgangs mit Chancen und Risiken. Die Programmgruppe entwickelt und erprobt Instrumente zur Vorsorge, Risikokommunikation und Konsensfindung, zum Management von Chancen und zur Kommunikation zwischen Wissenschaft und Gesellschaft.

Forschungszentrum Jülich
in der Helmholtz-Gemeinschaft

