
Präimplantationsdiagnostik, ja oder nein?

Förderung einer informierten und gut begründeten Entscheidung im Bereich Medizinethik mit einer multimedialen Lernumgebung im Museum

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Philosophie an der Ludwig-Maximilians-Universität
München

vorgelegt von
Martina Hänle
aus
München

München, im März 2012

Referent: Prof. Dr. Heinz Mandl

Korreferent: Prof. Dr. Dieter Frey

Tag der mündlichen Prüfung: 05. Juli 2012

Diese Arbeit entstand im Rahmen des WGL-Projekts „Lernen im Museum: Die Rolle von Medien für die Resituierung von Exponaten“ am Forschungsinstitut des Deutschen Museums München.

Inhaltsverzeichnis

1	Problemstellung und Ziele	7
1.1	Problemstellung	7
1.2	Zielsetzungen und Aufbau der Arbeit.....	10
2	Theoretische Grundlagen zur Präimplantationsdiagnostik (PID)	14
2.1	Einführung zu genetischen Untersuchungen in der Humanmedizin	14
2.2	Medizinisch-naturwissenschaftliche Grundlagen der PID	19
2.2.1	Verfahren.....	19
2.2.2	Anwendungsbereiche	25
2.2.3	PID-HLA zur Erzeugung eines Gewebespenders	26
2.3	Rechtliche Regelungen und Stellungnahmen zur PID.....	31
2.3.1	Regelung in Deutschland.....	31
2.3.2	Richtlinien und Empfehlungen in Deutschland.....	33
2.3.3	Ausgewählte internationale Regelungen	36
2.4	Ethische Analyse am Beispiel der PID-HLA	37
2.4.1	Moralischer Status des menschlichen Embryos	38
2.4.2	Diskussion der Argumente Pro und Contra PID-HLA.....	44
2.4.3	Ethische Stellungnahmen zur PID-HLA	58
2.5	Empirische Studien zu Wissen und Einstellungen zur PID in der BRD	59
3	Theoretische Grundlagen zur Entscheidungsfindung und Argumentation im Bereich Medizinethik.....	63
3.1	Einführung in die Medizinethik.....	63
3.2	Entscheidungsfindung und Argumentation im Bereich Medizinethik	69
3.2.1	Entscheiden und informelles Schlussfolgern.....	70
3.2.2	Informelles Schlussfolgern aus Sicht der Zwei-Prozess-Theorien.....	73
3.2.3	Argumentation.....	79
3.3	Empirische Studien zur Argumentationsqualität und zu Argumentationsmustern.....	87

3.4	Potentielle Einflussfaktoren auf die Entscheidungsfindung und Argumentationsqualität	89
4	Theoretische Grundlagen zur Konzeption der multimedialen Lernumgebung	92
4.1	Begriffsbestimmung Multimedia.....	92
4.2	Kognitions- und medienpsychologische Grundlagen zum Lernen mit neuen Medien.....	95
4.2.1	Kognitive Theorie zum multimedialen Lernen (CTML).....	96
4.2.2	Social Agency Theory	100
4.2.3	Parasoziale Interaktion (PSI) mit Medienfiguren	103
4.3	Gestaltung multimedialer Lernumgebungen	109
4.3.1	Problemorientiertes Lernen	110
4.3.2	Didaktische Gestaltung.....	114
4.3.3	Mediale Gestaltung.....	115
4.4	Feedback in multimedialen Lernumgebungen	117
4.4.1	Begriffsbestimmung Feedback	117
4.4.2	Wirkung von Feedback.....	119
4.4.2.1	Einflussfaktor Feedbackgestaltung	120
4.4.2.2	Einflussfaktor Feedbackrezeption.....	123
4.4.3	Feedback in Form einer sozialen Vergleichsinformation	125
4.5	Multimediale Lernumgebungen im Museum	132
4.6	Empirische Studien zur Effektivität von multimedialen Lernumgebungen zur Vermittlung von kontroversen Wissenschaftsthemen.....	142
5	Die multimediale Lernumgebung Dialogstation „Gentest“	151
5.1	Rahmenbedingungen	151
5.2	Ziele und Inhalte.....	152
5.3	Didaktische Gestaltung.....	157
5.4	Medial-technologische Gestaltung.....	158
5.5	Gestaltung der Feedbackinformation	160
5.6	Erste Evaluation der Lernumgebung.....	161
6	Zusammenfassung und Konsequenzen für die vorliegende Studie.....	169

7	Fragestellungen und Ableitung der Hypothesen	174
7.1	Analyse der parasozialen Interaktion mit der Medienfigur	174
7.2	Analyse des Entscheidungsverhaltens	175
7.3	Analyse der Argumentationsqualität	177
7.4	Analyse der Argumentationsmuster	179
7.5	Analyse von Einflussfaktoren auf das Entscheidungsverhalten und Argumentationsqualität.....	180
8	Methode	181
8.1	Stichprobe	181
8.2	Design	182
8.3	Verwendete Lernumgebung: Fallbeispiel PID-HLA.....	182
8.4	Operationalisierung der unabhängigen Variablen	189
8.4.1	Messzeitpunkt.....	189
8.4.2	Feedback.....	190
8.5	Versuchsablauf	192
8.6	Instrumente	196
8.6.1	Individuelle Lernvoraussetzungen.....	197
8.6.2	Realisierte Lernzeit	200
8.6.3	Manipulationskontrolle	200
8.6.4	Parasoziale Interaktion und Attraktivität der Medienfigur.....	202
8.6.5	Richtung, Sicherheit und Änderung der Entscheidungen.....	203
8.7	Erfassung der Argumentationsqualität und -muster	203
8.7.1	Kategorienschema zur Erfassung der Argumentationsqualität.....	205
8.7.2	Myside bias Index	209
8.7.3	Kategorienschema zur Erfassung der Argumentationsmuster.....	209
9	Ergebnisse.....	212
9.1	Überprüfung der experimentellen Voraussetzungen	212
9.1.1	Individuelle Lernvoraussetzungen.....	212
9.1.2	Realisierte Lernzeit	216
9.1.3	Manipulationskontrolle	217

9.2	Analyse der parasozialen Interaktion mit der Medienfigur	222
9.3	Analyse des Entscheidungsverhaltens	224
9.3.1	Richtung der Entscheidungen	224
9.3.2	Sicherheit der Entscheidungen	226
9.3.3	Änderung der Entscheidungen	229
9.4	Analyse der Argumentationsqualität	234
9.5	Analyse der Argumentationsmuster	242
9.6	Analyse von Einflussfaktoren auf das Entscheidungsverhalten und die Argumentationsqualität	246
10	Diskussion	251
10.1	Überprüfung der experimentellen Voraussetzungen	251
10.2	Parasoziale Interaktion mit der Medienfigur	254
10.3	Entscheidungsverhalten	255
10.4	Argumentationsqualität	262
10.5	Argumentationsmuster	266
10.6	Einflussfaktoren auf das Entscheidungsverhalten und die Argumentationsqualität	268
11	Gesamtdiskussion und Konsequenzen	271
11.1	Zusammenfassung und Diskussion der zentralen Ergebnisse	271
11.1.1	Lernwirksamkeit der multimedialen Lernumgebung	272
11.1.2	Einfluss der Feedbackinformation auf die Entscheidung und Argumentationsqualität ...	273
11.1.3	Bedeutung der Medienfigur und des Vorwissens im vorliegenden Lernkontext	274
11.2	Konsequenzen für die Forschung und Praxis	275
11.2.1	Konsequenzen für die Forschung	275
11.2.2	Konsequenzen für die pädagogische Praxis	279
12	Literatur	284
13	Abbildungs- und Tabellenverzeichnis	316
	Abkürzungsverzeichnis	321
	Anhang	322

1 Problemstellung und Ziele

1.1 Problemstellung

Vor über 30 Jahren, am 25. Juli 1978, erblickte Louise Brown in Oldham bei Manchester, England das Licht der Welt. Die Geburt von Louise sorgte damals weltweit für großes Aufsehen und leitete den Eintritt in eine neue Ära der Reproduktionsmedizin und Gentechnologie ein: Sie war das erste Kind, das erfolgreich mit der Methode der künstlichen Befruchtung, der In-vitro-Fertilisation (nachfolgend abgekürzt mit: IVF), außerhalb des Mutterleibes im Reagenzglas erzeugt und danach in die Gebärmutter ihrer leiblichen Mutter eingepflanzt wurde (Step toe & Edwards, 1978). Vier Jahre später wurde das erste dieser damals als Retortenbabys genannten Kinder in Deutschland im Universitätsklinikum Erlangen-Nürnberg geboren (Rauprich & Siegel, 2003).

Anfangs noch in der Öffentlichkeit heftig umstritten, legitimisierte sich die IVF sehr schnell als Behandlungsmethode für unfruchtbare Paare und gilt inzwischen als etabliert und gesellschaftlich akzeptiert (Ludwig, Diedrich, Ludwig & Felberbaum, 2006).

Die Nachfrage nach Kinderwunschbehandlungen wächst stetig; mittlerweile leben gemäß dem Deutschen IVF-Register (nachfolgend abgekürzt mit: DIR) über 130 000 nach einer IVF gezeugte Kinder in Deutschland (DIR, 2010). Der Erfinder der IVF, der Brite Robert Edwards, wurde für seine Pionierarbeit im Bereich der medizinisch assistierten Reproduktion im Jahr 2010 mit dem Nobelpreis für Medizin ausgezeichnet.

Ein weiteres bedeutsames Jahrhundertereignis bildete vor über 10 Jahren, am 26. Juni 2000, die Präsentation einer ersten Rohversion des entschlüsselten menschlichen Erbgutes. Dieser Meilenstein in der Geschichte der Humangenomforschung ermöglicht es u. a. mehr als 900 Erbkrankheiten molekulargenetisch, viele davon bereits im Mutterleib, zu diagnostizieren (vgl. HGQN, 2010; Höble, 2001a). Heute gehören genetische Untersuchungen wie die Pränataldiagnostik (nachfolgend abgekürzt mit: PND) am Embryo oder Fetus im Mutterleib zum medizinischen Alltag.

Inzwischen ist es mit dem Verfahren der Präimplantationsdiagnostik (nachfolgend abgekürzt mit: PID) auch gelungen, den nach einer IVF extrakorporal erzeugten Embryonen im Reagenzglas im frühen 8-16 Zellstadium ein bis zwei Zellen zu entnehmen und auf schwere Erbkrankheiten zu untersuchen. Embryonen, welche die Krankheitsanlage nicht aufweisen und damit „gesund“ sind, werden dann in die Gebärmutter der Frau eingesetzt, wohingegen Embryonen mit pathologischem Befund, die von der befürchteten Krankheit betrof-

fen sind, aussortiert und vernichtet werden (Kreß, 2009). Die PID ist aus ethischen und juristischen Gesichtspunkten eine sehr umstrittene neue Technologie, da sie unweigerlich mit einem hohen Verbrauch von menschlichen Embryonen verbunden ist und damit Fragen des bisherigen Menschenbildes und der Schutzwürdigkeit frühen menschlichen Lebens aufwirft.

Neben der Diagnose von Erbkrankheiten kann die PID auch zur Bestimmung des Gewebetyps von Embryonen eingesetzt werden, mit dem Ziel einen potentiell geeigneten Gewebespende für ein bereits existierendes Geschwisterkind mit einer therapiebedürftigen Erkrankung zu erzeugen (Robertson, 2003; Steinke & Rahner, 2009). In Deutschland ist eine PID, die dieses Anwendungsziel verfolgt, derzeit gesetzlich verboten, wird jedoch in anderen Ländern wie Großbritannien, Belgien, Frankreich oder den USA bereits praktiziert (Hennen & Sauter, 2004). Das erste Kind, das mittels PID im Labor gezeugt wurde, um das Leben seines Geschwisters zu retten, wurde am 29. August 2000 in den USA mit Adam Nash geboren. Seine Schwester Molly litt an der schweren, lebensbedrohlichen Blutkrankheit Fanconi-Anämie und konnte mithilfe einer Nabelschnurblutspende von Adam nach seiner Geburt geheilt werden (Boyle & Savulescu, 2001).

Die Geburt des ersten „Retterkindes“ Adam Nash sorgte damals weltweit für spektakuläre Schlagzeilen in den Medien und löste heftige ethische Debatten aus. In der ethischen Diskussion um das Für und Wider der PID aus Anlass zur Lebensrettung eines dritten stand insbesondere die Frage, ob die Aussicht auf Heilung eines kranken Kindes den hohen Verbrauch menschlicher Embryonen rechtfertigen kann. Viele Kritiker warfen den Eltern auch vor, Adam nicht um seiner selbst willen, sondern ausschließlich zum Zwecke der Gewebespende gezeugt zu haben. Diese ethisch fragwürdige Instrumentalisierung menschlichen Lebens zu fremden Zwecken wurde in den Medien häufig mit den Suggestivwörtern „menschliches Ersatzteillager“ oder „Designerbaby“ zum Ausdruck gebracht (Rehmann-Sutter, 2007).

Die rasanten Fortschritte im Bereich der Reproduktions- und Gentechnologie und die Anwendung dieser innovativen Technologien schaffen neue, ethisch brisante Entscheidungsdilemmata für die Gesellschaft und stellen auch den Staat vor diverse Regulierungsaufgaben (NEK-CNE, 2007). Die Gesellschaft und die in ihr lebenden Individuen sind aktuell vor die Aufgabe gestellt, die Chancen, aber auch die Risiken der Reproduktions- und Gentechnologie kritisch zu hinterfragen und reflektiert zu bewerten (Höble, 2003a).

In unserer immer stärker naturwissenschaftlich-technisch geprägten Welt erlangt naturwissenschaftliche Grundbildung (*scientific literacy*) zunehmend mehr an Relevanz (Schanze & Saballus, 2009). Naturwissenschaftliche Grundbildung bildet einen wesentlichen Bestandteil der Allgemeinbildung und soll u. a. das Individuum zu einer aktiven Teilhabe an

gesellschaftlicher Kommunikation und Meinungsbildung über die technische Entwicklung und Forschung befähigen (vgl. Gräber & Nentwig, 2002; KMK, 2004; Kolstø, 2001; OECD, 1999; Sadler & Zeidler, 2005).

Naturwissenschaftlich-technische Museen, die von jeher als Bildungsinstitution für die breite Öffentlichkeit konzipiert sind (Noschka-Roos & Teichmann, 2006), können einen wichtigen Beitrag zur naturwissenschaftlichen Grundbildung (*scientific literacy*) von Erwachsenen leisten (vgl. Schrader, Stadler & Körber, 2008). Für naturwissenschaftlich-technische Museen stellt sich jedoch momentan die Frage, wie kontroverse und gesellschaftlich relevante Wissenschaftsthemen im Rahmen von Ausstellungskonzepten überhaupt vermittelt werden können, um ihren Besuchern eine informierte Entscheidung und reflektierte Urteilsbildung ermöglichen zu können (Schwan, 2006).

Eine aktuelle Antwort auf diese Frage lautet *Public understanding of research* (nachfolgend abgekürzt mit: PUR): Statt der bloßen Vermittlung wissenschaftlicher Erkenntnisse und Forschungsergebnisse im Sinne eines *Public understanding of science* (nachfolgend abgekürzt mit: PUS), bedeutet PUR, ein reflektiertes Verständnis für die Arbeitsweise bzw. den Prozesscharakter der Forschung zu erzeugen. Ferner beinhaltet PUR offen zu legen, welche verschiedenen, auch gegensätzlichen Positionen innerhalb der *scientific community* existieren und welche komplexen gesellschaftlichen, ethischen und politischen Implikationen aus der aktuellen Forschung resultieren (vgl. Chittenden, Farmelo & Lewenstein, 2004; Durant, 2001; Hauser, 2005, 2006, 2009; Noschka-Roos & Teichmann, 2006; Schwan, 2006). Vor diesem Hintergrund stehen Museen vor der schwierigen Aufgabe, einen Dialog zwischen Wissenschaft und Öffentlichkeit zu initiieren, der die Ausstellungsbesucher aktiv mit ihren Positionen einbindet und zu einer fundierten Meinungs- und Urteilsbildung in Bezug auf kontroverse Wissenschaftsthemen anregen soll.

Einen innovativen Ansatz zur Lösung dieser Problemstellung stellen multimediale Lernumgebungen in Form von sog. Dialogstationen dar, die im Deutschen Museum innerhalb des „Zentrums neue Technologien“ (nachfolgend abgekürzt mit: ZNT) für eine Dauerausstellung zum Thema Nano- und Biotechnologie entwickelt wurden (vgl. Breitsameter, Hauser, Hauser & Kampschulte, 2009; Hauser, 2005, 2006).

Unter Dialogstationen werden im vorliegenden musealen Kontext konzeptionell völlig neue, interaktive Medieninstallationen verstanden, die konträre Meinungen von Betroffenen, Experten und Laien oder auch unterschiedliche Standpunkte zu brisant und kontrovers diskutierten Wissenschaftsthemen wiedergeben und zugleich den Austausch individueller Einstellungen/Meinungen erlauben (Hauser, 2005, 2006; Noschka-Roos, 2006; Zahn, 2006). Kernelement dieser Computerterminals bildet die Initiierung eines Meinungsaustauschs im Sinne eines Dialogs, der dem Besucher Gelegenheit gibt, seinen eigenen Stand-

punkt zum Thema zu äußern und andere existente, eventuell auch konträre, Meinungen zu evaluieren (Hauser, 2005, 2006; Knipfer, Zahn & Hesse, 2008a; 2008b).

Dialogstationen können als ein sinnvoller Ansatzpunkt zur Unterstützung der Meinungsbildung und zur Vermittlung von kontroversen Wissenschaftsthemen im Rahmen von Museumsausstellungen angesehen werden (vgl. Knipfer, 2009).

Am Deutschen Museum München wurde von der Kuratorin für Life Sciences, Dr. Birte Hauser und der Mediendesignerin Charlotte Kaiser eine multimediale Lernumgebung zur Förderung einer informierten und gut begründeten Entscheidung im Bereich Medizinethik entwickelt. Dabei handelt sich um die Dialogstation „Gentest“, die für den Einsatz in der Ausstellung „Nano- und Biotechnologie“ konzipiert wurde (Breitsameter et al., 2009; Gerber-Hirt & Noschka-Roos, 2012). Als Thema werden die ethischen Aspekte von prädiktiven und pränatalen genetischen Untersuchungen sowie dem Verfahren der PID behandelt.

Im Rahmen der Dialogstation wurde eine quasi persönliche Ansprache der Besucher umgesetzt (Hauser, 2006, 2009): Fiktive Personen, die in Form von sprechenden menschlichen Köpfen (*talking heads*) durch eine Halbre relief-Projektion realisiert wurden, sprechen den Besucher an und erzählen ihre eigene Geschichte zur Thematik. Dabei handelt es sich jeweils um ein typisches Fallbeispiel für den bewussten Einsatz oder die Ablehnung einer genetischen Untersuchung.

Jede dieser „Dilemmageschichten“ endet mit einer „Ja/Nein-Frage“, die den Besucher aktiv dazu auffordert, eine stellvertretende Handlungsentscheidung zu treffen, d. h. ob er an Stelle der Person in dieser Situation ebenso einen Gentest durchführen lassen würde oder nicht (sog. aktive Positionierung). Anschließend wird Feedback in Form einer Abstimmungsstatistik (sog. Konsensinformation) gegeben, aus der ersichtlich wird, wie sich die vorherigen Besucher im Durchschnitt entschieden haben.

Nach Erhalt dieser sozialen Vergleichsinformation in Form eines Umfrageergebnisses aller bisher abgegebenen Besucherentscheidungen können zum jeweiligen Fallbeispiel noch zusätzliche, vertiefende Informationen, z. B. in Form von ethischen Stellungnahmen von Experten wie dem Nationalen Ethikrat, die bei der Entscheidungsfindung helfen sollen, bei Bedarf von den Besuchern aufgerufen werden.

1.2 Zielsetzungen und Aufbau der Arbeit

Ausgehend von der skizzierten Problemstellung wird in der vorliegenden Arbeit die zentrale Zielstellung verfolgt, einen Beitrag zum pädagogisch-psychologischen Erkenntnisgewinn zur Förderung der Entscheidungsfindung und Argumentation im Bereich Medizin-

ethik mithilfe einer multimedialen Lernumgebung liefern. Anhand einer experimentellen Laborstudie soll analysiert werden, ob und inwieweit eine informierte und gut begründete Entscheidung (operationalisiert als Argumentationsqualität) im Bereich Medizinethik/PID durch eine multimediale Lernumgebung unterstützt und gefördert werden kann.

Den Kontext der Studie bildet die multimediale Lernumgebung Dialogstation „Gentest“, aus welcher als beispielhaftes und kontrovers diskutiertes medizinethisches Thema bzw. Untersuchungsgegenstand ein Fallbeispiel zur PID ausgewählt wurde.

Mit dieser Arbeit wird ferner versucht, einen ersten Beitrag zum Nachweis der Lernwirksamkeit dieser innovativen musealen Lernumgebung zu leisten. Ziel ist, die Effektivität einer von einer Kuratorin und Mediendesignerin konzipierten multimedialen Lernumgebung unter möglichst kontrollierten, d. h. labor-experimentellen Bedingungen im Hinblick auf die Förderung einer gut informierten und reflektierten Entscheidung zu überprüfen und zu erproben. Es soll untersucht werden, inwieweit die Arbeit mit der Lernumgebung zu einem Lernfortschritt im Sinne einer höheren Argumentationsqualität führt. Im Rahmen der hier präsentierten Studie wird damit auch der Frage nachgegangen, ob sich die multimediale Lernumgebung zur Vermittlung von kontroversen Wissenschaftsthemen im musealen Lernkontext am Beispiel der PID eignet.

Eine weitere zentrale Zielsetzung der Arbeit besteht darin, einen Beitrag zum pädagogisch-psychologischen Erkenntnisgewinn zur Wirkung von Feedback in Form einer sozialen Vergleichsinformation beizusteuern: Kann die Feedbackinformation darüber, wie andere sich entscheiden, die eigene Entscheidung zu einem medizinethischen Dilemma wie der PID verändern und die Argumentationsqualität fördern? Um den Einfluss der Feedbackinformation auf die Entscheidung und die Argumentationsqualität zu ermitteln, wurde das Feedback in Form einer sozialen Vergleichsinformation in Abhängigkeit von der getroffenen Entscheidung gezielt experimentell manipuliert. Variiert wurde die Kongruenz (Übereinstimmung) zwischen der eigenen Entscheidung und des Umfrageergebnisses über die durchschnittliche Gesamtentscheidung der anderen Versuchsteilnehmer.

Neben der systematischen Analyse der Argumentationsqualität und der Entscheidungsprozesse, die bei der Lösung der Dilemmageschichte zur PID ablaufen, sollen in dieser Untersuchung auch Erkenntnisse über die Rolle gewonnen werden, die verschiedenen Lernermerkmalen (z. B. Vorwissen) oder der Medienfigur (*talking head*) bei der Bearbeitung der multimedialen Lernumgebung zukommt.

Des Weiteren soll überprüft werden, ob und inwieweit die didaktische und mediale Gestaltung der musealen Lernumgebung basierend auf den Untersuchungsergebnissen hinsichtlich ihrer Lernwirksamkeit optimiert werden kann (Prenzel & Thoma, 2006).

Im Rahmen dieser Arbeit wird folglich sowohl ein genereller pädagogisch-psychologischer Erkenntnisgewinn als auch die Generierung von Anregungen für die (museums-) pädagogische Praxis angestrebt.

Die vorliegende Arbeit setzt sich aus einem theoretischen und empirischen Teil zusammen. Im Theorieteil werden in Kapitel 2 bis 6 die der empirischen Studie zugrundeliegenden theoretischen Grundlagen zur Präimplantationsdiagnostik (PID), zur Entscheidungsfindung und Argumentation im Bereich Medizinethik und zur Konzeption der multimedialen Lernumgebung erörtert. Auf dieser Basis wird dann die multimediale Lernumgebung Dialogstation „Gentest“ vorgestellt und es werden zusammenfassend Konsequenzen für die vorliegende Studie gezogen.

Kapitel 2 befasst sich zunächst mit den theoretischen Grundlagen zur Präimplantationsdiagnostik (PID), da dieses Thema in der hier präsentierten Studie als Untersuchungsgegenstand fungierte. Nach einer allgemeinen Einführung zu genetischen Untersuchungen in der Humanmedizin wird auf die naturwissenschaftlich-medizinischen Grundlagen der PID eingegangen. Danach wird ein zusammenfassender Überblick über rechtliche Regelungen und Stellungnahmen zur PID in Deutschland gegeben sowie ausgewählte internationale Regelungen vorgestellt. Im Anschluss daran folgt eine ausführliche ethische Analyse am Beispiel der PID zum Zwecke der Gewebespende, da diese Anwendungsindikation im Zentrum dieser Arbeit steht. Abschließend werden Befunde einer empirischen Studie zu Wissen und Einstellungen zur PID in der deutschen Allgemeinbevölkerung präsentiert.

Kapitel 3 widmet sich den theoretischen Grundlagen zur Entscheidungsfindung und Argumentation im Bereich Medizinethik. Nach einer kurzen Einführung in die Medizinethik wird im Speziellen auf die Entscheidungsfindung und Argumentation zu medizinethischen Themen eingegangen. Entscheidungsfindung wird in dieser Arbeit als informelles Schlussfolgern konzeptualisiert, das wiederum durch Argumentation erfasst werden kann. Auf dieser Grundlage werden im Anschluss ausgewählte Ergebnisse von empirischen Studien zur Argumentationsqualität und zu Argumentationsmustern zu kontroversen Wissenschaftsthemen dargestellt. Zum Abschluss dieses Kapitels werden mögliche Einflussfaktoren auf die Entscheidungsfindung und Argumentation abgeleitet und diskutiert.

Im Mittelpunkt von *Kapitel 4* stehen die theoretischen Grundlagen zur Konzeption der multimedialen Lernumgebung. Ausgangspunkt bildet zunächst eine Definition des Begriffes Multimedia. Es folgen Ausführungen zu den für diese Arbeit relevanten kognitions- und medienpsychologischen Grundlagen zum Lernen mit neuen Medien. Basierend hierauf wird auf die Gestaltung multimedialer Lernumgebungen aus didaktischer und medialer Sicht eingegangen. Im Anschluss daran werden Erläuterungen zum Feedback in multimedialen Lernumgebungen gegeben. Danach wird auf den konkreten Einsatz von multimedia-

len Lernumgebungen im informellen Lernort Museum eingegangen. Das Kapitel schließt mit der Darstellung von beispielhaften empirischen Befunden zur Effektivität von multimedialen Lernumgebungen zur Vermittlung von kontroversen Wissenschaftsthemen.

Kapitel 5 stellt ausführlich die multimediale Lernumgebung Dialogstation „Gentest“ vor, die als Kontext und Gegenstand der experimentellen Studie diente. Neben den Rahmenbedingungen werden die Ziele und Inhalte, die didaktische und mediale Gestaltung der Lernumgebung sowie die konkrete Gestaltung der implementierten Feedbackinformation beschrieben und erste Evaluationsergebnisse berichtet.

Zum Abschluss des Theorieteils werden in *Kapitel 6* aus den vorgestellten theoretischen Ansätzen und empirischen Befunden Konsequenzen für die vorliegende experimentelle Studie abgeleitet.

Der empirische Teil der Arbeit umfasst die Kapitel 7 bis 11 und ist folgendermaßen strukturiert:

In *Kapitel 7* werden die Fragestellungen der Studie präsentiert und aus den Ausführungen im Theorieteil Hypothesen abgeleitet und formuliert.

Kapitel 8 beschreibt die Methode der experimentellen Laborstudie. Es werden die Stichprobe und das Design der Studie sowie die verwendete Lernumgebung bzw. Experimentalumgebung dargestellt. Des Weiteren werden die Operationalisierung der unabhängigen Variablen und der Versuchsablauf dargestellt sowie die eingesetzten Erhebungsinstrumente zur Erfassung der Kontrollvariablen und der abhängigen Variablen beschrieben.

Kapitel 9 präsentiert die Ergebnisse der experimentellen Studie. Nach der Überprüfung der experimentellen Voraussetzungen werden die Ergebnisse bezüglich der Analyse der parasozialen Interaktion mit der Medienfigur, des Entscheidungsverhaltens, der Argumentationsqualität und der Argumentationsmuster dargestellt. Des Weiteren werden wichtige Einflussfaktoren auf das Entscheidungsverhalten und die Argumentation untersucht und analysiert.

In *Kapitel 10* werden die Untersuchungsergebnisse zusammengefasst und im Rückgriff auf die Theorie diskutiert.

Zum Abschluss der Arbeit werden in *Kapitel 11* in einer Gesamtdiskussion die zentralen Ergebnisse der Studie zusammengefasst und darauf aufbauend Konsequenzen für die zukünftige Forschung und für den Einsatz in der pädagogischen Praxis gezogen.

2 Theoretische Grundlagen zur Präimplantationsdiagnostik (PID)

Dieses Kapitel befasst sich sehr ausführlich mit dem gendiagnostischen Verfahren der PID, da dieses Thema als Untersuchungsgegenstand für diese Arbeit ausgewählt wurde. Die PID stellt aktuell ein besonders kontrovers diskutiertes ethisches Konfliktfeld in der Medizin dar, denn in keinem anderen medizinischen Bereich verläuft die Grenze zwischen Leben und Tod so nah nebeneinander wie bei der PID (Höble, 2003a). Insbesondere der Anwendungsbereich zur Erzeugung von sog. "Retterkindern", die einem kranken älteren Geschwisterkind für die Behandlung notwendige Stamm- oder Knochenmarkzellen liefern sollen (sog. PID-HLA), ist aus ethischer Sicht höchst umstritten und steht im Fokus dieser Arbeit. Um ein reflektiertes, ethisches Urteil bezüglich der Durchführung einer PID bzw. PID-HLA fällen zu können, ist es unerlässlich, sowohl die medizinisch-naturwissenschaftlichen Grundlagen als auch die aktuelle Rechtslage bezüglich dieses gendiagnostischen Verfahrens zu kennen (Höble, 2003a).

Zunächst wird eine kurze Einführung zu genetischen Untersuchungen in der Humanmedizin gegeben (Kapitel 2.1). Die anschließenden Kapitel befassen sich mit den deskriptiven Grundlagen der PID. Es folgen zum einen Ausführungen zu den medizinisch-naturwissenschaftlichen Grundlagen der PID (Kapitel 2.2), in denen das Verfahren, die verschiedenen Anwendungsbereiche und insbesondere die PID-HLA zur Erzeugung eines Gewebespenders vorgestellt werden. Zum anderen werden die rechtlichen Aspekte der PID behandelt (Kapitel 2.3), indem die Regelung in Deutschland, Richtlinien und Empfehlungen von Kommissionen und Gremien sowie ausgewählte internationale Regelungen präsentiert werden. Im Anschluss erfolgt eine ethische Analyse am Beispiel der PID-HLA zur Erzeugung eines Gewebespenders (Kapitel 2.4), in welcher der Frage nach dem moralischen Status von Embryonen nachgegangen wird, die wesentlichen Argumente für oder gegen eine PID-HLA diskutiert und diesbezüglich ethische Stellungnahmen von Kommissionen und Gremien präsentiert werden. Zum Abschluss der theoretischen Grundlagen zur PID wird eine empirische Studie zu Wissen und Einstellungen in der Bevölkerung zu diesem Verfahren vorgestellt (Kapitel 2.5).

2.1 Einführung zu genetischen Untersuchungen in der Humanmedizin

Als genetische Untersuchung definieren Schmitz & Wiesing (2008) allgemein „alle Untersuchungen, die Informationen über die genetische Ausstattung eines Menschen liefern“

(S. 14). Dementsprechend sind die Einsatzfelder der Gendiagnostik sehr vielfältig und umfassen beispielsweise neben dem Nachweis von Erbkrankheiten sowie Infektionskrankheiten, die Krebsvorsorge, Transplantationsanalysen zur Vermeidung oder Verringerung von Abstoßungsreaktionen oder die gerichtsmedizinische bzw. forensische Analytik (z. B. genetischer "Fingerabdruck" zur Ermittlung von Straftätern). Wenn in diesem Kapitel von einer genetischen Untersuchung, oder umgangssprachlich Gentest, die Rede ist, ist dabei die molekulare Analyse des menschlichen Erbgutes (engl.: *deoxyribonucleic acid*, nachfolgend abgekürzt mit: DNA) gemeint, die dem Nachweis einer Erbkrankheit dient (Müller, 2005).

Zur Analyse des menschlichen Humangenoms kommen unterschiedliche molekularbiologische Verfahren zur Anwendung, die jedoch an dieser Stelle, um den Rahmen der Arbeit nicht zu sprengen, nicht näher ausgeführt werden; für einen Überblick sei auf Holinski-Feder (2006) verwiesen.

Eine Diagnostik von genetisch bedingten Krankheiten kann einerseits bei Erkrankten durchgeführt werden, um diesen Personen Aufschluss über die Ursachen der Krankheit zu geben. Zum anderen können Gentests auch an nicht erkrankten Risikopersonen zur Ermittlung des individuellen Erkrankungsrisikos durchgeführt werden. Im letztgenannten Fall spricht man von einer *prädiktiven Gendiagnostik* (Berth, Dinkel & Balck, 2003).

Je nach Zeitpunkt der Untersuchung (postnatal, pränatal oder präimplantativ) können drei verschiedene gendiagnostische Verfahren in der Humanmedizin unterschieden werden (Kochte-Clemens, 2003):

(1) Postnatale *prädiktive genetische Untersuchungen* werden an Kindern und Erwachsenen durchgeführt und dienen der Ermittlung des individuellen Erkrankungsrisikos. Ziel ist die Identifikation bestimmter Gene bzw. Gendefekte, die im späteren Leben zu einer Erkrankung führen oder hierzu disponieren. Als Beispiele können hier Veranlagungen für gewisse Krebsarten wie Brust- oder Darmkrebs oder die Erbkrankheit Chorea Huntington angeführt werden.

(2) *Pränatale genetische Untersuchungen* werden *in vivo*, d. h. im Mutterleib, durchgeführt. Durch eine vorgeburtliche Pränataldiagnostik (PND) können bereits während einer bestehenden Schwangerschaft Fehlbildungen, Chromosomenabweichungen und zahlreiche Erbkrankheiten des Embryos oder Fetus festgestellt werden. Der Einsatz der PND in der Schwangerschaftsvorsorge ist inzwischen üblich und gilt als gesellschaftlich akzeptiert (Kreß, 2009). Zur PND zählt z. B. die Amniozentese, bei der zwischen der 15. und 18. Schwangerschaftswoche aus dem Fruchtwasser Zellen des Fetus entnommen werden können (vgl. Hengstschlager, 2006a). Kommt es bei einer PND zu einem pathologischen

Untersuchungsbefund, wird mit hoher Wahrscheinlichkeit die Schwangerschaft abgebrochen (NER, 2003).

(3) *Präimplantative genetische Untersuchungen*. Unter Präimplantationsdiagnostik (PID) versteht man ein gendiagnostisches Verfahren, das nach einer künstlichen Befruchtung und *in vitro*, d. h. am Embryo im Reagenzglas *vor* der Implantation in die Gebärmutter, durchgeführt wird. Da dieses gendiagnostische Verfahren im Mittelpunkt dieser Arbeit steht, wird die PID sehr ausführlich in den nächsten Kapiteln behandelt.

Genetische Untersuchungen sind in der Regel mit einer Reihe von Vor- und Nachteilen verbunden (Berth et al., 2003). Als Argumente für eine prädiktive Diagnostik werden u. a. die Schaffung von Sicherheit und Beseitigung von Ungewissheit, die teilweise gegebenen medizinisch-therapeutischen Möglichkeiten zur Vorsorge oder die bessere Lebensplanung und -gestaltung, z. B. bezüglich des Kindeswunsches oder der Berufs- oder Partnerwahl, gesehen (Zerres, 1993). Mögliche Argumente gegen eine genetische Untersuchung können Diskriminierungen und Stigmatisierungen (Arbeitsplatz, Versicherungen) oder die psychischen Belastungen für die Betroffenen und deren Familie sein (z. B. Angst, Depressionen oder Schuldgefühle). Da das durch eine genetische Untersuchung gewonnene Wissen irreversibel ist, kann es auch zu einer Einschränkung der Lebensqualität kommen (Faller 1997; Keller 2000). Insbesondere die möglichen Nachteile beim Abschluss von (Lebens-) Versicherungen stellen in den Medien ein intensiv diskutiertes Thema dar, auch wenn die Versicherer selbst bestreiten, auf die Ergebnisse von genetischen Untersuchungen zurückgreifen zu wollen (Regenauer, 2001).

Allerdings existieren, wie Berth et al. (2003) betonen, bislang keine allgemein gültigen Regeln, für welche Personen oder bei welchen Krankheiten eine genetische Untersuchung immer zu bestimmten Konsequenzen führt. Ob nun beispielsweise die Lebensqualität bei Inanspruchnahme eines Gentests eingeschränkt oder verbessert wird, hängt neben individuellen Persönlichkeitsfaktoren auch von krankheitsspezifischen Aspekten wie der Verfügbarkeit von Vorsorge- oder Behandlungsmaßnahmen ab. Für viele genetisch diagnostizierbare Erkrankungen, wie z. B. der Erbkrankheit Chorea Huntington, gibt es derzeit noch keine Heilungs- oder Vorsorgemöglichkeiten (Kreuz, 1996).

Chorea Huntington ist eine schwere neuropsychiatrische Erkrankung, die, wenn der entsprechende genetische Defekt diagnostiziert wird, relativ sicher im Durchschnitt zwischen dem 35. und 40. Lebensjahr ausbricht und nach ca. 15 Jahren immer zum Tode führt. Ein positives Testergebnis ermöglicht in diesem Fall einen „konstruktiven“ Umgang mit dieser Krankheit, denn es können sinnvolle Entscheidungen bezüglich der Lebens- und Familienplanung getroffen werden und die krankheitsfreie Zeit kann besser genutzt und genossen werden. Bei einem negativen Befund ist dagegen eine völlig andere Lebensplanung mög-

lich. Allerdings kann das Nichtwissen vom Ausbruch einer unheilbaren Krankheit unter Umständen auch mit einer besseren Lebensqualität verbunden sein als das Wissen über das zukünftige Schicksal (Höble, 2001a).

Die relativ sichere Vorhersagbarkeit der Chorea Huntington durch einen Gentest ist jedoch eher die Ausnahme. Für die meisten Krankheiten kann lediglich ein erhöhtes Risiko für das Auftreten der Krankheiten diagnostiziert werden. Aussagen über den tatsächlichen Eintritt, den Schweregrad oder die Verlaufsform der Krankheit sind nicht möglich (Kochte-Clemens, 2003). Außerdem stehen nur für sehr wenige erblich bedingte Erkrankungen, z. B. bei erblichem Brust- oder Darmkrebs, kurative Ansätze, die das Risiko zu erkranken, reduzieren können, zur Verfügung. Dies sind in der Regel sehr invasive und belastende Methoden, wie beispielsweise die prophylaktische Kol- oder Mastektomie, in der vorbeugend der Dickdarm oder beide Brüste zur Minimierung des Erkrankungsrisikos operativ entfernt werden (vgl. Eisinger et al., 2000).

Die Durchführung genetischer Testverfahren ist somit nicht nur mit Chancen, sondern auch mit zahlreichen Risiken verbunden, deren Identifikation für die ethische Bewertung eine zentrale Bedeutung zukommt (vgl. Heinrichs, 2006). Bei einer Entscheidung für oder gegen einen Gentest sollten deshalb vorher alle Vor- und Nachteile rational und reflektiert abgewogen werden (Höble, 2001a).

Die Inanspruchnahme einer genetischen Untersuchung unterliegt, wie bereits angesprochen, vielen Einflussfaktoren, eine bedeutende Rolle spielen in diesem Zusammenhang auch die Einstellungen gegenüber gendiagnostischen Verfahren. Im Allgemeinen wird der Einsatz gentechnischer Verfahren in der Medizin („rote Gentechnik“) von der Bevölkerung positiver bewertet als beispielsweise in der Landwirtschaft oder Lebensmittelherstellung („grüne Gentechnik“) (vgl. Götz, 2001; Hampel & Renn, 1999; Pfister, Böhm & Jungermann, 2000). Die Akzeptanz der Gentechnik hängt dabei wie Siegrist & Bühlmann (1999) feststellen konnten vom wahrgenommenen Nutzen ab: Medizinische Anwendungen wie die Gendiagnostik stoßen auf eine größere Akzeptanz als nichtmedizinische Anwendungen, da mit ihnen ein höherer Nutzen assoziiert wird.

Berth, Dinkel & Barth (2002) untersuchten in einer deutschlandrepräsentativen Studie die Akzeptanz und Einstellungen in Bezug auf genetische Untersuchungen. Sie befragten insgesamt 2076 Personen zu möglichen Vor- und Nachteilen sowie Befürchtungen im Hinblick auf die molekulargenetische Diagnostik für erbliche Erkrankungen. Insgesamt betrachtet zeigen die Ergebnisse, dass genetische Untersuchungen auf eine große Akzeptanz in der deutschen Allgemeinbevölkerung stießen, etwa zwei Drittel der Befragten standen ihnen befürwortend gegenüber. Trotz dieser positiven Grundhaltung wurden aber auch mögliche Nachteile von Gentests und Befürchtungen wie beispielsweise Schwanger-

schaftsabbrüche, Missbrauch der Daten, Gefahr der Eugenik oder Diskriminierung artikuliert. Als Determinanten der Akzeptanz zu genetischen Untersuchungen konnten religiöse Überzeugungen, Bildungsstand und Lebensalter identifiziert werden. In Abhängigkeit von diesen soziodemographischen Merkmalen konnten deutliche Unterschiede festgestellt werden, wobei die Religionszugehörigkeit am stärksten die Akzeptanz bestimmte: Personen, die einer Religionsgemeinschaft angehören, unterschieden sich sehr deutlich in ihren Einstellungen bezüglich genetischer Untersuchungen von konfessionslosen Personen. Konfessionsgebundene bewerteten Gentest deutlich kritischer als Konfessionslose, sie sahen weniger positive Aspekte und stimmten in höherem Maße den negativen Aussagen über genetische Untersuchungen zu. Ferner standen ältere Menschen und Personen mit einem niedrigeren Bildungsgrad genetischen Untersuchungen ablehnender und kritischer gegenüber. Ältere sahen mehr negative Aspekte als Jüngere, Personen mit einem höheren Bildungsgrad mehr positive und weniger negative Aspekte hinsichtlich genetischer Untersuchungen als niedrig Gebildete. Bezüglich des Geschlechts bestanden keine Unterschiede in den Einstellungen.

Im Vergleich zu den Ergebnissen internationaler Studien zeigte sich, dass genetische Untersuchungen in Deutschland insgesamt etwas kritischer beurteilt werden als beispielsweise in den USA (Shaw & Bassi, 2001), in Großbritannien (Human Genetics Commission, 2001) oder Finnland (Aro et al., 1997; Hietala et al., 1995; Jallinoja et al., 1998).

Hinsichtlich der Einstellungen zu gendiagnostischen Verfahren scheint auch das Wissen über Humangenetik eine Rolle zu spielen. Jallinoja & Aro (2000) konnten einen Zusammenhang zwischen dem allgemeinen Wissen über Genetik und den Einstellungen gegenüber Gentests nachweisen. Personen mit dem höchsten Wissen über Genetik waren genetischen Untersuchungen gegenüber gleichzeitig enthusiastischer und skeptischer eingestellt als diejenigen mit dem geringsten Wissen. Personen mit niedrigem Wissen hatten außerdem größere Schwierigkeiten, sich eine Meinung zu genetischen Untersuchungen zu bilden. Todt & Götz (1998) haben in einer psychologischen Studie die Interessen und Einstellungen von Jugendlichen an der Gentechnologie untersucht und dabei festgestellt, dass sich ein Interesse an diesem Themengebiet erst ab etwa 15 Jahren zeigt und der Interessenschwerpunkt überwiegend auf ethischen Aspekten und Anwendungen im Bereich der Umwelt und Medizin lag.

Im Hinblick auf die Akzeptanz und das Wissen zu genetischen Untersuchungen ist auch zu beachten, dass aufgrund der intensiven und kontroversen medialen Diskussion über die Möglichkeiten und Grenzen der Humangenetik die Meinungsbildung und die Kenntnisse über den Gegenstandsbereich massiv beeinflusst werden (Singer, Corning & Antonucci, 1999; Stockdale, 1999). Demzufolge kann davon ausgegangen werden, dass sich die Mei-

nung zu genetischen Untersuchungen sowie der Wissensstand in der Allgemeinbevölkerung in Zukunft noch verändern wird.

2.2 Medizinisch-naturwissenschaftliche Grundlagen der PID

2.2.1 Verfahren

Unter dem Begriff PID oder PGD (engl. *preimplantation genetic diagnosis*) wird die genetische Untersuchung von Embryonen *vor* der Implantation in die Gebärmutter der Frau verstanden (Hennen & Sauter, 2004). Eine konstitutive Voraussetzung für die Durchführung einer PID bildet die künstliche Befruchtung (IVF) der Eizelle mit einer Samenzelle in der Petrischale (Griesinger et al., 2003; Steck, 2001).

Ziel der PID ist es vor allem, Paaren mit einer Veranlagung für schwere genetisch bedingte Krankheiten, zu einem gesunden Kind zu verhelfen. Nach der Bestimmung der genetischen und chromosomalen Eigenschaften werden nur diejenigen Embryonen ausgesucht und transferiert, bei denen krankhafte Veränderungen im Erbgut mit großer Wahrscheinlichkeit ausgeschlossen werden können (Kollek, 2002). Embryonen mit pathologischem Befund werden in der Regel „verworfen“, d. h. vernichtet, indem die Versorgung eingestellt wird und die Embryonen dann absterben.

Die erste erfolgreiche klinische Durchführung einer PID fand 1990 in England statt (Handyside et al., 1990). Damaliges Anwendungsziel war die Geschlechtsbestimmung des Embryos vor dem Transfer, um die Übertragung einer über das X-Chromosom vererbten Krankheit zu verhindern.

Seitdem wurden gemäß den Daten der europäischen Fortpflanzungsmedizinergesellschaft ESHRE 4047 Kinder nach einer PID geboren (Harper et al., 2010). Das ESHRE PGD (European Society of Human Reproduction and Embryology Preimplantation Genetic Diagnosis) Konsortium ist ein freiwilliger Zusammenschluss vorwiegend europäischer, aber auch außereuropäischer Kliniken, die alle durchgeführten PID-Zyklen und deren Ausgang dokumentieren. Da nicht alle Zentren von der ESHRE-Datenerhebung vollständig erfasst werden, dürfte die aktuelle und genaue Zahl der „PID-Kinder“ weit höher liegen.

Die bislang zehn erschienenen ESHRE-Berichte dokumentieren eine zunehmende Anwendung der PID in der reproduktionsmedizinischen Praxis, die auf eine wachsende Akzeptanz und nicht zuletzt auch auf eine Ausweitung des Indikationenspektrums zurückzuführen ist. Der zuletzt veröffentlichte X. Bericht der ESHRE umfasst die Angaben von 57 PID-Zentren weltweit und deckt den Zeitraum des Jahres 2007 bis Oktober 2008 ab. Innerhalb

dieser Zeitspanne wurden insgesamt 5887 PID-Behandlungszyklen durchgeführt, die zu 1519 Schwangerschaften, zu 995 Geburten und aufgrund mehrerer Mehrlingsschwangerschaften letztendlich zu 1206 geborenen Kindern führten. Das Durchschnittsalter der Frauen, die eine PID durchführen ließen, lag bei 36 Jahren (Harper et al., 2010).

Nachfolgend wird die technische Durchführung einer PID näher beschrieben. Der Ablauf eines PID-Behandlungszyklus im Rahmen einer medizinisch assistierten Fortpflanzung setzt sich grob aus fünf aufeinander folgenden Schritten zusammen, die von der Erzeugung des Embryos bis zu dessen Transfer in die Gebärmutter reichen (vgl. Griesinger et al., 2003; Hengstschläger, 2006b; NER, 2003).

A) Hormonstimulation und Eizellgewinnung. Damit genügend Eizellen schnell heranreifen und zur Befruchtung zur Verfügung stehen, muss sich die Frau einer hormonellen Stimulationsbehandlung unterziehen, denn für gewöhnlich reift nur ein Ei in jedem Zyklus in den Eierstöcken heran. Sobald per Ultraschall genügend Eibläschen (Follikel) festgestellt wurden, werden die Eizellen durch ultraschallgesteuerte Punktion aus den Follikeln mit einer dünnen Nadel abgesaugt (sog. Follikelpunktion). Üblicherweise werden bei diesem operativen Eingriff im Rahmen einer IVF 8 bis 10 Eizellen gewonnen (Enquete-Kommission des Deutschen Bundestages, 2002; nachfolgend abgekürzt mit: EK). Bei einer PID werden in der Regel mehr Eizellen als bei der IVF ohne PID benötigt, da viele Embryonen aufgrund nicht erfolgreicher Biopsien oder nicht aussagekräftiger Untersuchungsergebnisse nicht transferiert werden können (Lanzerath, 2010). Die Anzahl der pro Zyklus gewonnenen Eizellen bei einer PID liegt im Schnitt bei etwa 12 (Harper et al., 2010).

B) Extrakorporale Befruchtung. Die durch die Follikelpunktion gewonnenen Eizellen können nun durch eine IVF oder intrazytoplasmatische Spermieninjektion (nachfolgend abgekürzt mit: ICSI) befruchtet werden. Bei der klassischen IVF werden die Eizellen mit dem aufbereiteten Sperma in einer Nährlösung in einem Reagenzglas zusammengebracht und ca. 16 bis 20 Stunden in einem Brutschrank bei Körpertemperatur inkubiert (Griesinger et al., 2003). Bei der extrakorporalen Befruchtung der Eizelle wird meist das ICSI-Verfahren der klassischen IVF vorgezogen. Bei der ICSI wird eine einzige, vorher bereits isolierte Samenzelle direkt in das Zytoplasma der Eizelle mithilfe einer Pipette injiziert. Der Vorteil dieser Methode besteht darin, dass das Risiko einer Kontamination der Eizelle mit genetischem Material von anderen, nicht an der Befruchtung beteiligten Spermien, die letztlich zu einer Verfälschung des Diagnoseergebnisses führen können, minimiert werden kann. Zudem ist ICSI auch bei schweren männlichen Fertilitätsstörungen möglich (Böcher, 2004).

C) Embryobiopsie (Entnahme). Der Befruchtungsvorgang ist mit der Verschmelzung der Zellkerne von Ei und Spermium und der Bildung eines diploiden Kerns der Zygote abge-

geschlossen. Dieser Vorgang wird als Furchung bezeichnet, die entstandenen Tochterzellen werden Blastomeren genannt. Auf das Zwei-Zellstadium folgen das Vier- und Acht-Zell-Stadium, nach etwa 3 bis 4 Tagen umfasst der Embryo etwa 16 Blastomeren. Die Embryo-biopsie wird in den meisten Zentren am 3. Tag nach der IVF durchgeführt. Zu diesem Zeitpunkt bestehen die Embryonen aus 6 bis 10 Blastomeren, von denen, je nach Entwicklungsstand des Embryos, zur Untersuchung des Erbmaterials ein bis maximal zwei Zellen entnommen werden. Durch die Biopsie im Blastomerenstadium wird die Entwicklung des Embryos nicht beeinträchtigt. Zu einem früheren Zeitpunkt wie etwa im Vier-Zell-Stadium kann die Embryonalentwicklung jedoch gestört werden (Griesinger et al., 2003). Für die Entnahme der Zellen muss der Embryo zunächst mit Hilfe einer Mikropipette angesaugt und am Ende der Pipette fixiert werden. Der Embryo ist zu diesem Zeitpunkt von einer Schutzhülle (*zona pellucida*) umgeben, die entweder chemisch mit Hilfe einer Säure, durch Laserbeschuss oder durch mechanische Punktion geöffnet wird. Anschließend werden durch das entstandene Loch mit einer feinen Kanüle ein oder zwei Zellen des Embryos abgesaugt.

Vor dem Hintergrund des Embryonenschutzgesetzes (nachfolgend abgekürzt mit: ESchG; siehe Kapitel 2.3.1), das die Entnahme und Zerstörung totipotenter Zellen verbietet, wird häufig eine Entnahme der Zellen im späteren sog. Blastozystenstadium vorgeschlagen (vgl. z. B. Bundesärztekammer, 2000; nachfolgend abgekürzt mit: BÄK). Denn die im früheren Blastomerenstadium entnommenen ein bis zwei Zellen können nach dem Stand der gegenwärtigen Forschung zu diesem Zeitpunkt noch totipotent sein, d. h. unter bestimmten Bedingungen könnten sich aus ihnen vollständige Organismen, d. h. eigenständige Embryonen entwickeln (vgl. Beier, 1999).

Deshalb wird in diesem Zusammenhang diskutiert, die Biopsie 5 bis 6 Tage nach der Befruchtung und damit im Blastozystenstadium (ab ca. 32 Zellen) durchzuführen. Diese Zellen sind nach der mehrheitlichen Auffassung der Forscher nicht mehr totipotent, sondern lediglich pluripotent, d. h. diese Zellen können sich nur noch zu verschiedenen Organen entwickeln. Da zu diesem Zeitpunkt bereits eine Differenzierung in Embryoblast (Embryonalgewebe) und Trophoblast (Nährgewebe) stattgefunden hat, kann die genetische Untersuchung an den Trophoblastzellen vorgenommen werden, die später nicht den Embryo, sondern die Plazenta bilden. Da keine embryonalen Zellen verwendet werden, erscheint diese Methode ethisch eher akzeptabel. Allerdings stellt sich die Biopsie in diesem späten Stadium als sehr schwierig heraus, da die Zellen zu diesem Zeitpunkt sehr klein sind und eng aneinanderkleben (Ziegler, 2004). Aufgrund dieser sog. Kompaktifizierung besteht eine größere Gefahr der Verletzung einzelner Zellen, deren Erbmaterial dann die Untersuchungsprobe verunreinigen und dadurch das Ergebnis verändern könnte. Bei einer Zellent-

nahme in diesem Stadium kommt es auch häufiger zu einer Zerstörung des Embryos als bei der Zellentnahme im Blastomerenstadium. Aufgrund dieser Nachteile wird die Blastozystenbiopsie wenig praktiziert (Griesinger et al., 2003).

D) Genetische Diagnostik. Die genetische Analyse der im Rahmen der Embryobiopsie entnommenen Zellen erfolgt je nach Fragestellung und Anwendungsindikation mittels verschiedener molekularbiologischer Methoden und dauert gewöhnlich zwischen 2 und 24 Stunden (vgl. Yaron, Gamzu & Malcov, 2004).

E) Embryotransfer. Nach erfolgter Diagnostik werden nur diejenigen Embryonen, welche die gewünschten genetischen Merkmale aufweisen, mithilfe eines Katheters in die Gebärmutter übertragen. Die restlichen Embryonen, bei denen das unerwünschte Merkmal nachgewiesen wurde, werden aussortiert und verworfen (Neuer-Miebach, 1999). Zur Vermeidung risikoreicher Mehrlingsschwangerschaften werden meist nur zwei Embryonen pro Behandlungszyklus transferiert (Harper et al., 2001). Bleiben nach der Gendiagnostik genetisch unauffällige Embryonen übrig, können diese überschüssigen Embryonen in manchen Ländern für eine spätere Übertragung durch definiertes Abkühlen auf ca. -180°C unter Zusatz von speziellen Gefrier- und Nährlösungen in flüssigem Stickstoff eingefroren werden (EK, 2002). Diese sog. Kryokonservierung von überzähligen Embryonen ist in Deutschland nicht erlaubt.

Abbildung 1 illustriert zusammenfassend den Ablauf einer PID an Blastomeren, der die fünf Schritte (A) Hormonstimulation und Eizellgewinnung, (B) extrakorporale Befruchtung, (C) Embryobiopsie, (D) genetische Diagnostik und (E) Embryotransfer umfasst:

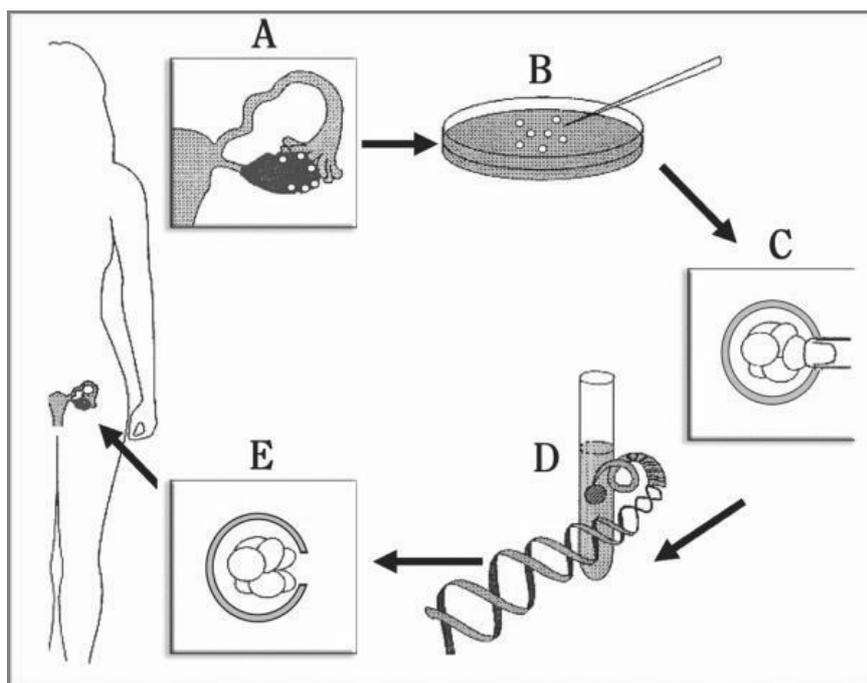


Abbildung 1: Ablauf einer PID an Blastomeren (Hengstschläger, 2006b, S. 10).

Zur PID zählt streng genommen auch das Verfahren der Polkörperdiagnostik (Präkonzeptionsdiagnostik; nachfolgend abgekürzt mit: PKD). Im Gegensatz zur PID wird bei der Polkörperanalyse die genetische Untersuchung nicht am 3 Tage alten Frühembryo (Embryobiopsie), sondern noch früher an den Polkörpern der Eizelle, vor der Verschmelzung der Kerne von Ei- und Samenzelle, durchgeführt (Schwinger, 2004). Die PKD ist in Deutschland gesetzlich erlaubt und ethisch weniger umstritten, da keine Embryonen verworfen werden müssen. Allerdings wird diese Diagnostik seltener angewendet, da sie nur Aufschluss über das Erbmateriale der Mutter gibt und eine eingeschränkte Aussagekraft besitzt (Griesinger et al., 2003). Vorliegende Arbeit befasst sich lediglich mit der PID nach Embryobiopsie.

Die Erfolgsaussichten und die Risiken der PID entsprechen weitgehend den Erfahrungen der IVF ohne PID (NER, 2003). Die Erfolgsrate im Sinne der sog. „Baby-take-home“-Rate, d. h. die Anzahl an lebend geborenen Kindern pro IVF-Zyklus, fällt relativ gering aus. In Deutschland lag laut Deutschem IVF-Register im Jahr 2010 die Wahrscheinlichkeit einer Lebendgeburt pro durchgeführten IVF-Zyklus lediglich bei etwa 17 % (DIR, 2010). Folglich bekommt höchstens nur jede 5. Frau ein Kind nach einer durchgeführten IVF-Behandlung. Viele Frauen müssen dementsprechend mehrere, wiederholte IVF-Zyklen durchlaufen, um überhaupt mit einem Kind aus der Behandlung entlassen zu werden (NER, 2003). Die Erfolgswahrscheinlichkeit hängt allerdings stark vom Alter der Frau und von anderen Faktoren wie der Anzahl und der Qualität der transferierten Embryonen ab.

Bezogen auf die PID gibt Kollek (2002) die Wahrscheinlichkeit, nach einem Embryotransfer schwanger zu werden und die Schwangerschaft über die 20. Woche hinaus zu halten, mit 15-20 %, im günstigsten Fall mit 30 % an. Demzufolge muss jede Frau, die eine PID durchführen lassen möchte, sich durchschnittlich drei bis sechsmal einer IVF/PID-Behandlung unterziehen.

Für die Frau wiederum ist die Durchführung einer IVF/ICSI mit erheblichen physischen und psychischen Belastungen verbunden (NER, 2003). Aufgrund der Hormonbehandlung kann es zu akuten Nebenwirkungen wie dem sog. ovariellen Überstimulationssyndrom (Ovarial Hyperstimulation Syndrome; nachfolgend abgekürzt mit: OHSS) kommen. Das OHSS macht sich im schwersten Stadium durch eine rasche Zunahme des Bauchumfanges bemerkbar, der durch eine äußerst schmerzhaft zystische Vergrößerung der Eierstöcke und durch massive Wasseransammlungen im Bauchraum bewirkt wird. Zusätzlich können schwerwiegende Komplikationen wie Thrombosen, akutes Leber-Nieren-Versagen oder Atemnot auftreten. In Einzelfällen sind auch Todesfälle berichtet worden (Kollek, 2002; NER, 2003; Steck, 2001). Bei der operativen Eizellentnahme und beim Transfer der Emb-

ryonen nach der Untersuchung können weitere Komplikationen, wie Blutungen oder Infektionen, auftreten, wobei dies in der Praxis aber nur sehr selten geschieht (NER, 2003).

Weitere Risiken der IVF/PID bestehen in der erhöhten Rate von Mehrlingsschwangerschaften und der Zerstörung oder Schädigung des Embryos bei der Entnahme der zu untersuchenden Zellen (vgl. Deutsches Referenzzentrum für Ethik in den Biowissenschaften (nachfolgend abgekürzt mit: DRZE), 2010; Harper et al., 2001). Einige Studien weisen auch auf ein erhöhtes Fehlbildungsrisiko und ein niedrigeres Geburtsgewicht von in vitro gezeugten Embryonen hin (DRZE, 2010; Kollek, 2002).

Zu nennen sind schließlich noch die hohen psychischen Belastungen der Frau bzw. des Paares, denn mit der Durchführung der PID sind zahlreiche Hoffnungen und Ängste verbunden:

ob genügend Eizellen reifen, ob die In-vitro-Befruchtung Erfolg hatte und die Embryonen sich entwickeln, ob und wie viele Embryonen bei der Biopsie geschädigt worden sind, ob die DNA- oder Chromosomenuntersuchung erfolgreich war, ob ungeschädigte, entwicklungsfähige Embryonen für den Transfer zur Verfügung stehen und schließlich die wichtigste Frage, ob eine Schwangerschaft eingetreten ist und aufrecht erhalten werden kann. (NER, 2003, S. 31-32)

Ferner besteht bei einer PID immer das Risiko von Fehldiagnosen, weil das Verfahren technisch gesehen sehr schwierig durchzuführen ist, denn für die Untersuchung stehen höchstens zwei Zellen zur Verfügung und der Test kann nicht wiederholt werden, da die Zellen dabei zerstört werden (Baruch, Kaufman & Hudson, 2008). Als weitere Gründe werden mangelnde Erfahrung der Untersucher oder Kontaminationen der UntersuchungsdNA durch Fremd-DNA genannt. Aktuelle Veröffentlichungen beziffern das Risiko für Fehldiagnosen zwischen weniger als 1 % bei der Geschlechtsbestimmung und 7 % bei einigen monogen bedingten Erbkrankheiten (Steinke & Rahner, 2009).

Aufgrund der unzureichenden Prognosesicherheit der PID wird betroffenen Paaren zur Absicherung des PID-Testergebnisses eine zusätzliche PND während der Schwangerschaft empfohlen und in fast der Hälfte der Fälle dann auch praktiziert. Bei pathologischem Befund wird in den meisten Fällen die Schwangerschaft abgebrochen (NER, 2003).

Die Kosten für die Durchführung einer PID sind relativ hoch und müssen im Regelfall von den Patienten selbst getragen werden. Nach Angaben des Brüsseler Reproduktionsmediziner Peter Platteau aus dem Jahr 2007 betragen in Belgien allein die Kosten für die künstliche Befruchtung pro Behandlungszyklus 3700 bzw. 4200 Euro (IVF bzw. ICSI); hinzu kommen zusätzliche Kosten für die Gendiagnostik mittels PID und für vorbereitende Maßnahmen bei ICSI mit PID, die mit jeweils 1400 Euro bzw. 600 Euro zu Buche schlagen (Kreß, 2009).

Zum Angebot der PID existieren verschiedene Alternativen für die betroffenen Paare. Beispielsweise stehen für genetische Hochrisikopaare medizinisch-technische Alternativen wie Methoden der PND (Amniozentese oder Chorionzottenbiopsie) offen. Außerdem können anstatt einer Inanspruchnahme einer PID auch soziale Alternativen wie der Verzicht auf genetisch eigene Kinder, Adoption und Pflegschaft oder die vorbehaltlose Annahme eines Kindes mit seiner genetischen Ausstattung in Betracht gezogen werden (EK, 2002).

2.2.2 Anwendungsbereiche

Ziel bei allen Anwendungsindikationen für eine PID ist zunächst die Erfüllung des Kinderwunsches (EK, 2002). Insbesondere der Wunsch nach einem gesunden Kind ist hierbei von enormer Bedeutung. Eine Ausnahme bilden Paare, die sich aufgrund einer eigenen Erbkrankheit selbst ein Kind wünschen, das unter der gleichen Krankheit leidet (z. B. Kleinwuchs oder Taubheit).

Zusammenfassend lassen sich vier verschiedene Zielgruppen im Hinblick auf eine Anwendung der PID unterscheiden (vgl. EK, 2002; Haker, 2002):

1. Sog. *Hochrisikopaare*, die zwar eine normale Fertilität aufweisen, jedoch genetisch schwer vorbelastet sind und für die ein hohes Risiko besteht, ein schwerkrankes Kind zu bekommen. Im Rahmen einer PID können etwa 200 Erbkrankheiten festgestellt werden (vgl. The Preimplantation Genetic Diagnosis International Society (PGDIS), 2008). Die wohl bekanntesten Beispiele für Erbkrankheiten, die im Rahmen einer PID diagnostiziert werden können, sind die neuronale Erkrankung Chorea Huntington, die Stoffwechselkrankheit Mukoviszidose oder chromosomale Störungen wie die Trisomie 21, bei der das 21. Chromosom dreifach vorliegt.
2. Paare, die an einer *Fruchtbarkeitsstörung* leiden. Mithilfe der PID lässt sich im Zuge einer Sterilitätstherapie die Erfolgsrate der IVF wesentlich verbessern, indem nach Tests ausgewählte Embryonen häufiger zu Schwangerschaften führen. Beim diesem sog. Aneuploidie-Screening werden die Embryonen routinemäßig auf Chromosomenstörungen überprüft, da diese als ursächlich für die o. g. Probleme gesehen werden (Hennen & Sauter, 2004).
3. Paare mit sog. *Altersrisiko*, bei denen infolge des mütterlichen Alters (über 35 Jahre) eine statistisch höhere Wahrscheinlichkeit besteht, ein Kind mit einer chromosomalen Störung zu bekommen und dann ein Aneuploidie-Screening durchgeführt wird.
4. Paare, welche die PID *ohne spezifische Krankheitsindikation* anwenden. In diesem Zusammenhang kann zwischen nicht-medizinischen Gründen (z. B. zur Geschlechtsselektion ohne Krankheitsbezug oder Auswahl erwünschter Eigenschaften) oder indirekt me-

dizinischen Gründen (z. B. Gewebespende für ein erkranktes Geschwisterkind) unterschieden werden.

Dem aktuellen X. Bericht des ESHRE PGD Konsortiums zufolge wird das Aneuploidie-Screening weltweit inzwischen am häufigsten durchgeführt (Harper et al., 2010): In fast zwei Drittel aller PID-Behandlungen (63.7 %; 3753 Zyklen) wurde ein Aneuploidie-Screening durchgeführt. Für 34.7 % der PID Untersuchungen (2042 Zyklen) lag als Indikation eine Erbkrankheit vor. Dem ESHRE-Report zufolge bildete die soziale Geschlechtsbestimmung mit 1.6 % (92 Zyklen) den geringsten Anteil der PID-Untersuchungen.

Um den Rahmen dieser Arbeit nicht zu sprengen, wird auf eine Darstellung der verschiedenen Indikationen für PID-Untersuchungen verzichtet, einen sehr guten Überblick diesbezüglich gibt beispielsweise die Monographie von Neubauer (2009).

Im Zentrum der Arbeit steht die sehr seltene und äußerst kontrovers diskutierte Anwendung der PID zur medizinischen Selektion von immunologisch geeigneten Embryonen als mögliche Gewebespende für ein älteres, erkranktes Geschwisterkind (sog. PID-HLA), die im nächsten Kapitel sehr ausführlich beschrieben wird.

2.2.3 PID-HLA zur Erzeugung eines Gewebespenders

Mit einer PID-HLA wird das Ziel verfolgt, einen Embryo auszuwählen, dessen Gewebe immunologisch mit einem bereits existierenden schwer erkrankten Geschwisterkinds verträglich ist, um für jenes als potentieller Gewebespende von Knochenmark oder Nabelschnurblut fungieren können (Neubauer, 2009; Robertson, 2003).

In diesem Zusammenhang spricht man auch von HLA-Typisierung, HLA-Matching, Gewebetypisierung (*tissue typing*) oder umgangssprachlich von der Erzeugung eines „Retterkindes“ (*saviour sibling*), „Rettergeschwisters“ oder „Designerbabys“.

Das HLA-System (engl. *human leukocyte antigen system*) ist ein erbliches System von Gewebsantigenen beim Menschen, die sich auf der Oberfläche von fast allen Gewebezellen nachweisen lassen. Der sog. Haupthistokompatibilitätskomplex umfasst eine Gruppe von Genen, die auf dem kurzen Arm des Chromosoms Nr. 6 lokalisiert sind. Diese genetischen Marker spielen eine Schlüsselrolle bei der Unterscheidung von körpereigenen und körperfremden Zellen durch das Immunsystem und sind von zentraler Bedeutung für die Akzeptanz eines Transplantats. Für die immunologische Gewebeverträglichkeit bei einer Transplantation von Zellen, Geweben oder Organen ist eine Übereinstimmung in den relevanten HLA-Merkmalen zwischen Spender und Empfänger entscheidend (Spender-Empfänger-Matching). Je größer die Kompatibilität der HLA-Marker, desto geringer ist die Gefahr

einer Transplantatabstoßung (vgl. Janeway et al., 2002; NEK-CNE, 2007; Rehmann-Sutter, 2007).

Der HLA-Typus variiert zwischen Individuen sehr stark, unter Geschwistern ist die Wahrscheinlichkeit einer HLA-Kompatibilität am größten (Burmester & Pezzutto, 1998). Die Wahrscheinlichkeit, dass zwei natürlich gezeugte Geschwister HLA-identisch sind, beträgt 25 % (NEK-CNE, 2007). Dies wird auch als *perfect match* bezeichnet und verspricht den besten Transplantationserfolg. Mit Hilfe einer PID-HLA kann die Wahrscheinlichkeit, einen immunologisch verträglichen Organspender zu erzeugen, erheblich gesteigert werden (NER, 2003). Gleichzeitig kann beim potentiellen „Retterkind“ auch die in der Familie aufgetretene Erbkrankheit ausgeschlossen werden. Die Wahrscheinlichkeit eines *perfect match* sinkt im Falle der doppelten Selektion bei einer autosomal-rezessiven Krankheit wie der Fanconi-Anämie von 25 % auf 18 % (de Wert, 2005).

Für die spezifische Auswahl nach immungenetischen Kriterien und dem Ausschluss der Erbkrankheit muss aufgrund „der kombinatorischen Möglichkeiten durch die gleichzeitige Berücksichtigung mehrerer Genorte“ (NER, 2003, S. 46) eine beträchtliche Anzahl von Embryonen (ca. 20-30) erzeugt werden. Nach der molekulargenetischen Bestimmung der HLA-Gene werden dann diejenigen Embryonen, die übereinstimmende Gewebemerkmale mit dem Geschwisterkind aufweisen und als Gewebespender geeignet sind, ausgewählt und in den Uterus der Frau transferiert (Rechitsky et al., 2004). Das HLA-Matching dient lediglich der Eignung als Gewebespender und ist für die Gesundheit des per PID ausgewählten Kindes selbst irrelevant (Rehmann-Sutter, 2007). Da es um die Therapie einer schweren Krankheit geht, zählt die PID-HLA zu den (indirekt) medizinischen Indikationen (EK, 2002; Haker, 2002). Allerdings ist der zu behandelnde Patient nicht mit dem per PID-HLA ausgewählten Individuum identisch (Rehmann-Sutter, 2007).

Bislang sind nur wenige Einzelfälle, in denen eine PID-HLA durchgeführt wurde, bekannt geworden (NER, 2003). Einer Studie von Baruch et al. (2008) zufolge bieten in den USA etwa ein Viertel (23 %) aller befragten IVF/PGD-Kliniken die PID-HLA an, die jedoch nur sehr selten, in etwa 1 % (43) aller Behandlungszyklen dann auch praktiziert wird. Der aktuelle ESHRE-Report bestätigt ebenfalls diese geringe Anzahl an durchgeführten Gewebetypisierungen; das HLA-Typing wurde lediglich in 3 % der Anwendungsfälle (180 Zyklen) durchgeführt (Harper et al., 2010). Die sehr niedrige Durchführungsquote der PID-HLA lässt sich u. a. darauf zurückführen, dass die Erbkrankheiten, die eine Gewebespende zur Therapie erfordern, äußerst selten auftreten und die HLA-Diagnostik sehr aufwändig ist (Steinke & Rahner, 2009).

In Anlehnung an Rehmann-Sutter (2007) soll nachfolgend exemplarisch eine typische Fallkonstellation zur PID-HLA beschrieben werden.

Fall Adam Nash. Die erste erfolgreiche Durchführung einer PID zur Selektion eines HLA-verträglichen Geschwisterkinds wurde erstmals im Jahr 2001 von der Arbeitsgruppe Verlinsky et al. publiziert. Adam Nash wurde im Rahmen einer IVF gezeugt und mithilfe der PID als passender Gewebespende für seine sechsjährige Schwester Molly ausgewählt, die mit der extrem seltenen und lebensbedrohlichen Krankheit Fanconi-Anämie geboren wurde (EK, 2002).

Die Fanconi-Anämie zählt zu den autosomal-rezessiven Erbkrankheiten (vgl. Buselmaier & Tariverdian, 2007). Bei einer autosomal-rezessiven Erbkrankheit müssen die entsprechenden Gene beider Chromosomen den Defekt aufweisen, d. h. beide Elternteile müssen Träger des Gendefekts sein, können aber selbst frei von Krankheitsanzeichen sein (sog. Heterozygoten). Für jedes der gemeinsamen Kinder liegt die Wahrscheinlichkeit, von der autosomal-rezessiven Erkrankung betroffen zu sein, bei 25 % (1:4).

Die Fanconi-Anämie führt zu einer aplastischen Anämie (aplastisch: nicht-bildend, Anämie: Blutarmut), einer Blutbildungsstörung, die durch eine stark verminderte oder völlig ausbleibende Bildung von roten und weißen Blutkörperchen sowie Blutplättchen gekennzeichnet ist (Deutsche Fanconi-Anämie-Hilfe e.V., 2005). Durch das Nachlassen der Blutproduktion leiden die betroffenen Kinder zunehmend mehr an Schwäche, Blutungen und Infektionen. Im Alter zwischen 4 und 10 Jahren entwickeln Fanconi-Anämie-Patienten meist ein lebensbedrohliches Knochenmarkversagen, das sich durch sinkende Blutbildwerte manifestiert und mit zahlreichen Komplikationen einhergeht.

Die schwer kranken Kinder sterben meistens im Alter von 15 bis 20 Jahren (Deutsche Fanconi-Anämie-Hilfe e.V., 2005; Hennen & Sauter, 2004). Daneben weisen Fanconi-Anämie-Patienten häufig angeborene körperliche Anomalien auf, wie beispielsweise Kleinwuchs oder Fehlbildungen der Arme und Daumen. Im Vergleich zur Durchschnittsbevölkerung besteht außerdem ein extrem erhöhtes Krebsrisiko, etwa 10-20 % der Fanconi-Anämie-Patienten erkranken an Leukämie (Deutsche Fanconi-Anämie-Hilfe e.V., 2005).

Der bislang einzig etablierte und langfristige Therapieansatz zur Heilung des Knochenmarkversagens bei der Fanconi-Anämie besteht in einer erfolgreichen Transplantation von Stammzellen des blutbildenden Systems aus dem Knochenmark oder Nabelschnurblut (sog. hämatopoietische Stammzellen) (Rehmann-Sutter, 2007). Ohne Knochenmarktransplantation liegen die Heilungschancen bei etwa 30 % (Hennen & Sauter, 2004). Wichtige Voraussetzung für eine hämatopoietische Stammzell-Transplantation ist das Vorliegen eines geeigneten Spenders, der eine möglichst hohe Ähnlichkeit im HLA-Typ aufweist.

Den idealen Spender und somit die erste Wahl zur Therapie der Fanconi-Anämie stellt ein gesundes HLA-identisches Geschwisterkind dar (Deutsche-Fanconi-Anämie-Hilfe e.V., 2005). In diesem Fall liegen die Überlebenschancen bei etwa 80-85 %. Allerdings verfügen nur 20-30 % der betroffenen Patienten über einen solchen Spender bzw. Donor.

Als zweite Wahl kommen gut passende unverwandte Fremdspender in Frage, da sich die Transplantationsergebnisse bei der Fremdspende zunehmend verbessern. Damit für möglichst viele betroffene Patienten ein geeignetes Transplantat zur Verfügung steht, ist in den letzten Jahrzehnten ein weltweit vernetztes System von Knochenmarkspenderdateien aufgebaut worden, in dem mittlerweile über 12 Millionen potentielle Knochenmarkspender registriert sind (Schütt & Bröker, 2009). Zur Ergänzung des Knochenmarkspendernetzes und aufgrund der erfolgreichen bisherigen Nabelschnurblut-Transplantationsergebnisse wurden in den letzten Jahren auch vermehrt Nabelschnurblutbanken eingerichtet. Bei vielen Erkrankungen kann auf diesem Wege bereits ein mit einem Geschwisterspender vergleichbarer Fremdspender gefunden werden, bei der Fanconi-Anämie war dies in der Vergangenheit jedoch noch nicht der Fall.

Im Fall Adam Nash fand sich für Molly kein HLA-genotypisch identischer Spender. Eine Transplantation aus genotypisch nicht-identischen Stammzellen hätte die Überlebenschancen deutlich vermindert. Daraufhin wurde die PID angewendet, um einen Embryo zu finden, der einerseits nicht von der Fanconi-Anämie betroffen war und andererseits auch mit Molly HLA-identisch war (Rehmann-Sutter, 2007).

Nach insgesamt fünf durchgeführten IVF-Zyklen über einen Zeitraum von 4 Jahren, in denen etwa 80-100 Embryonen erzeugt und 7 geeignete Embryonen transferiert wurden, wurde im Jahr 2000 der gesunde und HLA-identische Bruder Adam geboren. Im letztlich erfolgreichen 5. Behandlungszyklus wurde Adam aus 16 Embryonen per PID ausgewählt (EK, 2002). Bei der Geburt von Adam wurde sein Nabelschnurblut zurückbehalten, um als Quelle von HLA-identischen Stammzellen zu fungieren. Molly's Leben konnte daraufhin mittels einer Transfusion seines Nabelschnurblutes gerettet werden, ihre Blutbildung hatte sich nach der OP normalisiert (Franklin & Roberts, 2006).

Der Fall Adam Nash weist nach Rehmann-Sutter (2007, S. 120) drei Besonderheiten auf:

1. Selektion eines Embryos zunächst gegen die Krankheit, dann auf HLA-Kompatibilität.
2. PID-HLA stellte die einzige praktikable Alternative dar, um das Leben des erkrankten Geschwisterkindes zu retten.
3. Vom „Retterkind“ wurde lediglich das Nabelschnurblut verwendet.

Im Fall Adam Nash war die Gewinnung von transplantierbaren hämatopoietischen Stammzellen mit PID-HLA die am besten verfügbare Therapieoption, da kein HLA-kompatibler

verwandter Spender vorhanden war und die Verwendung von nicht-verwandten Stammzellen aus Spenderprogrammen oder Spenderdatenbanken entweder nicht verfügbar waren oder wesentlich schlechtere Prognosen boten (Rehmann-Sutter, 2007).

Die Entnahme von Nabelschnurzellen ist für den Spender mit keinerlei Risiko oder Schmerzen verbunden, da die Nabelschnur für gewöhnlich nach der Geburt mit der Plazenta weggeworfen wird. Bei der Entnahme von Knochenmark wird dagegen ein minimal-invasiver Eingriff unter Anästhesie durchgeführt, was in seltenen Fällen auch zu Komplikationen führen kann. Das bei der Knochenmarkspunktion entnommene Gewebe regeneriert sich von selbst, so dass das „Retterkind“ in der Regel keinen körperlichen Schaden davon nimmt. In beiden Fällen sind die entnommenen Zellen für das kranke Kind überlebenswichtig, jedoch nicht für das „Retterkind“ (Rehmann-Sutter, 2007).

Allerdings weist Rehmann-Sutter (2007) auch auf theoretisch mögliche Ausweitungen des Anwendungsbereiches einer PID-HLA hin, wie beispielsweise den Empfängerkreis der Spende, denn eine Erzeugung eines PID-HLA-Kindes ist auch zur Rettung für andere, erwachsene Angehörige wie eines Elternteils oder eines nahen Verwandten, denkbar. Wie Pennings & De Wert (2003) berichten, gab es bereits derartige Anfragen, ob auch andere nahe Verwandte als Geschwister Empfänger sein könnten.

Möglich wäre auch ein präventiver Einsatz des HLA-Matchings, in dem Eltern das prophylaktische Ziel verfolgen könnten, dass sich ihre Kinder bei Bedarf untereinander Stammzellen spenden könnten. Allerdings stehen in diesem Fall, wie Simon & Schenker (2005) bemerken, die hohen Kosten und Belastungen der PID mit dem wahrscheinlichen Nutzen in keinem tragbaren Verhältnis zueinander.

Eine weitere mögliche Ausweitung könnte auch in der Entnahme nachwachsender Organe (z. B. Teile der Leber) oder verzichtbarer Organe (z. B. einer Niere) bestehen. Diese Ausweitung wurde jedoch bisher noch nicht ernsthaft vorgeschlagen, geschweige denn durchgeführt. Ferner besteht die Gefahr wiederholter Transplantationen vom gleichen Spender, falls die Transplantation beim ersten Mal nicht anschlägt, das Transplantat nicht lange überlebt oder wenn die Gewinnung zu wenig ergiebig ist. Neben der Gefahr wiederholter Transplantationen weist Thomas (2004, zitiert nach Rehmann-Sutter, 2007) dazu auf eine Ausweitung der Praxis hin, denn von der Nabelschnurblutspende ist es nur ein kleiner Schritt zur Knochenmarkspende und dann zu invasiveren Verfahren.

Die beiden letztgenannten hypothetischen Szenarien und deren mögliche Folgen werden im Roman „Beim Leben meiner Schwester“ von Jodi Picoult (2004) thematisiert. Romanfigur Anna Fitzgerald wurde per PID als Gewebespende für ihre ältere und unheilbar an Leukämie erkrankte Schwester Kate erzeugt, da ihre Eltern und ihr Bruder Jesse als Spen-

der nicht in Frage kamen. Nach der Nabelschnurblutentnahme bei der Geburt folgen in den Jahren eine Lymphozyten-, Granulozyten- und sogar eine Knochenmarkspende, die ihrer Schwester aber immer nur zeitweise geholfen haben. Als schließlich Kates Nieren versagen, soll Anna im Alter von 13 Jahren auf Wunsch ihrer Mutter eine Niere für Kate spenden. Anna weigert sich und nimmt sich einen Anwalt und verklagt ihre Eltern auf das Recht über ihren eigenen Körper selbst zu bestimmen. Ob die in dem Roman beschriebene Zukunftsvision über diese möglichen Ausweitungsbereiche der PID-HLA tatsächlich eintreten könnte, ist bislang allerdings noch nicht vorherzusehen, da sich keines der bislang geborenen „Retterkinder“ im Alter der Romanfigur befindet.

Zusammenfassend ist bezüglich der Durchführung einer PID-HLA zur Therapie der Fanconi-Anämie festzuhalten, dass die Erfolgsraten dieses gendiagnostischen Verfahrens sehr gering sind, es extrem aufwendig und teuer ist, für die Mutter physisch und psychisch belastend und aus ethischer Sicht aufgrund des hohen Embryonenverbrauchs sehr problematisch ist (Deutsche Fanconi-Anämie-Hilfe e.V., 2005). Des Weiteren ist eine PID-HLA in Deutschland gesetzlich nicht zulässig (siehe Kapitel 2.3).

2.3 Rechtliche Regelungen und Stellungnahmen zur PID

Die Rechtslage im Hinblick auf die Zulässigkeit oder eines Verbots der PID ist in Europa und weltweit sehr unterschiedlich geregelt und hängt stark von der gesetzlichen Ausgestaltung des Embryonenschutzes des jeweiligen Landes ab, der in den einzelnen Ländern äußerst unterschiedlich ausgestaltet ist (vgl. DRZE, 2010).

Die in dieser Arbeit dargestellte Studie wurde im Jahr 2009 durchgeführt. Daher wird - bevor die gegenwärtige rechtliche Situation betrachtet wird, zunächst die zu dem damaligen Zeitpunkt geltende Rechtslage zur PID in Deutschland skizziert (Kapitel 2.3.1). Danach folgen ausgewählte Richtlinien und Empfehlungen von deutschen Ethikkommissionen und Gremien zur PID (Kapitel 2.3.2). Im Anschluss wird anhand ausgewählter Länder exemplarisch die recht unterschiedlich gestaltete Gesetzeslage zur PID im Ausland aufgezeigt (Kapitel 2.3.3).

2.3.1 Regelung in Deutschland

Im Jahr 2009 existierte keine klare einzelgesetzliche Regelung der PID in Deutschland (Lanzerath, 2010). Die Bestimmungen für die PID wurden im Rahmen des Embryonen-

schutzgesetzes (ESchG), das am 01.01.1991 in Kraft getreten ist, geregelt (siehe folgende Tabelle 1).

Der Geltungsbereich des ESchG auf die PID war dabei nicht explizit geregelt und unter Experten sehr umstritten (vgl. Middel, 2009). Diese Tatsache lässt sich u. a. darauf zurückführen, dass bei der Abfassung des Gesetzes im Jahr 1990 das Verfahren der PID noch nicht sehr weit entwickelt war und somit nicht berücksichtigt werden konnte.

Tabelle 1: Auszüge aus dem Embryonenschutzgesetz (ESchG) von 1991.

§ 1 Missbräuchliche Anwendung von Fortpflanzungstechniken

- (1) Mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe wird bestraft, wer
2. es unternimmt, eine Eizelle zu einem anderen Zweck künstlich zu befruchten, als eine Schwangerschaft der Frau herbeizuführen, von der die Eizelle stammt,
3. es unternimmt, innerhalb eines Zyklus mehr als drei Embryonen auf eine Frau zu übertragen,
5. es unternimmt, mehr Eizellen einer Frau zu befruchten, als ihr innerhalb eines Zyklus übertragen werden sollen.

§ 2 Missbräuchliche Verwendung menschlicher Embryonen

- (1) Wer einen extrakorporal erzeugten oder einer Frau vor Abschluss seiner Einnistung in der Gebärmutter entnommenen menschlichen Embryo veräußert oder zu einem nicht seiner Erhaltung dienenden Zweck abgibt, erwirbt oder verwendet, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.
- (2) Ebenso wird bestraft, wer zu einem anderen Zweck als der Herbeiführung einer Schwangerschaft bewirkt, dass sich ein menschlicher Embryo extrakorporal weiterentwickelt.

§ 8 Begriffsbestimmung

- (1) Als Embryo im Sinne dieses Gesetzes gilt bereits die befruchtete, entwicklungsfähige menschliche Eizelle vom Zeitpunkt der Kernverschmelzung an, ferner jede einem Embryo entnommene totipotente Zelle, die sich bei Vorliegen der dafür erforderlichen weiteren Voraussetzungen zu teilen und zu einem Individuum zu entwickeln vermag.

Unstrittig ist, dass eine PID mittels totipotenter Zellen im Blastomerenstadium aufgrund des ESchG verboten ist, da nach § 8 Abs. 1 jede einzelne entnommene totipotente Zelle einen Embryo im Sinne des Gesetzes darstellt, der durch diese Legaldefinition geschützt ist. Zudem untersagt das Gesetz die Entnahme totipotenter Zellen aus einem Embryo, da sie nicht zum Zweck der Erhaltung des Embryos dienen (§ 2 Abs. 1).

Lange und kontrovers diskutiert wurde hingegen, ob die PID an nicht-totipotenten, d. h. pluripotenten Zellen, mit dem ESchG vereinbar ist. Nach 20 Jahren des Meinungsstreits erlaubte der 5. Strafsenat des Bundesgerichtshofs (BGH) in einem Urteil vom 06.07.2010

die PID an pluripotenten Zellen bei Paaren mit einer Veranlagung zu schweren Erbschäden (BGH, 2010; NER, 2011) (zum Unterschied zwischen toti- und pluripotenten Zellen, siehe Kapitel 2.2.1).

Der BGH begründete seine Entscheidung damit, dass zwischen einem Verbot der PID und der rechtlichen Zulässigkeit eines Schwangerschaftsabbruchs ein Wertungswiderspruch besteht. In Deutschland ist eine Abtreibung in den ersten 12 Wochen zwar rechtswidrig, aber straffrei, wenn die schwangere Frau kein Kind haben möchte und eine Beratung nachweist (soziale Indikation, § 218 Abs. 1 StGB) oder wenn später Gefahr für das Leben oder die körperliche und seelische Gesundheit der Frau besteht (medizinische Indikation, § 218a Abs. 2) (von Loewenich, 2005). Wenn eine Behinderung oder schwere Erbkrankheit durch eine PND festgestellt wird, sind Abtreibungen auch noch nach der 12. Schwangerschaftswoche möglich, auch Spätabtreibungen nach der 22. Woche dürfen durchgeführt werden (Kreß, 2009).

Seit dem BGH-Urteil im Juli 2010 wurde die Diskussion über eine klare einzelgesetzliche Regelung der PID auf bundespolitischer Ebene wieder intensiv entfacht (Lanzerath, 2010). Am 07. Juli 2011 stimmten 326 von 594 Abgeordneten des Bundestags für einen Gesetzesentwurf, der die Anwendung der PID bei Paaren für zulässig erklärt, für deren Nachkommen eine hohe Wahrscheinlichkeit besteht, von einer schwerwiegenden Erbkrankheit betroffen zu sein oder wenn mit einer Fehl- bzw. Totgeburt zu rechnen ist. Um Missbrauch zu vermeiden, darf eine PID nur nach einem positiven Votum einer interdisziplinär zusammengesetzten Ethik-Kommission und nach einer vorhergehenden ausführlichen medizinischen und psycho-sozialen Beratung in wenigen lizenzierten Zentren durchgeführt werden. Das neue Gesetz zur Regelung der Präimplantationsdiagnostik (Präimplantationsdiagnostikgesetz - PräimpG) wurde im Bundesgesetzblatt am 24. November 2011 veröffentlicht.

Alle anderen Anwendungsindikationen wie beispielsweise die Selektion von Embryonen zum Zwecke der Gewebespende für ein erkranktes Geschwisterkind (PID-HLA) oder die soziale Geschlechtsauswahl ohne Krankheitsbezug sind weiterhin nach dem ESchG verboten und werden in Deutschland strafrechtlich verfolgt.

2.3.2 Richtlinien und Empfehlungen in Deutschland

In Deutschland gehen die Meinungen zur ethischen Vertretbarkeit der PID sehr weit auseinander, was sich auch in der gesellschaftlichen Akzeptanz der PID niederschlägt. Beispielsweise lehnen sowohl die katholische und evangelische Kirche als auch die Behinder-

tenverbände jegliche Form der PID ab, wohingegen die zuständigen medizinischen Fachgesellschaften, die Deutsche Gesellschaft für Gynäkologie und Geburtshilfe und die Gesellschaft für Humangenetik sich für eine eingeschränkte Zulassung der PID aussprechen (vgl. Böcher, 2004).

Aufgrund der Fülle an Veröffentlichungen erscheint eine Darstellung sämtlicher Stellungnahmen zur PID als nicht sinnvoll. Das Spektrum der vertretenen Positionen bezüglich der Legitimität der PID soll in diesem Kapitel anhand ausgewählter ethischer Stellungnahmen und Empfehlungen von Kommissionen und Gremien verdeutlicht werden. Nachfolgend werden die Richtlinien und Empfehlungen der Bundesärztekammer (2000), der Enquete-Kommission des Deutschen Bundestages (2002) und des Nationalen Ethikrates (2003) bzw. des Deutschen Ethikrats (2011) zitiert.

In ihrem Diskussionsentwurf zu einer „Richtlinie zur Präimplantationsdiagnostik“ vom 24. Februar 2000 unterbreitet die *Bundesärztekammer* einen Regelungsvorschlag zur PID, indem sie eine restriktive Anwendung vorsieht; demnach sollte die PID nur für Paare zugelassen sein, bei deren Nachkommen ein hohes Risiko für eine bekannte und schwerwiegende, genetisch bedingte Erkrankung besteht. Diese bekannten und schwerwiegenden, genetisch bedingten Erkrankungen werden im Diskussionsentwurf nicht explizit benannt, spätmanifestierende Krankheiten sollen aber von der Diagnose ausgeschlossen bleiben. Das Aneuploidie-Screening im Rahmen von IVF-Behandlungen wurde im Diskussionsentwurf der BÄK nicht befürwortet. Die Entnahme pluripotenter, d. h. nicht-totipotenter Zellen aus der Blastozyste zu diagnostischen Zwecken mit dem Ziel der Herbeiführung einer Schwangerschaft ist laut der der BÄK (2000) mit dem ESchG vereinbar (siehe Kapitel 2.3.1).

Im Gegensatz dazu spricht sich die Mehrheit von 16 Mitgliedern der *Enquete-Kommission* des 14. Bundestages „Recht und Ethik der modernen Medizin in ihrem Schlussbericht vom 14. Mai 2002 gegen eine Zulassung der PID aus und empfiehlt das im Embryonenschutzgesetz enthaltene Verbot der IVF zu diagnostischen Zwecken ausdrücklich im Hinblick auf die PID zu präzisieren (EK, 2002). Für die Mehrheit der Mitglieder der Enquete-Kommission hat der Embryo von Beginn an, d. h. ab dem Zeitpunkt der Verschmelzung von Ei- und Samenzelle, ein subjektives Recht auf Leben. Eine Minderheit von 3 Mitgliedern hält das gendiagnostische Verfahren hingegen für vertretbar und plädiert, die PID für hilfeschuchende Paare mit einem nachweisbar hohen genetischen Risiko in Anlehnung an die Indikationslage beim Schwangerschaftsabbruch unter besonderen Voraussetzungen in bestimmten Ausnahmefällen zuzulassen. Von der mehrheitlichen Gruppe wird als Gegenargument die Unmöglichkeit einer präzisen Indikationsbeschränkung für die Anwendung von PID angeführt.

Am 23. Januar 2003 legte der *Nationale Ethikrat* eine Stellungnahme zur PID vor, in der sich umgekehrte Mehrheitsverhältnisse bei den Voten bezüglich der Zulassung einer PID in Deutschland zeigen; während sich die Enquete-Kommission nach Mehrheitsvotum explizit für ein Verbot der PID aussprach, kam der Nationale Ethikrat mit einer Mehrheit von 15 (bei insgesamt 24 Mitgliedern) zu dem Schluss, dass durchaus legitime Anwendungsbereiche bestehen und sprach sich für eine eng begrenzte Zulassung der PID aus (NER, 2003). Die PID sollte ausnahmsweise zugelassen werden für genetische Hochrisikopaare, die mit dem Austragen eines Kindes mit einer schweren und nicht wirksamen therapierbaren Erkrankung oder Behinderung in einen existenziellen Konflikt geraten würden. Ein weiterer legitimer Anwendungsbereich für die PID besteht für Paare, die ein hohes Risiko zur Vererbung einer Chromosomenstörung tragen, bei denen der Embryo nicht die extrauterine Lebensfähigkeit erreichen würde. Analog zum Diskussionsentwurf der BÄK (2000) werden diese Krankheiten nicht genau definiert. Eine Zulassung der PID wurde auch für infertile Paare befürwortet,

wenn wissenschaftliche Untersuchungen bestätigen sollten, dass durch eine Untersuchung auf Chromosomenstörungen die Erfolgsrate der Sterilitätstherapie bei bestimmten Patientengruppen (z.B. erhöhtes Alter oder nach mehreren erfolglosen Behandlungszyklen ohne bekannte chromosomale Störung) signifikant gesteigert und die Anzahl der transferierten Embryonen mit dem Risiko von Mehrlingschwangerschaften verringert werden kann. (S. 106-107)

Darüber hinaus müsse eine angemessene Beratung der Paare hinsichtlich medizinischer, ethischer und psycho-sozialer Aspekte gewährleistet sein. Außerdem sollte die Durchführung einer PID nur wenigen widerruflich lizenzierten Zentren vorbehalten sein und durch geeignete Verfahrensvorschriften gesetzlich geregelt sein, die die Qualität der Anwendung, die Bindung an die Indikation, eine wissenschaftliche Begleitung und Auswertung sowie eine hinreichende Transparenz unter Wahrung der Schweigepflicht und des Datenschutzes sicherstellen. Ferner ist für eine zentrale Dokumentation und Kontrolle Sorge zu tragen.

Lediglich eine Minderheit von 7 Mitgliedern votierte für eine Beibehaltung und Präzisierung des im ESchG enthaltenen Verbots der assistierten Reproduktion (extrakorporalen Befruchtung) zu diagnostischen Zwecken und damit des Verbots der PID (NER, 2003). Dem Minderheitsvotum schließen sich 2 Mitglieder in einem ergänzenden Votum grundsätzlich an, vertreten jedoch die Ansicht, „dass in einem existenziellen Konfliktfall die zu treffende Gewissensentscheidung des Individuums frei sein muss und nicht durch ein staatliches Strafgesetz erzwungen werden kann“ (NER, 2003, S. 105). Alle Mitglieder des Nationalen Ethikrates schlagen vor, „alle mit der Reproduktionsmedizin zusammenhängenden wesentlichen Fragen in einem speziellen Fortpflanzungsmedizingesetz zu regeln“ (NER, 2003, S. 9).

Am 08. März 2011 veröffentlichte der *Deutsche Ethikrat* (nachfolgend abgekürzt mit: DER) eine aktuelle Stellungnahme zur PID, in der ebenfalls zwei alternative, sich einander widersprechende Vorschläge zu einer gesetzlichen Regelung der PID entwickelt wurden. 13 von 26 Ratsmitgliedern votieren für eine Zulassung der PID unter streng begrenzten Voraussetzungen, die im Wesentlichen den Kriterien der Stellungnahme aus dem Jahre 2003 entsprechen, 11 grundsätzlich dagegen; ein Ratsmitglied enthielt sich der Stimme und ein weiteres plädiert in einem Sondervotum für die generelle Zulassung der PID zur Identifikation von lebensfähigen Embryonen und für die Erstellung einer verbindlichen Indikationsliste (vgl. DER, 2011).

Für die vorliegende Arbeit ist die Stellungnahme des Nationalen Ethikrats aus dem Jahre 2003 relevant, da die experimentelle Studie im Jahr 2009 durchgeführt wurde.

2.3.3 Ausgewählte internationale Regelungen

Im Ausland ist die PID bereits zum medizinisch-technischen Standard geworden und wird bis auf wenige Ausnahmen wie die Schweiz, Österreich oder Irland durchgängig praktiziert (Hennen & Sauter, 2004; Kreß, 2009). Nachfolgend wird exemplarisch die relativ unterschiedliche rechtliche Situation in der Schweiz, Frankreich und in den USA vorgestellt.

In der *Schweiz* ist eine PID an Embryonen nach dem Fortpflanzungsmedizingesetz (FMedG Art. 5 III) grundsätzlich verboten (Lanzerath, 2010). Das Fortpflanzungsmedizingesetz erlaubt eine Selektion von Keimzellen nur, wenn die Gefahr der Übertragung einer schweren unheilbaren Krankheit auf die Nachkommen besteht (FMedG Art. 33). In der Schweiz ist die extrakorporale Erzeugung von Embryonen nur zur Überwindung der Unfruchtbarkeit eines Paares zulässig, wenn alle anderen Behandlungsmethoden versagt haben oder aussichtslos sind (FMedG Art. 5). Die Polkörperdiagnostik (PKD) wird dagegen als zulässig angesehen.

In *Frankreich* wird die Anwendung der PID seit 1994 durch das Bioethikgesetz (Loi No. 94-654 du 29 juillet 1994) reguliert. Demnach ist die PID nur für Paare erlaubt, die ein stark erhöhtes Risiko für ein Kind mit einer schweren genetischen Erkrankung haben, die zum Diagnosezeitpunkt als unbehandelbar und unheilbar gelten, wie beispielsweise die Chorea Huntington (Hennen & Sauter, 2004). Um die Diagnostik durchführen zu können, muss zuvor eindeutig bei einem Elternteil die für die Krankheit verantwortliche Anomalie festgestellt worden sein und eine speziell hierfür zugelassene Einrichtung aufgesucht werden. Außerdem müssen die Paare vorab mindestens 2 Jahre zusammenleben (Lanzerath, 2010). Die Neufassung des Bioethikgesetzes von 2004 (Loi No. 2004-800 du 6 août 2004

relative à la bioéthique) erlaubt in Ausnahmefällen eine PID zur Erzeugung eines „Retterkindes“, wie beispielsweise im Falle einer Fanconi-Anämie (Hennen & Sauter, 2004).

In den *USA* wird die PID sehr unterschiedlich geregelt, da die gesetzliche Regelung Aufgabe der einzelnen Bundesstaaten ist (DRZE, 2010). In einigen Staaten ist die PID explizit verboten (z. B. Florida, Louisiana, Maine, Minnesota und Pennsylvania). Andere Bundesstaaten erlauben die PID bei bestimmten medizinischen Indikationen (z. B. Massachusetts, Michigan, North Dakota, New Hampshire und Rhode Island). In der Mehrzahl der Bundesstaaten bestehen dagegen keine gesetzlichen Regelungen für die Zulassung und Durchführung der PID. In diesen Staaten gehen die Anwendungsgebiete weit über die medizinischen Indikationen hinaus. Neben der sozialen Geschlechtsselektion ist auch die Diagnostik von Anlagen für multifaktoriell bedingte Krankheitsanfälligkeiten, wie familiärer Brustkrebs oder Alzheimer, möglich. Das Testen auf derartige genetische Dispositionen mithilfe einer PID ist jedoch medizinisch gesehen äußerst fragwürdig und ethisch hochgradig umstritten, da die Wahrscheinlichkeit des Ausbrechens der Krankheit nur zum Teil von der Art der genetischen Abweichung abhängt und damit schwer vorhersagbar ist (Hennen & Sauter, 2004). Ferner ist in diesen Bundesstaaten auch die positive Selektion von genetisch bedingten Anomalien wie Taubheit oder eine HLA-Typisierung zur Erzeugung eines „Retterkindes“ zulässig (Hennen & Sauter, 2004).

Neben den USA und Frankreich ist eine PID-HLA auch in Großbritannien, Belgien und Kanada gesetzlich erlaubt (Steinke & Rahner, 2009).

2.4 Ethische Analyse am Beispiel der PID-HLA

Das HLA-Matching zählt neben der sozialen Geschlechtsselektion ohne Krankheitsbezug zu den kontroversesten und ethisch umstrittensten Anwendungsbereichen der PID und fungierte in dieser Arbeit als Untersuchungsgegenstand. Die folgenden Äußerungen richten sich daher primär auf die ethischen Aspekte der PID-HLA.

Mit einer PID werden, wie in Kapitel 2.2.2 beschrieben, sehr unterschiedliche Anwendungsziele verfolgt. In der ethischen Diskussion bezüglich der Zulässigkeit der PID geht es sowohl um die Wertung dieser Anwendungsziele selbst, als auch um die Frage, ob und wenn ja, welche dieser Anwendungsziele mögliche Verletzungen von Schutzansprüchen der beteiligten Embryonen rechtfertigen können (Lanzerath, 2010). Die ethische Legitimität und Hochrangigkeit dieser Anwendungsziele wird zwischen den Befürwortern und Kritikern der PID sehr kontrovers diskutiert (vgl. Schräer, 2009). Bei der Beurteilung der PID-

HLA müssen verschiedene medizin-, individual- und sozialetische Gesichtspunkte berücksichtigt werden (Kreß, 2009).

Im Rahmen dieses Kapitels wird zunächst der Frage nach dem moralischen Status des Embryos nachgegangen, indem die unterschiedlichen Positionen, die im medizinethischen Diskurs vertreten werden, ausführlich vorgestellt werden (Kapitel 2.4.1). Im Anschluss steht der aktuelle ethische Diskussionsstand zur Anwendungsindikation der PID-HLA im Mittelpunkt und es werden die wichtigsten Argumente, die am häufigsten für oder gegen die Erzeugung von „Retterkindern“ im ethischen Diskurs angeführt werden, skizziert (Kapitel 2.4.2). Abschließend werden ausgewählte ethische Stellungnahmen und Empfehlungen zur PID-HLA von Kommissionen und Gremien präsentiert (Kapitel 2.4.3).

2.4.1 Moralischer Status des menschlichen Embryos

Im Zentrum der medizinethischen Pro- und Kontra-Diskussion bezüglich der Zulässigkeit der PID steht zunächst die Frage nach der Schutzwürdigkeit und dem moralischen Status des frühen Embryos.

Wie im medizinisch-naturwissenschaftlichen Kapitel 2.2 bereits ausführlich dargelegt wurde, werden im Rahmen einer PID nach einer IVF gezeugte Embryonen genetisch untersucht und bei embryopathischen Befund oder sonstigen unerwünschten Eigenschaften vom Embryotransfer ausgeschlossen und in der Regel dann vernichtet. Aus diesem Grunde ist es für die Bewertung der Handlungsoptionen einer PID entscheidend, inwieweit mögliche Schutzansprüche des Embryos verletzt werden und welcher moralische Status den frühen menschlichen Embryonen innerhalb ihrer ontogenetischen Entwicklung in den ersten Tagen zugesprochen wird. Speziell im Hinblick auf eine PID-HLA muss eine Abwägung zwischen dem Hilfsanspruch eines bereits lebenden kranken Kindes mit dem Schutzanspruch von Embryonen erfolgen (Schräer, 2009).

Zur Bestimmung des moralischen Status von Embryonen ist das jeweils zugrunde gelegte ethische Schutzkonzept maßgebend, wobei sich zwei entgegengesetzte Grundpositionen unterscheiden lassen (Schräer, 2009). Vertreter der ersten Position schreiben dem menschlichen Embryo ein uneingeschränktes Lebensrecht vom Beginn der Befruchtung zu und vertreten somit ein absolutes Schutzkonzept. Die Vertreter der anderen, zweiten Position propagieren hingegen ein graduelles Schutzkonzept und sprechen dem Embryo eine Schutzwürdigkeit in abgestufter Weise zu, die vom Erreichen einer bestimmten Entwicklungsstufe abhängt und mit fortschreitender Entwicklung ansteigt. Als mögliche würderelevante Entwicklungszeitpunkte werden beispielsweise die Einnistung in die Gebärmutter

genannt, da ab diesem Zeitpunkt der Embryo erst wirklich entwicklungsfähig ist oder die Bildung des Primitivstreifens, da ab diesem Zeitpunkt die Mehrlingsbildung ausgeschlossen werden könne und somit die Individuation beendet ist. Nachfolgend werden die vier in der ethischen Debatte angeführten Hauptargumente der Befürworter eines hohen moralischen und schutzwürdigen Status von Embryonen dargestellt: Es handelt sich um das Speziesargument, das Kontinuumsargument, das Identitätsargument und das Potentialitätsargument (vgl. Damschen & Schönecker, 2003).

Diese vier Argumente werden häufig nach den Anfangsbuchstaben ihrer Schlagworte als SKIP-Argumente bezeichnet. Die vier Argumentationstypen werden dabei einzeln vertreten oder bestritten, teilweise aber auch in Kombination, da sie sich ergänzen und voneinander abhängen. Allen vier SKIP-Argumenten ist jedoch gemein, dass sie in der Konklusion münden, dass jeder menschliche Embryo mehr als ein bloßer Zellhaufen ist, Würde besitzt und die Würde des Menschen in Anlehnung an unser Grundgesetz (nachfolgend abgekürzt mit: GG) Art. 1 und 2 unantastbar ist. Diese Konklusion sichert wiederum in allen vier Argumenten den Schutz des Embryos und rechtfertigt das Tötungsverbot (Höble & Lude, 2004).

Nachfolgend werden alle vier SKIP-Argumente kurz skizziert und in eine syllogistische Form gebracht. Zu jedem SKIP-Argument werden in Anlehnung an Damschen & Schönecker (2003) sowohl Pro- als auch Contra-Position vorgestellt.

1. Das Speziesargument. Die Vertreter des absoluten Schutzkonzepts führen in der Regel, oft verbunden mit anderen Argumenten, das sog. Speziesargument an. Grundannahme des Speziesargument besteht darin, dass jeder Embryo biologisch der Spezies *homo sapiens* zugehörig ist und folglich wie alle Mitglieder dieser Spezies Würde besitzt und damit schutzwürdig ist. Embryonen haben dem Prinzip der Gleichbehandlung folgend das gleiche Grundrecht auf Leben wie geborene Menschen inne (Schockenhoff, 2003). Dies gilt ebenso für Neugeborene, Menschen mit einer schweren Behinderung und Komatöse.

Der logische Aufbau des Speziesarguments (S-Argument) lässt sich folgendermaßen darstellen (Damschen & Schönecker, 2003):

- (1) Jedes Mitglied der Spezies Mensch hat Würde aufgrund seiner natürlichen Artzugehörigkeit.
- (2) Jeder menschliche Embryo ist von Anfang an Mitglied der Spezies Mensch.
- (3) Also: Jeder Embryo hat Würde.

Gemäß den Erkenntnissen der modernen Entwicklungsbiologie und Genetik beginnt für Schockenhoff (2003) das individuelle Menschenleben mit der Verschmelzung von Ei- und Samenzelle, denn mit der damit verbundenen Konstituierung des neuen Genoms ist der

Schritt zu einem neuen Menschen vollzogen: „Von diesem Zeitpunkt an ist der Embryo sowohl artspezifisch (als Mensch) wie auch individualspezifisch (als dieser Mensch) festgelegt, ohne dass seine weitere Entwicklung Zäsuren aufweist, die dieses grundlegende Charakteristikum des individuellen Menschseins in Frage stellen oder verändern könnten“ (Schockenhoff, 2003, S. 26). Embryonen besitzen somit ab dem Zeitpunkt der Verschmelzung der Keimzellen Schutzwürdigkeit.

Kritiker wie Merkel (2003) distanzieren sich hingegen von der biologischen Zugehörigkeit als statusstiftende Einheit der Menschenwürde und wenden ein, dass das Spezies-Argument nicht haltbar ist und einen sog. naturalistischen Fehlschluss beinhaltet: „die (unmögliche) direkte Ableitung einer Norm aus einem Faktum“ (Merkel, 2001, S. 55). Ein naturalistischer Fehlschluss, auch Sein-Sollen-Fehlschluss genannt, liegt hier vor, da aus ausschließlich deskriptiven Sätzen eine normative Konklusion abgeleitet wird (Höble & Lude, 2004; Pfeifer, 2009). Konkret heißt das, dass von einem Sein des Menschen (Mitglied der Spezies) unzulässigerweise auf ein ethisches Sollen (Zurechnung von Menschenwürde) geschlossen wird. Aus dem reinen Faktum der Spezieszugehörigkeit können jedoch nicht normative Schlüsse gezogen werden, ohne weitere Annahmen zu treffen, auf die ein nur dem Menschen zustehendes Lebensrecht bzw. ein nur ihm gegenüber geltendes Tötungsverbot geschlossen werden kann.

Um das Speziesargument plausibler zu machen, ist es laut Merkel (2001) notwendig, zunächst bestimmte menschliche Eigenschaften sowie eine Norm zu benennen, die es moralisch gebietet, Wesen mit genau diesen Eigenschaften ein Recht auf Leben und Würde zu gewährleisten. Als Beispiele für diese Eigenschaften können nach Damschen & Schönecker (2003) Autonomie als Fähigkeit zur Zwecksetzung, moralische Autonomie im Sinne von Freiheit, kognitive Fähigkeiten wie Abstraktionsfähigkeit, Präferenzen als zukunftsorientierte Wünsche, Interessen oder Leidensfähigkeit angeführt werden. Insbesondere die Autonomie als die menschliche Fähigkeit zur freien, vernunftgeleiteten Selbstbestimmung gilt bereits seit Immanuel Kant als Grundlage des Würdeanspruchs und damit des Rechts auf Leben. Als weitere Eigenschaften werden der Besitz von Selbstbewusstsein (Singer, 1990) oder die Fähigkeit zur Selbstachtung (Nida-Rümelin, 2001) genannt. Der frühe Embryo weist aber in seinem aktuellen Zustand keine dieser Eigenschaften auf und demzufolge reiche der Hinweis auf die biologische Spezieszugehörigkeit laut Kritikern nicht aus, dem Embryo ein eigenes subjektives Recht auf Leben von Beginn an zuzuerkennen.

2. Das Kontinuumsargument. Eng verbunden mit der Fragestellung des Speziesarguments nach dem biologischen Beginn des menschlichen Lebens ist das sog. Kontinuumsargument. Ausgangspunkt dieses Argument ist, dass jeder Embryo sich kontinuierlich zu einem Menschen entwickeln wird und innerhalb der weiteren Entwicklung bis zum gebo-

renen Menschen keine moralisch relevanten und eindeutig definierbaren Einschnitte bestehen und demzufolge dem Embryo die gleichen Rechte wie einem geborenen Menschen einzuräumen sind (Honnefelder, 2003). Unabhängig von seinem Entwicklungsstadium komme jedem Embryo die gleiche Würde wie aktual existierenden Menschen zu, da der Embryo sich *als* Mensch und nicht *zum* Menschen entwickelt (Iglesias, 1984).

Der logische Aufbau des Kontinuumsarguments (K-Argument) lässt sich folgendermaßen darstellen (Damschen & Schönecker, 2003; Honnefelder, 2003):

- (1) Jedes menschliche Wesen, das aktual ϕ ist, hat Würde.
- (2) Jeder menschliche Embryo ist aktual ϕ , weil er ein menschliches Lebewesen ist, das sich, unter normalen Bedingungen, kontinuierlich (ohne moralrelevante Einschnitte) zu einem geborenen menschlichen Wesen entwickelt, das unzweifelhaft aktual ϕ ist.
- (3) Also: Jeder menschliche Embryo hat Würde.

Mit den ϕ -Eigenschaften sind die im Zusammenhang mit dem Speziesargument bereits genannten Eigenschaften oder Fähigkeiten gemeint, von denen man in der Regel annimmt, dass sie die Würde eines Menschen begründen und aufgrund derer wir ihre Träger unter normalen Umständen nicht töten dürfen.

Gegner des Kontinuumsarguments greifen insbesondere die zweite Prämisse des Arguments an und bestreiten die Kontinuität der Embryonalentwicklung. Den Kritikern zufolge sei es sehr wohl möglich, moralisch relevante Zäsuren innerhalb der Entwicklung zum geborenen Menschen auszumachen (Kaufmann, 2003). Als moralisch relevante Einschnitte in der Embryonalentwicklung werden beispielsweise die Nidation, in der sich die befruchtete Eizelle in die Gebärmutter einnistet oder die Herausbildung des Gehirns und damit das Einsetzen von Empfindungsfähigkeit und eines beginnenden Bewusstseins oder die Geburt gesehen. Ein Embryo im Blastozystenstadium, der beispielsweise sicher über keine Schmerzempfindung verfügt, darf deswegen moralisch anders behandelt werden als ein Neugeborenes. Laut den Kritikern sei ein graduelles, abgestuftes Schutzkonzept vernünftiger zu vertreten, da das Kontinuumsargument einen sog. Zirkelschluss beinhaltet: Bei dem Versuch wird die Aussage, dass es keine moralisch relevanten Einschnitte in der Embryonalentwicklung gebe, selbst als Voraussetzung verwandt. Die Argumentation dreht sich folglich im Kreis, da vorausgesetzt wird, was bewiesen werden möchte: Nämlich, dass es keine moralisch relevanten Einschnitte in der Embryonalentwicklung gebe (Kaufmann, 2003).

3. Das Identitätsargument. Dieses Argument hängt eng mit dem Kontinuumsargument zusammen und besagt, dass aus der unter moralrelevanter Hinsicht bestehenden Identität eines erwachsenen Menschen mit dem Embryo, aus dem er sich nach der Verschmelzung

von Spermien- und Eizelle entwickelt hat, und aus der Tatsache heraus, dass der Erwachsene eine Menschenwürde hat, auch auf die Würde des Embryos zu schließen sei (Enskat, 2003). Von Anfang an entwickelt sich dasselbe Lebewesen: ein Mensch mit einer unverwechselbaren Identität (Göbel, 2005; Honnefelder, 2002). Hätte beispielsweise eine Mutter ein Bild von der Zygote könnte sie später zu ihrem älteren Kind sagen: „Das bist du an deinem ersten Anfang.“ und das Kind könnte zu Recht behaupten: „Das bin ich!“. Dieses Beispiel von Göbel (2005) verdeutlicht, dass jede Person zu jedem Zeitpunkt und in jedem Entwicklungsstadium, sei es als Kind, Neugeborenes, Fötus oder Embryo, seine Identität erhalten sieht und auch als vorhanden betrachtet. Grundgedanke des Identitätsarguments besteht also darin, dass der zum Zeitpunkt der Befruchtung vorhandene Embryo mit dem aktuell existierenden Menschen identisch ist, so dass ihm auch die gleiche Würde zukommen müsse (Enskat, 2003).

Der logische Aufbau des Identitätsarguments (I-Argument) lässt sich folgendermaßen darstellen (Damschen & Schönecker, 2003; Stoecker, 2003):

- (1) Jedes Wesen, das aktuell ϕ ist, hat Würde.
- (2.1) Viele Erwachsene, die aktuell ϕ sind, sind mit Embryonen in moralrelevanter Hinsicht identisch.
- (2.2) Also: Die Embryonen, mit denen sie identisch sind, haben Würde.
- (2.3) Wenn irgendein Embryo Würde hat, dann alle.
- (3) Also: Jeder Embryo hat Würde.

Gegner des Identitätsarguments greifen insbesondere die zweite Prämisse an und führen in der Hauptsache zwei Erkenntnisse der Embryonalentwicklung an, nach denen man nicht zu jedem Zeitpunkt von einem biologischen Individuum sprechen könne. Zum einen wird aufgrund der Möglichkeit der Mehrlingsbildung bis zur Bildung des Primitivstreifens bezweifelt, ob tatsächlich eine Identität zwischen Embryo und Erwachsenem besteht (Stoecker, 2003). Bei der Entstehung von eineiigen Zwillingen ist jedes dieser Kinder mit dem Embryo identisch. In diesem Zusammenhang stellt sich die Frage, wie von einer Identität zwischen dem frühen Embryo und dem erwachsenem Menschen ausgegangen werden kann, wenn sich aus ein und demselben Zellhaufen zwei verschiedene Individuen entwickeln können? Wenn man behauptet, dass jedes Kind mit dem frühen Embryo identisch sei, so müssten demzufolge auch die Zwillingsgeschwister untereinander identisch sein.

Als zweites widerspruchskräftiges Argument führt Stoecker (2003) an, dass sich nur ein verhältnismäßig kleiner Teil des Embryos, etwa ein Drittel, tatsächlich zu einem Fötus und schließlich zu einem Kind entwickelt. Von einer faktischen Identität könne man vom Zeitpunkt der Kernverschmelzung nicht ausgehen, da nicht exakt feststellbar ist, welcher Teil des Embryos sich zum späteren eigentlichen Embryo entwickelt und welcher Teil später

die fötale Plazenta bildet. Den Gegnern des Identitätsarguments zufolge kann von einer Identität des frühen oder späten Embryos sowie seinen weiteren Entwicklungsstadien nicht ausgegangen werden, da es unklar bleibe, mit wem oder was man hier identisch sein sollte (Stoecker, 2003). Viel näher liegt es laut Stoecker (2003), „unsere Existenz dort anfangen zu lassen, wo wir unsere individuelle Gestalt bekommen haben“ (S. 142), also frühestens mit der Ausbildung der Keimscheibe am Ende der 1. Schwangerschaftswoche, vielleicht auch erst mit der Bildung des Primitivstreifens am Ende der 2. Woche.

4. Das Potentialitätsargument. Unter allen vier SKIP-Argumenten gilt das Potentialitätsargument als das einflussreichste und stärkste Argument (Damschen & Schönecker, 2003). Die dem Potentialitätsargument zugrunde liegende These konstatiert, dass jeder Embryo ein potentieller Mensch ist, da bereits die befruchtete Eizelle die volle Potentialität besitzt, sich zu einem geborenen Menschen, d. h. zu personalen Dasein und sittlichen Subjektsein, zu entwickeln (Wieland, 2003). Zwar eignen sich die aktuellen Eigenschaften menschlicher Embryonen nicht dafür, Menschenwürde und Tötungsverbot zu begründen, „aber seine erwartbaren künftigen Eigenschaften sind genau die, auf denen das allgemeine Menschenrecht auf Leben und Würde moralisch gründet“ (Merkel, 2001, S. 58). Potentiell habe der Embryo alle Eigenschaften, die ein geborener Mensch hat (Wieland, 2003). Aufgrund dieses von Anfang an vorhandenen spezifischen Potentials müsse dem Embryo dieselbe Schutzwürdigkeit wie geborenen Menschen zuerkannt werden und sein Leben darf demzufolge nicht zerstört werden. Beispielsweise töte man auch keinen Bewusstlosen, da er nur potentiell Bewusstsein habe.

Der logische Aufbau des Potentialitätsarguments (P-Argument) lässt sich folgendermaßen darstellen (Damschen & Schönecker, 2003):

- (1) Jedes Wesen, das potentiell ϕ ist, hat Würde.
- (2) Jeder menschliche Embryo ist ein Wesen, das potentiell ϕ ist.
- (3) Also: Jeder menschliche Embryo hat Würde.

Vertreter des Potentialitätsarguments legen somit den Schwerpunkt ihrer Argumentation auf die erste Prämisse. Von den Gegnern des Potentialitätsargument wird jedoch angeführt, dass die potentiellen Eigenschaften menschlicher Embryonen keine hinreichende Bedingung für einen moralisch wertvollen und schützenswerten Status sein können (Schöne-Seifert, 2003). Häufig wird hier das sog. Prinz-Charles-Argument oder Kronprinzenargument angeführt, das besagt, dass Prinz Charles zwar potentieller König ist, aber als Kronprinz noch lange nicht die Rechte eines Königs genießen kann. In Analogie dazu besitze der Embryo zwar die Potentialität, sich zu einer Person zu entwickeln, sei es aber im aktu-

ellen Zustand noch nicht und ihm können daher auch nicht dieselben Rechte zukommen wie einer geborenen Person (Schöne-Seifert, 2003; Schräer, 2009).

Ein weiterer Einwand der Kritiker bezieht sich auf den Unterschied in der Potentialität zwischen den befruchteten Eizellen und den Ei- bzw. Spermienzellen, auch Keimzellen oder Gameten, genannt (Schöne-Seifert, 2003). Gebe es keinen Unterschied in der Potentialität, so müsste man auch Gameten Würde und damit die volle Schutzwürdigkeit zuzusprechen; gebe es einen Unterschied, so bestehe er ausschließlich in einer gewissen Wahrscheinlichkeit, sich zu geborenen Menschen zu entwickeln. Ein weiteres Gegenargument formulierte der NER (2003). Bei einer strikter Anerkennung des Potentialitätsarguments und den daraus gezogenen Schlussfolgerungen dürften beispielsweise auch keine Nidationshemmer wie die Spirale oder die Pille danach zugelassen sein, da sie die Einnistung befruchteter Eizellen in die Gebärmutter Schleimhaut verhindern und damit zur Abtötung von Embryonen mit vollem Entwicklungspotential führen. Darüber hinaus verfügen Embryonen, bei denen letale Chromosomenstörungen vorliegen, überhaupt nicht über das Potenzial, sich zu einem vollständigen Menschen zu entwickeln, was auch für einen sehr hohen Prozentsatz bei der natürlichen Zeugung gilt (NER, 2003).

Mit der syllogistischen Wiedergabe der SKIP-Argumente sind keine Wahrheitsansprüche verbunden. Die logisch korrekte Form eines Arguments ist nur eine notwendige, aber keine hinreichende Bedingung für seine Wahrheit. Dennoch ist es wichtig, wie Damschen & Schönecker (2003) betonen, bei der Debatte um den moralischen Status des Embryos zu wissen, „welche logische Form ein Argument hat und um welche Prämissen gestritten wird“ (S. 2). Mit Ausnahme des Speziesarguments lassen sich die vier SKIP-Argumente in einen argumentativen Zusammenhang bringen: Beispielsweise kann man mit dem Begriff des Kontinuums den Begriff der Identität verbinden und diesen wiederum der Potentialität zurechnen (Damschen & Schönecker, 2003).

Zusammenfassend lässt sich feststellen, dass die Ansichten darüber, wie der moralische Status und der Schutz des menschlichen Embryos einzuschätzen sind, sehr stark divergieren (Schräer, 2009). Die Beurteilung des moralischen Status bestimmt jedoch maßgeblich, ob eine Abwägung zwischen dem Hilfsanspruch eines bereits existierenden kranken Geschwisterkindes und dem Schutzanspruch von Embryonen ethisch zu rechtfertigen ist.

2.4.2 Diskussion der Argumente Pro und Contra PID-HLA

Nachfolgend werden die im Zusammenhang mit der vorliegenden Studie wichtigsten ethischen Hauptargumente für oder wider PID-HLA diskutiert.

Die Auswahl von Embryonen nach immunologischen Merkmalen zur Gewebespende stellt dabei, wie Rehmann-Sutter (2007) betont, andere ethische Fragen als beispielsweise die PID zur Prävention einer erblichen Krankheit: Denn im Rahmen einer PID-HLA erfolgt die Selektion der Embryonen nicht (bzw. nicht nur) zur Verhinderung einer Erbkrankheit, sondern primär mit der Absicht, immunkompatibles Gewebe zu therapeutischen Zwecken für ein lebensbedrohlich erkranktes Geschwisterkind zu erzeugen. In diesen Fallkonstellationen steht nicht das neu gezeugte Kind und dessen In-die-Welt-Kommen im Vordergrund, sondern eine „fremdnützige“ medizinische Indikation, aufgrund derer die übrigen erzeugten Embryonen letztendlich vom Transfer ausgeschlossen und vernichtet werden (Schräer, 2009). Gegen diese Form der Selektion zu „fremdnützigen“ therapeutischen Zwecken richten sich gravierende ethische Bedenken, die nachfolgend diskutiert werden.

Der offensichtlichste und von Kritikern der PID-HLA am häufigsten vorgebrachte ethische Einwand betrifft die *Instrumentalisierung des „Retterkindes“* (Rehmann-Sutter, 2007). Denn im Falle der PID-HLA werde der Embryo nicht primär um seiner selbst willen geschaffen, sondern es wird das Ziel verfolgt, ein anderes Kind zu retten (Schräer, 2009). Wenn ein Kind speziell für diesen Zweck ausgewählt und erzeugt wird, finde eine ethisch nicht akzeptable Form der vollständigen Instrumentalisierung eines Menschen statt. Zusätzlich zur Zeugungsmotivation werde das Kind als Spender bei der späteren Gewebeentnahme in seiner körperlichen Integrität verletzt. Aus diesem Grunde wird dem Verfahren der PID-HLA vorgeworfen, dass es zu einer doppelten Instrumentalisierung des Spenderkindes führe: zum einen bei der Zeugung und zum anderen bei der späteren Entnahme der Stammzellen (NEK-CNE, 2007).

Dem kategorischen Imperativ Kants (1785) zufolge ist die vollständige Instrumentalisierung eines Menschen ethisch jedoch immer unzulässig, da dies einen Verstoß gegen die menschliche Würde bedeute. Eine vollständige Instrumentalisierung liegt dann vor, wenn ein Mensch ausschließlich auf den Status als Mittel zum Zweck reduziert wird, wie es die zweite Formel des kategorischen Imperativs von Immanuel Kant bestimmt: „Handle so, dass du die Menschheit, sowohl in deiner Person, als in der Person eines jeden andern, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchst“ (zitiert nach Pfeifer, 2009, S. 33). Im Falle der Selektion eines HLA-kompatiblen Embryos zur Gewebespende für ein bereits lebendes Geschwisterkind läge demnach eine vollständige Instrumentalisierung vor, da der Embryo letztendlich nur zu dem Zweck erzeugt wird, um ein anderes Kind zu retten (Schramme, 2002).

Demgegenüber argumentieren Befürworter der PID-HLA an dieser Stelle, dass bezogen auf das Spenderkind der Instrumentalisierungsvorwurf nicht notwendigerweise greife. Denn, dass das Kind von den Eltern als Spender gewünscht wird, schließt ja nicht aus, dass

es gleichzeitig auch um seiner selbst willen gewünscht und geliebt wird; selbst dann, wenn die Gewebespende im Ergebnis nicht geholfen hat (NER, 2003). Solange das „Retterkind“ auch als Mittel, gleichzeitig aber auch als Zweck behandelt wird, liege keine vollständige Instrumentalisierung im Sinne des kategorischen Imperativs von Kant vor (Rehmann-Sutter, 2007). Auch wenn der Anlass zu seiner Erzeugung die Erbkrankheit des Geschwisterkindes und der Versuch seiner Heilung war, ist es durchaus möglich, dass das Spenderkind unabhängig von seiner Funktion als Spender ein Wunschkind werden kann (NEK-CNE, 2007). In vielen Fällen kann auch bei natürlichen Schwangerschaften der Anlass zur Zeugung eines Kindes der Erfüllung bestimmter Zwecke dienen. Beispielsweise wenn das neue Kind eine Beziehung retten, den Fortbestand der Familie sichern oder als Erbe fungieren soll. Dies muss für das neue Kind nicht von Nachteil sein, da auch in diesen Fällen eine eigenständige Entwicklung einer Eltern-Kind-Beziehung möglich ist (NEK-CNE, 2007).

Einige Verfechter der PID zur Selektion von immunkompatiblen Embryonen zur Gewebespende nehmen jedoch den Instrumentalisierungsvorwurf sehr ernst und versuchen ihn zu vermeiden, indem sie eine Zwischenposition einnehmen (Schräer, 2009). Sie plädieren für eine Zulassung der PID-HLA, wenn der Kinderwunsch nicht primär durch die Krankheit des bereits lebenden Kindes motiviert ist, sondern durch *den Wunsch nach einem weiteren (gesunden) Kind*. Eine Einführung einer entsprechenden gesetzlichen Regelung, dass ein „Retterkind“ nur erzeugt werden darf, wenn sowieso ein zweites Kind gewünscht wird, wäre jedoch kaum praktikabel, da dieser Wunsch in der Regel nur sehr schwer überprüfbar ist. Ferner wäre eine solche Gewissensprüfung ein nicht legitimer Eingriff des Staates in die Privatsphäre der Eltern (NEK-CNE, 2007).

Prinzipiell kann der von Kritikern geäußerte Instrumentalisierungsvorwurf auch auf *die überzähligen und aussortierten Embryonen*, die bedingt durch den Auswahlprozess der PID nicht in den Mutterleib transferiert werden, angewendet werden (EK, 2002; NER, 2003). Generell ist die Durchführung einer PID mit der *Erzeugung und Verwerfung einer sehr großen Anzahl an Embryonen* verbunden.

Gegner der PID argumentieren, dass es sich um eine Selektion „unwerten Lebens“ und vor allem um die Tötung individuellen menschlichen Lebens handelt. Ob die Verwendung von Embryonen als bloßes genetisches Testmaterial und die Selektion nach bestimmten Merkmalen gegen das aus der Menschenwürde ableitbare Instrumentalisierungsverbot verstößt, hängt entscheidend davon ab, ob dem Embryo in vitro Menschenwürde und Lebensschutz zukommt (siehe Kapitel 2.4.1). Wenn dies dem Embryo zugestanden wird, dann muss er moralisch als Person behandelt werden, die ein Recht darauf hat, nicht auf den Status als Mittel reduziert zu werden und folglich nicht verworfen werden darf. Wenn dies dem Emb-

ryo nicht zugestanden wird, ist die Frage der Selektion eine Abwägung „guter“ Gründe (Bioethikkommission beim Bundeskanzleramt, 2004).

In diesem Zusammenhang wird auch häufig von Gegnern der PID-HLA angeführt, dass die ausschließliche Selektion nach HLA-Charakteristik die Erzeugung und Verwerfung einer beträchtlichen Anzahl an *gesunden Embryonen* nach sich zieht (NER, 2003; Rehmann-Sutter, 2007; Schrärer, 2009). Anders als im Falle der doppelten Selektion mit gleichzeitigem Ausschluss der Erbkrankheit, bei der auch „kranke“ Embryonen verworfen werden, sind die Embryonen in diesem Fall gesund, weisen jedoch nicht den gewünschten HLA-Typ auf. Zur Lösung dieses Dilemmas schlägt Devolder (2005) vor, Banken HLA-typisierter, kryokonservierter Embryonen einzurichten. Diese gesunden Embryonen könnten dann von Paaren, die einen HLA-identischen Spender benötigen, bei entsprechender Eignung adoptiert und ausgetragen werden. Möglich wäre auch eine Spende für die Embryonen- oder Stammzellforschung.

Neben dem Instrumentalisierungsvorwurf weisen Kritiker der PID-HLA vielfach auch auf mögliche *psycho-soziale Risiken für das „Retterkind“* hin (NEK-CNE 2007; Rehmann-Sutter, 2007). Ihrer Meinung nach können die psycho-sozialen Belastungen für das „Retterkind“ sehr groß sein, wenn es die Umstände seiner Erzeugung erfährt. Das Wissen, nur aufgrund der Erbkrankheit des Bruders oder der Schwester zu existieren und lediglich zum Zwecke seiner Heilung gezeugt worden zu sein, kann dazu führen, dass die Ausbildung einer eigenen Identität für das Kind schwierig sein kann.

Allerdings gibt es, wie Befürworter der PID-HLA anführen, bislang keinen empirischen Beleg in Bezug auf die vermuteten psycho-sozialen Beeinträchtigungen für das „Retterkind“, da aufgrund der erstmals im Jahr 2000 und auch seltenen Durchführung des Verfahrens diesbezüglich noch keine verlässlichen Studien durchgeführt werden konnten. Die unsystematisch gemachten Erfahrungen mit den ersten „Retterkindern“, die sich heute im Schulalter befinden, geben bislang keinen Anlass zur Beunruhigung. Insofern erweisen sich nach derzeitigem Erkenntnisstand die Befürchtungen psycho-sozialer Folgen für das „Retterkind“ als rein spekulativ und hypothetisch, sollten aber nach Ansicht von Ram (2006) und Rehmann-Sutter (2007) keinesfalls heruntergespielt werden.

Beispielsweise konnte Packman (1999) in einer Studie bezüglich psycho-sozialer Auswirkungen von Knochenmarkspende unter Geschwistern in Kalifornien feststellen, dass 5 von 21 Spendergeschwistern posttraumatische Stressreaktionen mittleren Grades aufwiesen, 2 sogar schweren Grades. Ein überraschendes Ergebnis der Studie war, dass sich die Knochenmarkspende unter Geschwistern auch auf die Nicht-Spender-Geschwister auswirkt; sieben Nicht-Spender-Kinder wiesen einen mittleren Grad von posttraumatischer Stressreaktion auf, 11 einen geringen Grad. Folglich erleben diese Kinder einen vergleichbaren

posttraumatischen Stress wie die Spender-Geschwister. Diese Ergebnisse machen deutlich, dass die Knochenmarkspende die ganze Familie betrifft. Eltern sollten nach Empfehlung von Packman (1999) ihren Kindern emotionale Unterstützung und Aufmerksamkeit zukommen lassen und wenn nötig auch professionelle psychologische Hilfen in Anspruch nehmen.

In einer anderen Studie untersuchten Baetens et al. (2005) die psychologischen Implikationen bei Paaren, in denen die Durchführung einer PID-HLA nicht erfolgreich verlief. Ein Misserfolg der Behandlung kann auftreten, wenn zu wenig Eizellen aufgrund des fortgeschrittenen mütterlichen Alters gewonnen werden konnten, wenn unter den erzeugten Embryonen keine HLA-kompatiblen oder keine HLA-kompatiblen als auch krankheitsfreie sind. Weitere Gründe für das Versagen der Behandlung kann die schlechte Qualität der Embryonen sein oder dass sich die transferierten Embryonen nicht einnisten oder dass die Schwangerschaft letztlich in einer Fehlgeburt endet (Rehmann-Sutter, 2007).

Die mit einem Misserfolg der Behandlung verbundenen psychischen Folgen für die betroffenen Familien sind laut Baetens et al. (2005) nicht zu unterschätzen. Jede erfolglose Behandlung, die nicht in einer Schwangerschaft endet, ist mit besonderem psychischem Stress verbunden, da jedes Mal noch neun Monate Schwangerschaft hinzugezählt werden müssen und gleichzeitig der Krankheitszustand des Kindes sich rapide verschlechtern kann. Gelingt die Spende und wird das kranke Geschwisterkind geheilt, kann dies positiv zum Selbstwertgefühl des Spender-Kindes beitragen. Wenn die Spende aber nicht gelingt, können Gefühle der Wut, der Schuld und der Scham beim Spenderkind auftreten (Baetens et al., 2005; Rehmann-Sutter, 2007). Diese Befunde untermauern nochmals den hohen Stellenwert einer kompetenten psychologischen Begleitung und Beratung der betroffenen Familien bei der Durchführung einer PID-HLA, da jederzeit die Gefahr besteht, dass familiäre und soziale Konflikte aufbrechen können (Kreß, 2009). Auch wenn die Verfechter der PID-HLA einwenden, dass die Eltern ihr Kind genauso lieben wie das andere bereits existierende Kind, ist es für diese sog. Ersatzteillager-Kinder trotz alledem schwierig, sich nicht als Überlebenskrücke zu fühlen (Spiewack, 2005).

Ob das Kind die Tatsache, dass es nur aufgrund seiner Gewebekompatibilität ausgewählt worden ist, als Belastung erlebt oder ob das Wissen, seinem Geschwister existenziell geholfen zu haben, es glücklich und stolz macht, hängt u. a. stark von der Art und Weise ab, wie die Familie mit der Situation umgeht (NER, 2003). Die Qualität der innerfamiliären Beziehungen spielen laut De Wert (2005) eine erhebliche Rolle, denn wenn das Kind fühlt, dass es genauso wie andere Kinder gewollt ist, dürfte nach seiner Ansicht mit keinen ernsthaften Problemen zu rechnen sein.

Befürworter der PID-HLA wenden dagegen an dieser Stelle ein, dass die *Lebensrettung eines kranken Kindes* schwerer wiege als das Eingehen potentieller psycho-sozialer Risiken beim „Retterkind“ (NEK-CNE, 2007). Der mögliche Schaden beim zukünftigen Kind würde durch den wahrscheinlichen Nutzen der Heilung des kranken Kindes bei weitem aufgewogen (Devolder, 2005; Shenfield et al., 2005).

Als Schwäche dieses utilitaristischen Arguments führt Rehmann-Sutter (2007) an, dass der Nutzen für das kranke Kind mit dem möglichen Schaden des zukünftigen „Retterkindes“ aufgerechnet wird, was wiederum mit einer Aberkennung der Menschenwürde dieses Individuums verbunden ist. Von den Befürwortern der PID-HLA wird zur Stützung dieser einfachen Nutzen-Schaden-Abwägung häufig auch das Argument des *unterstellten Altruismus des „Retterkindes“* genannt, das von einer hypothetischen Einwilligung (*hypothetical consent*) des zukünftigen Kindes ausgeht (Rehmann-Sutter, 2007). Laut den Befürwortern der PID-HLA sei es sehr wahrscheinlich, dass das „Retterkind“ später, wenn es alt genug wäre, der von seinen Eltern stellvertretend getroffenen Entscheidung zustimmen würde (Pennings, Schots & Libaers, 2002; Redmon, 1986).

Ferner machen einige Befürworter der PID-HLA geltend, dass Eltern aus deontologischer Sicht moralisch verpflichtet wären, ihr lebensbedrohlich erkranktes Kind zu retten, insbesondere dann, wenn keine anderen erfolgversprechenden Behandlungsoptionen verfügbar sind und ohnehin der Wunsch nach einem weiterem Kind besteht (Pennings et al., 2002; Rehmann-Sutter, 2007; Tsafirir, Shufaro, Simon & Laufer, 2005). Allerdings setzt das Argument einer postulierten *Pflicht der Eltern* zur Lebensrettung ihres Kindes voraus, dass einerseits die PID überhaupt vom staatlichen Gesundheitswesen zur Verfügung gestellt wird und andererseits der mit der PID-HLA verbundene Selektionsprozess von Embryonen für die Eltern ethisch akzeptabel und mit ihren persönlichen moralischen Vorstellungen vereinbar ist (Kreß, 2009; Rehmann-Sutter, 2009). Eltern, die beispielsweise den Embryo aus religiösen Gründen von Beginn der Verschmelzung von Ei- und Samenzelle als Menschen ansehen, können und wollen der postulierten Pflicht zur Lebensrettung nicht nachkommen.

Bei der komplexen ethischen Bewertung der PID bzw. PID-HLA muss neben der Individualenebene, in der insbesondere die Problemlage des betroffenen Paares und dessen Interessen im Vordergrund stehen, auch die Sozialebene, welche die gesellschaftlichen und politischen Konsequenzen miteinbezieht, berücksichtigt werden. Beide Ebenen hängen eng miteinander zusammen: Mit Individualentscheidungen sind stets weitere soziale Implikationen verknüpft, denn bei entsprechender Häufigkeit können individuelle Optionen und Entscheidungen auch Auswirkungen auf gesamtgesellschaftliche Trends, Einstellungen und schließlich gesetzgeberische Vorgaben haben. Individuell getroffene Entscheidungen wer-

den aber auch gleichzeitig von gesellschaftlichen Rahmenbedingungen, Einstellungen und Gesetzesvorgaben determiniert (Bioethikkommission beim Bundeskanzleramt, 2004; Schräer, 2009). Nachdem die Hauptargumente für oder gegen eine PID-HLA aus der individuelleethischen Perspektive vorgestellt wurden, soll nun im Folgenden die soziaethische Dimension der PID-HLA erörtert werden.

Gegner der PID befürchten weitreichende und unabsehbare gesellschaftliche Folgen mit einer Legalisierung der PID. Die Zulassung der PID werde die Türen für die Auswahl nach weiteren Wunscheigenschaften wie beispielsweise dem Geschlecht oder der Intelligenz irreversibel öffnen und im Ergebnis „zum Einfallstor für die Menschengzuchtung“ werden (NER, 2003, S. 144; Rehmann-Sutter, 2007). Diese Argumente werden vom Typ her auch als *Dammbbruchargumente*, „*slippery slope*“-Argumente oder *Argumente der schiefen Ebene* bezeichnet (vgl. Guckes, 1997; Walton, 1992).

Die allgemeine Form eines Dammbbrucharguments lässt sich nach Zoglauer (2004) folgendermaßen darstellen:

- 1) Aus A folgt B oder A führt über eine Kette von Zwischenschritten zu B.
- 2) B ist moralisch nicht wünschenswert.
- 3) Daher dürfen wir A nicht zulassen.

Demzufolge darf die PID nicht zugelassen werden, da sie geeignet ist, allgemein anerkannte Normen aufzuweichen und eine Entwicklung in Gang setzen würde, die auf lange Sicht gesehen zu einer moralisch inakzeptablen Praxis führen würde. Mit einer Zulassung der PID würde die Gefahr eines Dammbbruchs bestehen, welcher den ersten Schritt auf eine „schiefe Ebene“ darstellen könnte, was Schritt für Schritt weitere negative Konsequenzen zur Folge haben könnte (Zoglauer, 2004). Dammbbruchargumente verweisen somit auf Gefahren, die erst in Folge einer Handlung entstehen.

Auf das Beispiel der PID-HLA übertragen, könnte ein Dammbbruch folgendermaßen aussehen: Aufgrund „guter Gründe“, hier die Lebensrettung eines schwer erkrankten Kindes, wird die PID-HLA unter strengen gesetzlichen Auflagen zugelassen. Allerdings werden sich einige Paare mit Kinderwunsch mit dieser Regelung nicht zufrieden geben und eine Aufweichung der Norm fordern. Als Argument könnten sie anführen, wenn eine positive Selektion im Falle eines „Retterkindes“ nach nicht-krankheitsrelevanten Merkmalen zulässig ist, müssten auch ähnliche, positive Selektionen wie das Geschlecht zugelassen werden (Fischer, 2006; Zoglauer, 2004). Bei einer Zulassung der PID bestünde demnach die Gefahr, gesunde Embryonen unter dem Gesichtspunkt der positiven Optimierung nach beliebigen, erwünschten Merkmalen auszuwählen und so einen „Menschen nach Maß“ zu züchten bzw. ein „Designerbaby“ zu erzeugen (NER, 2003).

Einen Dambruch hin zu einem „Menschen nach Maß“ befürchten Kritiker auch bei der PID zum Ausschluss einer schweren Erbkrankheit. Aufgrund der wachsenden Nachfrage könnte es sehr problematisch werden, das Verfahren ausschließlich für Paare mit einer genetischen Disposition für schwere Erbkrankheiten zu erlauben und Paaren mit einer Veranlagung für weniger schwere Erbkrankheiten zu verweigern. Denn schon jetzt wird die Frage, ob eine Krankheit „schwer“ oder „nicht wirksam therapierbar“ sehr uneinheitlich beantwortet (NER, 2003). Auch in diesem Fall könnte als Konsequenz daraus gezogen werden, die PID auch für rein kosmetische Merkmale zu bewilligen (Maio, 2001).

Die Möglichkeit, gesunde Embryonen nach gewünschten Merkmalen auszuwählen, komme laut Kritikern einer eugenischen Selektion gleich und beschwöre damit die *Gefahr einer potentiellen neuen Eugenik* herauf. Der Begriff der Eugenik (griech. *eugenes* = wohlgeboren/von guter Geburt) wurde bereits 1883 von dem britischen Naturforscher Francis Galton, einem Vetter Charles Darwins, eingeführt und bezeichnet einen Wissenschaftszweig, der sich mit der Verbesserung des Erbgutes menschlicher Populationen beschäftigen sollte (Schmidt, 2003). In Deutschland wurde Eugenik vor dem Jahr 1945 auch als Rassen- oder Erbhygiene bezeichnet (Fuchs & Lanzerath, 1998). Im Allgemeinen werden zwei Formen oder Vorgehensweisen der Eugenik unterschieden (Schramme, 2002):

1. Die *positive* Eugenik, die darauf abzielt, die genetische Ausstattung des Menschen hinsichtlich bestimmter erwünschter Eigenschaften zu verbessern.
2. Die *negative* Eugenik, bei der bestimmte Krankheiten verhindert oder eliminiert werden sollen.

In der ersten Hälfte des 20. Jahrhunderts galt die Eugenik als eine etablierte Wissenschaft. Auf ihr gründeten zunächst die bis in die 30-er Jahre reichenden Sterilisations- und Einwanderergesetze in Europa und Nordamerika, hernach aber auch die rassistische Bevölkerungspolitik im Dritten Reich, welche die Sterilisation von hunderttausenden vermeintlich „erbgeschädigter“ Menschen und die Tötung von über 70 000 Behinderten und Kranken zur Folge hatte (Mieth, 2002). Seit dieser Zeit sind eugenische Maßnahmen und die damit verbundene Selektion in Verruf geraten und wecken schmerzliche Erinnerungen (Graumann, 2001; NER, 2003). Im Zusammenhang mit der PID wird als Folge einer möglichen Einführung eine „neue Eugenik“ befürchtet (Graumann, 2001), denn es werde „die schiefe Bahn zur Eugenik beschritten“ und „zudem ein Tabu gebrochen, das nach den NS-Untaten errichtet wurde“ (Jachertz, 2000, S. A-507).

Mit der PID sei gemäß den Kritikern erstmals eine effektive positive Eugenik möglich, da Eltern diejenigen Embryonen auswählen können, die am ehesten ihren Vorstellungen entsprechen und die von ihnen gewünschten Eigenschaften aufweisen. Diese positive Auswahl zwischen mehreren gezielt für diesen Zweck erzeugten Embryonen grenze die PID

auch von der PND ab, die ihrerseits mit einer negativen Eugenik einhergehe (EK, 2002). Stollorz (2000) fasst die Befürchtungen einer schleichenden Ausweitung in Richtung einer eugenischen Selektion wie folgt zusammen:

Die gezielte Selektion im Labor beschwört die Gefahr einer nützlichen, schmerzlosen und effizienten „neuen Eugenik“ herauf, bei der Wünschbares zur Norm wird. Zwar versucht die Ärzteschaft, den eugenischen Geist der PID in die Flasche strenger Indikation zu sperren. Doch mit Ethikkommissionen allein lässt sich der Dammbuch kaum stoppen. Wer kann einerseits einem Paar mit der Erbkrankheit Huntington die PID verweigern? Wer will andererseits verhindern, dass Paare den Embryonencheck in Rahmen einer künstlichen Befruchtung als Qualitätskontrolle nutzen, um ihre Chance auf ein gesundes Kind zu erhöhen? Die PID öffnet die Tür zur schönen neuen Welt des Baby-TÜVs - wenn auch zunächst nur einen Spalt breit. (S. 2)

Die Befürchtungen einer schleichenden Ausweitung der Indikationen der PID hin zu einem „Designerbaby“ und einer neuen Eugenik sind, wenn das Verfahren erst einmal in unserer Gesellschaft etabliert ist, sicherlich nicht leicht von der Hand zu weisen. Eine Tendenz zu „Designerbabys“ lässt sich bereits in den USA feststellen, wo eine Art Samen- und Eizellenmarkt existiert, bei dem die Käufer den Spender nach ihren gewünschten Kriterien wie dem Geschlecht, dem Aussehen, der körperlichen Konstitution, der Gesundheit oder des Intelligenzquotienten auswählen können (NER, 2003). Die Furcht vor einer möglichen Ausweitung hin zu einem „Designerbaby“ ist zum derzeitigen Zeitpunkt allerdings unbegründet, da bestimmte Eigenschaften wie Augen- oder Haarfarbe, Körpergröße oder Intelligenz im Rahmen einer PID aufgrund ihrer multifaktoriellen Vererbung (noch) nicht diagnostiziert werden können (NER, 2003; Steinke & Rahner, 2009).

Gegner der PID wenden zudem ein, dass die Begriffe der „Qualitätskontrolle“ oder „Designerbabys“ irreführend seien, da das Hauptziel einer PID nicht in der Selektion des besten Embryos bestünde, sondern im Ausschluss schwerer Erbkrankheiten, um genetisch vorbelasteten Paaren den Wunsch nach einem (gesunden) Kind zu erfüllen (NER, 2003). Auch im Falle einer PID-HLA erfolgt die Selektion im Gegensatz zu nicht-krankheitsrelevanten Wunscheigenschaften (z. B. Geschlecht) aus medizinischen Gründen, da es um die Therapie einer schweren Krankheit geht (Rehmann-Sutter, 2007). Außerdem könnten klar formulierte gesetzliche Regelungen einen wirksamen Schutz gegen möglichen Missbrauch bieten (NER, 2003; NEK-CNE, 2007).

Vielfach wird von Gegnern der PID, vor allem von Seiten der Behindertenverbände, angeführt, dass die PID zu einer steigenden *Diskriminierung und einer Stigmatisierung von Menschen mit Behinderungen*, chronisch Kranken und deren Familien führe und mit einem Verlust an gesellschaftlicher Solidarität mit behinderten Menschen einhergehe (vgl. Arndt & Obe, 2001; Bundesvereinigung Lebenshilfe e.V., 2010; Graumann, 2001). Bei einer

Legalisierung der PID wären Menschen, die mit einer Krankheit oder Behinderung leben, deren Diagnose zulässig wäre, mit einem staatlich tolerierten „Lebensunwert-Urteil“ konfrontiert, dass ihre eigene Existenz radikal in Frage stelle (Schockenhoff, 2005). Betroffene würden sich nach eigenem Bekunden als „fahrlässig nicht verhütete Unfälle“ zu betrachten haben (NER, 2003, S. 83). Ihre Existenz könnte nunmehr vermieden und vielmehr noch gezielt verhindert werden, dies stelle eine Abwertung behinderten Lebens und eine Diskriminierung von chronisch kranken und behinderten Menschen dar (Bundesvereinigung Lebenshilfe e.V., 2010). Im Falle einer Zulassung der PID sei auch damit zu rechnen, dass ein behindertenfeindliches Gesellschaftsklima begünstigt werde, da der Gesetzgeber die Verhinderung der Geburt behinderter und chronisch kranker Menschen für rechtmäßig erklärt (Bioethikkommission im Bundeskanzleramt, 2004; Graumann, 2001).

Befürworter der PID hingegen argumentieren, dass das Verfahren ethisch vertretbar sei, da es zur *Vermeidung konkreten Leids* sowohl für die Eltern als auch für das zukünftige Kind genutzt werde und hierdurch keineswegs bereits geborene Menschen mit bestimmten Krankheiten oder Behinderungen diskriminiert werden oder deren Menschenwürde in Frage gestellt wird (Lanzerath, 2010; Schräer, 2009; Schuh, 2000). PID-Befürworter räumen jedoch durchaus ein, dass für Menschen mit einer Behinderung die Konfrontation mit der Praxis der vorgeburtlichen Selektion eine Kränkung (Stigmatisierung) darstellen kann. Für Birnbacher (1999) stellt das Kränkungsargument sogar das stärkste Gegenargument gegen die vorgeburtliche Selektion dar, denn es geht „ausschließlich um die faktischen Gefühle der Betroffenen, nicht darum, ob diese berechtigt oder unberechtigt sind“ (S. 13).

Allerdings existieren bislang keine empirischen Belege im Hinblick auf eine kausale Verknüpfung zwischen der Nutzung von PND bzw. PID und einer Zunahme der Stigmatisierung und Diskriminierung von Menschen mit Behinderungen (EK, 2002).

Beispielsweise untersuchte van den Daele (2005) konkret die Auswirkungen der vorgeburtlichen Diagnostik auf die Benachteiligung und Diskriminierung behinderter Menschen und konstatiert, dass die verfügbaren empirischen Indikatoren bisher keine Anhaltspunkte für eine steigende Diskriminierung geben. In seinen Ausführungen befasst er sich zwar mit den Auswirkungen der PND, seine Ergebnisse lassen sich aber auch auf die PID übertragen. Zunächst falle laut van den Daele (2005) das selektive Potential der PID bei 900 Fällen pro Jahr neben der PND mit etwa 12 000 Fällen kaum ins Gewicht. Ferner führt er an, dass von den in Deutschland etwa 1,5 Millionen registrierten Fällen schwerer Behinderungen weniger als 10 % genetischen Ursprungs sind, von denen wiederum nur ein Teil überhaupt vorgeburtlich diagnostizierbar ist. Des Weiteren verweist van den Daele (2005) darauf, dass es sich empirisch nicht nachweisen lässt, dass PND und PID ein behindertenfeindliches Klima erzeugen und folglich zu einer Diskriminierung von behinderten Men-

schen führen. Eher das Gegenteil sei laut ihm der Fall, denn seit Jahrzehnten werden vor allem aufgrund der erheblichen Fortschritte der Behindertenpolitik die Rechte von Menschen mit Behinderungen kontinuierlich ausgebaut und mit steigendem finanziellem und personellem Aufwand gefördert. Wie Ergebnisse der Umfrageforschung belegen, ist die Zustimmung zur sozialen Integration behinderter Menschen und die Bereitschaft, mit ihnen zusammenzuleben, in der Bevölkerung hoch und hat in den letzten 30 Jahren deutlich zugenommen (NER, 2003; van den Daele, 2005). Im selben Zeitraum konnte allerdings auch eine Abnahme der Akzeptanz von Trisomie 21 in der Bevölkerung festgestellt werden. Bei einer pränatalen Diagnose eines Down-Syndroms ist die Abtreibungsrate in Deutschland sehr hoch, über 90 % der Schwangerschaften werden vorzeitig abgebrochen, obwohl Menschen mit dieser chromosomalen Konstitution oftmals eine hohe Lebensqualität zukommt (Bioethikkommission beim Bundeskanzleramt, 2004; van den Daele, 2005). Wie der Nationale Ethikrat bemerkt, spreche jedoch das gleichzeitige Vorliegen beider Entwicklungen gegen die Annahme, „dass die Zulassung der Selektion vor der Geburt zu einer Diskriminierung nach der Geburt führe“ (NER, 2003, S. 140).

Darüber hinaus könnte eine nationale Etablierung des Verfahrens neue soziale Zwänge für Paare mit Kinderwunsch mit sich bringen (EK, 2002). Für werdende Eltern könnte ein *gesellschaftlicher Druck* entstehen, die Geburt von behinderten oder kranken Kindern zu verhindern und damit die Entwicklung einer Eugenik von „unten“ fördern (Böcher, 2004; Schräer, 2009). Im Falle der Geburt eines behinderten Kindes könnten Eltern in Rechtfertigungsnot geraten, warum sie sich trotz genetischer Vorbelastung gegen eine vorgeburtliche Gendiagnostik entschieden haben (Maio, 2001). Wie die Bundesvereinigung Lebenshilfe e.V. (2010) in ihrem Positionspapier zur PID anführt, ist es heute bereits Realität, dass Eltern behinderter Kinder oft darauf angesprochen werden, „ob sie „es“ denn nicht gewusst hätten, ob es denn nötig sei „so“ ein Kind zu bekommen“ (S. 3).

Zudem könnte auf die Eltern möglicher kranker und behinderter Kinder auch ein ökonomischer Druck ausgeübt werden, die PID zum „Wohl der Allgemeinheit“ zu nutzen, um der Gesellschaft die höheren Kosten für die Betreuung und Pflege eines behinderten oder kranken Kindes zu ersparen (Lanzerath, 2010). Ein weiteres denkbare Szenario bestünde in der Verweigerung von Hilfeleistungen für die Eltern behinderter Kinder, mit der Begründung, dass sie deren Geburt doch hätten verhindern können (NER, 2003). Auf diese Weise entstünden möglicherweise Mechanismen, welche die in Verbindung mit der PID häufig zitierte reproduktive Autonomie der Eltern nicht erhöhen würde, sondern im Gegenteil eher noch einschränken oder erschweren könnte (Lanzerath, 2010; Schräer, 2009). Laut Kritikern sei es deshalb für die Gesellschaft unverzichtbar, die reproduktive Autonomie der

Eltern zu begrenzen, „um auch für ihre schwächsten Glieder Räume für die Selbstentfaltung offen zu halten“ (Schräer, 2009, S. 158).

Kritiker befürchten, dass auch Eltern, die dringend einen Gewebespender für die Therapie ihres kranken Kindes benötigen, einem deutlich spürbaren *moralischen Druck* ausgesetzt sind, die Technik der PID-HLA in Anspruch nehmen zu müssen (NEK-CNE, 2007). In einer Situation, in der keine anderen realisierbaren Optionen zur Heilung des kranken Kindes verfügbar sind, kann von einer echten Entscheidungsfreiheit der betroffenen Eltern kaum gesprochen werden: Für die Eltern gibt es selbst bei geringen Erfolgsaussichten eigentlich keine andere Wahl als das technologische Angebot der PID zu akzeptieren (Rehmann-Sutter, 2007).

Allein die Verfügbarkeit des Verfahrens der PID-HLA kann zu einem technologischen Imperativ werden, der geradezu nötige, ihm gemäß zu verfahren (Kreß, 2009). Baetens et al. (2005) konnten beispielsweise feststellen, dass ein wichtiger Einflussfaktor auf die Entscheidung, eine PID-HLA durchführen zu lassen, für die Paare im *anticipated decision regret* bestand. Der Begriff des *anticipated decision regret* geht auf Tymstra (1989) zurück und beschreibt den psychologischen Mechanismus von Patienten, einem technologischen Heilungsangebot nur deshalb zuzustimmen, um später nicht das Gefühl zu haben, eine möglicherweise entscheidende Behandlungsmöglichkeit verpasst zu haben und dies gegebenenfalls zu bereuen.

Der technologische Imperativ bildet demnach die Grundlage des *anticipated decision regret*. Für Eltern, die sich in der prekären Entscheidungssituation für oder gegen eine PID-HLA befinden, ist es enorm wichtig, zu wissen, dass sie alles in ihrer Macht stehende versucht haben, um ihr krankes Kind zu retten, um später keine Reue über eine möglicherweise verpasste Heilungschance empfinden zu müssen, wie folgendes Zitat aus der Studie von Baetens et al. (2005) belegt:

the parents needed to know that they had done everything possible to save their sick daughter. If they had done everything within their power to save the child, they felt they would not blame them afterwards. (p. 157)

Dem kranken Kind gegenüber gibt es für die Eltern letztlich keine Möglichkeit, eine PID-HLA abzulehnen, ohne dabei die Pflichten der guten Elternschaft zu verletzen (NEK-CNE, 2007). Die Eltern werden sich moralisch geradezu verpflichtet fühlen, so zu handeln.

Kritiker weisen noch auf *denkbare Alternativen* zu einer PID-HLA hin (NEK-CNE, 2007). Eine PID und die Erzeugung eines „Retterkindes“ würden sich erübrigen, wenn die Spendernetze für Knochenmark und der öffentlichen Nabelschnurblutbanken weiter ausgebaut werden würden. Die breite Verfügbarkeit von genügend Fremd Spendern brächte eine Ent-

Individualisierung des Problems mit, die den betroffenen Eltern die Entscheidungslast für oder wider PID-HLA ersparen würde (Kreß, 2009). Denn je mehr potentielle Spender von Knochenmark oder Nabelschnurblut vorhanden sind, desto höher ist die Wahrscheinlichkeit für alle HLA-Typen ein entsprechend geeignetes Transplantat zu finden (Rehmann-Sutter, 2007). Die bei einer Fremdspende geringfügig kleinere Erfolgswahrscheinlichkeit könnte zudem die mit einer PID-HLA verbundene Zeitverzögerung durch die IVF-Behandlungszyklen und die Schwangerschaft bei weitem aufwiegen. Ein weiterer Vorteil neben einer potentiellen Zeitersparnis bestünde darin, dass die Transplantation der hämatopoietischen Stammzellen zu einem therapeutisch optimalen Zeitpunkt durchgeführt werden könnte (Rehmann-Sutter, 2007).

Des Weiteren könnten die betroffenen Paare auch ein Kind auf natürlichem Wege zeugen und im späteren Verlauf der Schwangerschaft mit einer PND auf HLA-Merkmale testen. Diese Behandlungsalternative kann unter gewissen Umständen innerhalb eines limitierten Zeitraumes eine höhere Wahrscheinlichkeit, einen geeigneten Spender zu finden, versprechen (vgl. Shenfield et al., 2005).

Befürworter der PID argumentieren an dieser Stelle, dass die derzeitige gesellschaftliche Praxis der vorgeburtlichen Diagnostik *einen Wertungswiderspruch* beinhaltet, der aus ethisch und rechtlicher Sicht nicht zu rechtfertigen ist, da es für Paare mit einem bestimmten genetischen Risiko gesetzlich erlaubt ist, eine „Schwangerschaft auf Probe“ einzugehen und diese nach einer PND eventuell abubrechen. Es ist jedoch verboten, die genetische Untersuchung am Embryo vor seinem Transfer in die Gebärmutter durchzuführen (Haker, 2004; Maio, 2001; NEK-CNE, 2007; Schmidt, 2003). Bei einem Verbot der PID-HLA würden die betroffenen Eltern geradezu in die Alternative einer „Schwangerschaft auf Probe“, einer darauffolgenden PND und gegebenenfalls einer Abtreibung hineingezwungen (Rehmann-Sutter, 2007). Eine Abtreibung gehe jedoch mit einer größeren Belastung für die Frau einher und es wäre ethisch problematischer, einen Fetus in einem fortgeschrittenen Entwicklungsstadium, bei dem möglicherweise schon die Gehirnbildung und Stress- und Schmerzempfinden vorhanden sind, zu töten als einen Embryo im Frühstadium (Kreß, 2009; Rehmann-Sutter, 2007). Im Grunde stelle die für die PID notwendige „Zeugung auf Probe“ lediglich eine zeitliche Vorverlegung jener Selektion und Abtreibung dar, die ansonsten mit einer PND mehrere Wochen oder Monate später ohnehin stattfinden würde. Die PID könnte den Eltern und dem vorgeburtlichen Kind eine problematische, tragische und überaus belastende Abtreibung ersparen und stelle insofern das „kleinere Übel“ dar und wäre als solches tolerierbar (Kreß, 2009).

Der Nationale Ethikrat (2003) befürchtet als soziokulturelle Konsequenz einer Zulassung der PID auch *mögliche negative Auswirkungen auf das Menschenbild und das Selbstver-*

ständnis des Menschen. Die medizinische Diagnostik am Embryo und die Beurteilung seines moralischen Status berühren das unserer Verfassung zugrunde liegende Menschenbild und das Verständnis von Menschenwürde zutiefst (vgl. BÄK, 2000; Kreß, 2009). Da jeder Mensch selbst einmal ein Embryo war und auch alle zukünftig geborenen Menschen einmal Embryonen gewesen sein werden, wirft die PID generell die Frage nach dem Menschen und den Folgen für unser Menschenbild auf. Der Nationale Ethikrat gibt in diesem Zusammenhang zu bedenken, dass die assistierte Reproduktionsmedizin ein ganz neues Feld eröffnet hat,

denn sie verlegt die Zeugung neuen Lebens aus dem intimsten Bereich zweier Partner in das Labor, in dem ein Dritter die Befruchtung vornimmt. Aus einem der menschlichen Einflussnahme auf die Beschaffenheit des neuen Lebens weitgehend entzogenen Vorgang wird so ein medizinisch-technischer Vorgang. Aus Zeugung wird sozusagen Erzeugung. (NER, 2003, S. 99)

Diese Technisierung und Verdinglichung von Zeugung, Fortpflanzung und Schwangerschaft ist auch mit einem Wandel der Rolle des Arztes im Hinblick auf sein Handlungsfeld und seine Handlungsziele verbunden (Lanzerath, 2010). Der Arzt wird zum Dienstleister, bei dessen Tun nicht mehr die Heilung oder Linderung von Krankheiten im Vordergrund steht, sondern „Maßnahmen zur Produktionsbeschleunigung und zur Produktionssteigerung, zur Produktkontrolle, zur Aussonderung mangelhafter Erzeugnisse und zur Produktoptimierung sowie zur Vernichtung überschüssiger Produkte oder zu ihrer Aufbewahrung in als Biobanken bezeichneten Ersatzteillager“ (NER, 2003, S. 100).

Der Nationale Ethikrat befürchtet bei einer Zulassung der PID eine drohende Änderung des Menschenbildes, da menschliches Leben verdinglicht wird, der Unterschied zwischen Person und Sache undeutlich wird und sich auch schrittweise entsprechende Marktstrukturen entwickeln könnten (NER, 2003). Des Weiteren führt der Nationale Ethikrat mit Verweis auf Jürgen Habermas an, dass Menschen „sich nicht mehr als Freie, Gleiche und Ebenbürtige verstehen können, wenn die mit ihrer Herkunft verbundenen genetischen Merkmale und Eigenschaften nicht länger unverfügbar blieben, sondern fremder Verfügung und Planung unterlägen“ (NER, 2003, S. 100). Durch die PID wird der Mensch als Person infrage gestellt und zur prüfaren Ware gemacht, deren Eigenschaften fremdbestimmt festgelegt würden und sich nicht naturwüchsig ergeben: Eltern können mittels PID über die genetische Beschaffenheit und das Geboren- oder Nichtgeborenwerden ihrer Kinder nach ihren subjektiven Vorstellungen entscheiden. Dies sei laut Kritikern nicht mit dem Selbstverständnis des Menschen zu vereinbaren und lasse gravierende Auswirkungen auf das menschliche Selbstverständnis und das dem Achtungsgebot der Menschenwürde zugrunde liegende Menschenbild erwarten (NER, 2003).

Als weiteres Argument wird von PID-Befürwortern oft die *Gefahr einer Zunahme des sog. „PID-Tourismus“* ins benachbarte Ausland angeführt (NER, 2003). Betroffene Paare werden aufgrund des in Deutschland existierenden Verbots der PID lediglich dazu veranlasst, das Diagnoseverfahren schließlich im Ausland praktizieren. Das Tourismusargument überzeugt jedoch nicht, wollte man es anerkennen, müsste man zur Verhinderung jedes Ausweichens ins Ausland in Deutschland immer bereit sein, die weltweit „liberalste“ Regelung zu übernehmen, die in irgendeinem Land gilt (NER, 2003). Die Bioethikkommission des Bundeskanzleramtes (2004) führt sogar an, dass der Hinweis auf den „PID-Tourismus“ „als solcher noch kein ethisches Argument ist, sondern bloß der Verweis auf die nationale Begrenztheit eines rechtlichen Verbotes“ (S. 17). Nach inoffiziellen Schätzungen lassen jährlich etwa 50-100 Paare aus Deutschland eine PID im Ausland durchführen (NER, 2003). Ob mit einer Zunahme des PID-Tourismus deutscher Paare gerechnet werden kann, ist jedoch fraglich. Die Durchführung einer PID im Ausland ist sehr kostenintensiv, da zusätzlich zum teuren PID-Behandlungszyklus auch Verpflegung, Anreise und Unterkunft bezahlt werden müssen. Daher können sich nur ökonomisch besser gestellte Personen eine PID im Ausland leisten (Franklin & Roberts, 2006).

Insgesamt betrachtet lässt sich festhalten, dass in der ethischen Diskussion bezüglich der Legitimität der PID-HLA Argumente gegeneinander stehen, die für jeden Einzelnen unterschiedlich überzeugend sind (vgl. Schräer, 2009).

2.4.3 Ethische Stellungnahmen zur PID-HLA

Bezüglich der Frage, ob eine PID zur Auswahl eines geeigneten Gewebespenders für ein krankes Geschwister zugelassen werden soll, empfehlen die Stellungnahmen des Nationalen Ethikrates in Deutschland und der Bioethikkommission beim Bundeskanzleramt in Österreich mehrheitlich eine Ablehnung dieses Anwendungsbereiches.

Die *Bioethikkommission beim Bundeskanzleramt* begründete ihre Entscheidung damit, dass das zweite Kind lediglich Mittel zum Zweck zur Heilung des ersten Kindes ist. Diese ethisch nicht zu rechtfertigende Instrumentalisierung eines Menschen laufe dem Kindeswohl des zweiten Kindes zuwider (Bioethikkommission beim Bundeskanzleramt, 2004).

Der *Deutsche Ethikrat* sprach sich in seinen Stellungnahmen aus den Jahren 2003 und 2011 für eine eng begrenzte Zulassung der PID für genetisch vorbelastete Paare aus (siehe Kapitel 2.3.2) und lehnte die PID-HLA ab. Zugleich wies der Ethikrat aber daraufhin, dass es weiterer Erörterungen in der Gesellschaft bezüglich dieses Anwendungsbereiches bedarf

und die Gesichtspunkte der Instrumentalisierung der letztlich verworfenen Embryonen mit der Lebensrettung eines konkret hilfsbedürftigen Menschen abzuwägen sind (NER, 2003).

Die *Nationale Ethikkommission im Bereich der Humanmedizin* (nachfolgend abgekürzt mit: NEK-CNE) in der Schweiz hat im Jahr 2005 und 2007 zu den ethischen Aspekten der PID-HLA zur Auswahl eines Gewebespenders Stellung genommen. Im Verlauf der 2 Jahre hat sich die Haltung der Kommission bezüglich der PID-HLA differenziert. Im Jahr 2005 wurde der Einsatz der PID-HLA noch sehr eindeutig ohne Gegenstimmen abgelehnt (NEK-CNE, 2005). 2 Jahre später ist die Meinung der Kommission nach vertiefter Beschäftigung etwa hälftig geteilt, d. h. ein Teil der Kommission hat sich doch hinter diese Anwendungsmöglichkeit gestellt. Während dieser 2 Jahre hat die Kommission die Argumente pro und contra PID-HLA klarer herausarbeiten können.

In folgender Tabelle 2 werden die zwei etwa gleich stark vertretenen Positionen der Kommission bezüglich der ethischen Einschätzung der PID-HLA zusammenfassend gegenübergestellt (NEK-CNE, 2007). Ein Teil der Kommission lehnt die PID-HLA ab, da Abgrenzungsprobleme bestehen, etwa in Bezug auf die zu spendenden Zellarten und den Empfängerkreis und befürchtet eine Instrumentalisierung des „Retterkindes“ mit möglichen psycho-sozialen Risiken. Für den anderen Teil der Kommission wiegt die Rettung des Lebens dagegen schwerer als das „Eingehen gewisser Risiken psychosozialer Natur“. Außerdem könnten Eingrenzungen gesetzlich geregelt werden.

Die Vertreter beider Gruppen gelangen jedoch zudem zum Schluss, dass Eltern, die aufgrund des hierzulande geltenden Verbotes für die PID ins Ausland reisen, aus nachvollziehbaren und ehrenvollen ethischen Motiven handeln und somit keine moralischen oder ethischen Vorwürfe verdienen. Auch der Teil der Kommission, der die Legalisierung von PID zur Gewebetypisierung in der Schweiz ablehnt, begründet dies mit sozialetischen Überlegungen und stellt die individuelle Entscheidung der Eltern nicht in Frage.

2.5 Empirische Studien zu Wissen und Einstellungen zur PID in der BRD

Auch wenn sich mittlerweile zahlreiche Publikationen mit den medizinisch-naturwissenschaftlichen, rechtlichen und ethischen Aspekten der PID befassen, liegen bislang nur wenig deutschlandrepräsentative Studien zur Akzeptanz, den Einstellungen und dem Wissen zu diesem gendiagnostischen Verfahren vor.

Das Wissen und die Einstellungen zur PID in der deutschen Allgemeinbevölkerung wurden in einer im Jahr 2003 durchgeführten Studie erhoben (Brähler & Stöbel-Richter, 2004).

Tabelle 2: Positionen der Nationalen Ethikkommission der Schweiz bezüglich der ethischen Einschätzung der PID-HLA (NEK-CNE, 2007, S. 15-16).

<p>Ein Teil der Kommission lehnt PID-HLA aus folgenden ethischen Gründen ab:</p>	<p>Ein Teil der Kommission hält PID-HLA für ethisch vertretbar. Die Gründe dafür sind insbesondere:</p>
<ul style="list-style-type: none"> • Es ist schwierig, hämatopoietische Stammzellen gegenüber anderen Zelltypen, Gewebearten oder nachwachsenden Organen abzugrenzen. Zudem ist schwer zu begründen, weshalb nur eine Zellart auf diese Weise gespendet werden darf; die Praxis könnte sich ausweiten. • Es ist schwierig, den Empfängerkreis abzugrenzen; es ist schwer zu begründen, weshalb nicht auch andere enge Familienangehörige als Geschwister Empfänger sein könnten. • Es bestehen <i>psycho-soziale Risiken für das Spenderkind</i>, sowohl wenn die Behandlung gelingt als auch wenn sie misslingt. • Das Verfahren stellt eine <i>Instrumentalisierung des „Retterkindes“</i> dar. • Es könnte ein <i>moralischer Druck auf den Eltern</i> lasten, ein „Retterkind“ als ultima ratio zu erzeugen; es gibt für die Eltern dem kranken Kind gegenüber praktisch keine Möglichkeit abzulehnen, ohne die Pflichten der guten Elternschaft zu verletzen. Der Entscheid der Eltern lässt sich nachvollziehen. Das heißt aber nicht, dass man die äußeren Umstände nicht in Frage stellen kann, die den Entscheid als fürsorglichen Akt provozieren. • Es sind <i>Alternativen denkbar</i>: Ausweitung des Knochenmarkspendenetzes, Aufbau eines Systems öffentlicher Nabelschnurblutbanken. 	<ul style="list-style-type: none"> • Die <i>Rettung eines Lebens</i> wiegt schwerer als das Eingehen gewisser Risiken psychosozialer Natur. • Die Risiken sind nicht grundsätzlich andere als diejenigen, welche Eltern bei einer „natürlichen“ Fortpflanzung fraglos in Kauf nehmen. • Gesetze können die Eingrenzung bezüglich der zu spendenden Zellarten und bezüglich des Empfängerkreises gewährleisten.

Die untersuchte Stichprobe ($n = 2110$) setzte sich aus 929 Männern und 1181 Frauen im Alter von 18 bis 50 Jahren zusammen.

Wissen über PID. Insgesamt betrachtet fiel das Wissen über die PID in der deutschen Allgemeinbevölkerung sehr gering aus. Lediglich 30 % der Befragten wussten bereits etwas über die PID, 60 % hatten noch nie etwas von der Thematik gehört, 10 % konnten diesbezüglich keine Angabe machen. Der Bekanntheitsgrad der PID wird jedoch von Variablen wie dem Geschlecht, Alter, Bildung sowie der Stärke des aktuellen Kinderwunsches beeinflusst. Frauen, Personen im Alter von 31 bis 40 Jahren und Befragten mit Studienabschluss war die Thematik der PID häufiger bekannt als Männern, Befragten im Alter von 18 bis 30 und 41 bis 50 Jahren und Personen ohne Studienabschluss. Zudem zeigte sich, dass die Befragten umso häufiger schon einmal etwas über die PID gehört, gesehen oder gelesen hatten, je stärker der aktuelle Kinderwunsch bei diesen Personen ausgeprägt war. Von den Studienteilnehmern, die angaben, bereits etwas über die PID erfahren zu haben, schätzten lediglich 13 % ihr Wissen als eher gut bis sehr gut ein, 47 % als mittelmäßig und 40 % als eher schlecht bis sehr schlecht.

Interesse am Thema PID. Die überwiegende Mehrheit von 69 % der Teilnehmer an der Umfrage gab an, wenig bzw. gar kein Interesse an der Thematik zu haben. 25 % der Personen berichteten ein mittelmäßiges Interesse, lediglich 6 % der Befragten besaß ein starkes bzw. sehr starkes Interesse. Das Interesse am Thema PID hing allerdings vom Geschlecht, dem Alter und dem Bildungsabschluss ab. Frauen, Befragte im Alter von 18 bis 40 Jahren und Personen mit Studienabschluss berichteten ein höheres Interesse am Thema PID im Vergleich zu Männern, Befragten im Alter von 41 bis 50 Jahren und Personen ohne Studienabschluss.

Wissen zu diagnostischen Möglichkeiten von PID. Die Ergebnisse zu den eingeschätzten diagnostischen Möglichkeiten der PID zeigen, dass die Reichweite des Verfahrens im Hinblick auf die Diagnose von Krankheiten und Eigenschaften von den Befragten deutlich überschätzt wird. 79 % der Befragten gaben die Diagnose von schweren geistigen und körperlichen Behinderungen als Einsatzmöglichkeit der PID an. 54 % der Befragten nahmen an, dass mit Hilfe einer PID alle Arten von Krankheiten und Beeinträchtigungen diagnostiziert werden können. 23 % waren sogar der Meinung, dass zukünftige Körpermerkmale, wie Größe, Augen- und Haarfarbe feststellbar sind, 16 % hielten auch die Diagnose von Charaktereigenschaften, wie z. B. Aggressivität, für möglich. 41 % nahmen an, dass eine PID zur Diagnose des Geschlechts einsetzbar sei. Diese Ergebnisse zu den vermuteten Einsatzmöglichkeiten der PID verdeutlichen, dass kein exaktes Wissen bei den Befragten vorhanden ist, welche menschlichen Merkmale überhaupt eindeutig genetisch bestimmbar sind und folglich durch eine PID diagnostizierbar wären.

Befürwortung einer potentiellen Zulassung der PID vs. eigene Inanspruchnahme von PID in Bezug auf unterschiedliche Anwendungsbereiche. Bei der Frage, welche einzelnen Anwendungsindikationen der PID in Deutschland zulässig sein sollten, sprechen sich die Umfrageteilnehmer für eine sehr breite Zulassung des Verfahrens aus. Der großen Akzeptanz bezüglich einer potentiellen Zulassung der PID steht jedoch eine geringer ausgeprägte potentiell eigene Inanspruchnahme einer PID gegenüber. Beispielsweise befürworteten 70 % der Befragten häufiger eine Zulassung der PID zur Diagnose des Down-Syndroms in Deutschland, als sie sie selber in Anspruch nehmen würden (59 %). Die vorgefundene Diskrepanz zwischen der Zustimmung zur allgemeinen Nutzung und der antizipierten eigenen Inanspruchnahme der PID konnte auch schon in anderen medizinethischen Themenfelder wie beispielsweise der aktiven Sterbehilfe in der Bevölkerung festgestellt werden (vgl. Schröder, Schmutzer, Klaiberg & Brähler, 2003).

Zulassungskriterien von PID. Bezüglich der Frage nach einer allgemeinen Anwendungsindikation für eine Zulassung der PID in Deutschland, befürwortete die Mehrheit von 61 % der befragten Personen die Anwendung bei einem Verdacht auf eine spezifische Erkrankung, 31 % bei generellen Überprüfungen von genetischen Störungen. Lediglich 4 % der Befragten würden die Auswahl von Merkmalen ohne Krankheitsbezug befürworten, 24 % lehnen generell jegliche Anwendung der PID ab.

Mit der PID assoziierte Gedanken und Gefühle. Mit der PID verbinden die Befragten Gedanken und Gefühle wie „Ambivalenz“ (45 %), „Unsicherheit“ (44 %), „Hoffnung“ (31 %), „Angst“ (29 %) und „Zuversicht“ (22 %). Lediglich 9 % bzw. 6 % der Befragten gaben „Bevormundung“ bzw. „Freude“ an.

Beginn des menschlichen Lebens. Die Diskussion über die Zulassung der PID bringt auch die Frage nach dem Beginn des menschlichen Lebens mit sich. 30 % der Befragten legten den Beginn des menschlichen Lebens mit dem Zeitpunkt der Einnistung der befruchteten Eizelle in die Gebärmutter (Nidation) fest. 25 % waren der Meinung, dass menschliches Leben mit der Verschmelzung von Ei- und Samenzelle beginnt. 23 % der Studienteilnehmer gaben als relevanten Zeitpunkt den 2. Schwangerschaftsmonat, wenn beim Embryo die wesentlichen Organe ausgebildet sind, an. Für 4 % beginnt menschliches Leben erst mit der Geburt.

Zusammenfassend betrachtet lässt sich aus der vorgestellten Studie festhalten, dass das Wissen in der Allgemeinbevölkerung über die PID und deren diagnostischen Möglichkeiten sowie das Interesse am Thema PID äußerst gering ausfiel. Das Meinungsbild in der deutschen Bevölkerung zur Legalisierung der PID ist recht liberal: Mehrheitlich wird eher eine Legalisierung als ein Verbot der PID bevorzugt.

3 Theoretische Grundlagen zur Entscheidungsfindung und Argumentation im Bereich Medizinethik

In diesem Kapitel wird aus pädagogisch- und kognitionspsychologischer Perspektive auf die für diese Arbeit relevanten theoretischen Grundlagen zur Entscheidungsfindung und Argumentation im Bereich Medizinethik eingegangen.

Aus pädagogisch-psychologischer Sicht stellen ethische Dilemmata, wie sie im Zusammenhang mit der PID auftreten, schlecht strukturierte, komplexe Probleme ohne eindeutige Lösung dar (Jonassen, 2000; vgl. *socio-scientific issues*, Sadler & Zeidler, 2005). Als Erfolgskriterium zur Lösung solch komplexer Problemstellungen nennt Jonassen (2000) die Artikulation einer Präferenz mit einer gewissen Rechtfertigung. Die bei der Problemlösung ablaufenden Denk- und Entscheidungsprozesse können durch den Prozess des informellen Schlussfolgerns (*informal reasoning*) beschrieben werden. *Informal reasoning* kann durch Argumentation erfasst werden (Zohar & Nemet, 2002).

Im Folgenden wird zuerst eine Einführung in die Grundlagen der Medizinethik gegeben (Kapitel 3.1). Im Anschluss wird dann auf die Entscheidungsfindung und Argumentation im Bereich Medizinethik eingegangen (Kapitel 3.2) und es werden ausgewählte empirische Befunde zur Argumentationsqualität und zu Argumentationsmustern in Bezug auf kontroverse Wissenschaftsthemen präsentiert (Kapitel 3.3). Abschließend werden potentielle Einflussfaktoren auf die Entscheidungsfindung und Argumentation vorgestellt, im Speziellen das themenspezifische, inhaltsbezogene Vorwissen (Kapitel 3.4).

3.1 Einführung in die Medizinethik

Als Ethik (griech. *ethos* = Sitte) wird im Allgemeinen die „methodisch geleitete Reflexion auf die das menschliche Handeln bestimmenden moralischen Vorstellungen“ bezeichnet (Honnefelder, Korff & Kreß, 1998, S. 654). Die Ethik ist die Theorie von Moral und untersucht grundsätzlich Fragen nach dem moralisch richtigen Handeln des Menschen (Bögeholz et al., 2004; Höhle & Lude, 2004). Dabei werden nachvollziehbare Begründungen angestrebt, warum welches Tun erlaubt, geboten oder verboten ist.

Unter Moral werden individuelle oder kollektive Vorstellungen und Überzeugungen hinsichtlich gebotener, d. h. moralisch guten Handlungsweisen und unerlaubter, d. h. moralisch schlechten Handlungsweisen verstanden (Höhle & Lude, 2004; Pfeifer, 2003). Moral betrachtet Kohlberg (1964) daher als eine Menge von rationalen Prinzipien für Verhaltens-

entscheidungen. Dementsprechend geht es in der Ethik nicht darum, Moral zu predigen, sondern sie kritisch zu analysieren, zu reflektieren und rational zu begründen (Höble & Lude, 2004).

Die angewandte Ethik setzt sich mit ethischen Problemstellungen in allgemeinen Praxisfeldern wie beispielsweise der Bio- und Medizinethik auseinander und analysiert die grundlegenden Argumentationsstrukturen (Pfeifer, 2003).

Bioethik beschäftigt sich nach Höble & Lude (2004) „mit der kritischen Reflexion und ethischen Bewertung innovativer Techniken und Praktiken, die in der Gesellschaft kontrovers diskutiert werden“ (S. 24-25) und „den verantwortlichen Umgang des Menschen mit Leben (griech. *bios*) betreffen“ (S. 23). Die Bioethik schließt dabei Medizin-, Tier- und Umweltethik mit ein (vgl. Korff, Beck & Miktat, 2000). Zu den aktuellen Konfliktfeldern in der Tier- und Umweltethik zählen beispielsweise Fragen des Natur-, Umwelt- und Tierschutzes wie der Artenschutz, die Ressourcennutzung von Wasser und Boden, Tierversuche oder die Herstellung transgener Tiere.

Die Medizinethik im Speziellen beschäftigt sich mit der kritischen Reflexion und ethischen Bewertung der Legitimität von Eingriffen in menschliche Zeugungs-, Lebens- und Sterbeprozesse (Pfeifer, 2009). Zu den ethischen Konfliktfeldern in der Medizin zählen beispielsweise neben der PID die embryonale Stammzelltherapie, das Klonen, der Schwangerschaftsabbruch sowie die aktive Sterbehilfe oder die Organspende.

Ein wesentliches Kennzeichen von bio- und medizinethischen Dilemmasituationen besteht darin, dass bei deren Bearbeitung und Lösung meist nur zwei Handlungsoptionen möglich sind, die in einer kategorialen „Ja/Nein-Entscheidung“ bzw. der Festlegung einer eigenen „Pro/Contra-Position“ münden (Eggert, 2008; Pfeifer, 2003). Krüger & Vogt (2007) definieren dementsprechend ein Dilemma (griech. *di-lemma* = Doppelsatz, Zwiegriff) als

Eine Entscheidungssituation zwischen zwei gleich unangenehmen Möglichkeiten eines Alternativsatzes. Im moralischen Dilemma stehen sich zwei schlüssige Positionen, die auf unterschiedlichen ethischen Werten (z. B. „Recht auf Unversehrtheit“ und „Förderung des Wohlergehens“) gründen, unvereinbar gegenüber. Eine Entscheidung führt unabhängig vom gefälltten Urteil zwangsläufig dazu, einen Wert zu verletzen. (S. 255)

Übertragen auf das in dieser Arbeit als Untersuchungsgegenstand fungierende Thema der PID zur Gewebespende bestehen zwei Handlungsoptionen: Die erste Möglichkeit besteht darin, per PID einen zur Gewebespende geeigneten Embryo zur möglichen Heilung des schwer erkrankten Kindes herzustellen und dadurch den Tod zahlreicher Embryonen, die als Gewebespende nicht in Frage kommen, in Kauf zu nehmen. Die zweite Option besteht darin, das Leben der Embryonen, welche die nicht erwünschte Merkmalskombination auf-

weisen, zu schützen und das kranke Kind seinem natürlichen Schicksal zu überlassen und auf konventionellem Wege zu behandeln, was allerdings den frühen Tod des erkrankten Kindes und eine enorme psychische Belastung der Eltern zur Folge haben könnte. Weitere Alternativen existieren in der Erzeugung eines geeigneten Spenders auf natürlichem Wege oder der Suche nach passenden Stammzellspendern über Knochenmark- oder Nabelschnurblutdatenbanken.

Jedes moralische Dilemma beruht auf einer Konfliktsituation zwischen zwei oder mehreren Werten (Pfeifer, 2009). Werte stellen bewusste oder unbewusste allgemeine Zielorientierungen für menschliches Handeln dar, die dem menschlichen Dasein Sinn und Richtung geben. Diese idealen Orientierungsmuster werden dabei von den Menschen in unterschiedlichem Maße angestrebt oder geschätzt. Als Beispiele für Wertkategorien können Recht auf Leben, Freiheit, Gesundheit oder Bildung, aber auch Pflichten gegenüber Familie, Patienten, Staat oder persönliche Ideale wie Hilfsbereitschaft angeführt werden. Normen hingegen sind konkrete Handlungsorientierungen oder Handlungsvorschriften, die ausdrücken, dass eine bestimmte Handlung geboten, erlaubt oder verboten ist (z. B. „Du sollst nicht töten!“, „Du sollst helfen!“). Normen haben demgemäß die Funktion, die allgemeinen Werte zu konkretisieren (Pfeifer, 2009).

In dem Fallbeispiel zur PID-HLA werden zunächst zwei zentrale ethische Werte durch die Handlungsoptionen berührt, nämlich das Recht auf Leben des kranken Kindes und das Recht auf Leben der erzeugten Embryonen. Entscheiden sich die Eltern für eine PID-HLA, weil sie das Leben ihres schwer erkrankten Kindes retten wollen, wird dadurch der ethische Wert „Recht auf Leidverminderung/Heilung“ berührt und geschützt. Gleichzeitig wird bei dieser Entscheidung jedoch das Recht auf Leben der Embryonen verletzt, die aufgrund der nicht passenden Gewebekompatibilität verworfen werden. Im Falle der PID zur Gewebespende geraten die betroffenen Eltern in die Situation, ein neues Kind zeugen zu müssen, um mit dessen Nabelschnurblut- oder Knochenmarkzellen ihrem schwer erkranktem Kind das Leben zu retten. Auf der einen Seite verspricht die Anwendung der PID-HLA die Heilung ihres Kindes, auf der anderen Seite muss dafür die Tötung zahlreicher Embryonen in Kauf genommen werden.

Hieraus resultiert der Wertekonflikt „Recht auf Leben des schwer erkrankten Kindes“ vs. „Recht auf Leben der Embryonen“. Die Eltern befinden sich demzufolge in einem moralischen Dilemma zwischen der Hilfe für ihr krankes Kind und der Verwerfung und dem Verbrauch zahlreicher Embryonen (Hennen & Sauter, 2004). Unabhängig davon, für welche der beiden Handlungsoptionen sich die Eltern entscheiden, wird stets ein Wert verletzt. Entscheiden sich die Eltern keine PID durchführen zu lassen, verletzen sie möglicherweise ihre Pflicht zur Leidverminderung, das Recht auf Leben sowie das Recht auf Leidvermin-

derung ihres bereits geborenen kranken Kindes, da es wahrscheinlich sehr früh sterben muss. Bei der Entscheidung für eine PID könnte das Recht auf Selbstbestimmung und die Würde des „Retterkindes“ berührt werden, da es ausschließlich zum Zwecke der Gewebespende erzeugt wird und die Stammzellen ohne seine Zustimmung entnommen und genutzt werden. Durch die Verwerfung und das Absterben lassen der überzähligen Embryonen wird zudem das Recht auf Leben, Unversehrtheit und Würde dieser Embryonen verletzt.

Die aus einem Dilemma resultierenden Wertkonflikte versucht man in sorgfältigen Güter- (Werte-)Abwägungen zu lösen (Pfeifer, 2009). Im Falle einer PID zur Erzeugung eines Gewebespenders muss die Möglichkeit, einem schwerkranken Kind zu helfen, explizit in einer Güterabwägung mitberücksichtigt werden (Devolder, 2005). Ein Grundproblem von Güterabwägungen besteht jedoch in der Frage der Kommensurabilität, d. h. der Vergleichbarkeit von Werten: Wie wägt man beispielsweise die Menschenwürde und das Lebensrecht eines überzähligen Embryos gegen die mögliche Lebensrettung eines schwer erkrankten Kindes ab (Pfeifer, 2009)?

Zur ethischen Beurteilung der zwei Handlungsoptionen eines moralischen Dilemmas lassen sich viele verschiedene Argumente anführen, die wiederum auf unterschiedlichen Werthaltungen basieren und grob zwei Ethiktraditionen zugeordnet werden können, die bereits Ende des 18. Jahrhunderts entstanden (vgl. Höhle, 2001b; Höhle & Lude, 2004):

1. Der *deontologischen Ethik* (griech. *to deon*, die Pflicht), die auch als Pflichtethik bezeichnet wird (Kant, 1785).
2. Der *utilitaristischen Ethik* (lat. *utilitaris*, der Nutzen), die auch Nutzen- oder Folgenethik genannt wird (Bentham, 1789; Mill, 1871).

Die deontologische Ethik beurteilt die moralische Richtigkeit einer Handlung danach, ob sie moralischen Prinzipien, Werten oder Pflichten folgt, ungeachtet der Konsequenzen, die durch die Handlung verursacht werden. Der deontologische Argumentationstyp beurteilt die Phase vor einer Handlung und stützt sich dabei auf absolut gesetzte Werte und Prinzipien, die als unumstößlich gelten: Oberste Maxime ist die Würde des Menschen, welche Wesensmerkmale wie Individualität, Freiheit und Unverletzlichkeit beinhaltet. Ziel jeglicher Handlung sollte es demnach sein, die Würde des Menschen (aber auch anderer Lebewesen oder der Natur) und die ihnen zugesprochenen Rechte, wie beispielsweise das Recht auf Leben, zu wahren und zu schützen. Dementsprechend lehnen Vertreter des deontologischen Argumentationstyps technologische Eingriffe, welche die Menschenwürde verletzen, ab, ungeachtet dessen, ob die Folgen des Eingriffs dabei dem Wohl der Person dienen. Beispielsweise verbietet die deontologische Ethik die Nutzung von Embryonen zu Therapie Zwecken (Höhle, 2001b; Höhle & Lude, 2004).

Im Gegensatz zur deontologischen Ethik existieren für den Utilitarismus keine absoluten, allgemeingültigen Werte. Die utilitaristische Ethik beurteilt die Phase nach der Handlung und bestimmt die moralische Richtigkeit einer Handlung nach der Qualität der absehbaren Handlungsfolgen und nicht aus der Handlung selbst: Diejenige Handlung ist moralisch gut, deren Folgen für das Wohlergehen aller Betroffenen optimal sind (Höffe, 1992). Die Handlungsfolgen werden an ihrem Nutzen für alle von der Handlung betroffenen Menschen gemessen: Angestrebt wird das größtmögliche Glück für die größtmögliche Anzahl an Menschen. Für den Utilitarismus ist nicht das Glück oder Wohlergehen einzelner Individuen ausschlaggebend, sondern das allgemeine Wohlergehen und der kollektive Gesamtnutzen (Höffe, 1997). Als gute Folge einer Handlung gilt die Förderung des Wohlergehens in Form der Erfüllung menschlicher Interessen, Bedürfnisse und Wünsche wie beispielsweise gute Gesundheit.

Demzufolge befürworten Vertreter des utilitaristischen Argumentationstyps technologische Eingriffe, wie beispielsweise die Nutzung von Embryonen zu Therapiezwecken, wenn deren Folgen einen Nutzen für die Betroffenen haben, d. h. wenn dadurch das Leid der Mehrzahl der Beteiligten gemindert und das Glück bzw. Wohlergehen der Mehrzahl aller beteiligten Personen gesteigert wird (Höble, 2001b; Höble & Lude, 2004). Das Nutzen-Schaden-Kalkül im Hinblick auf die Handlungsfolgen bildet das zentrale Charakteristikum des utilitaristischen Argumentationstyps. Dies impliziert aber auch, dass ein utilitaristisches Urteil in der Regel zu einer Verletzung deontologischer Werte führt, wohingegen deontologische Urteile in der Regel die Folgen außer Acht lassen (Höble, 2001b).

Runtenberg (2001) weist daraufhin, dass in einer komplexen ethischen Argumentation beide Argumentationstypen nicht immer zweifelsfrei voneinander unterschieden werden können, häufig treten Mischformen auf. Beispielsweise verwenden Vertreter des deontologischen Argumentationstyp auch utilitaristische Argumente (und umgekehrt).

Höble (2003b) konnte in einer Studie, in der Schüler mit einem Dilemma zur embryonalen Stammzellforschung konfrontiert wurden, feststellen, dass zur Begründung des eigenen moralischen Urteils deontologische Argumente häufiger formuliert wurden als utilitaristische. Bei der Evaluation eines Lernprogramms von Götz (2001) zum Thema ethische Aspekte der Gentechnik zeigten sich dagegen konträre Befunde (siehe Kapitel 4.6): Die befragten Studenten führten zur Begründung ihrer Entscheidung häufiger utilitaristische Argumente an als deontologische. Allerdings hing die Verwendung der beiden Argumentationstypen stark vom jeweiligen präsentierten Dilemma, d. h. vom Thema bzw. Kontext ab.

Die in Kapitel 2.4.2 angeführten Argumente für und wider PID-HLA lassen sich auch den jeweiligen ethischen Grundpositionen der Deontologie und des Utilitarismus zuordnen. Zu den deontologisch geprägten Argumenten gegen die PID-HLA zählen beispielsweise die

Instrumentalisierung des „Retterkindes“ und die Tötung von Menschenleben bzw. die Verwerfung der überzähligen Embryonen. Als Argumente, die aus deontologischer Sicht für eine PID-HLA sprechen, können die Pflicht der Eltern zur Lebensrettung ihres kranken Kindes und das Recht auf Lebensrettung des letal erkrankten Kindes angeführt werden. Utilitaristische Argumente gegen die PID-HLA können beispielsweise die möglichen psychosozialen Folgen für das „Retterkind“ oder „slippery slope“-Argumente sein, welche mögliche gesellschaftliche Folgen der PID thematisieren (z. B. Missbrauchsgefahr). Utilitaristische Argumente für eine PID-HLA wären beispielsweise die Verminderung des Leids des letal erkrankten Kindes und eine geringere psychische Belastung der Eltern.

Zusammenfassend lässt sich festhalten, dass ein moralisches Dilemma stets eine Situation ist, in der keine Entscheidung unanfechtbar richtig ist: Es gibt kein „falsches“ oder „richtiges“ Urteil. Der Denkprozess darüber, ob eine Entscheidung richtig oder falsch ist, wird als moralisches Urteilen bezeichnet (Woolfolk, 2008). Da es in einem Dilemma keine klaren Lösungskriterien im Sinne von „falsch“ oder „richtig“ gibt, verlangt eine persönlich getroffene Entscheidung stets nach einer ethischen Rechtfertigung bzw. Begründung (Heitkamp, Borchardt & Witte, 2005). Ethisch nachdenken und argumentieren bedeutet demnach moralische Aussagen zu begründen (Bögeholz et al., 2004). Die Argumente, die zur Rechtfertigung einer moralischen Entscheidung angeführt werden, können dabei grob den beiden Ethiktraditionen Deontologie und Utilitarismus zugeordnet werden: Moralische Entscheidungen können demzufolge utilitaristisch oder deontologisch begründet werden.

In der Biologiedidaktik existieren zahlreiche Ansätze bzw. Modelle zur Förderung der moralischen Urteilsfähigkeit zu bio- oder medizinethischen Themen, die allesamt aus bzw. für die Unterrichtspraxis entwickelt wurden und in denen vor allem Dilemmasituationen bearbeitet werden (vgl. z. B. Bayrhuber, 1992; Dietrich, 2003; Dulitz & Kattmann, 1990; Höhle, 2001b; Meisert & Kierdorf, 2001).

Ein weiterer Schwerpunkt der biologiedidaktischen Lehr- und Lernforschung besteht momentan in der theoretischen Entwicklung und empirischen Überprüfung von Kompetenzstrukturmodellen, da der Kompetenzbereich Bewertung neben Fachwissen, Erkenntnisgewinn und Kommunikation einen expliziten Bestandteil der Bildungsstandards der KMK (2004) für das Fach Biologie der Mittelstufe bildet (Reitschert & Höhle, 2007). Ethische Bewertungskompetenz soll SchülerInnen dazu befähigen, an gesellschaftlichen, zum Teil kontrovers geführten Diskussionen zu bio- und medizinethischen Themen teilzunehmen, aktiv sich zu beteiligen, einen eigenen Standpunkt zu vertreten sowie ein reflektiertes und gut begründetes Urteil fällen zu können (vgl. KMK, 2004; Reitschert & Höhle, 2006). „Reflektieren bedeutet dabei zu hinterfragen, Positionen zu kontrastieren, Begriffe und Ideen zu analysieren“ (Bögeholz et al., 2004, S. 92). Reitschert (2009) entwickelte beispielsweise

im Rahmen ihrer Dissertation ein theoretisches Modell zur ethischen Bewertungskompetenz und überprüfte es empirisch anhand des bioethischen Themas der PID. Allerdings erscheinen die Ansätze zur Förderung der moralischen Urteilsfähigkeit und die Modelle zur ethischen Bewertungskompetenz im Hinblick auf das grundlegende Untersuchungsziel dieser Arbeit, der Unterstützung einer informierten und gut begründeten Entscheidung im musealen Lernsetting, wenig geeignet, da es sich bei Schulunterricht und der Entwicklung und Förderung von Kompetenzen grundsätzlich um langfristige Interventionen handelt. Multimediale Lernumgebungen in außerschulischen Lernfeldern wie dem Museum stellen hingegen in der Regel Interventionen von sehr kurzfristiger Dauer dar.

Im Rahmen der Arbeit wurde deshalb auf das aus dem angloamerikanischen Raum stammende pädagogisch-didaktische Konzept der *socio-scientific issues* zurückgegriffen, in denen Schüler „frei“ zu bio- und medizinethischen Themen argumentieren und die Argumentationen im Hinblick auf ihre Qualität und Struktur untersucht werden. Argumentieren, d. h. ein begründetes Urteil fällen zu können, stellt wiederum eine Teilkompetenz der ethischen Urteilskompetenz dar (Reitschert, 2009).

3.2 Entscheidungsfindung und Argumentation im Bereich Medizinethik

Bio- und medizinethische Themen werden in der englischsprachigen Literatur unter dem Begriff *socio-scientific issues* (nachfolgend abgekürzt mit: SSI) subsumiert. SSI lassen sich als kontroverse, gesellschaftlich, sozial und politisch relevante Themen oder Dilemmata beschreiben, die einen naturwissenschaftlichen Hintergrund aufweisen, jedoch nicht ausschließlich im Rückgriff auf naturwissenschaftliche Konzepte gelöst werden können.

SSI beinhalten neben gesellschaftlichen und politischen Komponenten häufig auch ethische und moralische Aspekte (vgl. Ratcliffe & Grace, 2003; Sadler, 2004b; Sadler, Barab & Scott, 2007; Sadler & Zeidler, 2003).

Als Beispiele für SSI können Wissenschaftsthemen angeführt werden, die aktuell den Gegenstand zahlreicher öffentlicher und kontroverser Debatten bilden: Sie reichen von der PID, der Pränataldiagnostik, der Gentherapie, dem Klonen oder der embryonalen Stammzellforschung bis hin zu Themen wie der Nanotechnologie, der Kernenergie, dem globalen Klimawandel, Fragen des Tierrechts, der Umweltverschmutzung oder der nachhaltigen Entwicklung.

Aufgrund ihres kontroversen Charakters erfordern diese lebensweltlich relevanten Themen im Schnittbereich zwischen Naturwissenschaften, Technik und Gesellschaft das Herausbil-

den eines eigenen Standpunkts (Bell & Lederman, 2003). Bei der Entscheidungsfindung müssen zahlreiche Bezüge hergestellt und es muss zwischen naturwissenschaftlichen, gesellschaftlich-politischen sowie ökonomischen Aspekten abgewogen werden.

Die Fähigkeit, diese kontroversen und in der Öffentlichkeit brisant diskutierten Themen bewerten zu können (*socio-scientific decision-making competence*) und aktiv an gesellschaftlichen Diskursen teilhaben zu können, bildet einen integralen Bestandteil von naturwissenschaftlicher Grundbildung (*scientific literacy*) (vgl. Gräber & Nentwig, 2002; KMK, 2004; Kolstø, 2001; OECD, 1999; Sadler & Zeidler, 2005). Personen sollten als „mündige“ Bürger in die Lage versetzt werden, informierte und reflektierte Entscheidungen über Themen treffen zu können, die möglicherweise später ihr Leben betreffen könnten (Dawson & Venville, 2009; Osborne, Erduran & Simon, 2004).

Die Lösung und Entscheidungsfindung bezüglich dieser komplexen Problemstellungen gestaltet sich jedoch häufig als sehr schwierig, „because they are open-ended, ill-structured, debatable problems subject to multiple perspectives and solutions“ (Sadler & Zeidler, 2004, p. 72). Der Denktyp, der bei der Lösung dieser schlechtstrukturierten, multiperspektivischen Problemstellungen auftritt, wird in der englischsprachigen Literatur als *informal reasoning* bezeichnet und kann durch Argumentation erfasst werden (vgl. z. B. Kuhn, 1991; Means & Voss, 1996; Perkins, Farady & Bushey, 1991; Sadler, 2004a; Voss et al., 1986; Zohar & Nemet, 2002).

3.2.1 Entscheiden und informelles Schlussfolgern

Entscheiden bezeichnen Betsch, Funke & Plessner (2011) als „einen Denkprozess, bei dem die Person zwischen mindestens zwischen zwei Alternativen (Optionen) eine Wahl trifft“ (S. 3). Um zu einer Entscheidung zu gelangen, muss eine Entscheidungssituation zunächst identifiziert und durch eine Informationssuche mögliche Handlungsoptionen generiert werden. Die identifizierten Handlungsoptionen werden dann mithilfe der Verwendung zahlreicher Entscheidungsstrategien miteinander verglichen, bewertet und eine Handlungsalternative zur Umsetzung der Entscheidung ausgewählt und implementiert (vgl. Betsch & Haberstroh, 2005; Jungermann, Pfister & Fischer, 2005).

Denken definiert Shaw (1996) als: „the process of constructing and evaluating arguments“ (p. 52). Argumente bestehen aus Konklusionen (Schlussfolgerungen) und mehreren Aussagen, den Prämissen, die der Unterstützung und Begründung der Konklusion dienen, indem sie Gründe oder Belege für die Wahrheit derselben angeben (Astleitner, 1998; Halpren, 1989).

In der Psychologie wird Denken (*reasoning*) als Schlussfolgerungsprozess verstanden, der auf zwei Wegen erfolgen kann: deduktiv (*formal*) oder induktiv (*informal*). Deduktives Denken folgt den Regeln der formalen Logik und wird deshalb auch als logisches Schlussfolgern bezeichnet. Kennzeichnend für das deduktive Denken ist, dass aus einer oder mehreren wahren Prämissen mit Hilfe logischer Regeln wahre Konklusionen über einen Einzelfall abgeleitet werden, d. h. es werden Schlussfolgerungen vom Allgemeinen auf das Besondere gezogen (Beller & Spada, 1996; Mietzel, 2005).

Beim deduktiven Denken werden mentale Modelle (bildliche Vorstellungen) aus den Prämissen konstruiert und daraus mögliche Schlussfolgerungen abgeleitet (vgl. Theorie der mentalen Modelle; Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). Diese mentalen Modelle repräsentieren mögliche Gesamtsituationen, die durch die Prämissen beschrieben werden. Um die vorläufige Schlussfolgerung auf ihre Allgemeingültigkeit zu überprüfen, wird versucht, alternative mentale Modelle zu bilden, welche die Schlussfolgerung verifizieren oder falsifizieren. Gemäß dem Fall, dass kein falsifizierendes Modell generiert werden kann, wird die Schlussfolgerung als valide akzeptiert.

Im Gegensatz zur Deduktion gründet induktives Denken nicht auf den Regeln der Logik. Ein Charakteristikum induktiver Schlüsse ist, dass diese unter Unsicherheit erfolgen und nur wahrscheinliche, aber keine logisch sicheren Schlüsse zulassen. Beim induktiven Denken werden aus gegebenen Einzelaussagen allgemeine Aussagen bzw. Gesetzmäßigkeiten abgeleitet, d. h. es werden Schlussfolgerungen vom Besonderen auf das Allgemeine gezogen. Eine solche Generalisierung kann auch als Vorgang der Bildung von Hypothesen verstanden werden, die anschließend geprüft und verändert werden können (Beller & Spada, 1996; Mietzel, 2005).

Zohar & Nemet (2002) definieren *informal reasoning* im Kontext einer Studie zu moralischen Dilemmata im Bereich Humangenetik folgendermaßen:

It involves reasoning about causes and consequences and about advantages and disadvantages, or pros and cons, of particular propositions or decision alternatives. It underlies attitudes and opinions, involves ill-structured problems that have no definite solution, and often involves inductive (rather than deductive) reasoning problems. (p. 38)

Informal reasoning schließt sowohl deduktives als auch induktives Denken mit ein, wobei letzteres prävalenter ist (Green, 1994). Im Gegensatz zum deduktiven Schlussfolgern, bei welchem der Problemraum gut definiert ist, wird informelles Schlussfolgern typischerweise bei schlechtstrukturierten Problemen angewendet (Evans & Thompson, 2004). Informelles Schlussfolgern schließt die Entwicklung und Bewertung von Standpunkten zu komplexen Themen, zu denen keine eindeutigen Lösungen existieren, mit ein (Sadler, 2004a).

Bei Aufgaben, die informelles Schlussfolgern erfordern, werden in der Regel nicht immer alle Prämissen expliziert und dementsprechend können die Konklusionen der Argumente oft nur sehr schwer erkennbar sein. Darüber hinaus wird informelles Schlussfolgern häufig in Situationen angewendet, in denen sowohl Gründe für als auch gegen die Konklusion existieren, wie es bei der Entscheidungsfindung recht häufig der Fall ist (Shaw, 1996). Formelles und informelles Schlussfolgern teilen aber auch Gemeinsamkeiten, beide werden als ein rationaler Prozess der Entwicklung und Bewertung von Argumenten beschrieben (Kuhn, 1993).

Die Identifikation, Konstruktion und Evaluation von Argumenten bildet wiederum einen zentralen Bestandteil von kritischem Denken (Astleitner, 1998; Ennis, 1987). Unter kritischem Denken (*critical thinking*) versteht Facione (1990) die bewusste zielgerichtete, selbstregulative Urteilsbildung, die Interpretation, Analyse, Bewertung und Schlussfolgerung beinhaltet.

Dick (1991) hat anhand ausgewählter Ergebnisse empirischer Untersuchungen eine Taxonomie kritischen Denkens entwickelt und aufgestellt. Nach dieser Taxonomie besteht kritisches Denken aus fünf zentralen Bestandteilen bzw. Fähigkeiten und Dispositionen (Dick, 1991; Astleitner, 1998):

1. Dem Identifizieren der Bestandteile eines Arguments, wie beispielsweise aufgegriffene Themen, Schlussfolgerungen und deren Begründungen.
2. Der Analyse von Argumenten, d. h. die Offenlegung ihrer zugrundeliegenden Annahmen oder von vorhandenen Unklarheiten.
3. Der Berücksichtigung von äußeren Einflüssen, wie Werten (Auffassungen über Handlungen oder Einstellungen bezogen auf Menschen, Dinge oder Ziele) oder Autoritäten (übergeordnete Personen oder Institutionen).
4. Dem wissenschaftlich-analytischen Schlussfolgern, das Kausalitäten (ursächliche Zusammenhänge) erforscht und Daten auf der Basis statistischer Überprüfungen überprüft.
5. Dem logischen Schlussfolgern, bei dem Analogien, vor allem aber deduktiven und induktiven Schlüssen eine hohe Bedeutung zukommt.

Kritisches Denken umfasst Fähigkeiten und Dispositionen wie Argumente zu identifizieren, zu analysieren und zu evaluieren, die Glaubwürdigkeit einer Quelle und von Stellungnahmen zu beurteilen und diese gegebenenfalls in Frage zu stellen sowie das Erstellen und Bewerten von Deduktionen, Induktionen und Stellungnahmen, um letztlich begründete Schlussfolgerungen ziehen zu können (Dick, 1991; Ennis, 1987; Facione 1990).

Demzufolge kann kritisches Denken als eine notwendige Voraussetzung für gutes informelles Schlussfolgern erachtet werden. Gutes informelles Schlussfolgern zeichnet sich durch Kohärenz, interne Konsistenz und die Fähigkeit zur Betrachtung mehrerer Perspek-

tiven aus, d. h. dass auch Gegenargumente zur eigenen Entscheidung miteinbezogen werden, um zu einer Neubewertung seiner eigenen Entscheidung zu gelangen (Kuhn, 1991; Voss & Means, 1996).

Formelles und informelles Schlussfolgern wurden in der psychologischen Forschung lange Zeit isoliert voneinander untersucht, in jüngster Zeit wurde versucht, diese beiden Forschungsrichtungen miteinander zu verbinden (Galotti, 1989), indem beispielsweise die Theorie der mentalen Modelle (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) auch als Rahmenmodell zur Erklärung informellen Schlussfolgerns herangezogen wurde (Evans & Thompson, 2004).

3.2.2 Informelles Schlussfolgern aus Sicht der Zwei-Prozess-Theorien

In der Kognitions- und Sozialpsychologie werden zwei unterschiedliche kognitive Systeme für das menschliche Denken, Urteilen und Entscheiden für bedeutsam gehalten (vgl. Chaiken & Trope, 1999).

Nach dem Zwei-Prozess-Modell von Evans (2002, 2003) basiert menschliches Denken auf zwei getrennten, unterschiedlichen kognitiven Systemen, die er als System 1 (implizit) und System 2 (explizit) bezeichnet. Die Nutzung von System 1 erfolgt unbewusst, ist pragmatisch und kontextualisiert, wohingegen System 2 auf einer bewussten und kontrollierbaren Ebene operiert und logisch/abstraktes Denken miteinschließt. Die Denkprozesse in System 1 erfolgen schnell, parallel und automatisch, in System 2 finden dagegen langsame, sequentielle und bewusst kontrollierte Prozesse statt, die hypothetisches Denken erlauben und Aufmerksamkeitsressourcen benötigen, die jedoch aufgrund der limitierten Kapazität des Arbeitsgedächtnisses begrenzt sind. System 1 stützt sich auf bereits vorhandenes Vorwissen und persönliche Erfahrungen, weswegen die Nutzung von einfachen Entscheidungsregeln (Heuristiken) in der Regel auf System 1 zurückgeführt werden kann.

Um zu einer persönlichen Entscheidung oder einem Urteil in Bezug auf eine schlechtstrukturierte Problemstellung (z. B. SSI) zu gelangen, muss ein Lernender auf seine Fähigkeiten im informellen Schlussfolgern zurückgreifen. Die unbewusste Nutzung von System 1 ermöglicht es zunächst dem Lernenden auf der Basis seines Vorwissen und seiner vergangenen Erfahrungen schnell ein initiales mentales Modell in Bezug auf das Problem zu bilden. Der Lernende kann nun entweder ausschließlich auf der Basis von System 1 seine persönliche Entscheidung intuitiv treffen oder zusätzlich bewusst System 2 nutzen, um eine rationale Schlussfolgerung zu ziehen. Die bewusste Nutzung von System 2 erlaubt hypothetisches Denken, so dass das anfängliche mentale Modell revidiert werden kann und neue mentale Modelle konstruiert werden können (Wu & Tsai, 2007).

Aus Sicht der Zwei-Prozess-Modelle kann der Prozess des informellen Schlussfolgerns nach Wu & Tsai (2007, p. 1166) folgendermaßen modelliert werden (siehe Abbildung 2):

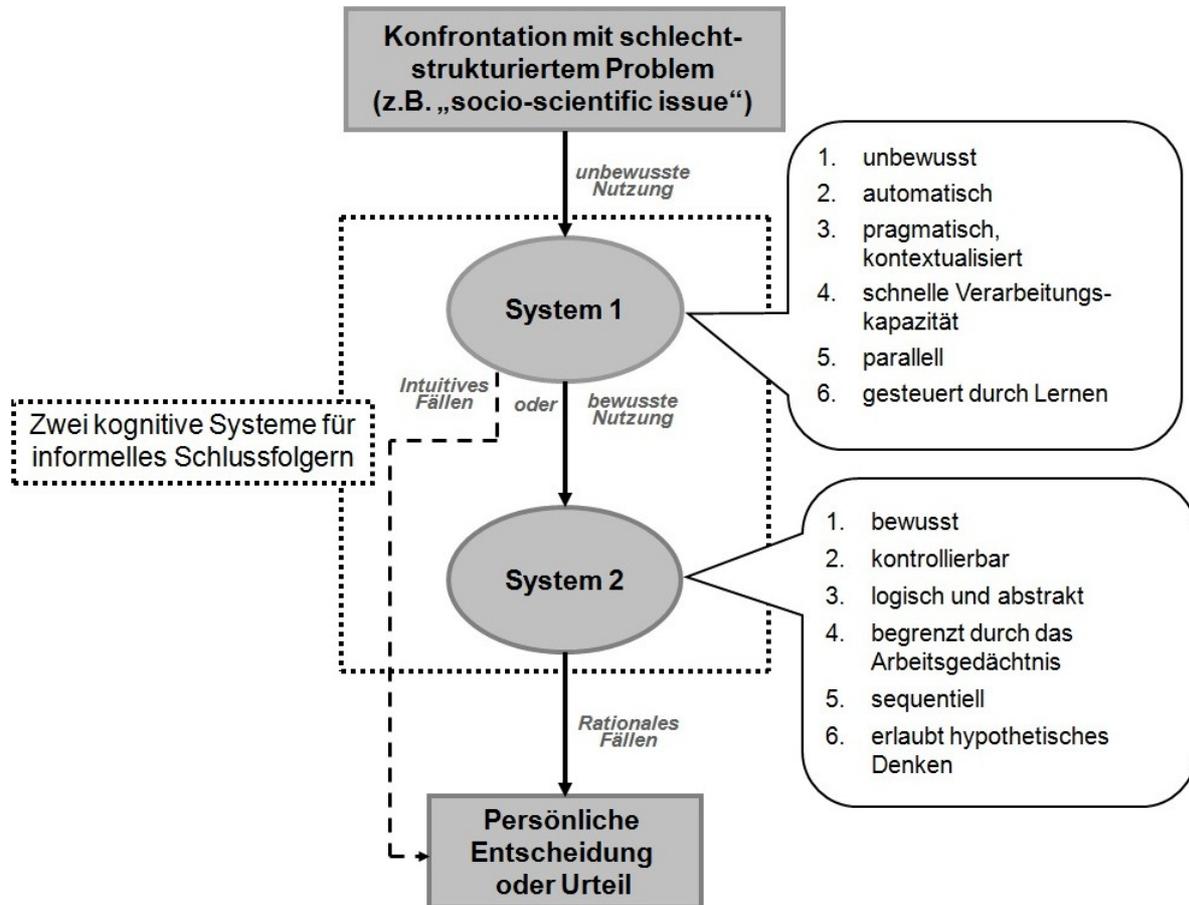


Abbildung 2: Der Prozess des informellen Schlussfolgerns aus Sicht der Zwei-Prozess-Theorien (nach Wu & Tsai, 2007, p. 1166).

Innerhalb der Kognitions- und Sozialpsychologie existieren zahlreiche dieser Zwei-Prozess-Modelle, für deren Bezeichnung jeweils verschiedene Begriffspaare verwendet werden, beispielsweise „peripher vs. zentral“ (Petty & Cacioppo, 1986), „heuristisch vs. systematisch“ (Chaiken, Liberman & Eagly, 1989), „impulsiv vs. reflektiv“ (Strack & Deutsch, 2004), „intuitiv vs. deliberativ“ (Hogarth, 2005), oder „intuitiv-experientiell - analytisch-rational“ (Epstein, Pacini, Denes-Raj & Heier, 1996).

Zwei besonders prominente Modelle stellen die Persuasionstheorien *Elaboration Likelihood Model* (ELM) (Petty & Cacioppo, 1986) und das *Heuristic-Systematic Model* (HSM) (Chaiken et al., 1989) dar. Persuasion (lat. *persuadere*, „überzeugen“, „überreden“) definiert Bohner (2003) als „Einstellungsbildung oder -änderung, gewöhnlich in Reaktion auf Argumente und/oder andere Informationen über das Einstellungsobjekt“ (S. 276). Eine Einstellung ist eine psychologische Tendenz, die dadurch zum Ausdruck gebracht wird, dass ein bestimmtes Objekt mit einem gewissen Grad an Zustimmung oder Ablehnung bewertet wird (Eagly & Chaiken, 1993).

Beide Theorien unterscheiden zwei idealtypische und qualitativ unterschiedliche Informationsverarbeitungsprozesse des kognitiven Systems voneinander: einen langsamen, mühevollen und kognitiv aufwändigen Prozess (zentrale bzw. systematische Verarbeitung) sowie einen schnellen, mühelosen und "oberflächlichen" Prozess (periphere bzw. heuristische Verarbeitung). Welcher Informationsverarbeitungsprozess eingeschlagen wird, hängt dabei von der Fähigkeit und Motivation des Rezipienten ab.

Die Fähigkeit zu einer elaborierten bzw. systematischen Verarbeitung von Informationen und Argumenten wird determiniert durch das themenspezifische Vorwissen (Wood, 1982; Wood, Kallgreen & Preisler, 1995), die Komplexität und Verständlichkeit der vorliegenden Informationen (Hafer, Reynolds & Obertynski, 1996) und die zur Verfügung stehenden kognitiven Kapazitäten (Petty, Wells & Brock, 1976). Die Motivation zu einer aufwändigen Verarbeitung ist dagegen abhängig von dem themenspezifischen Interesse bzw. dem persönlichen Bezug zum Thema (Involvement), der situativ angestrebten Urteilssicherheit (Chaiken et al., 1989) und individuellen Dispositionen wie beispielsweise dem Kognitionsbedürfnis, d. h. der Freude am Denken (*need for cognition*; Cacioppo & Petty, 1982) oder dem Bedürfnis nach Bewertung (*need to evaluate*; Jarvis & Petty, 1996).

Bei geringer Motivation und/oder niedriger Kapazität bestimmen leicht zu verarbeitende periphere bzw. heuristische Informationen (*cues*) die Informationsverarbeitung. Heuristische Cues können Merkmale des Kommunikators sein, wie beispielsweise der Expertisegrad, die Glaubwürdigkeit, Sympathie oder Attraktivität (Petty, Cacioppo & Goldman, 1981; Petty, Cacioppo & Schumann, 1983; Priester & Petty, 1995), aber auch Merkmale der Kommunikation wie beispielsweise die Modalität (Chaiken & Eagly, 1978, 1983) oder die Anzahl und Ausweitung der Argumente (Petty & Cacioppo, 1984; Wood et al., 1985). Die über den heuristischen Pfad gebildeten oder geänderten Einstellungen, Urteile oder Entscheidungen sind in der Regel relativ instabil.

Unter heuristischer Verarbeitung werden, um die Gültigkeit einer Schlussfolgerung zu beurteilen, weniger die Inhalte oder Argumente einer Kommunikation überprüft, sondern einfache Entscheidungsregeln angewendet (Chaiken et al., 1989):

When processing heuristically, people focus on that subset of available information that enables them to use simple inferential rules, schemata, or cognitive heuristics to formulate their judgments and decisions. (p. 213)

Heuristiken werden durch entsprechende Hinweisreize (*cues*) ausgelöst, die entweder dem Kommunikator oder der Kommunikation zugeordnet werden können. Als Beispiele für einfache Entscheidungsregeln, die sich auf den Kommunikator beziehen, können die Sympathieheuristik (*people agree with people they like*; Chaiken, 1980) oder die Expertenheuristik (*expert's statements can be trusted*; Chaiken et al., 1989) genannt werden. Aber auch

Aspekte der Kommunikation können Heuristiken aktivieren wie die Anzahl an Argumenten (*length implies strength*; Wood et al., 1985; *the more arguments the better*; Petty & Cacioppo, 1984) oder Mehrheitsmeinungen (Konsensinformationen, z. B. in Form von Umfrageergebnissen, *consensus implies correctness*; Axson, Yates & Chaiken, 1989; Erb, 1998; siehe Kapitel 4.4.3).

Eine systematische Verarbeitung erfordert im Gegensatz zur heuristischen Verarbeitung einen hohen kognitiven Aufwand, da entscheidungsrelevante Informationen oder Argumente detailliert und kritisch überprüft werden, indem u. a. Vorwissen aktiviert, Bezüge hergestellt und Bewertungen erstellt werden. Systematische Verarbeitung definieren Chaiken et al. (1989) als

a comprehensive, analytic orientation in which perceivers access and scrutinize all informational input for its relevance and importance to their judgments task, and integrate all useful information in forming their judgments. (p. 212)

Eine systematische Verarbeitung wird jedoch nur dann initiiert, wenn sowohl eine hohe Motivation als auch eine hohe Fähigkeit beim Rezipienten vorliegt. Die über den systematischen Weg gebildeten oder geänderten Einstellungen, Urteile oder Entscheidungen sind relativ stabil.

Wie Chaiken et al. (1989) betonen, können heuristische und systematische Prozesse auch parallel ablaufen. Häufig wird bei einer systematischen Verarbeitung die Wirkung heuristischer *cues* aufgehoben oder dramatisch abgeschwächt (*attenuation*). Heuristische und systematische Verarbeitung können jedoch auch voneinander unabhängige, d. h. additive Effekte auf die Urteilsbildung bzw. eine Entscheidung haben (*Additivität*). Eine weitere Möglichkeit des Zusammenwirkens beider Modi ist die verzerrte Verarbeitung von entscheidungsrelevanten Informationen und Argumenten.

Darüber hinaus postulieren Chaiken et al. (1989) drei Typen von Motivation, die bei der Verarbeitung von Informationen auftreten und zu kognitiv und motivational bedingten Verzerrungen führen können.

Menschen sind bestrebt zu einer objektiv richtigen und korrekten Einstellung zu gelangen (Richtigkeitsmotivation; *Accuracy-Motiv*). Personen mit einer hohen Richtigkeitsmotivation versuchen die Meinungen anderer Personen zu verstehen und suchen in einer balancierten oder sogar selbstkritischen Art und Weise nach Informationen (Lundgren & Prislín, 1998).

Menschen streben auch nach Konsistenz in der eigenen Wissensstruktur (vgl. Festinger, 1957) und haben das Bedürfnis, die eigene Einstellung gegen konträre Informationen zu schützen und zu verteidigen, indem sie systematisch standpunktunterstützende Informatio-

nen bevorzugen und standpunktwidrig sprechende Informationen vernachlässigen (Verteidigungsmotivation, Lundgren & Prislín, 1998; siehe Kapitel 3.2.3, *confirmation bias*: Bestätigungsfehler). Das *Defense-Motiv* wird insbesondere dann ausgelöst, wenn eine Einstellung wichtig und in das Selbstkonzept einer Person eingebunden ist (Chen & Chaiken, 1999):

Defense Motivation refers to the desire to hold attitudes and beliefs that are congruent with one's perceived material interest or existing self-definitional attitudes and beliefs. (p. 77)

Selbstwert-relevante Einstellungen und Überzeugungen können eigene Werte (z. B. Gerechtigkeit), die soziale Identität (z. B. Beruf) oder persönliche Merkmale (z. B. Intelligenz) sein.

Die Eindrucks motivation beschreibt hingegen das Bedürfnis, Einstellungen zu bilden, die auf soziale Akzeptanz stoßen. Das *Impression-Motiv* wird aktiviert, wenn soziale Aspekte einer Situation salient sind und Rezipienten mit ihrer Einstellung soziale Konsequenzen verbinden und nach sozialer Akzeptanz suchen. In diesem Fall kann ein *Konfirmations-bias* in Richtung der Meinung der anderen bei der Informationssuche auftreten (Lundgren & Prislín, 1998). Unabhängig davon, welches Motiv in der Persuasionssituation dominiert, kann die Informationsverarbeitung zur Bildung einer korrekten, selbstdienlichen und sozial akzeptierten Einstellung systematisch oder heuristisch erfolgen.

Von zentraler Bedeutung für das HSM ist auch das sog. *Suffizienz-Prinzip*, demzufolge eine motivationale Determinante der systematischen Verarbeitung von einstellungsrelevanten Informationen die Diskrepanz von angestrebter (erwünschter) und tatsächlicher Sicherheit in Bezug auf die Gültigkeit der eigenen Einstellung bildet (Chaiken et al., 1989). Gemäß dem Suffizienzprinzip steigt die Wahrscheinlichkeit einer systematischen Verarbeitung, wenn die Diskrepanz zwischen der angestrebten, also der subjektiv als hinlänglich empfundenen Sicherheit und der tatsächlichen, in einer bestimmten Situation empfundenen Sicherheit größer wird. Diese Diskrepanz kann beispielsweise auftreten, wenn Einflüsse wie der persönliche Bezug bzw. die persönliche Bedeutsamkeit eines Einstellungsobjektes zu einer Erhöhung der angestrebten Sicherheit oder zu einer Reduktion der tatsächlichen Sicherheit der Einstellung führen.

Die Zwei-Prozess-Theorien wurden neuerdings von Haidt (2001) auch auf moralisches Urteilsverhalten übertragen. Im Gegensatz zu den rein rationalistischen Ansätzen in der Moralpsychologie (z. B. Kohlberg, 1964, 1971) betont Haidt (2001) den Einfluss emotionaler und intuitiver Reaktionen auf das moralische Urteilsverhalten. Häufig reagieren Menschen auf moralisch relevante Sachverhalte mit spontanen Gefühlen wie beispielsweise Ablehnung oder Empörung, aber auch Zustimmung (Reitschert & Höhle, 2007).

Haidt (2001) unterscheidet deshalb in seinem *social intuitionist model* zwischen einem intuitiven und einem reflektierendem System und zwei parallel arbeitenden kognitiven Prozessen, die an der Entstehung moralischen Urteilens beteiligt sind: Automatische und unbewusste Verarbeitungsprozesse, die er als Intuitionen bezeichnet (*intuitiv*) und absichtliches, bewusstes Nachdenken (*reflektierend*).

Gemäß Haidt (2001) entstehen die meisten moralischen Urteile nicht durch rationale Denkprozesse, sondern sind das Ergebnis unbewusster, intuitiver Informationsverarbeitungsprozesse. Moralische Intuitionen und Emotionen in Form schneller und automatischer Evaluationen von moralischen Ereignissen als „gut vs. schlecht“ gehen dabei dem moralischen Urteil voraus und lösen es spontan aus. Das rationale, bewusste Denken findet oft erst im Nachhinein in Form einer Post-hoc-Rechtfertigung statt, in der nach Pro-Argumenten für das intuitiv bereits getroffene Urteil gesucht wird. Argumente, die der eigenen Entscheidung widersprechen, werden dabei typischerweise abgeschwächt oder ignoriert. Demzufolge wird die am Anfang intuitiv getroffene Entscheidung häufig beibehalten und die Entscheidung lediglich im Nachhinein gerechtfertigt (Haidt, 2001; Kagan, 1984). Moralische Entscheidungen werden somit oftmals auch spontan und intuitiv getroffen. In Bezug auf komplexe und moralische Situationen nimmt Intuition eine wichtige Funktion in Entscheidungsprozessen ein (vgl. auch Betsch, 2004; Epstein et al., 1996; Gigerenzer, 2007).

Allen Zwei-Prozess-Modellen der Informationsverarbeitung ist gemein, dass sie zwischen zwei unterschiedlichen Verarbeitungsmodi unterscheiden: einem schnellen, mühelosen und assoziativen Prozess sowie einem langsamen, mühevollen und systematischen Prozess. Erstgenannte Informationsverarbeitungsprozesse laufen unbewusst immer ab, wohingegen systematische und bewusste Verarbeitung von der Fähigkeit (z. B. themenspezifisches Vorwissen) und der Motivation (z. B. themenspezifisches Interesse) abhängt.

Die Moduswahl kann jedoch auch von individuellen Präferenzen für einen Modus abhängen, die relativ zeitstabile Vorlieben darstellen und als Motivation verstanden werden können, auf der Basis eines bestimmten Modus zu entscheiden (Betsch, 2004). Beispielsweise konnten Epstein et al. (1996) feststellen, dass Menschen sich darin unterscheiden, ob sie ihre Entscheidung bevorzugt aufgrund von Gefühlen (*intuitiv*) oder aufgrund von bewussten, planvollen Überlegungen (*reflektiv*) treffen. Betsch (2004) unterscheidet dementsprechend zwischen einer Präferenz für Deliberation und einer Präferenz für Intuition und entwickelte und validierte ein Messinstrument für diese zwei voneinander unabhängigen Dimensionen. Die Skala zur Präferenz für Deliberation erfasst individuelle Unterschiede in der Vorliebe für Entscheidungsstrategien, die auf Kognitionen basieren, die Skala zur Präferenz für Intuition hingegen affektbasierte Entscheidungsstrategien.

3.2.3 Argumentation

Eine Entscheidung sollte, unabhängig davon, welche Entscheidung letztendlich getroffen wurde, nachvollziehbar und gut begründet dargelegt werden (Sadler & Donnelly, 2006), d. h. die Lernenden sollten geeignete Argumente zur Bestimmung ihres Standpunktes formulieren können (Kuhn, 1991). Informierte, gut begründete und reflektierte ethische Entscheidungen müssen sich demnach auf Argumente stützen (Bögeholz et al., 2004).

Die häufige Nutzung der Begriffe Argument und Argumentation in dieser Arbeit erfordert zunächst eine Definition. Was ist ein Argument? Von den zahlreich in der Literatur vorhandenen Definitionen werden drei an dieser Stelle eingeführt. Für Habermas (1995) ist ein Argument „die Begründung, die uns motivieren soll, den Geltungsanspruch einer Behauptung oder eines Gebots bzw. einer Bewertung anzuerkennen“ (S. 15). Kuhn (1991) beschreibt ein Argument als „assertion with accompanying justification“ (p. 12), Means & Voss (1996) geben folgende Definition: „an argument is a conclusion supported by at least one reason“ (p. 141).

Gemäß diesen Definitionen bestehen Argumente stets aus Daten und Begründungen (Toulmin, 1958), die eine zuvor aufgestellte Behauptung, Entscheidung oder Schlussfolgerung rechtfertigen sollen (Erduran et al., 2004; Foong & Daniel, 2010). Argumentationen stellen sprachliche Äußerungen eines oder mehrerer miteinander verknüpfter Argumente dar und dienen dazu, „Behauptungen zu begründen oder Entscheidungen zu rechtfertigen“ (Bayer, 2007, S. 18).

Folgendes Beispiel illustriert ein Argument, das von einem Studenten im Rahmen dieser Untersuchung als Antwort auf die Frage: „Würden Sie an ihrer Stelle für eine Präimplantationsdiagnostik nach Amerika gehen?“ formuliert wurde:

Ich würde nach Amerika gehen, da ich denke, dass man alles Menschenmögliche tun sollte, um sein Kind, das man ja über alles liebt, zu heilen. [65b3_t1_Ja].

Die Konklusion dieses Arguments ist „ich würde nach Amerika gehen“. Es wird durch das Statement gerechtfertigt: „da ich denke, dass man alles Menschenmögliche tun sollte, um sein Kind, das man ja über alles liebt, zu heilen“. Dieses Statement dient demzufolge als Rechtfertigung oder Begründung für die Konklusion.

Die Grundstruktur eines Arguments kann der philosophischen, formalen Logik zufolge in einer syllogistischen Form wiedergegeben werden. Zur Veranschaulichung stellt folgende Tabelle 3 die formale Grundstruktur eines Arguments im theoretischen Syllogismus anhand der wohl bekanntesten Aristotelischen Schlussfigur (Modus Barbara) dar (Bayer, 1999; Pfeifer, 2009). Aus den zwei deskriptiven, d. h. auf Fakten beruhenden Prämissen

„Alle Menschen sind sterblich“ und „Sokrates ist ein Mensch“ wird die logische Schlussfolgerung gezogen, dass Sokrates sterblich ist.

Tabelle 3: Formale Grundstruktur eines Arguments im theoretischen Syllogismus (Bayer, 1999; Pfeifer, 2009).

1. Prämisse	Alle Menschen sind sterblich.
2. Prämisse	Sokrates ist ein Mensch.
3. Konklusion (Schlussfolgerung)	Also ist Sokrates sterblich.

Die Richtigkeit der Konklusion basiert dabei auf der Richtigkeit der Prämissen. Deshalb ist grundsätzlich zu prüfen, ob die Prämissen stimmen, d. h. wahr sind und ob die anhand der Prämissen gezogene Schlussfolgerung richtig bzw. zwingend ist (Pfeifer, 2009).

In Bezug auf bio- und medizinethische Dilemmata wird dagegen ein normativer Syllogismus herangezogen, der sowohl deskriptive als auch zusätzliche normative Prämissen enthält (Mittelsten-Scheid & Höhle, 2008). Eine plausible und schlüssige ethische Argumentation setzt einerseits umfassendes, fallbezogenes Fachwissen und andererseits das Hinzuziehen relevanter Werte und Normen voraus (Zoglauer, 2005). Tabelle 4 veranschaulicht ein mögliches Argument in Form des normativen (praktischen) Syllogismus am Beispiel der Diskussion um die PID (Dietrich, 2004, 2005; Höhle & Reitschert, 2006, S. 105).

Tabelle 4: Normativer Syllogismus am Beispiel der PID (Dietrich, 2004, 2005; Höhle & Reitschert, 2006, S. 105).

Normativer Syllogismus	Beispiel PID
1. Deskriptive Prämisse	Menschliches Leben muss von Beginn an unbedingten Schutz erhalten.
2. Relevante normative Prämisse	Bei der PID werden bestimmte Embryonen verworfen und nicht geschützt.
3. Normative Schlussfolgerung	Also darf die PID nicht durchgeführt werden.

Die deskriptive Prämisse beschreibt dabei den Sachverhalt, wohingegen die normative Prämisse die für den Sachverhalt relevante Norm benennt (Dietrich, 2004, 2005). In einem praktischen Syllogismus wird aus einer normativen Prämisse und (mindestens) einer nicht-normativen, deskriptiven Prämisse ein gemischt-normativ-deskriptiver Schluss gezogen. Die normative Schlussfolgerung ist ein wesentliches Kennzeichen ethischen Argumentierens. Die syllogistische Rekonstruierbarkeit bildet eine elementare Voraussetzung zur Überprüfung ethischer Argumente und Urteile (Pfeifer, 2009).

Die im praktischen Syllogismus gemachten normativen Aussagen können streng genommen weder wahr noch falsch sein, in erster Linie geht es darum, ob sie gültig sind, also Geltung für sich beanspruchen. Diese Gültigkeit muss sich dann letztlich in einem argumentativen Diskurs begründen lassen. Die logisch korrekte Form eines Arguments ist demnach nur eine notwendige, aber keine hinreichende Bedingung für dessen Gültigkeit.

Allerdings ist die Anwendung normativer Syllogismen zur Begründung normativer Aussagen sehr begrenzt, da sie grundsätzlich nur analytische Schlüsse zulassen (Pfeifer, 2003). Außerdem argumentieren Menschen in Alltagssituationen erfahrungsgemäß nicht in dieser klaren syllogistischen Form (Kuhn, 1991; Reitschert & Höble, 2006). Ein Charakteristikum von Alltagsargumentation ist beispielsweise, dass im Gegensatz zur formalen Logik mit einer normativen Aussage begonnen wird, um diese dann nachfolgend zu begründen (Mittelsten-Scheid, 2009). Dieses Phänomen wird von Bayer (2007) als „Spitzenformulierung“ bezeichnet.

In der informellen Logik wurde von Stephen Toulmin (1958, 1996) ein Schema zur Analyse von Argumentationen entwickelt, das geeignet ist, „alltagssprachliche“ Argumentationen strukturell zu erfassen, beschreiben und erläutern zu können.

Das Toulmin-Schema wird sehr häufig in Studien im Bereich des *socio-scientific decision-making* zur Klassifikation der Argumentationsqualität verwendet (vgl. z. B. Dawson & Venville, 2009; Erduran et al., 2004; Jiménez-Aleixandre, Rodriguez & Duschl, 2000; Osborne et al., 2004; Sadler & Fowler, 2006).

Argumentation umfasst nach dem Modell von Toulmin (1958, 1996) sechs Komponenten:

1. Eine Behauptung (Konklusion, *claim*), „deren Tauglichkeit wir zu begründen versuchen“ (Toulmin, 1996, S. 89).
2. Tatsachen (Fakt, Daten, *data*), „die wir als Begründung für die Behauptung heranziehen“ (Toulmin, 1996, S. 89) und der Stützung der Konklusion dienen.
3. Eine Schlussregel (*warrant*), die im Sinne einer Begründung als hypothetische, brückenartige Aussage den Übergang zwischen Datum und Behauptung spezifiziert.
4. Eine Stützung der Schlussregel (*backing*) durch den Verweis auf weitere Versicherungen, „ohne die die Schlußregeln selbst weder zulässig noch geläufig wären“ (Toulmin, 1996, S. 94) und demzufolge angibt, ob die Schlussregel als zulässig gelten kann.
5. Einen Modaloperator (*qualifier*), der den Grad der Sicherheit der Behauptung angibt und unter entsprechenden Vorbehalten einschränkt.
6. Ausnahmebedingungen (*rebuttal*), bei deren Vorhandensein der Geltungsbereich der Behauptung unter Verweis auf das Vorliegen bestimmter Bedingungen eingeschränkt wird.

Übertragen auf eine Entscheidungssituation stellt nach Toulmin's Modell (1958, 1996) die getroffene Entscheidung die Konklusion (*claim*) dar, die durch einen Fakt bzw. eine Tatsache (*data*) begründet wird. Bei diesem einfachen Argument kann nun die Wahrheit des Datums oder die Relation zwischen Datum und Konklusion angezweifelt werden. Der Zusammenhang zwischen Datum und Konklusion kann durch eine Schlussregel (*warrant*) gerechtfertigt werden. Diese Schlussregel lässt sich allerdings bezweifeln. Zur Stützung (*backing*) der Schlussregel könnte auf bestimmte Gesetze und rechtliche Vorkehrungen verwiesen werden.

Abbildung 3 stellt das Argumentationsstrukturmodell von Toulmin (1958, 1996) graphisch dar:

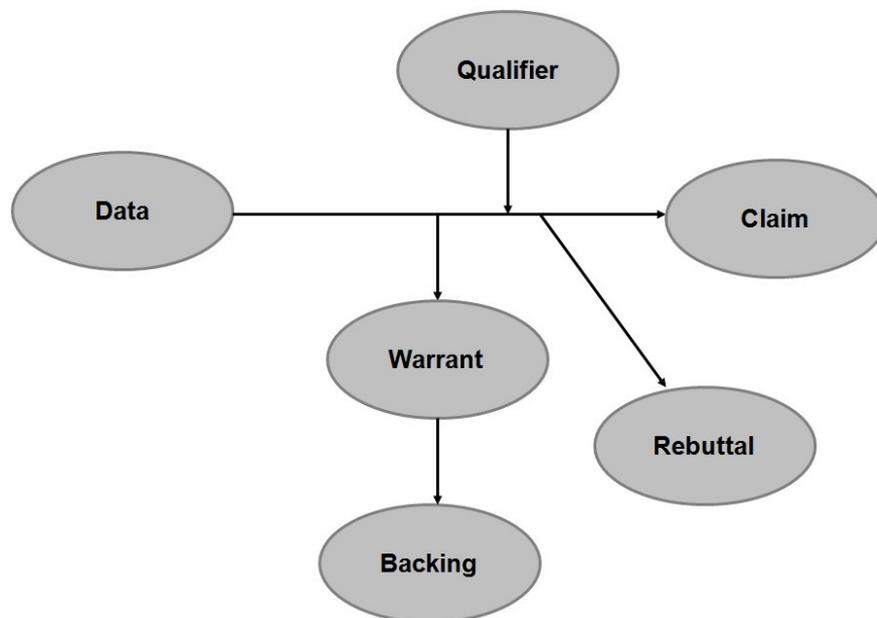


Abbildung 3: Toulmin's Argumentschema (1958, 1996).

Toulmin (1996) gibt folgendes Beispiel zur Veranschaulichung des Zusammenwirkens der sechs Komponenten: Die Konklusion, dass Harry ein britischer Staatsbürger ist (*claim*) wird durch die Tatsache, dass Harry auf den Bermudas geboren wurde (*datum*) begründet. Bei diesem einfachen Argument kann nun die Wahrheit des Datums oder die Relation zwischen Datum und Konklusion angezweifelt werden. Der Zusammenhang zwischen Datum und Konklusion kann durch eine Schlussregel (*warrant*), z. B. „wer auf den Bermudas geboren wurde, ist im Allgemeinen britischer Staatsbürger“, gerechtfertigt werden. Da sich diese Schlussregel bezweifeln lässt, könnte zur Unterstützung (*backing*) der Schlussregel auf bestimmte Gesetze und rechtliche Vorkehrungen verwiesen werden, die bestimmen, dass auf den Bermudas geborene Personen britische Staatsbürger sind.

Zum flexiblen Einsatz bei komplexeren Begründungszusammenhängen erweitert Toulmin (1996) sein Argumentationsschema um zwei Komponenten, indem er Modaloperatoren (*qualifier*) („vermutlich ist Harry deshalb britischer Staatsbürger“) und Ausnahmebedingungen (*rebuttals*) („es sei denn, beide Eltern sind Ausländer oder er wurde durch Einbürgerung Amerikaner“) einführt. Modaloperatoren schränken den Grad der Sicherheit der Behauptung durch die Voranschaltung von „vermutlich“ oder „wahrscheinlich“ ein, wohingegen Ausnahmebedingungen den Geltungsbereich der Behauptung auf das Vorliegen bestimmter Bedingungen eingrenzen.

Das Skelett seines Argumentationsschemas (*data, claim, warrant, backing*) überträgt Stephen Toulmin auch auf normative Aussagen (Pfeifer, 2003). Aus der Konklusion wird bei normativen Aussagen ein *claim*, aus der deskriptiven Prämisse ein *datum* und aus der normativen Prämisse werden *warrants*, die ihrerseits wiederum durch *backings* weiter begründet werden. Als *warrants* kommen überwiegend Normen, moralische Daumenregeln, Sitten oder Gewohnheiten in Frage, als *backings* auf einer abstrakteren Ebene moralische Prinzipien oder Ethiktheorien. Pfeifer (2009, S. 21) führt folgendes Beispiel an: Die Behauptung (*claim*), dass „X moralisch schlecht gehandelt hat“ kann durch die Tatsache (*data*) „X hat gelogen“ begründet werden. Der Übergang vom Datum zur Konklusion kann durch die Rechtfertigung (*warrant*) „Lügen ist moralisch schlecht“ sichergestellt werden. Die Rechtfertigung bedarf ihrerseits einer Stützung (*backing*) wie beispielsweise dem Hinweis auf die schlechten Folgen des Lügens.

Der Philosoph Detlef Horster (1999, zitiert nach Pfeifer, 2003, S. 155) fasst die Art und das didaktische Gewicht des Toulminischen Prüfverfahrens in Bezug auf moralische Entscheidungen sehr treffend zusammen:

Moralische Entscheidungen eines jeden Menschen müssen letzten Endes vor der Stützregel (*backing*) Bestand haben. Zuvor muss jede Rechtfertigungsregel (*warrant*) überprüft werden; sodann müssen die Daten (*data*) eben die zur Verfügung stehenden Sachverhalte bewertet werden. Für das moralische Problem, wie denn abstrakt-universelle Prinzipien, z. B. der kantische Imperativ, in der konkreten Handlungssituation zur Anwendung gebracht werden können, scheint mir das Toulmin-Schema nach wie vor unübertroffen.

In Argumentationen müssen, wie Toulmin (1996) selbst betont, nicht immer alle Komponenten expliziert werden, beispielsweise werden oft die Schlussregel (*warrant*) und Stützung (*backing*) nicht oder nur auf Nachfrage genannt. Von vielen Forschern wird außerdem darauf hingewiesen, dass in Argumentationen häufig die einzelnen Bestandteile nur sehr schwer zu erkennen sind, was insbesondere die Unterscheidung von Tatsachen (*data*), Schlussregeln (*warrants*) und Stützungen (*backings*) betrifft (vgl. Chang & Chiu, 2008; Dawson & Venville, 2009; Stein & Albro, 2001).

Aus diesem Grund wird in vielen Studien zur Analyse der Argumentation deshalb die Unterscheidung von Tatsachen (*data*), Schlussregeln (*warrants*) und Stützungen (*backings*) aufgegeben; stattdessen werden diese drei Argumentationselemente häufig in eine einzige Kategorie unter den Begriff Begründungen (*grounds* oder *reasons*) zusammengefasst (vgl. z. B. Erduran et al., 2004; Osborne et al., 2004; Sadler & Fowler, 2006; Voss & Means, 1991).

Die Qualität von Argumentationen kann sowohl mithilfe inhaltlicher Kriterien (z. B. hohe Relevanz der Fakten für die Behauptung bzw. Entscheidung) als auch mithilfe struktureller Kriterien ermittelt werden (vgl. z. B. Schwarz, Neuman, Gil & Ilya, 2003; Zohar & Nemet, 2002). Eine hohe strukturelle Qualität der Argumentation liegt beispielsweise vor, wenn besonders viele und unterschiedliche Argumentationselemente des Toulmin-Schemas (1958) verwendet werden (vgl. z. B. Osborne et al., 2004; Schwarz et al., 2003; Zohar & Nemet, 2002).

Häufig wird die Argumentationsqualität quantitativ anhand der Anzahl der Begründungen gemessen, indem jeder Begründung ein statistischer Wert zugewiesen wird. Die Argumentationsqualität steigt dabei mit der Anzahl der Begründungen. Beispielsweise würde ein Argument, das keine Begründungen enthält, mit 0 kodiert werden und somit das qualitativ unterste Niveau einer Argumentation darstellen. Höhere Stufen der Argumentation setzen hingegen das Artikulieren von Begründungen (*data*, *warrants*, *backings*) voraus. Auf diese Art und Weise kann die Qualität von Argumentationen abgebildet werden.

Neben der Anzahl der angeführten Argumente können Argumentationen in Bezug auf die Qualität auch daraufhin analysiert werden, ob gegnerische Argumente berücksichtigt und mögliche Gegenpositionen bzw. -argumente diskutiert und durch Widerlegungen (*rebuttals*) entkräftet werden (vgl. Knipfer, 2009; Kuhn, 1991; Kuhn, Shaw & Felton, 1997; Sadler & Zeidler, 2005).

Gegenargumente definieren Voss et al. (1986) als „reasons that lead to the contradiction of the conclusion or negate supportive reasons” (p. 223), Means & Voss (1996) als “an argument with a reason that supports the contradiction of a given claim, sometimes signaled by a phrase such as “on the other hand” (p. 141). Demzufolge stellen Begründungen oder Widersprüche, die sich gegen eine Schlussfolgerung oder Entscheidung richten bzw. Argumente, die der eigenen Entscheidung widersprechen, mögliche Gegenpositionen dar, die wiederum durch Widerlegungen (*rebuttals*) entkräftet werden können.

Allerdings konnte in zahlreichen Studien festgestellt werden, dass einseitige Argumente einer ausgewogenen, zweiseitigen Argumentation oft vorgezogen werden: Personen können typischerweise in der Regel mehr Argumente zur eigenen Position generieren als Ge-

genargumente, selbst dann, wenn explizit zur Nennung von Gegenargumenten aufgefordert wird (vgl. z. B. Baron, 1995; Kuhn, 1991; Toplak & Stanovich, 2003; Voss et al., 1986; West, Toplak & Stanovich, 2008). Dieses Phänomen wird als *myside bias* (Perkins, 1989) bezeichnet.

Im Rahmen von Untersuchungen zur Dissonanztheorie (siehe Kapitel 4.4.3) konnte auch vielfach nachgewiesen werden, dass Personen nach Entscheidungen eine Bestätigungstendenz bei der Auswahl von Informationen (*confirmation bias*) (Wason, 1968) aufweisen, d. h. sie bevorzugen Informationen oder Argumente, die ihre gewählte Entscheidungsalternative unterstützen und vernachlässigen Informationen, die der eigenen Entscheidung widersprechen (vgl. z. B. D'Alessio & Allen, 2002; Frey, 1981, 1986; Jonas, Schulz-Hardt, Frey & Thelen, 2001).

Personen tendieren neben einer konfirmatorischen Informationssuche auch zu einer verzerrten Aufnahme von neuen Informationen und Argumenten: Jene werden häufig nicht unvoreingenommen, sondern in Abhängigkeit von einer bereits bestehenden Meinung bzw. Entscheidungspräferenz bewertet (*prior belief effect*; Edwards & Smith, 1996). Dabei werden Informationen oder Argumente, die konsistent zur eigenen Entscheidung sind, systematisch für glaubwürdiger, entscheidungsrelevanter und überzeugender gehalten als Informationen, die der eigenen Entscheidung widersprechen. Das Phänomen der meinungs- und präferenzkonsistenten Informationsbewertung konnte bereits viele Male in Studien empirisch bestätigt werden (vgl. z. B. Ditto & Lopez, 1992; Ditto et al., 1998; Edwards & Smith, 1996; Greitemeyer, Schulz-Hardt & Frey, 2003; Jonas et al., 2003; Koehler, 1993; Lord, Ross & Lepper, 1979; Traut-Mattausch, Schulz-Hardt, Greitemeyer & Frey, 2004).

Edwards & Smith (1996) führen die unterschiedliche Evaluation von präferenzinkonsistenten und -konsistenten Argumenten auf quantitative Testungsunterschiede zurück: Präferenzkonsistente Argumente werden nur oberflächlich per Augenschein-Validität geprüft, so dass Schwachstellen der Argumente nicht erkannt und die Argumente akzeptiert werden. Präferenzinkonsistente Argumente werden hingegen detaillierter und kritischer geprüft, so dass infolge dieser tiefen Überprüfung die Wahrscheinlichkeit steigt, dass Schwachstellen der Argumente erkannt werden und eine kognitiv aufwändige Suche nach Gegenargumenten in Gang gesetzt wird.

Ein geringer *confirmation* und *myside bias* werden als zentrale Indikatoren für kritisches Denken und reflektiertes Urteilen angesehen (vgl. Baron, 1995; Knipfer, 2009; Toplak & Stanovich, 2003; West et al., 2008).

Nachfolgend sollen exemplarisch noch zwei Kategorienschemata zur Klassifikation der Argumentationsqualität näher vorgestellt werden. Sadler & Fowler (2006) definieren in

einer Studie bezüglich gentechnologischer Dilemmata zu den Themen Gentherapie und Klonen fünf Kategorien. Den theoretischen Ausgangspunkt zur Klassifikation der Argumentationsqualität bildete das Toulmin-Schema (1958). Das unterste Niveau der Argumentationsqualität, Stufe 0, stellte keine Rechtfertigung der Entscheidung dar (*no justification*). Stufe 1 wurde vergeben, wenn die eigene Entscheidung ohne eine valide Begründung gerechtfertigt wurde (*justification with no grounds*). Als Begründungen wurden Tatsachen (*data*), Schlussregeln (*warrants*) oder Stützungen (*backings*) akzeptiert. Statements, in denen die eigene Entscheidung mit einer validen Begründung gerechtfertigt wurde, wurden hingegen mit den Werten „2“ und „3“ kodiert. Im Unterschied zu den Antworten der Kategorie Stufe 2, die nur eine einzige Begründung enthielten (*justification with simple grounds*), wurden bei Antworten der Stufe 3 mehrere und elaboriertere Begründungen gegeben (*justification with elaborated grounds*). Die letzte und höchste Stufe 4 wurde an Statements vergeben, in denen die Studienteilnehmer nicht nur begründete Rechtfertigungen für ihre eigene Entscheidung abgaben, sondern auch (mindestens) eine Gegenposition berücksichtigten (*justification with elaborated grounds and a counterposition*).

Foong & Daniel (2010) schlagen hingegen ein sechsstufiges Kategorienschema zur Klassifikation der Argumentationsqualität vor. Antworten, in denen keine Entscheidung getroffen wurde (*no claim*), wurden mit dem Wert „0“ kodiert. Stufe 1 wurde vergeben, wenn eine Entscheidung ohne eine valide Begründung oder Widerlegung getroffen wurde (*no valid grounds or rebuttal*). Eine Entscheidung, die sich auf ein oder zwei Begründungen stützte, aber keine Widerlegung enthielt (*one or two grounds, but without rebuttal*), wurde mit dem Wert „2“ kodiert, bei drei oder mehr Begründungen mit „3“ (*three grounds or above, but without rebuttal*). Die höheren Niveaus der Argumentationsqualität setzten Widerlegungen (*rebuttals*) voraus, bei Niveau 4 wurden ein oder zwei Begründungen mit einer Widerlegung genannt (*one or two grounds with rebuttal*), beim höchsten Niveau 5 drei oder mehr Begründungen mit einer Widerlegung (*three grounds or above with rebuttal*).

Sadler & Zeidler (2005) konnten in einer qualitativen Studie bei Studenten drei typische Argumentations-/Begründungsmuster in Bezug auf die Entscheidungsfindung zu sechs gentechnologischen Dilemmata zu den Themen Gentherapie und Klonen rekonstruieren: rationale, emotionale und intuitive Argumentationsmuster (*rationalistic, emotive and intuitive informal reasoning patterns*). Beim rationalen Argumentationsmuster (*reason-based considerations*) wurde die Entscheidungsfindung durch rationale und logische Kalkulationen geleitet, indem beispielsweise deontologische und utilitaristische Prinzipien angewendet, Kosten-Nutzen-Analysen durchgeführt sowie Vor- und Nachteile und Chancen und Risiken abgewogen wurden. Die beiden anderen vorgefundenen Argumentationsmuster, emotional und intuitiv, waren ausschließlich affektiver Natur. Beim emotionalen Argu-

mentationsmuster (*care-based considerations*) lösten die Studenten die Szenarien auf der Basis von Emotionen wie Sympathie und Empathie, indem sie beispielsweise die Notlage, Gefühle und das Leid der in den Szenarien beteiligten Personen bei der Lösung der Themen beachteten. Das intuitive Argumentationsmuster (*immediate reactions, gut feeling*) basierte auf einer unverzüglichen, positiven oder negativen Reaktion auf den Kontext des Szenarios, indem die Entscheidungsfindung durch ein rational nicht erklärbares „Bauchgefühl“ geleitet wurde (vgl. Haidt, 2001; Gigerenzer, 2007). Der Unterschied zwischen einer emotionalen und intuitiven Argumentationsweise besteht darin, dass emotionale Argumentationsmuster sich stets auf reale Personen oder fiktive Charaktere richten, wohingegen intuitive Muster persönliche Reaktionen oder Antworten zu spezifischen Kontextaspekten darstellen.

3.3 Empirische Studien zur Argumentationsqualität und zu Argumentationsmustern

In einer Vielzahl von in den USA durchgeführten Studien konnte festgestellt werden, dass die Argumentationsqualität von Schülern im Bereich SSI relativ gering ist. Die meisten Schülerargumentationen weisen nach den strukturellen Kriterien nach Toulmin (1958) nur eine sehr geringe Qualität auf; häufig werden Entscheidungen nur sehr einfach unter Zuhilfenahme eines Argumentationselements (z. B. *data, warrant, backing*) begründet (vgl. z. B. Osborne et al., 2004; Sadler & Fowler, 2006). Jiménez-Aleixandre et al. (2000) konnten beispielsweise zeigen, dass Schüler, wenn sie Argumente über Genetik generieren, sich eher auf das Erstellen von detaillierten Behauptungen (*claims*) konzentrieren, diese aber nicht mit Tatsachen (*data*) oder Schlussregeln (*warrants*) belegen.

Im Hinblick auf die Argumentationsmuster konnten Sadler & Zeidler (2005) an Schülern ($n = 30$) feststellen, dass viele Schüler eine Kombination der drei verschiedenen Argumentationsmuster zeigten, deren Verwendung jedoch stark vom Kontext des Szenarios abhing. Eine stärkere Kontextabhängigkeit konnte insbesondere für das emotionale und intuitive Argumentationsmuster nachgewiesen werden. Beispielsweise variierte die Anzahl der Schüler, die eine emotionale Argumentationsweise verwendeten, in den drei Szenarien zum Thema Gentherapie zwischen 10 % (3) und 80 % (24), wohingegen die Anzahl in den drei Szenarien zum Thema Klonen mit 50-70 % (15-21) relativ konstant blieb. Allem Anschein nach wurden durch das Thema Klonen mehr emotionale Reaktionen ausgelöst als durch das Thema Gentherapie. Die Verwendung des intuitiven Argumentationsmuster erstreckte sich von 3 % (1) zu 50 % (15) und trat beim Thema Klonen häufiger auf als beim Thema Gentherapie. Die Nutzung des rationalen Argumentationstyps variierte nur wenig

von 73 % (22) zu 100 % (30) und wies demzufolge eine niedrigere Kontextabhängigkeit auf als die beiden anderen affektiven Argumentationsweisen.

Über alle sechs Szenarien hinweg, wurde das rationale Argumentationsmuster bei der Lösung der gentechnologischen Dilemmata von den Schülern am häufigsten geäußert (88 %; 158), gefolgt vom emotionalen (47 %; 85), am wenigsten wurde das intuitive Argumentationsmuster genutzt (25 %; 45). Sadler & Zeidler (2005) konnten jedoch feststellen, dass intuitives Argumentieren einen signifikanten Faktor bei der Lösung einiger Szenarien bildete. Die intuitiven, vielfach negativen Reaktionen determinierten häufig die endgültige individuelle Entscheidung und wurden oft als erste Antwort gegeben, d. h. gehen dem rationalen und emotionalen Argumentieren voraus. Sadler & Zeidler (2005) konnten auch zeigen, dass ältere Schüler überwiegend das rationale Argumentationsmuster verwenden, während jüngere Schüler vor allem intuitiv argumentieren.

Ein weiteres wichtiges Ergebnis der Studie war, dass die Mehrheit von 28 der 30 Schüler (98 %) explizit die moralischen und ethischen Implikationen, die aus der Anwendung der Technologien des Klonens und der Gentherapie resultieren, erkannten. Die Wahrnehmung und Berücksichtigung dieser moralischen Aspekte determinierte allerdings nicht notwendigerweise die Entscheidungsfindung der Schüler (Sadler & Zeidler, 2005).

Dawson & Venville (2009) untersuchten in einer Studie die Argumentationsqualität und Argumentationsmuster von 30 australischen Schülern zum Thema Gentechnologie. Die Schüler wurden in halbstrukturierten Interviews mit einer Gruppengröße von 2-3 Studenten zu ihrem Verständnis und zu ihren Ansichten zu verschiedenen biotechnologischen Themen befragt (z. B. Klonen, genetische Untersuchungen auf Erbkrankheiten, Vaterschaftstests, forensische Analytik, Produktion und Konsum gentechnisch veränderter Früchte). Diese Themen wurden ausgewählt, da sie alle kontrovers und brisant diskutierte SSI darstellen, die ein gewisses Verständnis über Genetik und Biotechnologie erfordern. Die Schüler wurden in den Gruppeninterviews nicht explizit dazu aufgefordert, Argumente für oder gegen ihre Position zu nennen. Die Argumentationsmuster wurden mithilfe von zwei Kodierern nach Sadler & Zeidler (2005) den Kategorien rational, emotional und intuitiv zugeteilt. Zur Erfassung der Argumentationsqualität wurde ein ähnliches Kategorienschema wie in der Studie von Sadler & Fowler (2006) verwendet.

Die Studienergebnisse konnten zeigen, dass im Gegensatz zu den Befunden von Sadler & Zeidler (2005), die eine Prädominanz des rationalen Argumentationstyps feststellten, am häufigsten intuitiv argumentiert wurde, gefolgt von einer emotionalen und dann rationalen Argumentationsweise. Als mögliche Erklärung dieser konträren Befunde führen die Autoren dieser Studie an, dass Themen wie Klonen oder forensische Analysen unter Umständen mehr intuitive und emotionale Reaktionen auslösen als rationale. Darüber hinaus nehmen

sie an, dass die von Sadler & Zeidler (2005) untersuchten Studenten möglicherweise über bessere Fähigkeiten in der Artikulation ihrer emotionalen und rationalen Denkweisen verfügten. Eine mögliche theoretische Erklärung liefert das „*Threshold Model of Content Knowledge Transfer*“ (Sadler & Donnelly, 2006; Sadler & Fowler, 2006; siehe Kapitel 3.4): Es ist durchaus denkbar, dass die Schüler nicht über genügend inhaltliches Wissen verfügten, um rational argumentieren zu können. In Bezug auf die Argumentationsqualität konnte festgestellt werden, dass die Mehrzahl der Schüler insgesamt, d. h. über alle Themen hinweg, auf Stufe 2 argumentierten. Viele Schüler verwendeten demzufolge nur sehr einfache Begründungen, um ihre Entscheidungen in Bezug auf biotechnologische Dilemmata zu rechtfertigen.

Sowohl Sadler & Zeidler (2005) als auch Dawson & Venville (2009) weisen darauf hin, dass Überlappungen der Argumentationsmuster vorhanden sind, d. h. emotionale und intuitive Argumentationsweisen können nicht immer eindeutig voneinander getrennt werden und sollten daher zusammengefasst dargestellt werden.

3.4 Potentielle Einflussfaktoren auf die Entscheidungsfindung und Argumentationsqualität

Aus Sicht der psychologischen Forschung und im Speziellen der Persuasionsforschung lassen sich folgende kognitive und motivationale Lernvoraussetzungen ableiten, welche die Entscheidungsfindung, die Informationsverarbeitung und damit auch die Argumentationsqualität potentiell beeinflussen können (siehe Kapitel 3.2.2): das themenspezifische Vorwissen (z. B. Wood, 1982; Wood et al., 1995), der persönliche Bezug (Involvement) bzw. das Interesse und die (Vor-)Einstellungen zum Thema (Petty & Cacioppo, 1986) sowie die Präferenz für Intuition und Deliberation (Betsch, 2004). Auch die situativ angestrebte und tatsächliche Sicherheit der Entscheidung (Chaiken et al., 1989), die parasoziale Interaktion mit einer Medienfigur (interpersonales Involvement) und die Attraktivität der Medienfigur bzw. des Kommunikators (Hartmann & Klimmt, 2005) können potentiell die Entscheidungsfindung beeinflussen.

Nachfolgend wird für diese Arbeit eine besonders relevante Einflussgröße auf die Entscheidung und Argumentation vorgestellt und diskutiert: das themenspezifische, inhaltsbezogene Vorwissen. Die Erkenntnis, dass das Vorwissen eine zentrale Rolle beim Lernen spielt, gilt in der pädagogisch-psychologischen Forschung als unbestritten (vgl. z. B. Ausubel, 1968; Dochy, 1992; Duit, 1999; Gagné, 1965). Zwischen dem Vorwissensniveau und dem Lernerfolg besteht ein positiver Zusammenhang (Bereiter, 1995). Das Vorwissen gilt demnach als ein wichtiger Prädiktor für den Lernerfolg.

Grundsätzlich kann auch davon ausgegangen werden, dass inhaltsbezogenes Fachwissen eine wichtige Voraussetzung für hochwertige Argumentationen ist (Means & Voss, 1996; Perkins & Salomon, 1989). Die aktuelle Befundlage zum Einfluss des inhaltlichen Fachwissens auf die Argumentationsqualität und Argumentationsmuster im Bereich SSI ist jedoch bislang recht uneinheitlich.

Sadler & Zeidler (2004) konnten in einer Studie an 30 Studenten einen positiven Zusammenhang zwischen dem Vorwissen über Humangenetik auf die Argumentationsqualität feststellen: Bei hohem humangenetischen Fachwissen fiel die Argumentationsqualität signifikant höher aus als bei niedrigem inhaltlichem Wissen. Das Vorwissen hatte jedoch keinen Einfluss auf die verwendeten Argumentationsmuster (rational, emotional und intuitiv).

In einer anderen Studie von Sadler & Donnelly (2006) zeigten sich hingegen konträre Befunde. Die Autoren dieser Studie untersuchten den Einfluss des Vorwissens und der moralischen Urteilsfähigkeit auf die Argumentationsqualität in Bezug auf gentechnologische Themen und verwendeten dabei einen Methodenmix: Die 56 Schüler, die an der Studie teilnahmen, bearbeiteten zunächst einen Vorwissenstest über Humangenetik und den Defining Issue Test (DIT) von Rest (1979) zur Messung der moralischen Urteilsfähigkeit. Im Anschluss daran bearbeiteten und lösten die Schüler drei kontroverse gentechnologische Szenarien zum Thema Klonen und Gentherapie und wurden im Rahmen von halbstrukturierten Interviews befragt. Anhand multipler Regressionsanalysen konnten keine statistisch signifikanten und bedeutsamen Zusammenhänge zwischen dem themenspezifischen Vorwissen über Humangenetik, der moralischen Urteilsfähigkeit und der Argumentationsqualität identifiziert werden. Sowohl das inhaltliche Vorwissen über Humangenetik als auch die moralische Urteilsfähigkeit determinierten nicht substantiell die Argumentationsqualität. Die quantitativ vorgefundenen Ergebnisse in Bezug auf das Vorwissen konnten auch in einer qualitativen Auswertung der Interviewtranskripte bestätigt werden: Die Schüler wendeten nur äußerst selten ihr Vorwissen über Humangenetik zur Lösung der gentechnologischen Szenarien an.

Sadler & Donnelly (2006) nehmen jedoch an, dass möglicherweise ein nicht-linearer Zusammenhang zwischen dem inhaltlichen Fachwissen über Humangenetik und der Argumentationsqualität bestehen könnte und schlagen deshalb ein Schwellenmodell, das *Threshold Model of Content Knowledge Transfer* (nachfolgend abgekürzt mit: TMCKT) vor. Grundannahme des TMCKT ist, dass bei Studenten erst mit einem hohen Vorwissen-niveau und einer vermutlich höheren Verständnisebene das Fachwissen an Bedeutung für die Argumentation gewinnt, indem dieses dann auch adäquat genutzt werden kann. Dem TMCKT entsprechend vermuten Sadler & Fowler (2006), dass die in ihrer Studie untersuchten Schüler unterhalb dieser bestimmten Schwelle lagen und somit nicht über die ent-

sprechende Vorwissensstufe verfügten und deshalb eine schlechtere Argumentationsqualität aufwiesen.

Einige Grundannahmen des TMCKT konnten Sadler & Fowler (2006) in einer Studie bereits empirisch bestätigen. Die 45 untersuchten Schüler und Studenten wiesen drei unterschiedliche Vorwissensniveaus auf: Die erste Gruppe bildeten Schüler, die über ein variables Vorwissen über Humangenetik verfügten (*high school students*), die zweite Gruppe bestand aus Studenten nicht-naturwissenschaftlicher Studienrichtungen mit wenig Vorwissen (*college non-science majors*) und die dritte Gruppe aus Studenten naturwissenschaftlicher Fächer mit fortgeschrittenem Genetik-Wissen (*college science majors*). Zusätzlich bearbeiteten alle Studenten einen Vorwissenstest über Genetik. Die Argumentationsqualität wurde anhand einer fünfstufigen Skala erhoben (siehe Kapitel 3.2.3). Wie die Ergebnisse zeigen konnten, wiesen die Studenten mit einem höheren Vorwissen (*science majors*) eine signifikant höhere Argumentationsqualität auf als Schüler mit einem niedrigeren Vorwissen (*non-science majors* bzw. *high-school students*). Die beiden Gruppen mit einem niedrigen Vorwissen unterschieden sich statistisch nicht bedeutsam in der Argumentationsqualität.

Auch Chang & Chiu (2008) konnten zeigen, dass Studenten mit einem höheren Vorwissen (*science majors*) eine höhere Argumentationsqualität aufwiesen als Studenten mit einem geringeren Vorwissen (*non-science majors*). Die recht heterogene Befundlage zum Zusammenhang zwischen Fachwissen und Argumentationsqualität führen beide Autoren auf die doch sehr unterschiedlichen, in den verschiedenen Studien verwendeten Themen zurück, die häufig nur wenig persönlichen Bezug zum alltäglichen Leben der Studierenden oder Schülern aufweisen.

In einer Studie von Götz (2001) zeigte sich ein klassischer ATI-Effekt: Der Einfluss des Lernprogramms zum Thema ethische Aspekte der Gentechnik auf den Lernerfolg wurde vom Vorwissen der Lernenden moderiert, es profitierten vor allem vorwissensschwächere Lernende von der Lernumgebung.

4 Theoretische Grundlagen zur Konzeption der multimedialen Lernumgebung

In diesem Kapitel werden aus pädagogisch-psychologischer Sicht die theoretischen Grundlagen zur Konzeption der in dieser Arbeit untersuchten multimedialen Lernumgebung beschrieben.

Ausgehend von einer Definition des Begriffes Multimedia (Kapitel 4.1) werden die für diese Arbeit relevanten kognitions- und medienpsychologischen theoretischen Grundlagen zum Lernen mit neuen Medien vorgestellt (Kapitel 4.2). Dabei handelt es sich um die kognitive Theorie zum multimedialen Lernen, die Social Agency Theory sowie den Ansatz der parasozialen Interaktion, der sich explizit mit der Auseinandersetzung mit virtuellen Figuren befasst. Darauf aufbauend wird auf die Gestaltung multimedialer Lernumgebungen aus didaktischer und medialer Sicht eingegangen (Kapitel 4.3). Als theoretische Basis für die Ableitung von didaktischen Gestaltungskriterien dient der Ansatz des problemorientierten Lernens. Die Grundlage zu den medialen Gestaltungskriterien bilden aktuelle Befunde und Erkenntnisse aus der Forschung mit neuen Medien. Im Anschluss daran werden die theoretischen Grundlagen zum Feedback in multimedialen Lernumgebungen beschrieben (Kapitel 4.4). Ausgehend von einer Begriffsbestimmung werden zwei zentrale Einflussfaktoren auf die Wirkung von Feedback betrachtet: die Feedbackgestaltung und -rezeption. Danach folgen Erläuterungen zum Feedback in Form einer sozialen Vergleichsinformation. Im Mittelpunkt steht dann der konkrete Einsatz von multimedialen Lernumgebungen im informellen Lernort Museum (Kapitel 4.5). Schließlich werden empirische Befunde zur Effektivität von multimedialen Lernumgebungen zur Vermittlung von kontroversen Wissenschaftsthemen präsentiert (Kapitel 4.6).

4.1 Begriffsbestimmung Multimedia

Der Begriff „Multimedia“ wird sehr uneinheitlich gebraucht, in der Literatur existiert eine Fülle unterschiedlichster Definitionen, häufig wird Multimedia auch unter den Begriff „neue Medien“ subsumiert und beide Termini synonym verwendet. Zunächst kann der Begriff „Multimedia“ nach rein technisch orientierten, nach technischen und anwendungsbezogenen Aspekten und nach Aspekten von Multimedia definiert werden.

Aus einer rein technischen Dimension definiert Hornung (1994) „Multimedia“ als „Integration von Text, Grafik, Pixelbildern, Video und Audio“ (S. 2). Schaumburg & Issing

(2004) verstehen unter Multimedia „die computergestützte Integration vormals getrennter Einzelmedien wie Schaubild, Buch, Film, Tonband, Video etc. über eine einzige Nutzerschnittstelle, das Computerterminal“ (S. 718). Nach Steinmetz (2000) stellt Multimedia eine Kombination aus zeitabhängigen (kontinuierlichen) Medien (z. B. Audio, Video und 2D- und 3D-Animationen) und zeitunabhängigen (diskreten) Medien (z. B. Text, 2D- und 3D-Grafik) dar.

Klimsa (2002) weist jedoch darauf hin, dass eine Definition des Begriffs „Multimedia“ aus einer rein technischen Perspektive zu kurz greift und um den Aspekt der Anwendung bzw. des Nutzungskontextes ergänzt werden muss. Beispielsweise stellen ein PC mit Tonausgabe und integriertem CD-ROM-Laufwerk oder ein CBT (Computer Based Training), in dem neben Text auch Bilder sowie Grafiken integriert werden, noch lange kein Multimediasystem dar. Dementsprechend erweitert Klimsa (2002) die technische um eine anwendungsbezogene Dimension, die explizit den Nutzungskontext und die Funktionalität von Multimedia mitberücksichtigt. Multimedia ist nach Klimsa (2002) „ein Konzept, das nicht nur digitale Medien, sondern auch die gesamte technische und die anwendungsbezogene Dimension integriert“ und seinen Einsatz „in den Bereichen Information, Kooperation, Lernen und Unterhaltung“ findet (S. 6). Die technische Dimension differenziert sich in einen Medienaspekt und einen Integrations- und Präsentationsaspekt aus. Der Medienaspekt umfasst analog zur Definition von Steinmetz (2000) die Kombination von zeitabhängigen und zeitunabhängigen Medien, der Integrations- und Präsentationsaspekt beinhaltet Interaktivität, Multitasking (gleichzeitiges Ausführen mehrerer Prozesse) und Parallelität (parallele Medienpräsentation). Bei der anwendungsbezogenen Dimension unterscheidet Klimsa (2002) zwischen Datenbanksystemen, Kommunikationssystemen, Hypermediasystemen, spezifischen Autorensystemen sowie Multimediamiwerkzeugen und Systemen der virtuellen Realität.

Weidenmann (2002) bemängelt jedoch an beiden oben dargelegten Definitionen von Steinmetz (2000) und Klimsa (2002), dass Codierung, Modalität und Medien nicht immer einwandfrei voneinander getrennt werden können. In der Definition von Klimsa (2002) werden beispielsweise Codalität und Modalität beide unter den Medienaspekt subsumiert. Deshalb schlägt Weidenmann (2002) eine Ausdifferenzierung des Begriffs „Multimedia“ vor, indem er zwischen dem *Medium*, definiert als „Objekte, technische Geräte oder Konfigurationen, mit denen sich Botschaften speichern und kommunizieren lassen“ (S. 46), der *Kodierung*, den verwendeten Formaten bzw. Symbolsystemen zur Informationsdarstellung, und der *Modalität*, den durch die Informationsvermittlung angesprochenen Sinneskanal bzw. Sinnesorgan (auditiv, visuell usw.) beim Lerner, unterscheidet. Unter *multimedialen* Informationsangeboten wird die integrierte Präsentation mehrerer auf unterschiedliche

Speicher- und Präsentationstechnologien verteilte Medien verstanden, beispielsweise wenn eine Lernumgebung neben einem PC auch einen Beamer verwendet. *Multicodale* Informationsangebote weisen hingegen unterschiedliche Symbolsysteme bzw. Codierungen auf (textuell, bildlich, numerisch). Eine multicodale Lernumgebung liegt vor, wenn mehr als ein Zeichensystem verwendet wird, beispielsweise ein mit Bildern illustrierter Text oder eine Grafik, die mit Text und Zahlen beschriftet ist. Wenn mehrere Sinnesmodalitäten beim Nutzer angesprochen werden, spricht man von einem *multimodalen* Angebot, z. B. wenn in einer Lernumgebung Videos eingesetzt werden, in denen Informationen nicht nur visuell, sondern auch akustisch dargeboten werden.

Tabelle 5 fasst die Dimensionen zur Beschreibung medialer Angebote von Weidenmann (2002) mit Beispielen nochmals zusammen.

Tabelle 5: Dimensionen zur Beschreibung medialer Angebote (Weidenmann, 2002, S. 47).

	Mono-...	Multi-...
Medium	<u>Monomedial:</u> <ul style="list-style-type: none"> • Buch • Videoanlage • PC und Bildschirm 	<u>Multimedial:</u> <ul style="list-style-type: none"> • PC + CD-ROM-Player • PC + Videorekorder
Codierung	<u>Monocodal:</u> <ul style="list-style-type: none"> • nur Text • nur Bilder • nur Zahlen 	<u>Multicodal:</u> <ul style="list-style-type: none"> • Text mit Bildern • Grafik mit Beschriftung
Sinnesmodalität	<u>Monomodal:</u> <ul style="list-style-type: none"> • nur visuell (Text, Bilder) • nur auditiv (Rede, Musik) 	<u>Multimodal:</u> <ul style="list-style-type: none"> • audiovisuell (Video, CBT-Programme mit Ton)

Die von Weidenmann (2002) vorgeschlagene Beschreibung medialer Angebote ist jedoch nicht nur auf computerbasierte Medien beschränkt, sondern kann auf alle Arten von Medien angewendet werden: Auch ein Lehrbuch kann multicodal sein, wenn neben Texten auch Abbildungen enthalten sind, und ein Informationsangebot eines Lehrers kann multimodal sein, wenn er ein Tafelbild zur gleichzeitigen Erklärung eines Sachverhaltes anfertigt (Brünken, Seufert & Leutner, 2008).

Computerbasierte Medien besitzen jedoch spezifische Vorzüge gegenüber herkömmlichen Medien, da sie neue technische Möglichkeiten zur Generierung von multicodaler und multimodaler Präsentation von Lerninhalten eröffnen, die traditionelle Medien weitestgehend nicht erfüllen können, beispielsweise können Lehrbücher nicht multimodal dargeboten

werden. In vielen kognitionspsychologischen Untersuchungen hat sich die Verwendung multipler Repräsentationsformen, in denen Informationen in unterschiedlichen Sinnesmodalitäten (*Multimodalität*; z. B. visuell und akustisch) und Kodierungsformen (*Multicodalität*; z. B. verbal und bildlich) präsentiert werden, vielfach als sehr lernförderlich erwiesen (vgl. Brünken et al., 2008). Mithilfe von computerunterstützten Systemen kann auch die dynamische Darstellung von Sachverhalten und Gegenstandsbereichen in Form von Animationen und Simulationen bedeutend leichter realisiert werden als mit anderen Medien (vgl. Lewalter, 1997).

Ein weiterer Vorzug computerbasierter Medien besteht in der *Adaptivität*, denn Lerninhalte können auf ökonomische Weise individuell und flexibel an bestimmte Lernvoraussetzungen (z. B. Vorwissen) angepasst werden (vgl. Leutner, 2002). Des Weiteren kann durch eine Reihe von Eingriffs- und Steuerungsmöglichkeiten dem Lerner ermöglicht werden, aktiv mit dem System zu interagieren und unmittelbare Rückmeldungen (Feedback) zu seinen Aktionen zu erhalten. *Interaktivität* zwischen den Lernenden und dem Programm lässt sich auf vielfältige Weise durch die Implementierung folgender Merkmale erzielen: Lernende können selbstständig auf bestimmte Informationen und Lerninhalte zugreifen, zwischen verschiedenen Beispielen und Aufgaben wählen, Fragen, Übungsaufgaben oder Probleme bearbeiten (Ja/Nein- und Multiple Choice-Antwortmöglichkeiten) oder Zusatzinformationen aufrufen (vgl. Haack, 2002; Niegemann et al., 2004). Zu guter Letzt haben neue Medien auch Motivationsvorteile, da Lernende in der Regel motivierter sind, mit neuen Medien zu arbeiten als mit klassischen Texten (Brünken et al., 2008).

Computerunterstützte, multimediale Lernumgebungen ermöglichen somit eine aktive, selbstgesteuerte sowie eine individuelle und bedürfnisorientierte Beschäftigung mit den Lerninhalten. Allerdings resümiert Schnotz (2001) in Bezug auf das Lernpotential multimedialer Lernumgebungen, dass eine Vielfalt von technischen Medien und eine Vielzahl von unterschiedlichen Darstellungsformen sowie Sinneserfahrungen nicht notwendigerweise auch zu einem höheren Lernerfolg führen als traditionelle Formen der Wissensvermittlung. Vielmehr geht es darum, die verwendeten technischen Medien, die Repräsentationsformen und die angesprochenen Sinnesmodalitäten so zu arrangieren, dass Lernprozesse positiv unterstützt werden (vgl. Kerres, 2001; Schnotz, 2001).

4.2 Kognitions- und medienpsychologische Grundlagen zum Lernen mit neuen Medien

In diesem Kapitel werden kognitions- und medienpsychologische Grundlagen zum Lernen mit neuen Medien präsentiert, die in Bezug auf die Konzeption der in dieser Arbeit unter-

suchten multimedialen Lernumgebung von besonderer Relevanz sind: Die kognitive Theorie zum multimedialen Lernen von Mayer (2001, 2005a), die Social Agency Theory von Mayer (2005c) und der medienpsychologische Ansatz der parasozialen Interaktion, der sich explizit mit der Auseinandersetzung mit virtuellen Figuren auf kognitiver und affektiver Ebene beschäftigt (Hartmann, Schramm & Klimmt, 2004a; 2004b). Aus allen drei Theorien lassen sich verschiedene konkrete Designprinzipien für die Gestaltung von multimedialen Lernumgebungen sowie virtuellen Figuren (z. B. in Form von *talking heads*) ableiten.

4.2.1 Kognitive Theorie zum multimedialen Lernen (CTML)

Mayer (2001, 2005a) entwickelte ausgehend von der Theorie der dualen Kodierung von Paivio (1986), dem Arbeitsgedächtnismodell von Baddeley (1999) sowie der Cognitive Load Theorie (Chandler & Sweller, 1991; Sweller, 1999) ein kognitionspsychologisches Modell zur Beschreibung der Informationsverarbeitungsprozesse beim Wissenserwerb mit neuen Medien: die kognitive Theorie des multimedialen Lernens (Cognitive Theory of Multimedial Learning, nachfolgend abgekürzt mit: CTML).

Im Kontext seiner Theorie definiert Mayer (2005a) Multimedia „as the presentation of material using both words and pictures” (p. 2). „Wörter“ umfassen für Mayer (2005a) Material, das in verbaler Form präsentiert und sich auf geschriebenen oder gesprochenen Text bezieht. Der Begriff „Bilder“ bezieht sich hingegen auf Material in piktorialer Form und umfasst sowohl ruhende Bilder wie Fotos, Grafiken, Illustrationen, Pläne oder Karten als auch bewegte Bilder wie Videos und Animationen.

Multimediales Lernen fasst Mayer (2001) dabei als dual kodiertes Lernen oder als Lernen mittels zweier Kanäle auf: „Thus, what i call as multimedial learning is more accurately called as dual-code or dual-channel learning” (p. 3). Demnach muss beim Lernen mit neuen Medien zwischen der Kodierungsform (Bild oder Text) und der Modalität der präsentierten Information (visuell oder auditiv) unterschieden werden.

Die CTML basiert auf folgenden drei zentralen Grundannahmen der menschlichen Informationsverarbeitung (Mayer, 2005b):

1. Der Existenz von zwei verschiedenen Verarbeitungskanälen für visuell/piktoriell und auditiv/verbal präsentierte Informationen (duale Kodierung).
2. Der limitierten Kapazität des Arbeitsgedächtnisses, da nur in jedem Kanal ein begrenztes Ausmaß an Informationen verarbeitet werden kann.
3. Der Annahme der aktiven menschlichen Informationsverarbeitung, die darauf basiert, dass Lernende sich aktiv mit der präsentierten multimedialen Botschaft be-

schäftigen, um ein kohärentes mentales Modell ihrer vorhandenen Erfahrungen zu konstruieren.

In Rahmen der CTML werden drei Gedächtnisspeicher unterschieden: das Sensorische Gedächtnis (*sensory memory*), das Arbeitsgedächtnis (*working memory*) und das Langzeitgedächtnis (*long-term memory*).

Abbildung 4 gibt die CTML wieder und veranschaulicht das Zusammenspiel zwischen den drei Gedächtnisspeichern bei der Verarbeitung einer eingehenden multimedialen Botschaft.

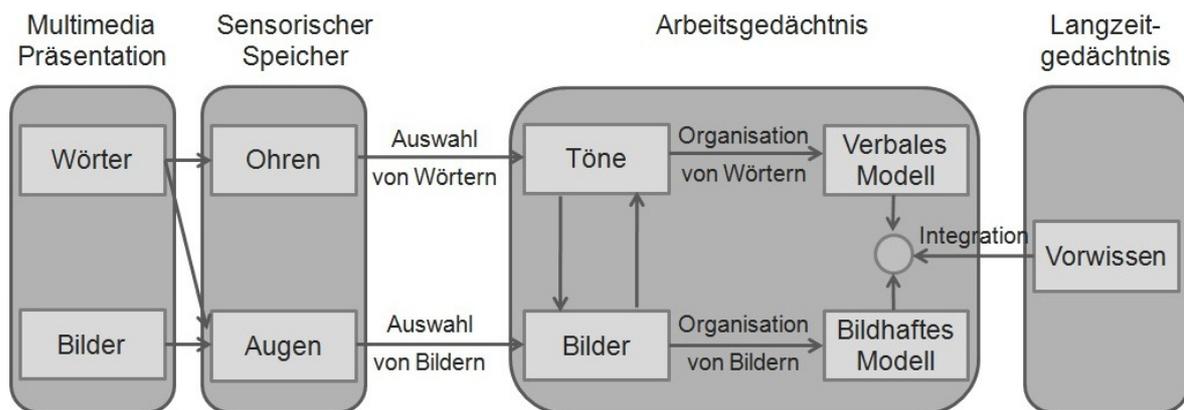


Abbildung 4: Cognitive Theory of Multimedial Learning (Mayer, 2001; Übersetzung von Zumbach, 2007).

Der obere Kanal stellt die Verarbeitung von auditiv/verbalen Repräsentationen dar, wohingegen der untere Kanal die Verarbeitung von visuell/piktoralem Material illustriert.

Dem Lernenden wird eine multimedial präsentierte Information in Form von Bildern und Wörtern präsentiert. Mit Hilfe der Augen werden Bilder aufgrund ihrer visuellen Modalität über den visuellen Kanal, Wörter (Töne und gesprochener Text) hingegen mit den Ohren über den auditiven Kanal aufgenommen und können dann im Sensorischen Gedächtnis für einen sehr kurzen Zeitraum gespeichert werden. Die Informationsverarbeitung für bildliche und textuelle Informationen sowie für visuelle und auditive Informationen erfolgt getrennt.

Einen Sonderfall bildet die Verarbeitung visuell präsentierter Sprache, also geschriebenem Text oder On-Screen-Text, der über beide Kanäle erfolgt (siehe Abbildung 4). Das visuell dargebotene verbale Material wird zunächst mit den Augen erfasst und als visuelles Abbild in das Sensorische Gedächtnis eingespeist. Durch mentales Aussprechen des entsprechenden Wortmaterials kann diese piktorale Verschlüsselung im Arbeitsgedächtnis in eine auditiv-verbale Repräsentation umgewandelt und so in das verbale Modell integriert werden. Visuell präsentiertes verbales Material wird demzufolge auf einem komplexeren Weg verarbeitet als auditiv dargebotenes verbales Material, da sowohl die kognitiven Kapazitäten des auditiv/verbalen als auch visuell/piktoralen Kanals beansprucht und belastet werden.

Die zentralen Informationsverarbeitungsprozesse finden im Arbeitsgedächtnis statt, das durch eine begrenzte Kapazität charakterisiert ist und gleichzeitig nur eine gewisse Informationsmenge bewältigen kann. In Anlehnung an Miller (1956) wird von einer Größenordnung von 7 ± 2 Informationseinheiten (= *Chunks*) ausgegangen. Im Arbeitsgedächtnis finden nach Mayer (2005b) fünf verschiedene kognitive Prozesse statt, die beim Lernen in einer multimedialen Lernumgebung auftreten können und nicht an eine lineare Abfolge gebunden sind, sondern in zahlreichen Iterationen (Wiederholungen) durchlaufen werden können.

Aufgrund der begrenzten Verarbeitungskapazität des Arbeitsgedächtnisses ist es notwendig, eine (1) *Auswahl von relevanten Wörtern* und (2) *Bildern* innerhalb des Lernmaterials zu treffen, um eine erste auditive bzw. visuelle Repräsentation des Sachverhaltes innerhalb des verbalen bzw. piktorialen Arbeitsgedächtnisses zu generieren.

Durch die (3) *Organisation der ausgewählten Wörter* werden in einem nächsten Schritt Verbindungen zwischen den ausgewählten Wörtern hergestellt, um ein kohärentes verbales Modell im Arbeitsgedächtnis zu erzeugen, wie beispielsweise einfache, sinnstiftende Verbindungen in Form von Ursache-Wirkungs-Ketten. In Analogie dazu werden bei der (4) *Organisation der ausgewählten Bilder* ebenfalls kohärente einfache piktoriale Modelle gebildet.

Den fünften und nach Mayer (2005b) wahrscheinlich entscheidenden Schritt in der Verarbeitung multimedialen Lernmaterials stellt die (5) *Integration der wort- und bildbasierten Repräsentationen* dar, in denen Verbindungen zwischen den verschiedenen verbalen und piktorialen Modellen sowie dem Vorwissen des Lernenden aus dem Langzeitgedächtnis aufgebaut werden. Diese Integrationsprozesse erfordern höchste kognitive Kapazität und können sowohl im visuellen als auch im verbalen Arbeitsgedächtnis stattfinden. Der Lernende kann zur Koordination des Integrationsprozesses zwischen den beiden mentalen Modellen sein Vorwissen nutzen. Auf diese Weise kann neues Wissen konstruiert und im Langzeitgedächtnis in Form von Schemata abgespeichert werden.

Die obig dargestellten Informationsverarbeitungsprozesse als auch die „kognitive Architektur“ des Lerners sind nach Mayer (2005b) bei der Gestaltung von multimedialen Lernumgebungen zu berücksichtigen, um effektives Lernen zu ermöglichen. Da beide Informationsverarbeitungskanäle in ihrer Kapazität begrenzt sind, sollte bei der Konzeption von multimedialen Lernmaterial generell darauf geachtet werden, dass bei der Verarbeitung von Informationen möglichst beide Kanäle angesprochen werden und kognitive Überlastung (*cognitive overload*) vermieden wird (Chandler & Sweller, 1991; Mayer & Moreno, 2003). Nach Mayer (2005b) erwerben Personen mehr Wissen, wenn sowohl visuelles als auch verbales Material präsentiert wird, da auf diese Weise der Lernende zwei statt nur

einer mentalen Repräsentation zu einem Sachverhalt generieren kann. Diese beiden verschiedenen mentalen Modelle können dann integriert und mit dem Vorwissen verbunden werden. In diesem Fall verfügt der Lerner dann über zwei mentale Repräsentationen eines Sachverhaltes, die Lerninhalte sind „doppelt kodiert“ (Brünken et al., 2008; vgl. Paivio, 1986).

Zahlreiche Forschungsergebnisse zur kognitiven Theorie des multimedialen Lernens von Mayer (2005a) haben zur Formulierung von basalen Designprinzipien geführt, die in der Praxis bei der effektiven Gestaltung von multimedialen Lernumgebungen berücksichtigt werden sollten und nachfolgend skizziert werden.

(1) Multimedia-Prinzip (*multimedia principle*): Es werden bessere Lernleistungen erzielt, wenn sowohl verbale als auch bildliche Informationen präsentiert werden, also wenn zu einem Text geeignete Bilder präsentiert werden anstatt nur den Text allein.

(2) Räumliches und (3) zeitliches Kontiguitätsprinzip (*spatial and temporal contiguity principle*): Lernen mit Medien ist effektiver, wenn Text und korrespondierende Bilder räumlich nah beieinander statt weiter entfernt, und gleichzeitig statt nacheinander am Bildschirm präsentiert werden (*Effekt der räumlichen und zeitlichen Nähe*). Eine simultane Präsentation von verbalen und bildhaften Informationen entlastet das Arbeitsgedächtnis, denn eine getrennte Darstellung würde die kognitiven Kapazitäten des Lerners überstrapazieren.

(4) Modalitätsprinzip (*modality principle*): Es wird erfolgreicher gelernt, wenn Erläuterungen von Bildern und Animationen in gesprochener Sprache und nicht durch geschriebenen Text präsentiert werden. Dies entlastet die Kapazität des visuellen Kanals, da bei gleichzeitiger Darbietung von geschriebenem und gesprochenem Text beide Darbietungsformate miteinander konkurrieren und die visuelle Aufmerksamkeit zwischen den zwei Quellen geteilt werden muss (*Effekt der geteilten Aufmerksamkeit* oder *split-attention effect*; Chandler & Sweller, 1992).

(5) Kohärenzprinzip (*coherence principle*): Es werden bessere Lernleistungen erzielt, wenn auf zusätzliches Lernmaterial, das nicht zwingend zur Erreichung der Lernziele benötigt wird, verzichtet wird. Das Präsentieren von zu vielen interessanten, aber irrelevanten Zusatzmaterialien wie Geschichten, Bildern und Hintergrundmusik kann das Lernen der eigentlich wichtigen Informationen beeinträchtigen (*seductive detail effect*).

(6) Redundanzprinzip (*redundancy principle*): Medienbasiertes Lernen ist effektiver, wenn doppelte und unnötige Informationen vermieden werden, also geschriebener Text nicht gleichzeitig vorgelesen wird, denn Grafiken mit gesprochenem Text führen zu besseren Lernergebnissen als Grafiken mit gesprochenem und geschriebenem Text.

(7) Prinzip der individuellen Unterschiede (*pretraining principle*): In Studien zu diesem Modell wurde festgestellt, dass multimediale Lernumgebungen unterschiedlich auf Lernende mit unterschiedlichen Lernvoraussetzungen wirken. Beispielsweise profitieren Lernende mit geringem Vorwissen in höherem Maße von der Anwendung der oben genannten Gestaltungsprinzipien.

(8) Signalisierungsprinzip (*signaling principle*): Es wird besser gelernt, wenn Hinweise gegeben werden, welche die Organisation wesentlicher Lernelemente hervorheben und erleichtern.

(9) Segmentierungsprinzip (*segmenting principle*): Die Präsentation von multimedialen Botschaften in Form von lerngerechten Teileinheiten ist lernförderlicher als die Darbietung in einer durchgängigen Einheit. Lernende sollten zudem in ihrer eigenen Geschwindigkeit die Einheiten bearbeiten können.

(10) Personalisierungsprinzip (*personalization principle*): Es werden bessere Lernleistungen erzielt, wenn Text in einem dialogorientierten Sprachstil statt in einem konventionellen, formalen dargeboten wird (z. B. direkte Anrede der Benutzer in der ersten oder zweiten Person: „Wie würden Sie entscheiden? Siehe Kapitel 4.2.2).

4.2.2 Social Agency Theory

Die Überlegenheit personalisierter Texte und Filme wird meist mithilfe der Social Agency Theory (Mayer, 2005c) erklärt. Grundannahme dieser Theorie ist, dass soziale Hinweisreize (*social cues*) wie beispielsweise eine umgangssprachliche Formulierung, Stimmen oder Bilder eines Sprechers unweigerlich eine soziale Reaktion beim Lernenden auslösen (*social responses*). Die Aktivierung einer sozialen Reaktion durch soziale Hinweisreize führt wiederum zu einer größeren kognitiven Anstrengung beim Lernenden, präsentierte Informationen vertieft zu verarbeiten und infolge dessen zu besseren Lernergebnissen.

Die erste Annahme, dass soziale Hinweisreize zu einer sozialen Reaktion führen, wird durch Befunde im Rahmen der Forschung zur Media Equation Theory von Reeves & Nass (1996) gestützt. Grundaussage dieser Theorie ist, dass Menschen sich gegenüber Medien genauso verhalten wie gegenüber realen Personen und den Computer als sozialen Interaktionspartner anerkennen. Beispielsweise konnte experimentell gezeigt werden, dass die Evaluation eines Computerprogrammes höflicher (im Sinne von positiver) ausfällt, wenn der Fragebogen am gleichen Rechner präsentiert wird und nicht an einem anderen Rechner oder als Paper-Pencil-Version. Demzufolge verhalten sich Menschen gegenüber Compu-

tern höflich und wenden Höflichkeitsregeln aus der zwischenmenschlichen Kommunikation an (Nass, Moon & Carney, 1999).

Die zweite Annahme, dass eine soziale Reaktion die aktive, kognitive Verarbeitung steigert, wird durch Untersuchungen von Grice (1975) belegt, der zeigen konnte, dass bei sozialen Kommunikationsprozessen der Zuhörer annimmt, dass der Sprecher versucht, etwas Sinnvolles zu sagen und informative, genaue, relevante und prägnante Informationen vermitteln will. Gemäß dem Grice'schen Kooperationsprinzip strengt sich der Zuhörer wiederum mehr an, den Sprecher zu verstehen und aus dem Gesagten Sinn zu stiften, da Zuhörer und Sprecher eine implizite Vereinbarung haben, so zu verfahren. Lernende verstehen demzufolge eine Lernsituation nicht als reine Informationsübermittlung, sondern als soziale Kommunikation und aktivieren soziale Konversationschemata wie das Kooperationsprinzip (Grice, 1975).

Die dritte Annahme, dass eine höhere aktive, kognitive Verarbeitung zu besseren Lernergebnissen führt, lässt sich aus den Grundannahmen der CTML ableiten (vgl. Mayer, 2005b).

Abbildung 5 fasst das aufgeführte Erklärungsmodell zur Social Agency Theory zusammen und veranschaulicht, wie soziale Hinweisreize ein tieferes Verständnis der instruktionalen Botschaft bzw. des zu vermittelnden Lerninhalts fördern bzw. *nicht fördern* (Mayer, 2005c; Rey, 2009, S. 84):

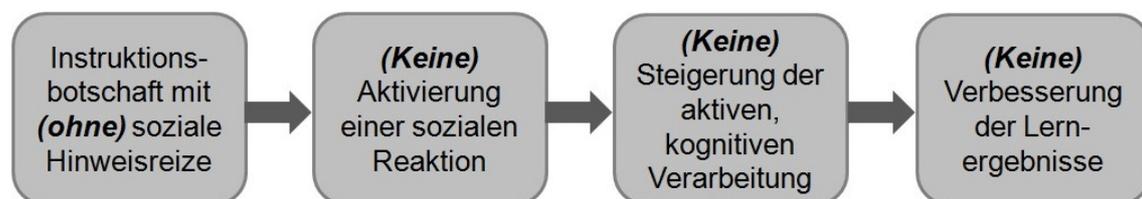


Abbildung 5: Schematische Darstellung des Einflusses vorhandener bzw. **fehlender** sozialer Hinweisreize auf den Lernprozess in multimedialen Lernumgebungen im Rahmen der Social Agency Theory (Mayer, 2005c; Rey, 2009, S. 84).

Aus der Social Agency Theory können laut Mayer (2005c) folgende drei soziale Gestaltungsprinzipien abgeleitet werden:

(1) Personalisierungsprinzip (*personalization principle*): Es werden bessere Lernleistungen erzielt, wenn der Text in einer multimedialen Lernumgebung nicht in einem formalen, sondern in einem umgangssprachlichen, dialogorientierten Stil formuliert ist. Um einen personalisierten Text umzusetzen, können hauptsächlich zwei Techniken eingesetzt werden: Zum einen (1) können Formulierungen in der ersten und zweiten Person („ich“ und „du“) statt in der dritten Person („man“) verwendet werden, die den Lernenden direkt ansprechen. Zum anderen (2) können Sätze ergänzt werden, indem der Verfasser direkte

Kommentare für den Lernenden bereithält (z. B. „Vergleiche nun das rotierte Bild mit der schematischen Darstellung“). Das Personalisierungsprinzip konnte bereits in einigen experimentellen Untersuchungen bestätigt werden (Mayer, Fennell, Farmer & Campbell, 2004; Moreno & Mayer, 2000, 2004).

(2) Prinzip der Stimme (*voice principle*): Es wird besser gelernt, wenn auditiver Text in einer multimedialen Botschaft von einer menschlichen Stimme anstatt von einer computer-generierten Stimme präsentiert wird, da eine menschliche Stimme vertrauter, persönlicher und attraktiver wirkt. Bessere Lernergebnisse zeigen sich auch, wenn in multimedialen Lernumgebungen akzent- und dialektfreie Stimmen verwendet werden anstatt Stimmen mit Akzent oder Dialekt. Das Stimmen-Prinzip konnte bereits empirisch belegt werden (Atkinson, Mayer & Merrill, 2005; Mayer, Sobko & Mautone, 2003).

(3) Prinzip des Bildes (*image principle*): Es wird angenommen, dass bessere Lernleistungen erzielt werden, wenn der Sprecher in einer multimedialen Präsentation auf dem Bildschirm zu sehen ist, als wenn er nicht zu sehen ist. Zu diesem Zweck werden häufig sog. pädagogische Agenten eingesetzt, die mit dem Lerner kommunizieren, ihn durchs Programm führen, Feedback, Hinweise, Beispiele, Demonstrationen und Erklärungen geben. Die Annahme, dass das Aufzeigen einer Person einen Einfluss auf das Lernen hat, konnte bislang allerdings in experimentellen Studien nicht eindeutig geklärt werden (vgl. Moreno, 2005). In vier Studien setzten Moreno, Mayer, Spires & Lester (2001) einen pädagogischen Agenten namens Herman-the-Bug in Form einer Comic-Figur ein. Im ersten Experiment zeigte sich, dass Lernende, die Herman-the-Bug auf dem Bildschirm sahen, schlechtere Transfertestleistungen aufwiesen als Lernende, die nur seine Stimme hörten. In einem zweiten Experiment konnte dagegen nachgewiesen werden, dass Lernende etwas besser im Transfertest abschnitten, wenn Herman's Bild zusätzlich zu seinen Textkommentaren visuell auf dem Bildschirm zu sehen war als Lernende, denen lediglich der Text präsentiert wurde. In zwei folgenden Studien wurde anstatt einer animierten Comic-Figur ein Video eines menschlichen Kopfes (*human talking head*) verwendet. Es zeigte sich, dass die Leistungen im Transfertest bei der Versuchsbedingung, die das Video des talking heads erhielt, etwas höher waren als bei der Gruppe, die nur die menschliche Stimme hörten. Der uneinheitlichen Befundlage zufolge wird somit nicht notwendigerweise besser gelernt, wenn der Sprecher in einer multimedialen Präsentation auch zu sehen ist. Mayer (2005c) selbst nimmt an, dass möglicherweise die Sichtbarkeit einer Person ein seduktives Detail (*seductive detail effect*) darstellen könnte, das den Lerner ablenken kann und in Anlehnung an die Cognitive Load Theory (Chandler & Sweller, 1991; Sweller, 1999; Sweller, van Merriënboer & Paas, 1998) zu extrinsischer kognitiver Belastung (*Extraneous Load*) führen könnte (Harp & Mayer, 1998).

Den lernförderlichen Einfluss der filmischen Personalisierung von Ausstellungsinhalten im musealen Lernkontext konnte Töpper (2009) erstmals in einer Kombination von Feld- und Laborstudien nachweisen. Diese Studien werden ausführlich in Kapitel 4.5 vorgestellt.

4.2.3 Parasoziale Interaktion (PSI) mit Medienfiguren

Ein medienpsychologischer Ansatz, der sich explizit mit der sozialen Auseinandersetzung auf kognitiver und emotionaler Ebene mit Pädagogischen Agenten oder Medienfiguren befasst, stellt das Konzept der parasozialen Interaktion (nachfolgend abgekürzt mit: PSI) dar (vgl. Giles, 2002; Horton & Wohl 1956; Hartmann et al., 2004a; Vorderer, 1996).

Dieser kommunikationswissenschaftliche Ansatz geht auf die zwei amerikanischen Soziologen Horton & Wohl (1956) zurück, die feststellten, dass Fernsehzuschauer die auf dem Bildschirm agierenden Personen nicht nur passiv beobachten, sondern auf sie ähnlich wie in einer realen Kommunikationssituation reagieren und mit ihnen interagieren.

Dieses Phänomen bei personenzentrierten TV-Sendungen bezeichneten Horton & Wohl (1956) als *parasoziale* Interaktion, da die Fernsehakteure die Zuschauer nicht sehen und auf sie reagieren können; die Zuschauer wiederum können es, ihre Reaktionen erreichen jedoch aufgrund des fehlenden Rückkanals die Fernsehfiguren nicht und „prallen gewissermaßen an der Mattscheibe ab“ (Wirth & Schramm, 2005, S. 547). Aufgrund der fehlenden Reziprozität besteht lediglich „die Illusion eines Face-to-Face-Kontakts zu den auf dem Bildschirm agierenden Personen“ (Gleich, 1997, S. 36; Horton & Wohl, 1956).

Aus Sicht der Sozialpsychologie lässt sich PSI somit grob als asymmetrische Interaktionsform bezeichnen (Hartmann et al., 2004b). Schramm (2008) bezeichnet PSI auch als „Einbahnstrassenkommunikation“, da es sich im Gegensatz zur realen, personalen Interaktion um eine einseitige Interaktion handelt, die aber auch enorme Vorteile mit sich bringen kann: Denn Verhaltensweisen, die im sozialen Setting normalerweise negativ sanktioniert werden würden oder sich verbieten, bleiben im parasozialen Setting ohne Folgen (Töpper, 2009).

In der PSI-Forschung wird eine medial vermittelte Figur als *Persona* bezeichnet. Als *Persona* kommen generell alle Personen in Frage, die in Medien auftauchen, also nicht nur Schauspieler, Moderatoren, Nachrichtensprecher, Musiker oder Sportler, sondern auch ganz „normale Menschen“ und sogar artifizielle Personen, wie z. B. Trickfilmfiguren (Hartmann et al., 2004b; Schweiger, 2007).

Die differenzierteste Konzeptualisierung von PSI findet sich im Zwei-Ebenen-Modell parasozialer Interaktionen von Hartmann et al. (2004a) wieder, das nachfolgend detailliert

vorgestellt wird. Hartmann et al. (2004a) modellieren PSI als multidimensionales Konstrukt, das durch bestimmte Eigenschaften der Medienfigur (z. B. Obtrusivität, Persistenz und Adressierung des Publikums) sowie bestimmten Eigenschaften der Rezipienten (z. B. Motivation, Relevanz) dynamisch beeinflusst wird (dazu später mehr).

Eng verwandt mit dem Konstrukt der PSI, aber dennoch davon abzugrenzen, ist das Konzept der parasozialen Beziehungen (nachfolgend abgekürzt mit: PSB), das über die einzelne Begegnung mit einer Persona hinausgeht. PSB stellen dauerhafte Bindungen an eine Medienfigur dar, die sich im Laufe der Zeit aus den PSI entwickeln, indem als Ergebnis der Auseinandersetzung mit der Persona ein Beziehungsschema aufgebaut wird (Hartmann et al., 2004a; Wirth & Schramm, 2005). Beispielsweise kann ein Rezipient im Laufe der Jahre das Gefühl bekommen, einen TV-Showmaster gut zu kennen und mit ihm befreundet zu sein (Schweiger, 2007). Im Gegensatz zur PSB bezieht sich PSI stets auf den Augenblick der Rezeptionssituation und ist somit von eher kurzer Dauer. Hartmann et al. (2004a) verstehen bereits die „Auseinandersetzung mit einer Persona nach der ersten Eindrucksbildung bzw. Wiedererkennung“ (S. 30) als die eigentliche PSI.

In ihrem Modell vertreten Hartmann et al. (2004a) die Position, „dass zu einer Persona immer (irgendwie geartete) PSI-Prozesse ablaufen, man also mit einer „anwesenden“ Medienperson nicht *nicht* parasozial interagieren kann“ (S. 30). Die Intensität der Auseinandersetzung eines Zuschauers mit einer Persona bewegt sich dabei auf einem Kontinuum zwischen „minimal“ bis „sehr stark“, wobei die konkrete Ausprägung während eines Rezeptionsprozesses schwanken kann. In Anlehnung an medien- und sozialpsychologische Theorien postulieren Hartmann et al. (2004a) zwei prototypische Intensitätsausprägungen der PSI, daher die Bezeichnung „Zwei-Ebenen-Modell“: eine oberflächlich-schwache Beschäftigung mit einer Medienperson (*Low-Level-PSI*) und eine intensiv-starke Auseinandersetzung mit derselben (*High-Level-PSI*).

Neben der Intensität unterscheiden Hartmann et al. (2004a) in ihrem Modell drei verschiedene PSI-Dimensionen, die ihrer Struktur nach in (1) perzeptiv-kognitive, (2) affektive und (3) konative PSI-Teilprozesse differenziert werden, denen sich wiederum verschiedene Unterdimensionen zuordnen lassen (siehe folgende Tabelle 6). Die einzelnen Subdimensionen werden hier an dieser Stelle nicht vollständig erörtert, eine detaillierte Erläuterung der einzelnen Subdimensionen geben Hartmann et al. (2004a).

Die erste, *perzeptiv-kognitive* PSI-Dimension umfasst Aspekte der personabezogenen Informationsaufnahme und -verarbeitung und jegliche Form von Denkprozessen, die sich auf die Persona beziehen. Unter der Bedingung von *High-Level-PSI* zeichnet sich die Medienrezeption durch eine aktiv-gezielte personabezogene Informationsaufnahme (1) und intensiven Anstrengungen des Verstehens der Situation und Handlungen der Persona (2) sowie

einer tiefgehenden Anknüpfung der personabezogenen Informationen an das bisherige Vorwissen über die Persona (3) aus. Bei einer intensiv-starken Auseinandersetzung mit einer Medienfigur wird eine gründliche Bewertung der Persona und ihre Handlungen (4) vorgenommen, intensiv über die nächsten Handlungen der Persona in der Zukunft (5) nachgedacht sowie ein Bezug zwischen Persona und Selbst (6) in Form von sozialen Vergleichen (vgl. Festinger, 1954) hergestellt. Bei einer *Low-Level-PSI* hingegen erfolgt die Informationsaufnahme „beiläufig“ und unkontrolliert (1), die Handlungen der Persona und die aufgenommenen Informationen werden nur oberflächlich nachvollzogen und mit den Gedächtnisinhalten verknüpft (2, 3). Es finden keine oder nur oberflächlich-schwache Bewertungen der Persona und ihrer Handlungen (4), keine antizipierenden Beobachtungen (5) und keine sozialen Vergleiche (6) statt.

Tabelle 6: PSI-Dimensionen: drei PSI-Teilprozesse und ihre Unterdimensionen (Hartmann et al., 2004a).

Teilprozess	Subdimensionen
(1) Perzeptiv-Kognitiv	(1) Personabezogene Informationsaufnahme (2) Verstehen der Situation und Handlungen der Persona (3) Verknüpfung der Aussagen/Handlungen der Persona mit eigenen Gedächtnisinhalten (4) Bewertung der Persona und ihrer Handlungen (5) Überlegungen über die nächste Zukunft der Persona (6) Herstellung eines Bezugs zwischen Persona und Selbst
(2) Affektiv	(1) Sympathie/Antipathie (2) Empathie/Counterempathie (3) Emotionsauslösung (Ansteckung, Induktion)
(3) Konativ	(1) Nonverbale (motorische, mimische, gestische) Verhaltensweisen (2) (Para-)Verbale Verhaltensweisen

Unter die zweite, *affektive* PSI-Dimension werden alle emotionalen Reaktionen subsumiert, die sich auf die Medienfigur richten bzw. durch sie ausgelöst werden. Dazu zählen positive und negative Emotionen gegenüber der Medienfigur wie Sympathie/Antipathie oder Empathie/Counterempathie und Emotionen, die durch die Persona ausgelöst werden. Bei den emotionalen PSI-Teilprozessen sind die Ausprägungen für High- und Low-Level-PSI schwer abzugrenzen. Hartmann et al. (2004a) nehmen grundsätzlich jedoch an, dass High-Level-PSI in intensiven emotionalen Reaktionen resultiert, wohingegen sich bei Low-Level-PSI kaum emotionale Reaktionen in Bezug auf die Persona zeigen.

Die dritte, *konative* PSI-Dimension bezieht sich auf von außen sichtbare und hörbare Verhaltensweisen, die sich auf die Persona richten. Hierzu zählen verbale Äußerungen (z. B.

Begrüßen der Persona, Beschimpfungen, Zurufe) und non-verbale Verhaltensweisen (z. B. Zuwinken, Zurücklächeln). Bei einer oberflächlich-schwachen Auseinandersetzung mit der Medienfigur (Low-Level-PSI) werden vermutlich Verhaltensäußerungen dieser Art nicht auftreten, wohingegen unter High-Level-Bedingungen eher damit zu rechnen ist.

Wie Hartmann et al. (2004a) bemerken, müssen nicht auf allen drei Dimensionen messbare Phänomene vorliegen, um letztendlich von PSI-Prozessen sprechen zu können. Konative PSI-Prozesse finden beispielsweise eher selten in der alltäglichen Rezeption statt. Des Weiteren wird angenommen, dass PSI-Prozesse durch das Medialitätsbewusstsein, also dem Bewusstsein, sich in einer medial vermittelten Situation zu befinden, „eingerahmt“ und strukturell geformt werden.

Die Motivation, sich dezidierter mit einer Medienfigur auseinanderzusetzen, hängt sowohl von den Eigenschaften der Persona bzw. ihrer Darstellung als auch von den Eigenschaften der Rezipienten ab. Als besonders relevante Determinanten der Intensität von PSI sind in diesem Zusammenhang Obtrusivität, Persistenz und Art der Adressierung einer Persona sowie die Motivation des Publikums zu nennen (Hartmann et al., 2004a, 2004b).

Generell besitzen reale Personen in non-fiktiven Medienangeboten ein höheres PSI-Potential als Schauspieler in fiktiven Rollen, denn „je weniger künstlich (‘Artifizialität‘), je menschenähnlicher (‘Anthropomorphismus‘) eine Persona ist und je glaubwürdiger ihr Aussehen und Verhalten sind, desto eher eignet sie sich für PSI-Prozesse“ (Schweiger, 2007, S. 125). Damit eine PSI überhaupt stattfinden kann, muss die Medienfigur eine hohe Bildschirmpräsenz aufweisen, die durch die Faktoren *Obtrusivität* und *Persistenz* determiniert wird (Hartmann et al., 2004b). Die Aufdringlichkeit einer Persona in einer Szene wird als Obtrusivität bezeichnet, wohingegen die Persistenz sich auf die Dauer ihres Auftritts bezieht. Mediale Figuren müssen demzufolge auffällig und über einen längeren Zeitraum in Szene gesetzt werden, damit sich das Publikum überhaupt mit ihnen auseinandersetzen kann.

Eine weitere, wichtige Determinante der Intensität von PSI-Prozessen und damit der Herstellung der „Illusion eines Face-to-Face-Kontaktes“ bildet die vom Rezipienten empfundene *Adressierung* durch eine Medienfigur, also ob sich die Rezipienten durch jene persönlich angesprochen fühlen: Je höher die Adressierungsleistung der Persona ist, desto stärker sollte die PSI ausgeprägt sein (Hartmann et al., 2004a).

Nach Hartmann et al. (2004b) determinieren drei Schlüsselreize die Adressierungsleistung: die dargestellte räumliche Distanz der Persona, die non-verbale Bezugnahme und die verbale Bezugnahme. Zunächst sollte eine geringe *räumliche Distanz* zwischen Persona und Rezipient realisiert werden, um räumliche Nähe und Unmittelbarkeit zu suggerieren. Denn

je geringer die räumliche Distanz ist, desto höher ist die Wahrscheinlichkeit, dass der Rezipient sich adressiert fühlt, also das Gefühl hat, persönlich angesprochen zu sein. Die Persona kann auch *non-verbal* auf den Zuschauer Bezug nehmen, indem sie sich ihm über die Körperhaltung oder die Blickrichtung im Sinne eines „Augengrußes“ zuwendet. Im erstgenannten Fall findet eine indirekte, im letztgenannten Fall eine direkte Adressierung statt. Eine direkte Adressierung liegt auch vor, wenn eine Persona über *verbale Bezugnahme* den Zuschauer beispielsweise anspricht und begrüßt, wie es häufig bei Nachrichtensprechern oder Moderatoren zu beobachten ist (z. B. „Guten Abend, liebe Zuschauer“). Demzufolge dürfte die Wahrscheinlichkeit für eine intensive PSI bzw. High-Level-PSI erheblich steigen, wenn eine Persona sich dem Zuschauer zuwendet, Blickkontakt suggeriert und ihn direkt anspricht.

Einen weiteren wichtigen Einflussfaktor auf die Intensität der PSI stellt die *Motivation* des Publikums dar, denn nur wenn ein Rezipient hinreichend motiviert ist, dürfte es zu einer tiefergehenden Auseinandersetzung mit medial dargestellten Person kommen. Insbesondere die vom Rezipienten wahrgenommene *Attraktivität* der Medienfigur sollte sich positiv auf die Motivation und die Intensität der PSI-Prozesse auswirken, denn attraktive Figuren werden aus evolutionären Gründen eher wahrgenommen und beachtet (Hartmann & Klimmt, 2005; Schweiger, 2007).

Attraktivität stellt ein mehrdimensionales Konstrukt dar, das sich auf physische Aspekte wie das äußere Erscheinungsbild im Sinne von „hübsch“ bzw. „hässlich“ (*physical attractiveness*) beziehen kann, aber auch auf psychische Aspekte wie den Handlungserfolg (*task attractiveness*) und die charakterlich-soziale Attraktivität (*social attractiveness*), die sich eher in Urteilen wie „erfolgreich“ bzw. „erfolglos“ äußern (Hartmann & Klimmt, 2005; McCroskey & McCain, 1974). Hartmann et al. (2004b) gehen davon aus, dass je nach Neigung und subjektivem Urteil des Rezipienten sowohl sehr attraktive als auch sehr unattraktive Medienfiguren das Interesse auf eine tiefer gehende Beschäftigung wecken und folglich zu intensiveren PSI führen können.

Mit einer intensiveren PSI ist ebenfalls zu rechnen, wenn die Persona, mit der agiert wird, für den Rezipienten *persönlich relevant* ist. Als weitere PSI-relevante Eigenschaften des Rezipienten nennen Hartmann et al. (2004b) *Einsamkeit* und das *Bedürfnis nach sozialem Nutzen*. Menschen, die über wenig soziale Kontakte verfügen und Menschen, die ein Medium aufgrund einer starken sozialen Nützlichkeitsersparung nutzen, sind möglicherweise empfänglicher für PSI. Allerdings konnten diese beiden Annahmen bislang noch nicht eindeutig empirisch bestätigt werden. Folgende Tabelle 7 fasst die PSI-relevanten Eigenschaften von Medienpersonae und die PSI-relevanten Nutzereigenschaften nochmals zusammen.

Tabelle 7: PSI-relevante Eigenschaften von Medienpersonae und Mediennutzern/Rezipienten (Hartmann et al., 2004a; 2004b).

Medienpersonae		Mediennutzer/Rezipienten
<i>Stabil</i>	<i>Dynamisch</i>	
• Attraktivität	• Obtrusivität	• Motivation/Persönliche Relevanz
• Artifizialität	• Persistenz	• Einsamkeit
• Anthropomorphismus	• Adressierung	• Bedürfnis nach sozialem Nutzen

Hartmann et al. (2004a) weisen daraufhin, dass ihre vorgeschlagene Modellierung der PSI erhebliche Parallelen zum Involvement-Konzept aufweist, das eine wichtige Rolle in Persuasionstheorien spielt (vgl. Chaiken & Trope, 1999; siehe Kapitel 3.2.2). Unter Involvement wird der persönliche Bezug oder die persönliche Wichtigkeit (Relevanz) im Hinblick auf Themen, Objekte und Ereignisse verstanden. Stärker involvierte Individuen sind in der Informationsverarbeitung aktiver als Individuen mit einem geringen Involvement. Die Intensitätsausprägung der High-Level-PSI entspricht dabei starkem Involvement, das zu einer intensiven Beschäftigung und Elaboration mit einem Gegenstand führt, Low-Level-PSI dagegen geringem Involvement. Parasoziale Interaktionen definieren Hartmann et al. (2004a) daher

als ein vom Bewusstsein der medialen Vermitteltheit geprägtes interpersonales Involvement von Rezipientinnen und Rezipienten mit einer Medienperson, das sich in perceptiv-kognitiven, affektiven und konativen Teilprozessen und Erlebensweisen manifestiert kann und dessen Intensität im Rezeptionsverlauf dynamischen Schwankungen unterliegen kann. (S. 37)

PSI kann demzufolge als ein mehr oder weniger starkes interpersonales Involvement mit einer Medienfigur aufgefasst werden (Wirth & Schramm, 2005). Eine erhöhte PSI sollte demnach zu elaborierten kognitiven Informationsverarbeitungsprozessen führen. Töpper (2009) konnte beispielsweise feststellen, dass Filme im Ausstellungskontext, die Informationen durch eine interviewte Person narrativ präsentierten, den Besucher stärker motivierten, sich kognitiv und affektiv mit den vermittelten Wissensinhalten auseinanderzusetzen.

Auf der Grundlage des Zwei-Ebenen-Modells wurde ein Erhebungsinstrument zur Messung von PSI entwickelt (vgl. Schramm & Hartmann, 2008). Die Autoren empfehlen zur Vermeidung kognitiver Verzerrungen einen postrezeptiven Befragungszugang, der nach Möglichkeit unmittelbar nach der Rezeptionssequenz vorgenommen werden sollte (Hartmann & Schramm, 2006). Einige Grundannahmen ihres Modells konnten Hartmann & Klimmt (2005) bereits in einer Studie empirisch bestätigen. Beispielsweise stellten sie fest, dass die den Rezeptionsprozess begleitenden PSI zu einer Medienfigur umso stärker aus-

fallen, je eher diese als attraktiv empfunden wird und je eher sich die Zuschauer von ihr adressiert fühlen. Eine attraktive Medienfigur verringerte auch das Medialitätsbewusstsein, also das Gefühl sich nur in einer medialen Interaktionssituation zu befinden. Darüber hinaus führten intensivere PSI zu einer positiveren Gesamtbewertung der rezipierten Sendung, wobei die affektiven PSI-Prozesse am stärksten zur positiven Beurteilung des rezipierten Unterhaltungsformates beitrugen. Die in dieser Studie verwendeten Subskalen zur Operationalisierung der PSI-Teilprozesse erwiesen sich als ausreichend reliabel und die Bestätigung einiger aus dem Modell abgeleiteten Hypothesen lassen laut den Autoren auf eine erste Konstruktvalidierung der PSI-Skalen schließen.

Töpfer (2009) konnte experimentell in Rahmen einer Vorstudie belegen, dass die Präsentation von personalisierten Filmen durch narrative Betroffeneninterviews tatsächlich zu einer höheren affektiven und kognitiven PSI mit den interviewten betroffenen Personen führte als bei neutralen Filmversionen mit einem Sprecher aus dem Off. Die Sprecher der personalisierten Filme wurden auch insgesamt als attraktiver und wärmer bewertet als die Off-Sprecher der neutralen Filmversion. Die personalisierten Filmversionen wurden gegenüber den neutralen auch als interessanter bewertet. Töpfer (2009) konnte damit die unterschiedlichen Wirkungen von personalisierten und neutralen Filmen auf emotionaler, als auch auf kognitiver Ebene empirisch belegen (siehe Kapitel 4.5).

Aus dem Zwei-Ebenen-Modell parasozialer Interaktionen lassen sich konkrete Gestaltungsempfehlungen für die Darstellung medial vermittelter Figuren ableiten, um eine intensive Auseinandersetzung mit ihnen und dem zu vermittelnden Lerninhalt zu fördern: Medienfiguren sollten auffällig und lange in Szene gesetzt werden, möglichst attraktiv sein und den Zuschauer indirekt und direkt adressieren.

4.3 Gestaltung multimedialer Lernumgebungen

Um die Potenziale neuer Medien realisieren zu können, setzt der Einsatz neuer Medien eine geeignete und angemessene instruktionale Gestaltung der Lernumgebung voraus (Fischer, Mandl & Todorova, 2009; Kopp & Mandl, 2009; Reinmann & Mandl, 2006).

Wie bisherige Forschungsergebnisse belegen, erfolgt der Einsatz von multimedialen Lehr- und Lernangeboten jedoch noch zu wenig theorie- und konzeptgeleitet sowie zielgerichtet (Reinmann-Rothmeier & Mandl, 2000). Für einen erfolgreichen Wissenserwerb ist insbesondere eine adäquate didaktische Aufbereitung der Lerninhalte entscheidend (Kopp, Dvorač & Mandl, 2003). Bevor auf die didaktischen und medialen Gestaltungskriterien eingegangen wird, wird im folgenden Kapitel zunächst der Ansatz des problemorientierten Lernens vorgestellt.

4.3.1 Problemorientiertes Lernen

Problemorientiertes Lernen zeichnet sich vor allem durch zwei zentrale Aspekte aus. Zum einen steht das Lösen von Problemen, mit welchen sich die Lernenden aktiv und eigenverantwortlich auseinandersetzen können, im Mittelpunkt des Wissenserwerbs, zum anderen werden die Problem- oder Entscheidungssituationen so realitätsnah wie möglich dargestellt, um die Komplexität der Praxis weitgehend abbilden zu können (Gräsel, 1997). Ein weiteres Kennzeichen bildet die Kombination von aktivem, selbstbestimmtem Lernen und instruktionaler Unterstützung (Reinmann, 2005).

Problemorientierte Lernumgebungen integrieren Elemente kognitivistischer und konstruktivistischer Lernumgebungen und bieten Lernenden sowohl instruktionale Unterstützung als auch eine aktive Wissenskonstruktion (Reinmann & Mandl, 2006). Die Balance zwischen Instruktion und Konstruktion (Linn, 1990) in Abhängigkeit von den Lernvoraussetzungen und dem Lerngegenstand bildet den zentralen Ausgangspunkt problemorientierter Lernumgebungen (Mandl, Kopp & Dvorak, 2004; Reinmann & Mandl, 2006).

Problemorientierte Lernumgebungen basieren auf einer konstruktivistischen Auffassung vom Lernen, die sich mit folgenden sechs Prozessmerkmalen beschreiben lässt (Gerstenmaier & Mandl, 1995; Reinmann & Mandl, 2006):

- (1) *Lernen ist ein aktiver Prozess.* Lernen wird nicht als passive Aufnahme von Informationen verstanden, sondern ist nur über die eigenständige und aktive Beteiligung des Lernenden möglich. Demzufolge ist ein gewisses Maß an Lernmotivation und situatives Interesse eine unabdingbare Voraussetzung für erfolgreiche Lernprozesse (vgl. Krapp, 1992; Schiefele, 1996).
- (2) *Lernen ist ein selbstgesteuerter Prozess.* Lernen ohne jegliche Selbststeuerung ist kaum denkbar. Die Auseinandersetzung mit einem Inhaltsbereich erfordert von dem Lernenden stets die Steuerung und Kontrolle des eigenen Lernprozesses.
- (3) *Lernen ist ein konstruktiver Prozess.* Lernen baut auf bereits vorhandenen Wissensstrukturen auf, die auf der Basis individueller Erfahrungen interpretiert werden. Ohne einen solchen individuellen Erfahrungs- und Wissenshintergrund und eigene Interpretationen können im Prinzip keine kognitiven Prozesse stattfinden, die zu einer dauerhaften Veränderung des Wissens und Könnens führen.
- (4) *Lernen ist ein emotionaler Prozess.* Leistungsbezogene und soziale Emotionen beeinflussen das Lernen stark. Während des Lernprozesses sollten Lernende positive Emotionen wie etwa Freude empfinden, denn Angst und Stress erweisen sich als sehr hinderlich für

das Lernen. Insbesondere auch im Hinblick auf die Lernmotivation sind emotionale Aspekte wesentlich (vgl. Pekrun & Schiefele, 1996).

(5) *Lernen ist ein situativer Prozess.* Lernen erfolgt stets kontextgebunden, es ist immer an einen spezifischen Kontext oder an eine bestimmte Situation gebunden, die den Interpretationshintergrund für die Bewertung der Lerninhalte liefern und konkrete Lernerfahrungen ermöglichen. Lernen findet immer im Rahmen einer bestimmten Lernumgebung statt.

(6) *Lernen ist ein sozialer Prozess.* Lernen ist kein individueller Prozess, sondern schließt auch immer soziale Komponenten mit ein. Schulisches, aber auch außerschulisches Lernen wird dabei auf unterschiedlichen Ebenen durch soziale Komponenten beeinflusst, denn Lernen findet zum einen stets in einem soziokulturellen Umfeld und zum anderen meist in Interaktion mit anderen statt.

Von zentraler Bedeutung im Zusammenhang mit dem Konstruktivismus ist die *Situated cognition-Bewegung* (Greeno, 1989; Lave & Wenger, 1991; Resnick, 1987; Rogoff, 1990), eine Forschergruppe, die besonders die Kontextgebundenheit und damit Situiertheit von Denken, Handeln und Lernen und die aktive Wissenskonstruktion durch den Lernenden betont. Seit Ende der 1980er Jahre wurden auf der Grundlage des konstruktivistischen Lernbegriffs mehrere verschiedene Instruktionsmodelle zum *situierten Lernen* entwickelt. Ausgangspunkt bildete das Problem, dass das im schulischen Unterricht theoretisch erworbene Wissen häufig nicht zur Lösung praktischer Probleme übertragen werden kann (Resnick, 1987). Dieses Phänomen der mangelnden Anwendbarkeit von erworbenem Wissens wird auch als „träges Wissen“ bezeichnet (*inert knowledge*, Withehead, 1929; vgl. Renkl, 1996).

Die prominentesten und am häufigsten zitierten Ansätze situierten Lernens stellen die verankerte Instruktion (*Anchored Instruction*), die Theorie der kognitiven Flexibilität (*Cognitive Flexibility Theory*) und die kognitive Meisterlehre (*Cognitive Apprenticeship*) dar, die allesamt in den USA entwickelt wurden und nachfolgend kurz vorgestellt werden.

(1) „*Anchored Instruction*“ (Cognition and Technology Group at Vanderbilt (nachfolgend abgekürzt mit: CTGV), 1997). Zentrales Kennzeichen dieses Ansatzes ist, wie der Name bereits nahelegt, der Einsatz von „narrativen Anker“ in Form von authentischen und komplexen Problemsituationen, deren Funktion es ist, „träges Wissen“ zu vermeiden. Die Verankerung des Lernens in einer komplexen Problemsituation wird begründet mit der Situations- und Kontextgebundenheit des Wissens.

Folgende sieben Prinzipien sind für die Gestaltung von Ankerreizen charakteristisch (vgl. CTGV, 1997; Reinmann & Mandl, 2006): Die authentischen Problemsituationen werden

multimedial über Video präsentiert (*videobasiertes Präsentationsformat*) und sind in eine Geschichte bzw. eine Erzählhandlung eingebunden (*narratives Präsentationsformat*). Auf diese Weise kann Interesse und ein persönlicher Bezug zum Thema hergestellt sowie die Identifikation der Lernenden mit dem Protagonisten der Geschichte erleichtert werden. Ein weiteres zentrales Gestaltungsprinzip besteht darin, dass die Lernenden die in den Geschichten angesprochenen Probleme zunächst eigenständig identifizieren und schließlich lösen. Nach der Bearbeitung sollte den Lernenden zusätzlich die Möglichkeit geboten werden, ihre eigene Lösung mit einer Beispiellösung, die ebenfalls im Anker enthalten sein sollte, zu vergleichen (*generatives Lernformat*). Weitere Prinzipien sind, dass alle für die Problemlösung notwendigen Informationen in die Geschichte bzw. das Video unauffällig eingebettet sind (*Prinzip der eingebetteten Daten*) und die Problemsituation weitgehend der Komplexität einer realen Situation entspricht (*Problemkomplexität*). Damit die Lernenden ihre Kenntnisse flexibel anwenden können, werden jeweils zwei ähnliche Geschichten zur gleichen Thematik dargeboten (*Paare verwandter Abenteuer*). Die Geschichten sind zudem so konstruiert, dass sie Verknüpfungen zu verschiedenen anderen Disziplinen herstellen, um den Transfer auf andere Problemstellungen zu erleichtern (*fächerübergreifende Elemente*).

Die Idee der „Anchored Instruction“ wurde zunächst in mehreren Bildplatten für den Mathematik-Unterricht umgesetzt. Die „Jasper Woodbury Problem Solving Series“ bestehen aus insgesamt 12 videobasierten 12-20 minütigen Abenteuergeschichten aus der Lebenswelt von Kindern und Jugendlichen, die jeweils mit einem komplexen Problem enden. Die Aufgabe der Lernenden besteht darin, dem Protagonisten dieser Geschichten, Jasper, zu helfen, indem sie das jeweilige Problem identifizieren und lösen. In einer der Abenteuergeschichten findet beispielsweise ein Wildhüter einen verletzten Adler in einem Reservat, der für seine medizinische Versorgung von der Fundstelle in eine Tierklinik transportiert werden muss. Um eine Lösung zur Rettung des Adlers zu finden, müssen die Lernenden etliche mathematische Kenntnisse und Fertigkeiten erwerben und anwenden wie beispielsweise den Pythagoras-Satz (Reinmann & Mandl, 2006). Empirische Untersuchungen konnten belegen, dass Schüler, die mit Hilfe videobasierter Ankergeschichten lernten, beim mathematischen Problemlösen und beim Transfer auf andere Kontexte Schülern ohne Intervention deutlich überlegen waren (Bransford et al., 1990; CTGV, 2007).

(2) Der instruktionspsychologische Ansatz der „*Cognitive Flexibility Theory*“ (Jacobson & Spiro, 1992; Spiro & Jehng, 1990) stammt aus der Expertiseforschung (Gruber & Mandl, 1996). Die Bezeichnung „Cognitive Flexibility“ bringt das zentrale Merkmal dieses Ansatzes bereits zum Ausdruck: Eine Aufgabe soll unter verschiedenen Perspektiven betrachtet und bearbeitet werden, um so eine möglichst hohe kognitive Flexibilität bei den

Lernenden zu erzielen (Spiro, Feltovich, Coulson & Anderson, 1989; Spiro, Feltovich, Jacobson & Coulson, 1991, 1992; Spiro et al., 1987).

Unter kognitiver Flexibilität wird die Fähigkeit verstanden, aus verschiedenen Elementen Wissensrepräsentationen zu generieren, was zu einer erhöhten Anwendbarkeit des Wissens führt. Um den Aufbau flexibler und multipler Wissensrepräsentationen zu unterstützen, eignen sich besonders Falldarstellungen und die Technik des „Landscape Criss-Crossing“, bei der dasselbe Konzept zu verschiedenen Zeiten in verschiedenen Kontexten unter veränderter Zielsetzung und aus verschiedenen Perspektiven beleuchtet wird (Spiro & Jehng, 1990). Durch die Auseinandersetzung mit Problemen aus verschiedenen Perspektiven kann die Wahrscheinlichkeit erhöht werden, dass das erworbene Wissen auch auf andere Kontexte und Situationen transferiert werden kann.

Allerdings sind ausführliche Fallbeispiele, die nur einen Blickwinkel berücksichtigen, für diese Art der Instruktion nur wenig geeignet, stattdessen empfehlen die Vertreter dieses Ansatzes kurze und vereinfachte Fälle, sog. Mini-Cases, zu verwenden, die eine Bearbeitung mehrerer Fälle ermöglichen und damit das Bilden einfacher Analogie-Schlüsse vermeiden (Spiro & Jehng, 1990; Gräsel, 1997).

Der Ansatz der „Cognitive Flexibility-Theory“ eignet sich vor allem zur Förderung des fortgeschrittenen Wissenserwerbs in komplexen und wenig strukturierten Gebieten, wie beispielsweise in einigen Anwendungsfeldern im Bereich der Medizin, Psychologie, Wirtschaftswissenschaften und Technik (Reinmann & Mandl, 2006). Diese Domänen zeichnen sich durch eine hohe Komplexität und starke Irregularität der Probleme aus, die eine enge Vernetzung des domänenspezifischen Wissens erfordern, da sich für die Problemlösung weder der Abruf analoger Probleme noch die Anwendung gespeicherter Schemata eignen (Gräsel, 1997).

(3) Im Rahmen des „*Cognitive Apprenticeship*“-Ansatzes werden Prinzipien der traditionellen Handwerkslehre auf den Erwerb kognitiver Fertigkeiten und Fähigkeiten übertragen (Brown, Collins & Duguid, 1989; Collins, Brown & Newman, 1989). Der Grundgedanke dieses Lehr-Lern-Ansatzes besteht darin, Lernende in Anlehnung an das Meister-Lehrlings-Modell der Handwerkslehre zunächst gezielt bei der Lösung von realen Problemstellungen anzuleiten und zu unterstützen, damit sie zunehmend eigenständiger arbeiten können. Das Lernen beginnt zunächst an globalen Problemstellungen und wird im Laufe des Lernprozesses zunehmend komplexer und unterschiedlicher gestaltet.

Nach Collins et al. (1989) können folgende sieben Methoden der instruktionalen Unterstützung bzw. Phasen unterschieden werden: Modeling, Coaching, Scaffolding, Fading, Articulation, Reflection und Exploration. Beim (1) *Modeling* demonstriert zunächst ein

Experte die erfolgreiche Bearbeitung einer realen Problemstellung, indem er sein Vorgehen genau verbal expliziert und damit als kognitives Modell für den Lernenden dient. Nach dem kognitiven Modellieren folgt das (2) *Coaching*, eine Phase, in der die Lernenden selbstständig Problemlöseaufgaben bearbeiten und dabei vom Experten beobachtet und bei Bedarf von ihm betreut und unterstützt werden. Gemäß dem Fall, dass der Lernende die Aufgaben nicht allein bewältigen kann, hilft ihm der Experte beim (3) *Scaffolding* durch das Geben von Tipps, Hilfen, Hinweisen und Rückmeldungen. Während der Phase (4) *Fading* werden die Hilfestellungen vom Lehrenden graduell, d. h. schrittweise nach und nach, gänzlich ausgeblendet. In den nachfolgenden Phasen der (5) *Articulation* und (6) *Reflection* werden die Lernenden instruiert, ihre Denkprozesse und Problemlösestrategien zu artikulieren und ihre ablaufenden Lernprozesse im Austausch mit dem Experten und anderen zu vergleichen und zu reflektieren. In der letzten Phase der (7) *Exploration* lösen die Lernenden eigenständig Problemstellungen, nachdem sich der Experte komplett ausgeblendet hat.

Die im Rahmen des „Cognitive Apprenticeship“-Ansatzes spezifizierten Formen der instruktionalen Unterstützung machen deutlich, dass der Lernende mit authentischen und komplexen Problemen nicht allein gelassen wird, sondern in Abhängigkeit von seinem aktuellem Kompetenzgrad verschiedene Grade und Arten der äußeren Anleitung und Unterstützung erfährt. Multimediale Lernumgebungen basieren sehr häufig auf dem Methodenrepertoire des „Cognitive Apprenticeship“-Ansatzes (Mandl, Gruber & Renkl, 2002).

4.3.2 Didaktische Gestaltung

Aus den drei doch recht unterschiedlichen situierten Ansätzen können nach Gerstenmaier & Mandl (1995) und Reinmann & Mandl (2006) vier zentrale problemorientierte Gestaltungsprinzipien für Lernumgebungen abgeleitet werden, die eine *gemäßigte* Form des Konstruktivismus darstellen, in welcher Lernen aktiv, konstruktiv, selbstgesteuert, situativ, und sozial verläuft und eine begleitende Anleitung und Unterstützung nicht ausgeschlossen wird:

(1) *Authentizität und Anwendungsbezug*. Lernumgebungen sind möglichst so zu gestalten, dass sie den Umgang mit authentischen Situationen und realen Problemstellungen ermöglichen. In diesem Zusammenhang spielen besonders authentische Fälle bzw. Fallbeispiele (Zumbach, Haider & Mandl, 2008) oder narrative Ankergeschichten eine große Rolle. Die authentischen Rahmengeschichten und Fälle sollten für die Lernenden bedeutsam, also relevant oder interessant sein, so kann Interesse erzeugt, die Motivation gesteigert und ein hoher Anwendungsbezug beim Lernen gesichert werden. Die Fallgeschichten können

schriftlich (Winkler & Mandl, 2002), videobasiert (CTGV, 1997) oder im Rahmen von computerunterstützten Lernumgebungen (Gräsel, 1997) präsentiert werden.

(2) *Multiple Kontexte und Perspektiven.* Um zu verhindern, dass neu erworbene Kenntnisse oder Fertigkeiten auf eine bestimmte Situation fixiert bleiben, sollte eine Lernumgebung dieselben Lerninhalte aus mehreren verschiedenen Zusammenhängen und/oder aus verschiedenen Perspektiven anbieten. Den Lernenden wird so ermöglicht, spezifische Inhalte in verschiedenen Situationen und aus mehreren Blickwinkeln zu betrachten. Auf diese Weise werden die Abstrahierung und der Transfer von Wissen auf neue Problemstellungen gefördert, so dass Wissen unter verschiedenen situativen Bedingungen flexibel abgerufen und angewendet werden kann.

(3) *Soziale Lernarrangements.* Lernen sollte nicht ausschließlich als individueller Prozess erfolgen. Dementsprechend sollte die Lernumgebung möglichst oft Phasen des kooperativen Lernens und Problemlösens in Lerngruppen und des gemeinsamen Lernens und Arbeitens von Lernenden und Experten ermöglichen. Auf diese Weise können neben kognitiven auch soziale Kompetenzen zur Kommunikation und Kooperation sowie die Integration der Lernenden in eine Expertenkultur gefördert werden (Lave & Wenger, 1991).

(4) *Instruktionale Anleitung und Unterstützung.* Lernen in problemorientierten Lernumgebungen verlangt nach instruktionaler Anleitung und Unterstützung, da der selbstgesteuerte und soziale Umgang mit komplexen Aufgaben leicht zur Überforderung führen kann, insbesondere bei Lernenden mit ungünstigen Lernvoraussetzungen. Die Lernumgebung ist deshalb so zu gestalten, dass Lernende bei der eigenständigen Bearbeitung komplexer Problemstellungen gezielt unterstützt werden, indem das erforderliche Wissen bereitgestellt und erworben wird. Als unterstützende instruktionale Maßnahmen können beispielsweise die Methoden des „Cognitive Apprenticeship“-Ansatzes eingesetzt werden.

Die vorgestellten problemorientierten Gestaltungsprinzipien lassen sich sehr gut in computerunterstützten, multimedialen Lernumgebungen umsetzen. In einer Metastudie, die auf 43 empirischen Studien basierte, konnte die Effektivität problemorientierten Lernens in Bezug auf den Wissenserwerb bzw. die Wissensanwendung bereits nachgewiesen werden (vgl. Dochy, Segers, Van den Bossche & Gijbels, 2003).

4.3.3 Mediale Gestaltung

Bei der Gestaltung von multimedialen Lernumgebungen sollte neben dem didaktischen Design auch darauf geachtet werden, dass die Lerninhalte in medial adäquater Weise aufbereitet werden (vgl. Mayer, 2005a; Niegemann, 2001). Aus diesem Grund müssen gewis-

se Vorüberlegungen in Bezug auf die Integration medialer Elemente, wie z. B. Text, Bild, Ton, Video oder Animation, in eine computerunterstützte Lernumgebung angestellt werden (Kopp et al., 2003). Im Folgenden sollen die Elemente Text, Ton und Video und 3-D-Animationen näher betrachtet werden.

Texte. In multimedialen Lernumgebungen stellen Texte meist ein zentrales Element dar. Die mediale Aufbereitung von Texten in einer multimedialen Lernumgebung ist jedoch nicht unproblematisch, da die Bereitschaft Lernender eine Passage am Bildschirm mehrmals zu lesen, bis die Zusammenhänge verstanden wurden, wesentlich geringer ist als beim Lesen von Printtexten (Bruns & Gajewski, 2000). Da das Lernen mit Texten am Bildschirm anstrengender ist als das Lernen mit einem Buch, muss das Element Text in einer computerunterstützten Lernumgebung neuen Gestaltungskriterien unterliegen (Euler, 1992; Bruns & Gajewski, 2000; Lang & Pätzold, 2002; Niegemann, 2001). Es sollte darauf geachtet werden, die Textlänge bzw. -menge möglichst gering zu halten, da Texte am Bildschirm deutlich langsamer gelesen werden, die Leser schneller ermüden und die Konzentration und Aufmerksamkeit nachlassen. Eine entscheidende Rolle spielen auch das Textlayout und Texthervorhebungen, da sie dem Lernenden bei der Strukturierung, Organisation und Verarbeitung wichtiger Inhalte helfen. Darüber hinaus sollten Texte auch in einer sachlogisch aufgebauten, kurzen, präzisen und einfach formulierten Form präsentiert sowie abwechslungsreich und interessant gestaltet werden (Kopp et al., 2003).

Ton. Die Integration von Tönen kann auf Lernende sehr motivierend und ansprechend wirken, weswegen immer häufiger Sprechertexte in computerunterstützten Lehr-/Lernangeboten verwendet werden. Beim Einsatz von Tonelementen ist es jedoch wichtig, dass der Nutzer diese individuell steuern kann (Bruns & Gajewski, 2000).

Video und 3-D-Animation. Videofilme eignen sich besonders für die Darstellung von Bewegungs- und Arbeitsabläufen und von sozialen Interaktionen. Mithilfe von 3-D-Animationen lassen sich dagegen insbesondere komplexere Prozesse und Konstruktionen darstellen. Allerdings sollten bereits bei der Planung von Video-Elementen die erheblichen Produktionskosten und Aufwendungen für die entsprechend ausgestatteten Computer und die teilweise sehr langen Übertragungszeiten berücksichtigt werden. Eine günstigere und schnellere Alternative stellen beispielsweise Flash-Animationen dar (Bruns & Gajewski, 2000).

Eine effektive multimediale Lernumgebung zeichnet sich demzufolge durch eine didaktisch sinnvolle Kombination der verwendeten Medien aus. In der Praxis sollten bei der effektiven Gestaltung von multimedialen Lernumgebungen ebenfalls Designprinzipien, die sich aus der kognitiven Theorie des multimedialen Lernens nach Mayer (2005a), der Social

Agency Theory von Mayer (2005c) und dem Ansatz der PSI mit virtuellen Figuren (Hartmann et al., 2004a) ableiten lassen, mitberücksichtigt werden (siehe Kapitel 4.2).

Interessant im Zusammenhang mit der Gestaltung von multimedialen Lernumgebungen ist, dass nicht alle Lernenden gleichermaßen von der Verwendung computerunterstützter Lernumgebungen profitieren: In einigen Studien konnten sog. Aptitude-Treatment-Interaktions-Effekte (ATI-Effekte; Cronbach & Snow, 1977) festgestellt werden, d. h. dass eine erfolgreiche Nutzung dieser Lernumgebungen von Merkmalen der Lernenden wie z. B. dem domänenspezifischen Vorwissen abhängt (Mandl, Gruber & Renkl, 2002; siehe auch Kapitel 4.2.1, CTML: Prinzip der individuellen Unterschiede).

In der Regel profitieren Lernende mit geringem Vorwissen stärker von einer guten instruktionalen Gestaltung des multimedialen Lernmaterials als Lernende mit einem hohen Vorwissen (Weidenmann, 2002; Mayer, 1997, 2005a). Auch können Lernende, die eher zu einer bildhaften Verarbeitung von Informationen neigen, von einer stärker auf Visualisierung hin konstruierten Lernumgebung mehr profitieren als Personen, die Informationen eher textuell verarbeiten (Plass, Chun, Mayer & Leutner, 1998).

4.4 Feedback in multimedialen Lernumgebungen

In der vorliegenden experimentellen Studie wird u. a. in Anlehnung an die Befunde von Knipfer (2009; siehe Kapitel 4.6) auch der Fragestellung nachgegangen, welche Wirkung Feedback in Form einer sozialen Vergleichsinformation auf das Entscheidungsverhalten und die Argumentationsqualität und damit die Lernwirksamkeit der multimedialen Lernumgebung hat. Nachfolgend wird daher auf die theoretischen Grundlagen zum Feedback in multimedialen Lernumgebungen eingegangen.

Ausgehend von einer Begriffsbestimmung (Kapitel 4.4.1) werden zwei zentrale Einflussfaktoren auf die Wirkung von Feedback betrachtet (Kapitel 4.4.2): die Feedbackgestaltung und -rezeption. Danach folgen Erläuterungen zum Feedback in Form einer sozialen Vergleichsinformation, das im Rahmen der untersuchten multimedialen Lernumgebung verwendet wurde (Kapitel 4.4.3).

4.4.1 Begriffsbestimmung Feedback

Feedback stellt aus Sicht der Pädagogik und Psychologie für die Lernleistung eine zentrale Komponente in Lehr-Lernprozessen dar (Krause, Stark & Mandl, 2003). Unter Feedback wird generell

jede Art von Rückmeldung verstanden, die den Lernenden über die Richtigkeit seiner Antwort bzw. seiner Aufgabenlösung informiert (Mory, 2004) oder die dem Lernenden inhaltliche und/oder strategische Hilfen und Informationen zu dessen Bearbeitungsprozess zur Verfügung stellt (Lipowsky, 2009, S. 87).

In Lehr-Lernkontexten gibt Feedback Aufschluss darüber, inwieweit ein Lernender mit seiner erbrachten Leistung (Ist-Zustand) das anvisierte Lernziel (Soll-Zustand) erreicht hat bzw. wie groß die Diskrepanz zwischen Ist- und Soll-Zustand ist (Krause, 2007). Feedback erfüllt damit eine diagnostische Funktion (Fischer & Mandl, 1988), indem es Informationen über die Erreichung von Lernzielen bereitstellt und dem Lernenden ermöglicht, seine Selbsteinschätzung zu validieren (Krause et al., 2003).

Aus kommunikationspsychologischer Sicht wird Feedback definiert als die Botschaft eines Senders, der Feedbackquelle, an einen oder mehrere Feedbackempfänger (Schulz v. Thun, 1981).

Ilgen, Fisher & Taylor (1979) unterscheiden drei *Feedbackquellen*: (1) andere Personen, wie z. B. Lehrende oder Vorgesetzte, Gleichgestellte oder Untergebene, (2) das Aufgabenfeld, indem Feedback Bestandteil der Aufgabe selbst oder in die Aufgabe integriert sein kann (z. B. automatische Feedbackgabe in computergestützten Lernumgebungen). Eine weitere Feedbackquelle stellt die (3) Selbstbeurteilung dar, deren Nutzen und Effizienz vom Selbstvertrauen und von der Erfahrung der Person in dem betreffenden Aufgabenbereich abhängt. In Lehr-/Lernsituationen kommt dem aufgaben- bzw. themenspezifischen Vorwissen eine besondere Bedeutung für die Selbstbeurteilung zu. Als *Feedbackempfänger* kommen entweder einzelne Personen oder Gruppen in Betracht, folglich thematisiert Feedback je nach Aggregationsniveau individuelle oder kollektive Aspekte.

Feedback hat in Lehr-Lernkontexten kognitive, metakognitive und motivationale Funktionen (Krause, 2007). Eine zentrale kognitive Aufgabe von Feedback besteht beispielsweise darin, Fehler, Fehlkonzepte und Wissenslücken aufzuzeigen (Fischer & Mandl, 1988). Feedback fördert auch die bewusste Auseinandersetzung mit den Lerninhalten und deren Reflexion und kann auf metakognitiver Ebene Verständnis- bzw. Kompetenzillusionen vorbeugen. Aus motivationaler Sicht kann Feedback das Erleben von Selbstwirksamkeit und Kompetenz unterstützen (vgl. Deci & Ryan, 1993) sowie die Überzeugung fördern, dass sich der Lernaufwand lohnt. Krause (2007) weist jedoch in Anlehnung an Vroom (1964) darauf hin, dass eine klare Trennung der kognitiven, metakognitiven und motivationalen Feedbackfunktionen empirisch kaum möglich ist.

Das Geben von Rückmeldungen bzw. von Feedback basiert auf unterschiedlichen lerntheoretischen Ansätzen (vgl. Kopp & Mandl, in Druck; Mory, 2004; Musch, 1999). Über mehrere Jahrzehnte dominierten in der Feedbackforschung die behavioristischen Ansätze, die

Feedback mit Verstärkung von erwünschtem Verhalten bzw. mit der Hemmung von unerwünschtem Verhalten gleichsetzen: Durch positives, bestätigendes Feedback wird die Auftretenswahrscheinlichkeit von korrekten Antworten erhöht, bei negativem Feedback hingegen verringert (Krause, 2007).

Ab den 70er Jahren rückte mit den kognitiven Ansätzen die informative Komponente mehr in den Vordergrund und Feedback wurde als *Informationsquelle* zur Korrektur von Fehlern angesehen. Feedback liefert diagnostische Informationen für die Handlungsregulation: Die durch die Rückmeldung gegebenen Informationen ermöglichen den Lernenden durch einen Ist-Soll-Vergleich, die Diskrepanzen zwischen erbrachter und angestrebter Leistung zu analysieren, mögliche Fehlerquellen zu identifizieren und zu korrigieren (Krause, 2007).

Aus einer konstruktivistischen Perspektive (Gerstenmaier & Mandl, 1995) wird Feedback als *Angebot* verstanden, das von den Lernenden in unterschiedlicher Art und Weise zur Wissenskonstruktion genutzt werden kann. Die Nutzung des angebotenen Feedbacks zur Optimierung der eigenen Lernprozesse und des Lernerfolgs hängt dabei wesentlich von der Motivation der Lernenden ab (Kopp & Mandl, in Druck).

In dieser Arbeit wird Feedback auf der Basis kognitiver und konstruktivistischer Überlegungen als Informationsangebot verstanden, das vom Empfänger unterschiedlich intensiv verarbeitet werden kann, aber nicht automatisch genutzt werden muss.

4.4.2 Wirkung von Feedback

In mehreren Metaanalysen konnte die Lernwirksamkeit von Feedback gegenüber keinerlei Feedback bereits empirisch belegt werden (vgl. z. B. Azevedo & Bernard, 1995; Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Clariana, 1993; Hattie & Timperley, 2007; Mory, 2004; Schimmel, 1988).

Bei näherer Betrachtung der zahlreichen Forschungsergebnisse zur Lernwirksamkeit von Feedback zeigt sich allerdings eine äußerst heterogene Befundlage. Studien ergaben nicht nur positive Effekte von Feedback auf die Lernleistung, sondern zum Teil auch keine oder sogar negative Auswirkungen (vgl. Bangert-Drowns et al., 1991; Jacobs, 2002; Kluger & DeNisi, 1996). Auch die Frage, welche Feedbackform am lernförderlichsten ist, konnte die Forschung bislang nicht eindeutig klären (vgl. z. B. Azevedo & Bernard, 1995; Bangert-Drowns et al. 1991; Clariana, 1993; Dempsey, Driscoll & Swindell, 1993; Kulhavy & Stock, 1989; Mory, 2004).

Als Erklärungsansätze für die recht inkonsistente Befundlage zur Lernwirksamkeit der verschiedenen Feedbackformen werden zum einen die Feedbackgestaltung auf Seiten der

Feedbackquelle und zum anderen die Feedbackrezeption auf der Seite der Feedbackempfänger herangezogen.

4.4.2.1 Einflussfaktor Feedbackgestaltung

Im Hinblick auf die Gestaltung von Feedback sind mehrere verschiedene Aspekte zu berücksichtigen, die im Folgenden erörtert werden (vgl. Krause, 2007).

Bezugsnormorientierung. Feedback über Lernfortschritte kann unter Anlegung von drei Vergleichsmaßstäben erfolgen (Rheinberg, 2006): (1) den Lernzielen (kriterienbezogener Vergleich), die im Allgemeinen die Basis für das Geben von Rückmeldungen bilden, (2) den früher erzielten Leistungen eines Lernenden (intraindividueller Vergleich) oder (3) den Leistungen anderer (interindividueller Vergleich). Diese drei Vergleichsmaßstäbe werden nach Heckhausen (1974) als sachliche, individuelle und soziale Bezugsnormen bezeichnet. Während die sachliche Bezugsnorm Informationen über das Erreichen des Lernziels gibt, erfolgt bei der sozialen Bezugsnorm ein Vergleich der individuellen Leistung mit der Leistung von anderen Lernenden, bei der individuellen Bezugsnorm wird hingegen die früher erzielte Leistung desselben Individuums verglichen (Krause, 2007). Nach Heckhausen (1989) können durch die Art und Weise der Rückmeldungen beim Wissenserwerb auch Prozesse der Selbstattribution, die sich z. B. durch Stolz bei erfolgreichem Lernen und durch Scham über Fehler und Versagen äußern, beeinflusst werden (Musch, 1999).

Einige Untersuchungsergebnisse zeigen, dass bei Leistungsrückmeldungen die sachliche und insbesondere die individuelle Bezugsnorm der sozialen vorzuziehen ist (z. B. Rheinberg, 2006; Rheinberg & Krug, 1999). Rückmeldungen mit individueller Bezugsnorm wirken sich besonders positiv auf die Lernmotivation und das Selbstkonzept aus, da die Leistungsbewertung an den eigenen Lernfortschritten und nicht im sozialen Vergleich zu den Leistungen anderer Lernender erfolgt. Auf diese Weise erleben auch schwächere Lernende motivierende Lernfortschritte. Feedback, das auf der sozialen Bezugsnorm basiert, kann hingegen Leistungsdruck erzeugen und kontrollierend erlebt werden und ungünstige Auswirkungen in Bezug auf motivationale Aspekte nach sich ziehen. Im Vergleich zur sozialen Bezugsnorm vergrößern Rückmeldungen auf der Basis der individuellen Bezugsnorm außerdem die Hoffnung auf Erfolg, verringern die Furcht vor Misserfolg und ziehen eine realistischere Einschätzung der Zielsetzungen nach sich (Rheinberg, 2001, 2006; Rheinberg & Krug, 1999).

Ein wesentlicher Vorteil der sachlichen Bezugsnorm besteht darin, dass durch das Aufzeigen der Ist- und Soll-Diskrepanz konkrete Anknüpfungspunkte für die Feedbackverarbeitung geboten werden und aufgezeigt wird, inwieweit weitere Lernanstrengungen nötig

sind. Wie Hattie & Timperley (2007) in einer Metaanalyse feststellen konnten, erzielte die sachliche Bezugsnorm verglichen mit anderen Feedbackformen die höchsten Effekte.

Aber auch Rückmeldungen, die auf einer sozialen Bezugsnorm basieren, können durchaus motivierend und lernwirksam sein, wenn sie in einer informierenden Weise gegeben werden (Rheinberg, 2001, 2006). Gemäß der Theorie der sozialen Vergleichsprozesse von Festinger (1954) stellt der Vergleich mit anderen ein menschliches Grundbedürfnis dar und ermöglicht es, die eigenen Stärken und Schwächen realistisch einzuschätzen und zu validieren (siehe Kapitel 4.4.3).

Da jede der drei Bezugsnormen mit verschiedenen Vor- und Nachteilen verbunden ist, erscheint im Hinblick auf die Feedbackgestaltung ein kombiniertes Ansprechen der verschiedenen Bezugsnormen als sinnvoll (Heckhausen, 1974; Rheinberg, 2001, 2006).

Angesprochene Ebene: Aufgabe, Motivation und Selbst. In ihrer Metastudie unterscheiden Kluger & DeNisi (1996) drei unterschiedlichen Hierarchieebenen von rückgemeldeten Informationen, die sich auf das (1) Aufgabenlernen (*task-learning-processes*), die (2) Aufgabenmotivation (*task-motivation-processes*) und (3) Metaprozesse (*meta-task-processes*) beziehen können, wobei letztere auch das Selbst des Lernenden umfassen. Indem der Aufmerksamkeitsfokus durch die Art der Feedbackgestaltung auf eine der drei Ebenen gelenkt wird, können Aufmerksamkeitsprozesse der Lernenden günstig oder ungünstig beeinflusst werden. Gemäß Kluger & DeNisi (1996) ist es im Hinblick auf die Lernwirksamkeit von Feedbackinterventionen günstig, wenn sich Rückmeldungen konkret auf die im Mittelpunkt stehende Aufgabe beziehen. Ungünstig ist es dagegen, wenn die Aufmerksamkeit von der Aufgabe zur Metaprozessebene, also zum Selbst, hingelenkt wird. Hier besteht die Gefahr, dass im Falle einer negativen Abweichung Feedback schnell als bedrohlich und verunsichernd empfunden werden und ungünstige emotionale Prozesse, wie z. B. Scham, auslösen kann.

Informationsgehalt. In der pädagogisch-psychologischen Feedbackforschung werden in Abhängigkeit von der Komplexität und der Elaboriertheit verschiedene Feedbackformen unterschieden (vgl. z. B. Clariana, 2000; Dempsey et al., 1993; Jacobs, 2002; Kulhavy & Stock, 1989; Schimmel, 1988). Einfache Rückmeldeformen informieren beispielsweise den Lernenden darüber, ob seine aktuelle Aufgabenlösung richtig oder falsch ist (*Knowledge of results*, KOR) oder wie die korrekte Antwort lautet (*Knowledge of correct response*, KCR). Im Gegensatz dazu gehen komplexere und elaboriertere Rückmeldeformen über die Mitteilung der korrekten Antwort bzw. KOR und KCR hinaus. Elaboriertes, auch informatives Feedback genannt, beinhaltet zusätzliche Erläuterungen, warum eine Lösung bzw. Antwort richtig oder falsch ist (Dempsey et al., 1993; Jacobs, 2002). Die wichtigste Form elaborierten Feedbacks stellt für Jacobs (2002) die „detaillierte Beschrei-

bung des Lösungsweges in Form eines Musterlösungsansatzes mit einer entsprechenden Begründung“ (S. 6-7) dar.

Kulhavy & Stock (1989) unterscheiden drei Formen elaborierten Feedbacks, das sich auf die Aufgabe (*task-specific*), die Instruktion (*instruction-based*) und über die Instruktion hinausgehende Aspekte (*extra-instructional*) bezieht: (1) die aufgabenspezifische Elaboration, die weitestgehend dem KCR entspricht und die korrekte Lösung der Aufgabe präsentiert, (2) die instruktionsbasierte Elaboration, die sich auf das ursprüngliche Lernmaterial bezieht und u. a. Erklärungen der Aufgabenlösung, Unterstützung von Lösungsregeln, Verbesserung von Fehlern beinhaltet und (3) die extrainstruktionale Elaboration, die zusätzliche, bislang noch nicht dargebotene Informationen, wie z. B. neue Beispiele, zur Verfügung stellt.

Eine Reihe von Forschungsergebnissen deutet grundsätzlich auf eine höhere Lernwirksamkeit von elaboriertem Feedback gegenüber einfachen Feedbackformen wie KOR und KCR hin (vgl. z. B. Bangert-Drowns et al., 1991; Collins, Carnine & Gersten, 1987; Fischer & Mandl, 1988; Huth, 2004; Musch, 1999; Narciss, 2006; Pridemore & Klein, 1991). Die Überlegenheit von informativem, elaboriertem Feedback gilt insbesondere dann, wenn es um das Erlernen komplexer Inhalte, wie z. B. dem anwendungsbezogenen Wissenserwerb im Bereich empirischer Forschungsmethoden, geht. Beim Erlernen von reinem Fakten- und Konzeptwissen genügt in der Regel eine einfache richtig/falsch-Rückmeldung (vgl. Krause et al., 2003).

Auch der *Zeitpunkt der Feedbackgabe* hat einen Einfluss auf die Wirkung von Feedback. Rückmeldungen können entweder sofort nach einer Fragenbeantwortung oder verzögert mit einem zeitlichem Abstand gegeben werden. Feedback, das unmittelbar im Anschluss an eine bearbeitete Aufgabe erfolgt, besitzt eher *formativen* Charakter, da es für die Bearbeitung einer nachfolgenden Aufgabe genutzt werden kann, wohingegen Feedback, das erst im Anschluss an eine Lernsequenz gegeben wird, eher *summativen* Charakter hat. Sofortiges Feedback erwies sich in einer Metaanalyse von Kulik & Kulik (1988) über 53 Studien gegenüber einer verzögerten Rückmeldung als lernwirksamer, vor allem dann, wenn das Feedback eine Wiederholung des Lernstoffs und damit eine zusätzliche Lernphase mit sich bringt.

Im Hinblick auf eine motivierende Wirkung ist auch entscheidend, ob Feedback *bestätigend/kritisch* oder *informierend/kontrollierend* gegeben wird. Bestätigendes Feedback wird in der Regel als glaubwürdiger und erwünschter eingestuft sowie besser akzeptiert als kritisches Feedback. Studien konnten auch zeigen, dass Feedback, wenn es in einer informierenden und nicht-kontrollierenden Weise gegeben wird, das Kompetenz- und Autonomieerleben der Lernenden unterstützt und die intrinsische Motivation fördert, wohingegen

kontrollierendes Feedback jene eher beeinträchtigt (vgl. Deci, Koester & Ryan, 2001; Deci & Ryan, 1993; Sansone, 1986). Informierendes Feedback unterstützt Lernende beim selbstgesteuerten Lernen und zielt darauf ab, die Kompetenz der Lernenden zu steigern, wohingegen bei kontrollierendem Feedback gute Leistungen eingefordert werden und infolgedessen Leistungsdruck aufgebaut wird, der sich ungünstig auf motivationale Aspekte auswirken kann.

Auch die Art der Feedbackübermittlung, also ob es *mündlich* oder *schriftlich* erfolgt, hat einen Einfluss auf die Wirkung von Feedback (Krause, 2007). Schriftliches Feedback hat gegenüber einer mündlichen Rückmeldung den Vorteil, dass es in einem individuellen Tempo rezipiert und bei Bedarf auch nochmals gelesen werden kann. Allerdings sind bei schriftlichem Feedback meist keine direkten Rückfragen möglich. Bei der Gestaltung von schriftlichem Feedback sollte daher besonders auf Verständlichkeit (vgl. Langer, Schulz von Thun & Tausch, 1999) geachtet werden (z. B. Einfachheit, Gliederung/Ordnung, Kürze/Prägnanz). Ebenfalls sollten gewisse *Feedbackregeln* beachtet und eingehalten werden (vgl. Fengler, 1998). Beispielsweise wird empfohlen, dass Feedback möglichst beschreibend und nicht wertend sowie konkret auf begrenztes Verhalten und nicht allgemein auf die ganze Person bezogen sein sollte.

In computerunterstützten Lernumgebungen kann schriftliches Feedback auch *automatisch* und in standardisierter Form gegeben werden, was den Vorteil hat, dass die Feedbackgabe sofort und auf eine ökonomische Weise erfolgt und die Aufmerksamkeit der Lernenden weniger leicht abgelenkt wird als dies bei einer mündlichen Feedbackgabe der Fall ist. Der Aufmerksamkeitsfokus bleibt eher der Aufgabe zugewandt und nicht auf dem „Selbst“.

4.4.2.2 Einflussfaktor Feedbackrezeption

Einen weiteren zentralen Einflussfaktor auf die Wirkung von Feedback stellt neben der Feedbackgestaltung die Art der Feedbackrezeption durch den Feedbackempfänger dar (vgl. Krause, 2007). Damit überhaupt eine lernförderliche Wirkung von Feedback eintreten kann, muss in Anlehnung an die kognitivistischen und konstruktivistischen Ansätze die Feedbackinformation vom Feedbackempfänger verarbeitet und aktiv zur Wissenskonstruktion genutzt werden (Mory, 2004).

In der Feedbackforschung wird häufig davon ausgegangen, dass die Feedbackverarbeitung quasi automatisch nach dem Geben der Rückmeldung erfolgt (Hancock, Thurman & Hubbard, 1995). Studienergebnisse zeigen jedoch, dass die Feedbackverarbeitung häufig suboptimal und nicht in der vom Feedbackgeber intendierten Art und Weise erfolgt (vgl. z. B. Hancock et al., 1995; Krause, 2002; Stark, 2001; Stark & Mandl, 2001a). Um positive

Wirkungen zu erreichen, muss Feedback bewusst rezipiert und bei der weiteren Aufgabenbewältigung bzw. dem weiteren Wissenserwerb berücksichtigt werden (Bangert-Drowns et al., 1991). Bangert-Drowns et al. (1991) bezeichnen diesen Prozess auch als *mindful reception*, also die bewusste und aufmerksame Rezeption.

Befunde von Hancock et al. (1995) weisen daraufhin, dass die Feedbackrezeption und -nutzung vor allem von den kognitiven, metakognitiven und motivational-emotionalen *Lernvoraussetzungen* des Feedbackempfängers und seiner *Aufmerksamkeit* und *Konzentration* beim Empfang des Feedbacks abhängt. Den wichtigsten kognitiven Einflussfaktor auf die Wirkung von Feedback bildet das Vorwissen des Lernenden im jeweiligen Inhaltsbereich (Alexander, Kulikowich & Schulze, 1994). Zwei weitere situationsspezifische kognitive Faktoren stellen die Aufmerksamkeit und die Konzentration dar. Nur wenn dem Feedback Aufmerksamkeit und Beachtung geschenkt wird, kann es verarbeitet werden (Kluger & DeNisi, 1996). Die Verarbeitung von gegebenem Feedback ist auch abhängig von der Konzentration bzw. von auftretenden Ermüdungserscheinungen des Lernenden (Hancock et al., 1995). Entscheidend ist in diesem Zusammenhang auch, ob der Lernstoff und die Aufgabe verstanden wurden (Kulhavy, Yekovich & Dyer, 1976).

Auch die *Sicherheit (Konfidenz)* der Lernenden im Hinblick auf die Korrektheit der Antwort kann die Wirkung von Feedback beeinflussen (Kulhavy & Stock, 1989; Musch, 1999). Bei geringer Antwortsicherheit ist der Nutzen von Feedback eher gering: War die Antwort richtig, können Lernende von dieser positiven und bestätigenden Rückmeldung nur unwesentlich profitieren. Bei hoher Antwortsicherheit hängt der Nutzen dagegen von der Korrektheit der Antwort ab: Eine falsche Antwort trotz hoher Sicherheit beinhaltet eine Überraschung und der Lernende befasst sich länger mit dem Feedback, um nach den Ursachen des Fehlers zu suchen.

In computerunterstützten Lernumgebungen gewinnen Rückmeldungen aufgrund des höheren Anteils an selbstgesteuertem Lernen zunehmend mehr an Bedeutung (Kopp & Mandl, in Druck). Da computerbasierte Lernprogramme häufig darauf abzielen, den Wissenserwerb zu ermöglichen bzw. zu vertiefen, ist es für Lernende generell sehr wichtig, zu erfahren, inwiefern der eigene Lernprozess angemessen abläuft oder Veränderungen erfordert. Bei einer automatischen Feedbackgabe stellt sich auch die Frage nach der *Adaptivität* von Feedback, also inwiefern Rückmeldungen antwortabhängig gestaltet werden können. Sales (1993) unterscheidet in diesem Kontext „adaptiertes“ und „adaptives“ Feedback. Während adaptiertes Feedback alle Lernenden unabhängig von Leistungsunterschieden nach der Aufgabenbearbeitung erhalten (z. B. in Form von Musterlösungen), vermag adaptives Feedback zwischen unterschiedlich erfolgreichen Lernenden zu differenzieren, z. B. indem weniger erfolgreiche Lernende ein ausführlicheres Feedback bekommen (Krause, 2007).

Als Resümee kann zusammenfassend betrachtet festgehalten werden, dass im Hinblick auf eine lernwirksame Gestaltung von Feedback in computergestützten Lernumgebungen mehrere relevante Dimensionen berücksichtigt werden sollten. Die Feedbackgestaltung hängt davon ab, welche Funktionen es im Lehr-Lernszenario erfüllen soll und letztendlich auch von ökonomischen Überlegungen.

An dieser Stelle muss allerdings auch einschränkend angemerkt werden, dass sich die pädagogisch-psychologische Feedbackforschung vor allem mit Aufgaben beschäftigt, die Aussagen über „richtig“ oder „falsch“ erlauben (z. B. Problemlöseaufgaben zum Thema Korrelationsrechnung; vgl. Krause, 2007). In dieser Untersuchung wurde jedoch eine Dilemmageschichte zum Thema PID-HLA von den Lernenden bearbeitet, zu der keine objektiven Lösungskriterien im Sinne von „falsch“ oder „richtig“ existieren (siehe Kapitel 3.1). Fehlen objektive Bewertungsstandards zur Überprüfung der Richtigkeit der eigenen Entscheidung, wird in der Regel zur Bewertung der Korrektheit der eigenen Entscheidung der Vergleich auf der sozialen Ebene gesucht (vgl. Festinger, 1954).

Dementsprechend wurde im Rahmen der multimedialen Lernumgebung Dialogstation „Gentest“ Feedback in Form einer sozialen Vergleichsinformation gegeben (siehe Kapitel 5). Mit Hilfe dieser sog. Konsensinformation können die Lernenden ihre eigene Entscheidung mit der durchschnittlichen Gesamtentscheidung anderer Personen vergleichen und validieren. Nachfolgend wird daher speziell der theoretische Hintergrund zum Feedback in Form einer sozialen Vergleichs- bzw. Konsensinformation beleuchtet.

4.4.3 Feedback in Form einer sozialen Vergleichsinformation

Unter einer Konsensinformation wird in der Sozialpsychologie generell eine Information verstanden, die Aufschluss darüber gibt, ob andere Menschen sich demselben Stimulus gegenüber genauso verhalten wie der Agierende (Aronson, Wilson & Akert, 2008).

Konsens im Speziellen versteht Erb (1998) als Information darüber, „inwieweit Meinungsübereinstimmung in Bezug auf beliebige Objekte besteht, also „die anderen“ ein bestimmtes Einstellungsobjekt evaluieren“ (S. 156). Eine Konsensinformation gibt Aufschluss darüber, ob eine bestimmte Einstellung, Meinung oder Entscheidung von einer Minderheit (Minorität) oder einer Mehrheit (Majorität) vertreten wird. Die Meinungsübereinstimmung (Konsens) fällt dabei bei einer Minderheitsposition stets numerisch kleiner aus als bei einer Mehrheitsposition (Erb & Bohner, 2002).

Eine typische Konsensinformation stellen beispielsweise die Ergebnisse von Meinungsumfragen dar, die eine bedeutende Möglichkeit sozialen Einflusses darstellen, da Menschen

Umfrageergebnisse zu ihrer eigenen Urteilsbildung nutzen (vgl. Thoben & Erb, 2010). Wenn das Ergebnis einer Meinungsumfrage in eine Lernumgebung als Feedbackkomponente implementiert wird, handelt es sich demzufolge um eine Situation sozialen Einflusses. Sozialen Einfluss definieren Hewstone & Martin (2007) als „Veränderung von Einstellungen, Überzeugungen, Meinungen, Werten bzw. Verhaltensweisen infolge der Tatsache, dass man mit den Einstellungen, Überzeugungen, Meinungen, Werten bzw. Verhaltensweisen anderer Menschen konfrontiert ist“ (S. 360).

Aus der Theorie der sozialen Vergleichsprozesse von Festinger (1954) ist bekannt, dass Menschen ihre eigenen Meinungen und Entscheidungen mit denen der anderen vergleichen, da jeder Mensch das Bedürfnis hat, seine Meinungen und Fähigkeiten zu bewerten: „There exists, in the human organism, a drive to evaluate his opinions and abilities“ (p. 117). Aus diesem Motiv heraus ergibt sich auch ein Bedürfnis des Individuums,

subjektive Gewissheit über die Korrektheit der eigenen Einschätzung über Meinungen und Fähigkeiten zu erlangen und falsche Informationen und Kognitionen zu vermeiden. Je geringer die subjektive Gewissheit ausgeprägt ist (je unsicherer sich also eine Person ist), desto weniger ist das Bedürfnis befriedigt und desto stärker wird somit die Vergleichsmotivation ausfallen (Frey, Dauenheimer, Pargé & Haisch, 1993, S. 87).

Das Ausmaß, inwieweit sich Menschen mit anderen vergleichen, hängt neben der subjektiven Gewissheit aber auch von individuellen Präferenzen und Dispositionen wie beispielsweise der sozialen Vergleichsorientierung (Gibbons & Buunk, 1999) oder der Selbstüberwachung, im Speziellen der Orientierung an anderen (Snyder, 1979) ab. Vergleichsinformationen werden insbesondere auch dann gesucht, wenn sie für die eigene Entscheidung wichtig bzw. nützlich sind. Für den sozialen Vergleich werden in der Regel meist Menschen herangezogen, die in Bezug auf ihre Fähigkeiten und Meinungen als ähnlich wahrgenommen werden (Ähnlichkeitshypothese) (Frey et al., 1993). Wenn Personen Meinungsdiskrepanzen im Vergleichsprozess feststellen, so versuchen sie in der Regel, die aufgetretenen Diskrepanzen zu reduzieren (siehe später: Dissonanztheorie).

Wie zahlreiche Befunde aus der Sozialpsychologie belegen, bewirkt die Konfrontation mit einer Mehrheitsmeinung in vielen Fällen eine Anpassung der individuellen Meinung an die Meinung der Majorität (Konformität), da der herrschende Konsens dem Einzelnen eine Orientierungshilfe bietet, an ihn möglicherweise herangetragene Erwartungen offenbart und festlegt, was als „normal“ und akzeptiert gilt: Denn hoher Konsens impliziert Validität (Erb, 1998; Kelley, 1967; Mackie, 1987).

Zwei Arten des sozialen Einflusses können hier wirksam werden (Deutsch & Gerard, 1955): Zum einen der *normative soziale Einfluss*, der Menschen aufgrund ihres Bedürfnisses nach Anerkennung und Akzeptanz konform gehen lässt, zum anderen der

informationale soziale Einfluss, der auf dem Bedürfnis gründet, mit dem eigenen Urteil richtig zu liegen. In eindeutigen Situationen führt die auf normativen Einfluss beruhende Konformität häufig zur öffentlichen Compliance („Folgsamkeit“) mit den Überzeugungen der Gruppe, aber nicht unbedingt zur privaten Akzeptanz. In mehrdeutigen Situationen dagegen, in denen eine große Unsicherheit über die Angemessenheit eines Verhaltens herrscht, steigt der informationale Einfluss der Mehrheit, denn je unsicherer Menschen sind, desto mehr verlassen sie sich auf die anderen (Allen, 1965; Baron, Vandello & Brunzman, 1996; Tesser, Campbell & Mickler, 1983). Im Sinne des „Prinzips der sozialen Bewährtheit“ (*social proof*) dient das Verhalten der Mehrheit als Orientierungshilfe für die Beurteilung der Angemessenheit des eigenen Verhaltens in einer gegebenen Situation („Was alle tun, ist richtig“), insbesondere dann, wenn die anderen als ähnlich betrachtet werden (Cialdini, 2007). Informativer Einfluss führt in der Regel zu Konversion, d. h. privater und öffentlicher Konformität.

Dass Individuen dazu neigen, ihr individuelles Verhalten, ihre Einstellungen, Urteile oder Entscheidungen an die Position der Mehrheit anzupassen, d. h. konform zu gehen, konnte bereits vor über 60 Jahren Solomon Asch (1956) nachweisen. In seinem mittlerweile schon als klassisch zu bezeichnenden Experiment saßen die Versuchspersonen (nachfolgend abgekürzt mit: Vpn) zu mehreren an einem Tisch und wurden gebeten, laut ihr Urteil abzugeben, welche von drei visuell dargebotenen und unterschiedlich langen Vergleichslinien die gleiche Länge aufwies wie eine gesondert dargebotene Referenzlinie. Die zu lösende Aufgabe in dieser Studie war so einfach, dass in einer Kontrollbedingung so gut wie keine Fehler auftraten. Ganz anders jedoch sahen die „Fehlerquoten“ in den Experimentalgruppen aus, in denen die Vpn ihr Urteil abzugeben hatten, nachdem zuvor eine Mehrheit angeblich anderer Vpn, die tatsächlich aber Konföderierte des Versuchsleiter waren, absichtlich und übereinstimmend falsche Urteile abgaben. Die Ergebnisse waren bemerkenswert: Die Urteile der „echten“ Versuchspersonen glichen sich in etwa einem Drittel der Fälle denen der falsch urteilenden Mehrheit an.

Die Befunde von Asch (1956) konnten inzwischen in vielen darauffolgenden Experimenten repliziert werden (vgl. Metaanalyse zu 133 Studien von Bond & Smith, 1996). Insgesamt betrachtet konnte die sozialpsychologische Forschung eindrucksvoll belegen, dass hoher Konsens, auch bei offensichtlich fehlerhaften Urteilen, zu einer starken Beeinflussung individueller Urteile führt (z. B. Jones & Gerard, 1967).

Zur Erklärung der Befunde zum sozialen Einfluss von Mehrheiten und Minderheiten wurden in der Sozialpsychologie mehrere unterschiedliche psychologische Ansätze entwickelt, deren ausführliche Diskussion allerdings über den Rahmen dieser Arbeit hinausgehen würde; einen Überblick zu den Theorien zum sozialen Einfluss durch Minderheiten und Mehr-

heiten geben beispielsweise Erb & Bohner (2002). Für diese Arbeit wurde aus diesen Theorien der persuasionstheoretische Ansatz gewählt, der im Gegensatz zu den anderen Theorien nicht die komplexe Interaktion innerhalb von Kleingruppen untersucht, sondern auf die individuellen kognitiven Prozesse fokussiert, die bei der Informationsverarbeitung ablaufen.

Als Rahmenmodell zur Erklärung der Wirkung von Konsensinformationen bei der Verarbeitung persuasiver Kommunikation kann das HSM von Chaiken et al. (1989) herangezogen werden (Erb, 1998; vgl. Persuasionstheorien, siehe Kapitel 3.2.2). Information über Konsens kann im HSM „als persuasiver Hinweisreiz auf die soziale Akzeptanz der vertretenen Position konzeptualisiert werden“ (Erb, 1998, S. 160). Die Wirkung von Konsensinformationen auf individuelle Urteile und Entscheidungen hängt dabei von der Motivation (Involvement/ persönliche Relevanz) und der Fähigkeit des Rezipienten ab.

Heuristisch verarbeitende Individuen, bei denen nur eine geringe Motivation und Fähigkeit zur systematischen Auseinandersetzung mit den Inhalten und Argumenten einer Botschaft vorliegt, lehnen häufig die Minderheitsposition ab und stimmen einer Mehrheitsmeinung unter der einfachen Anwendung der Konsensheuristik zu (*consensus implies correctness*): Die Mehrheitsmeinung wird ungeprüft in das eigene Urteil übernommen (vgl. z. B. Axsom et al., 1987; De Vries, De Dreu, Gordijn & Schuurman, 1996; Hazlewood & Chaiken, 1990). Denn hoher Konsens impliziert ein valides Urteil, da die meisten Menschen annehmen, dass die Mehrheit wahrscheinlich die zutreffende Position repräsentiert (*objective consensus*; Mackie, 1987) und die eigene Einstellung von einer Mehrheit geteilt wird (*false consensus effect*; Ross, Green & House, 1977; *social projection*; Allport, 1924). Unter heuristischer Verarbeitung kann eine Konsensinformation somit direkt Urteile und Entscheidungen determinieren.

Bei Personen mit hoher Motivation hat eine Konsensinformation dagegen häufig nur einen geringen Einfluss auf das Urteil. Unter systematischer Verarbeitung wird die Wirkung einer Konsensinformation auf die Urteilsbildung oft aufgehoben und dramatisch abgeschwächt, da Informationen und Argumente unter hohem kognitiven Aufwand gründlich und detailliert geprüft werden (*attenuativer Effekt*; vgl. z. B. Maheswaran & Chaiken, 1991). Aber auch bei systematischer Verarbeitung kann Konsens unter bestimmten Bedingungen Urteile beeinflussen, zum einen wenn die Mehrheitsmeinung mit dem eigenen Urteil übereinstimmt und zum anderen wenn die Konsensinformation verlässlich (reliabel) erscheint.

Bei hochmotivierten und systematisch verarbeitenden Individuen wird die Übereinstimmung (Kongruenz) einer Mehrheitsmeinung mit der eigenen, selbstständig gebildeten Mei-

nung als ergänzender Hinweis auf die Richtigkeit (Validität) des eigenen Urteils interpretiert (*additiver Effekt*; vgl. z. B. Maheswaran & Chaiken, 1991).

Wie Hazlewood & Chaiken (1990, zitiert nach Eagly & Chaiken, 1993) in einer Studie feststellen konnten, hängt der Einfluss von heuristischen Hinweisreizen bei Personen mit hoher Motivation zur systematischen Verarbeitung auch von der wahrgenommenen Reliabilität der Heuristiken ab. Den an dieser Untersuchung teilnehmenden Studierenden wurde zunächst mitgeteilt, dass bestimmte Prüfungsarten entweder im nächsten Jahr (hohe Relevanz) oder im nächsten Jahrzehnt (geringe Relevanz) eingeführt werden sollen. Den Studierenden wurde zusätzlich das Ergebnis einer Meinungsumfrage präsentiert, dass 80 % von den befragten Studierenden dem Prüfungsvorschlag zustimmten (vs. nicht zustimmten). Die Reliabilität dieser Konsensinformation wurde anhand der Samplegröße variiert (10 vs. 1000 Studierende). Die Ergebnisse konnten zeigen, dass niedrig involvierte Personen dem Prüfungsvorschlag unabhängig von der Samplegröße mehr zustimmten als hochinvolvierte Personen. Diese Personen wendeten zur Einstellungsbildung offenbar die Konsensheuristik an (*consensus implies correctness*), ohne die Reliabilität der präsentierten Information genau zu überprüfen, wohingegen Personen mit hoher Motivation diese berücksichtigten. Basierte die Konsensinformation dagegen auf einem kleinen Sample, wurde die Einstellungsbildung bei der Gruppe mit hoher Relevanz nicht beeinflusst, beruhte die Konsensinformation jedoch auf einem großen Sample, waren die Einstellungen gegenüber dem Examensvorschlag positiver. Die Information, dass 80 % von 1000 Studierenden dem Prüfungsvorschlag zustimmten, beeinflusste demzufolge nicht nur niedrigmotivierte, sondern auch hochmotivierte Personen.

In einer anderen Studie variierten Giner-Sorolla & Chaiken (1997) die Kongruenz einer Konsensinformation in Form einer Meinungsumfrage in Abhängigkeit von den bestehenden Einstellungen zu schriftlichen Examensprüfungen (inkongruent vs. kongruent). Bei Kongruenz deckte sich das Gesamtergebnis der Meinungsumfrage mit den ein paar Wochen zuvor erhobenen Einstellungen zu schriftlichen Examensprüfungen, wohingegen bei Inkongruenz das Umfrageergebnis im Konflikt zu den bestehenden Einstellungen stand. Als Ergebnis der Analyse der Gedankenauflistungen konnte festgestellt werden, dass bei Inkongruenz die Reliabilität der Konsensinformation geringer eingeschätzt wurde als bei Kongruenz (*the poll was invalid*). Demzufolge kann verzerrte systematische Verarbeitung unter dem Defense-Motiv auch dazu dienen, die bestehende eigene Einstellung gegen konträre Informationen zu schützen und zu verteidigen.

Mackie (1987) konnte in einer Studie feststellen, dass Versuchspersonen, die mit einer zu ihrer eigenen Meinung inkongruenten Konsensinformation konfrontiert wurden, ihre Meinung daraufhin änderten und die in der Konsensinformation vertretene Mehrheitsmeinung

annahmen und mindestens eine Woche beibehielten. Knipfer (2009) konnte ebenfalls feststellen, dass inkongruentes, konfligierendes Feedback im Vergleich zu den beiden Bedingungen ohne und mit kongruentem Feedback zu einer signifikanten Änderung des Gesamturteils in Richtung der Majorität führte. Zudem konnte bei der Lernbedingung, die konflikthafte Feedback erhielt, ein geringerer *myside bias* im Argument Repertoire, eine bessere Integration von Gegenargumenten in der Begründung der eigenen Position und eine bessere Diskussion von Gegenargumenten nachgewiesen werden als bei den beiden Kontrollgruppen (siehe Kapitel 4.6).

Die Präsentation einer Konsensinformation, die der eigenen Entscheidung widerspricht, löst in der Regel kognitive Dissonanz aus (Festinger, 1957; siehe Abbildung 6).

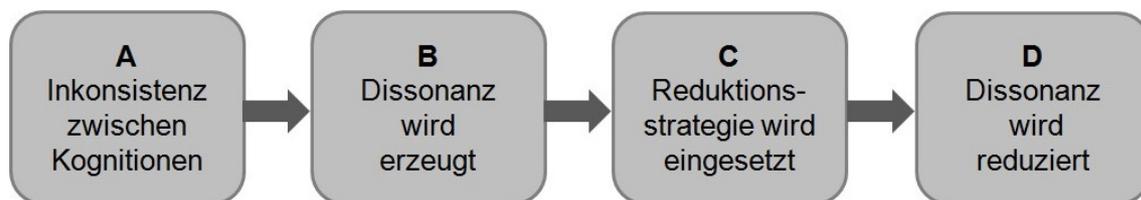


Abbildung 6: Schema von Festinger's (1957) Prozessmodell der Dissonanz (Devine et al., 1999, p. 298).

Grundannahme der Theorie der kognitiven Dissonanz von Festinger (1957) besteht darin, dass Menschen nach Konsistenz (Kongruenz, Konsonanz, Gleichgewicht) und Harmonie in ihrem kognitiven System streben und versuchen, diese aufrechtzuerhalten und wiederherzustellen.

Festinger (1957) geht davon aus, dass (A) eine Inkonsistenz zwischen Kognitionen (B) kognitive Dissonanz erzeugt. Unter Kognitionen werden dabei Meinungen, Attitüden, Werthaltungen, Glaubensweisen, Wissenseinheiten, etc. verstanden. Die Kognitionen eines Menschen können in einer konsonanten oder in einer dissonanten Beziehung stehen. Konsonante Kognitionen sind miteinander vereinbar (z. B. „Ich rauche viel“ und „Rauchen beruhigt mich“), dissonante dagegen nicht (z. B. „Ich rauche“ und „Rauchen schadet meiner Gesundheit“). Kognitionen gelten dann als dissonant, wenn sie für eine Person nicht miteinander vereinbar sind, d. h. sie müssen nicht tatsächlich in einem logischen bzw. kausalen Widerspruch zueinander stehen. Die Höhe der kognitiven Dissonanz wird nach Festinger (1957) durch das Verhältnis der dissonanten zu den konsonanten Kognitionen und durch die Wichtigkeit der beteiligten Kognitionen determiniert. Unter kognitiver Dissonanz wird ein physiologischer Spannungs- und Erregungszustand (*arousal*) verstanden, der psychisch unangenehm erlebt wird (*psychological discomfort*) und bei Menschen eine Motivation erzeugt, ihn zu reduzieren, um einen Gleichgewichtszustand in ihrem kognitiven System wiederherzustellen. Zur Reduktion der kognitiven Dissonanz (D) können verschie-

dene Strategien (C) eingesetzt werden (vgl. Devine et al., 1999; Festinger, 1957; Frey & Gaska, 1993). Es können (vgl. Peus, Frey & Stöger, 2006),

- neue konsonante Kognitionen hinzugefügt werden,
- bestehende dissonante Kognitionen subtrahiert werden (z. B. durch Ignorieren, Vergessen, Verdrängen)
- vorhandene Kognitionen durch andere Kognitionen ersetzt werden (dissonante Kognitionen werden abgezogen, wobei gleichzeitig konsonante Kognitionen hinzugefügt werden)

Es werden stets die Konditionen geändert, die den geringsten Änderungswiderstand besitzen. Nach Festinger ist eine Kognition besonders änderungsresistent, wenn sie (1) mit vielen anderen Kognitionen in konsonanter Beziehung steht und nur zu wenigen in dissonanter sowie (2) außerpsychische Realität repräsentiert.

Um kognitive Dissonanz nach der Konfrontation mit einer präferenzinkonsistenten Feedbackinformation zu reduzieren, kann eine Person z. B. ihre Entscheidung in Richtung der in der Information vertretenen Entscheidung (Majorität) ändern, die Informationsquelle abwerten (etwa als unglaubwürdig, inkompetent oder nicht relevant), den Inhalt der Information verzerren (passend machen) oder soziale Unterstützung bzw. bestätigende Informationen für die eigene Entscheidung suchen (vgl. Frey & Gaska, 1993).

Die Dissonanztheorie von Festinger (1957) postuliert ferner, dass nach fast jeder Entscheidung kognitive Dissonanz entsteht, da die gewählte Entscheidungsalternative in der Regel mit Nachteilen und die nicht-gewählte Option mit Vorteilen verbunden ist. Die Nachteile der gewählten und die Vorteile der nicht-gewählten Alternative stehen im Widerspruch zur getroffenen Entscheidung und werden vom Entscheider als dissonant erlebt.

Die entstandene Dissonanz kann entweder durch eine Revision der Entscheidung oder durch eine Änderung der Kognitionen in Bezug auf die Attraktivität der verschiedenen Alternativen reduziert werden. Da die Revision einer getroffenen Entscheidung meist mit einem hohen Aufwand verbunden ist, wird Dissonanz nach Entscheidungen typischerweise dadurch reduziert (vgl. Frey & Gaska, 1993; Peus et al., 2006), dass

- (1) die gewählte Alternative aufgewertet und die nicht gewählte Alternative abgewertet wird; dieser sog. „*spreading-apart*“ Effekt (Brehm, 1957) stellt nach Festinger (1957) die effektivste und häufigste Art der Dissonanzreduktion dar,
- (2) die Sicherheit darin geändert wird, dass die richtige Entscheidung getroffen wurde oder
- (3) bevorzugt mit der Entscheidung konsonante Informationen gesucht werden, welche die Richtigkeit der eigenen Entscheidung bestätigen und dissonante, die sie in Frage stellen, vermieden werden (Bestätigungsfehler: *confirmation bias*, vgl. D'Alessio & Allen, 2002; siehe Kapitel 3.2.3).

Je irreversibler und wichtiger die Entscheidung ist und je weniger kognitive Überlappung die Entscheidungsalternativen haben, umso stärker ist die danach auftretende Dissonanz bzw. der „*spreading-apart*“ Effekt ausgeprägt. Festinger (1957) geht des Weiteren von einem kurvilinearen Zusammenhang zwischen der Stärke der Dissonanz und dem Ausmaß der Selektivität der Informationssuche sowie der Informationsvermeidung aus. Je höher die Dissonanz, desto mehr wird die Person nach der Entscheidung versuchen, die gewählte Alternative durch die selektive Suche nach unterstützenden Informationen abzusichern. Gemäß dem Fall, dass die Dissonanz so stark ausgeprägt ist, dass jene auf diese Weise nicht mehr adäquat reduziert werden kann, wird die Person im Extremfall sogar bevorzugt nach dissonanten Informationen suchen, um ihre erste Entscheidung, falls möglich, zu revidieren (Frey et al., 2010).

Zur indirekten Erfassung von Dissonanz bzw. des erlebten psychologischen Unbehagens (*psychological discomfort*) entwickelten Elliot & Devine (1994) einen Affektfragebogen und testeten ihr „Dissonanzthermometer“ in einer Reihe von experimentellen Studien. In diesem Affektfragebogen werden sowohl negative Gefühle (z. B. beunruhigt, unangenehm) thematisiert, die in einem Zusammenhang mit Erleben von Dissonanz stehen sollten, als auch positive Gefühle (z. B. glücklich) und negative Gefühle (z. B. schuldig), die nicht mit dem Erleben von Dissonanz assoziiert sein sollten. In den zu diesem Affektfragebogen durchgeführten Experimenten konnte nachgewiesen werden, dass es sich bei Dissonanz tatsächlich um einen als aversiv erlebten psychischen Zustand handelt und die mit der Dissonanz erlebten unangenehmen Gefühle zur Dissonanzreduktion motivieren, um den negativen Affekt zu verringern. Auch in neueren Studien konnten die Befunde von Elliot & Devine (1994), dass Dissonanz unangenehm erlebt wird und einen motivierenden Faktor zur Dissonanzreduktion darstellt, repliziert werden (Harmon-Jones, 2000).

4.5 Multimediale Lernumgebungen im Museum

Zu den klassischen Aufgaben von Museen zählen das Sammeln, Bewahren, Erforschen und Ausstellen, wobei seit den 1970er Jahren die Präsentations- und Vermittlungsaufgabe und damit die Bildungsfunktion immer stärker in den Vordergrund gerückt ist (Hense, 1990; KMK, 1996; Schwan et al., 2008). Aufgrund der Forderung nach lebenslangem Lernen gewinnen informelle Lernorte wie Museen in den letzten Jahren zunehmend mehr an Bedeutung (Anderson, 1999; Chadwick, 2004; Dohmen, 2001; Harms & Krombaß, 2008).

Eine zentrale Bildungsaufgabe von naturwissenschaftlich-technischen Museen stellt die Wissenschaftskommunikation dar: Museen sollen heutzutage nicht nur im Sinne eines PUS naturwissenschaftlich-technische Fakten und „Lehrbuchwissen“ vermitteln, sondern auch

im Sinne eines PUR einen Dialog zwischen Gesellschaft und Wissenschaft fördern und die gesellschaftliche Relevanz sowie den Prozesscharakter der Forschung aufzeigen. Es geht dabei darum, ein vertieftes Verständnis für die Arbeitsweise der Forschung zu erzeugen und aufzuzeigen, welche verschiedenen, auch konträre, Positionen innerhalb der *scientific community* existieren und welche gesellschaftlichen, ethischen und politischen Implikationen daraus resultieren (vgl. Chittenden et al., 2004; Durant, 2001; Hauser, 2005, 2006; Lewalter & Noschka-Roos, 2009; Teichmann & Noschka-Roos, 2006).

Insbesondere die Nano- und Gentechnologie stellen aktuell besonders gesellschaftsrelevante Forschungsfelder dar, da diese Schlüsseltechnologien des 21. Jahrhunderts uns alle betreffen und mit großen Chancen, aber auch zahlreichen Risiken verbunden sind und mit ihren Zukunftsvisionen Ängste und Hoffnungen wecken.

Die Lernumgebung Museum zielt in diesem Zusammenhang darauf ab, situationales Interesse bei den Besuchern an solch kontroversen Ausstellungsthemen zu wecken, zur Beschäftigung mit diesen gesellschaftlich relevanten und komplexen Sachverhalten anzuregen und die Besucher dabei zu unterstützen, sich eine Einstellung/Meinung zu diesen Sachverhalten bilden zu können oder diese zu reflektieren (Lewalter, 2009). Ausstellungen in Museen sollten so geplant und gestaltet werden, dass sie möglichst positive kognitive und motivationale Lernprozesse bei den Besuchern anstoßen und auch zum nachhaltigen Lernen anregen, d. h. auch zu einer späteren Auseinandersetzung mit dem Ausstellungsthema nach dem Museumsbesuch und außerhalb des Museums führen (Prenzel, 2009).

Zur Ausübung ihrer Bildungsfunktion verwenden Museen eine Vielzahl an Medien, die von klassischen Texten, Schaubildern über dreidimensionale Modelle (sog. Dioramen) bis hin zum Einsatz neuer Medien reichen. Aufgrund des technologischen Wandels werden multimediale Lernumgebungen mit steigender Tendenz im Museum eingesetzt: Mittlerweile bilden digitale Medien einen festen Bestandteil von Museumsausstellungen (Reussner, Schwan & Zahn, 2007; Scheersoi, 2006; Schwan et al., 2008).

Multimediale Lernumgebungen eignen sich hervorragend zur Vermittlung komplexer Wissenschaftsthemen (Azevedo, 2005). Insbesondere Themen wie die Gen- oder Nanotechnologie stellen die Gestalter von Ausstellungen aufgrund ihrer immanenten Komplexität und Unanschaulichkeit derzeit vor besondere Herausforderungen und erfordern neue Präsentationsformen in Ausstellungen (Gerber-Hirt & Noschka-Roos, 2012; Schwan et al., 2008).

Darüber hinaus interessieren sich Besucher nicht mehr nur allein für die Vermittlung der einzelnen Techniken, sondern auch verstärkt für die mit der Anwendung verbundenen Chancen und Risiken (Hauser, Noschka-Roos & Teichmann, 2007). Um die interessierte Öffentlichkeit in die Lage zu versetzen, sich in einer Ausstellung einen Überblick über die

verschiedenen Standpunkte zu kontrovers und brisant diskutierten Wissenschaftsthemen zu verschaffen und auf dieser Grundlage begründete Meinungen und Urteile bilden zu können, bietet sich der Einsatz digitaler Medien deshalb besonders an (Schwan et al., 2008).

Im Gegensatz zu formalen Bildungsinstanzen wie Schule und Hochschule findet im Museum informelles Lernen statt, das durch einen hohen Grad an Freiwilligkeit und durch eine Vielzahl von Wahlmöglichkeiten (Selbststeuerung) gekennzeichnet ist. Falk & Dierking (2000) bezeichnen das Lernen im informellen Museumssetting daher auch als „*free-choice learning*“. Ein Museumsbesuch zeichnet sich in der Regel dadurch aus, dass er in der Freizeit, aus Eigeninitiative und damit freiwillig stattfindet: Es existiert keine Anwesenheitspflicht und kein Lehrplan oder Curriculum mit extern vorgegebenen Lernzielen, vielmehr bestimmen die Besucher selbst Zeit, Richtung und die Art der Besichtigung (Hein, 1998; Noschka-Roos, 2001). Nach Csikszentmihalyi & Hermanson (1999) und vielen Museumsfachleuten gilt die intrinsische Motivation als ein wichtiger Motor für das Lernen im Museum.

Im Folgenden werden nach Lewalter & Noschka-Roos (2009) situative Merkmale der informellen Lernumgebung Museum beschrieben. Das zentrale Kennzeichen naturwissenschaftlich-technischer Museen besteht typischerweise in der Präsentation von authentischen Originalobjekten (Exponaten) oder Modellen (vgl. Paris, 2002). Um möglichst effektive Lernprozesse zu initiieren, werden die verwendeten Objekte im Ausstellungskontext so arrangiert, dass sie den Besucher zu einer möglichst selbstständigen und intensiven Auseinandersetzung mit den Inhalten anregen. Insbesondere interaktive Ausstellungselemente eröffnen hohe Freiheitsgrade in Bezug auf die Auseinandersetzung mit dem präsentierten Informationsangebot, indem sie vielfältige Zugangsmöglichkeiten bereitstellen.

Die Darstellung von Sachverhalten unter multiplen Perspektiven bietet zahlreiche Anknüpfungspunkte an eigene Erfahrungen und bereits vorhandene Wissensbestände. Aufgrund der hohen Authentizität der im Museum verwendeten Originalobjekte wird dem Besucher das Erkennen der Relevanz der präsentierten Inhalte für bestimmte Lebensbereiche oder Praxisfelder sowie eine flexible Anwendung des neu erworbenen Wissens erleichtert. Indem verschiedene Informationshierarchien genutzt werden können, können die Besucher selbst den Intensitätsgrad der Auseinandersetzung mit einem Thema in Abhängigkeit von ihrem individuellen Vorwissen und Interesse bestimmen.

Zu guter Letzt stellt ein Museumsbesuch auch immer ein soziales Ereignis dar, denn er findet häufig nicht allein, sondern in Begleitung der Freunde oder Familienmitglieder statt, die sich dann im Museum gemeinsam mit den einzelnen Ausstellungselementen beschäftigen. Aus diesem Grunde werden einzelne Ausstellungselemente oftmals so konzipiert, dass sie auch kooperative Lernprozesse ermöglichen.

Eine einzelne Ausstellungseinheit sollte auch so gestaltet sein, dass sie unabhängig vom Gesamtkontext der Ausstellung funktioniert. Die museale Lernumgebung sollte Neugier (Loewenstein, 1994) und situationales Interesse am Ausstellungsthema wecken (Krapp, 1992), verständlich sein und zu möglichst einer intensiven und langen Auseinandersetzung mit dem dargestellten Inhalt anregen (Schwan, 2009; Schwan et al., 2008).

Damit entsprechen die oben erwähnten situativen Merkmale der Lernumgebung „Museum“ den wesentlichen Annahmen des *gemäßigten Konstruktivismus*, in dem die Auseinandersetzung mit dem Lerngegenstand aktiv, konstruktiv, selbstgesteuert, situativ und sozial verläuft (vgl. Hein 1998; Reinmann & Mandl, 2006; siehe Kapitel 4.3.1). Demzufolge liegen im Museum aus lernpsychologischer Sicht äußerst günstige Rahmenbedingungen für den Erwerb von vertieftem und flexibel anwendbarem Wissen vor (Schwan, 2005).

Allerdings zeigen jedoch zahlreiche Befunde aus der Besucherforschung, dass diese günstigen Rahmenbedingungen nicht notwendigerweise auch zu einem vertieften und flexiblen anwendbaren Wissen führen, denn häufig steht einer überwiegend intrinsisch motivierten Auseinandersetzung mit dem musealen Angebot auch ein sehr stark selektives und auch zeitlich eingeschränktes Nutzungsverhalten gegenüber (Schwan, 2005): Die überwiegende Mehrheit der Museumsbesucher geht eher unsystematisch und selektiv durch eine Ausstellung, dementsprechend wird in der Regel nicht jedes einzelne Ausstellungsobjekt betrachtet, sondern nur eine beschränkte Anzahl an Exponaten ausgewählt, wobei die Auseinandersetzung oft nur sehr kurz und oberflächlich erfolgt und die Zuwendungsdauer selten 2 bis 3 Minuten überschreitet (vgl. z. B. Hein, 1998; Morris Hargreaves McIntyre, 2004; Schwan et al., 2008; Serrell, 1997). Besucher bewegen sich somit mit einer gewissen Zeitökonomie durch die Museumsausstellungen (Noschka-Roos, 2001).

Daneben spricht eine Ausstellung generell ein sehr heterogenes Publikum, z. B. im Hinblick auf Alter oder Bildungsstand, an, denn Museen werden von den unterschiedlichsten Menschen besucht (Schwan, 2005). Insbesondere in Bezug auf das Vorwissen, Motivation und Interesse über das Ausstellungsthema und die Exponate unterscheiden sich die Besucher von Museen sehr stark (Falk & Dierking, 2000; Falk & Storksdieck, 2005). Dieser Heterogenität des Zielpublikums und dem zeitlich eingeschränkten und selektivem Nutzungsverhalten sollte bei der Konzeption einer Medienstation Rechnung getragen werden, wenn man erreichen möchte, dass die museale Lernumgebung potentiell von allen Besuchern genutzt wird.

Um Unterschiede in Vorwissen, Motivation, Interesse und Informationskompetenz der Besucher kompensieren zu können, werden deshalb beim informellen und selbstgesteuerten Lernen im Museum besonders hohe Anforderungen an eine besucherorientierte Präsentation und Vermittlung der Ausstellungsinhalte gestellt (Schwan, 2005, 2009). Wie empiri-

sche Studien zeigen konnten, kann durch eine angemessene Konzeption der Ausstellung eine quantitativ und qualitativ deutlich bessere Auseinandersetzung mit den Inhalten bei den Besuchern erzielt werden. Digitale Medien, insbesondere die Verwendung innovativer Formen, eignen sich dabei sehr gut zur Unterstützung von Informationsverarbeitungs- und Wissenserwerbsprozessen in Museen und Ausstellungen (vgl. Geser & Pereira, 2004; Sandifer, 2003; Schwan, 2005). Eine Adaption an die Heterogenität des Museumspublikums kann durch den Einsatz digitaler Medien beispielsweise in Form von computerunterstützten Informationsterminals angemessen unterstützt werden, indem diese Terminals den Besuchern ergänzende, erläuternde oder vertiefende Informationen zum Ausstellungsthema bereitstellen, die ihnen erlauben, „eine individualisierte Themenvertiefung zu realisieren, die ihrem Vorwissen und ihren Interessen entspricht“ (Schwan, 2005, S. 50).

Neben Informationsterminals werden neue Medien in Museen und Ausstellungen auch sehr häufig zur Erschließung eines Objekts oder einer ganzen Objektsammlung eingesetzt oder das Medium fungiert selbst als (digitales) Exponat (vgl. Noschka-Roos, 2006).

In naturwissenschaftlich-technischen Museen kommt zunehmend auch das Erzählen von Geschichten zur Wissensvermittlung zum Einsatz (vgl. Töpfer, 2009). Storytelling im Museum kann sowohl mithilfe traditioneller Medien (z. B. Text) als auch mit neuen Medien (z. B. Video) umgesetzt werden, im letztgenannten Fall spricht man von Digital Storytelling (vgl. Mangold & Woletz, 2007). Als Beispiel für die Umsetzung von Digital Storytelling im Museum kann die in Kapitel 5 beschriebene Dialogstation „Gentest“ angeführt werden, in der betroffene Personen ihre Geschichte zur Gentestproblematik erzählen.

Im Hinblick auf die Konzeption von Computerstationen im Museum gibt Scheersoi (2006) folgende Empfehlungen: Zum einen sollte eine Medienstation aus inhaltlicher Sicht dem Besucher spannende, eventuell auch unerwartete Erfahrungen ermöglichen und einen persönlichen Bezug zum Thema herstellen, indem der Besucher erfährt, warum das Ausstellungsthema für ihn relevant ist. Die äußere Form betreffend, sollte eine Computerstation über eine einfache und klare Bedienung verfügen, in einem ansprechenden und attraktiven Design gestaltet sein und eine kurze „Spieldauer“ aufweisen, damit das Programm von vielen Besuchern ohne lange Wartezeiten genutzt werden kann.

Mit der Frage, ob und was im Museum gelernt wird, befasst sich die Besucherforschung. Generell ist die konkrete Lernwirkung einer Museumsausstellung sehr schwierig zu untersuchen, da Faktoren wie die Qualität der Ausstellungsinhalte, Objekte und des Ausstellungsdesigns nur sehr schwer empirisch messbar sind (Graf, 2003). Zudem tritt die Lernwirkung im Sinne von Veränderungen in kognitiven und motivationalen Strukturen der Ausstellungsbesucher häufig erst nach dem Besuch und außerhalb des Museums auf. Solch ein nachhaltiges Lernen geschieht beispielsweise, wenn über die im Museum gemachten

Erfahrungen mit der Familie oder Freunden später gesprochen, vertiefende Literatur oder Medienbeiträge zum Ausstellungsthema rezipiert werden (Graf, 2003; Prenzel, 2009).

Die Besucherforschung kann in folgende fünf Gegenstandsfelder untergliedert werden (Graf & Noschka-Roos, 2009; Klein, 2003; Noschka-Roos, 1994):

(1) Die *Besucherstrukturanalyse*, die das Ziel verfolgt, ein detailliertes Bild vom Publikum zu gewinnen, um angemessene Maßnahmen für die Öffentlichkeitsarbeit oder die Ausstellung ableiten und bestimmen zu können. Erhoben werden beispielsweise soziodemographische Daten wie Alter, Geschlecht, Sozialstatus, Ausbildung oder die Besuchsmotivation, also wer kommt, warum, wann, wie lange, wie oft, usw.

(2) Die *Besucherbeobachtung* befasst sich mit der systematischen Erfassung von Reaktions- und Interaktionsweisen von Museumsbesuchern, ihrem Besuchsverlauf und wo und wie lange sich die Besucher aufhalten, usw.

(3) Die *experimentelle Forschung*, in welcher das Ausstellungsarrangement oder einzelne isolierte Elemente gezielt verändert werden, um ihre Wirkung auf den Besucher zu untersuchen. „Diese gezielte Veränderung kann aufgrund hypothetischer Vorannahmen erfolgen oder der empirischen Erprobung alternativer Ausstellungs-elemente dienen“ (Noschka-Roos, 1994, S. 166).

(4) Die *Evaluation* als Gegenstandsfeld der Besucherforschung beschäftigt sich mit der Überprüfung der Effektivität einer Ausstellung. In der Entwicklungsphase einer Ausstellung kommt die formative Evaluation zum Einsatz, in welcher die Zielsetzung verfolgt wird, Planungs- und Entscheidungshilfen beispielsweise zur Optimierung der Ausstellung bzw. einzelner Sequenzen zu liefern. Demgegenüber zielt die summative Ausstellungsevaluation auf die abschließende Überprüfung der erwünschten Lerneffekte ab. Um die Effektivität einer Ausstellung zu überprüfen, werden in empirischen Untersuchungen zum Besucherverhalten zwei Parameter als essentiell angesehen (Shettel, 1968; Noschka-Roos, 1994). Einerseits die *attracting power*, das Attraktivitätspotential, das als die Stopzahl der Besucher an der Ausstellungseinheit definiert wird, andererseits die *holding power*, welche die Verweilzeiten der Ausstellungsbesucher misst. Durch die Erfassung des Besucherverhaltens (z. B. Reaktionen oder Aussagen) lassen sich ebenfalls Rückschlüsse auf die Qualität der Ausstellung ziehen.

(5) *Theoretische Ausführungen und Quellenmaterialien*, die an dieser Stelle nicht näher ausgeführt werden.

Als grundlegendes Ergebnis von Besucherstrukturanalysen lässt sich festhalten, dass jede Museumsart eine charakteristische Besucherstruktur in Bezug auf die Begleitsituation und soziodemographischen Merkmale zeigt (Graf, 2003; Graf & Noschka-Roos, 2009; Klein, 1990). Im Jahr 1998 und 1999 wurde beispielsweise im Deutschen Museum eine Besucherstrukturanalyse durchgeführt, in denen 754 Interviews mit Besuchern geführt wurden

(Klein, 2000). Die Ergebnisse zeigen, dass das Publikum des Deutschen Museums recht jung ist. Das Durchschnittsalter der Besucher liegt bei etwa 35 Jahren, würde man die Kinder unter 15 Jahren miteinfassen, würde sich wahrscheinlich das durchschnittliche Alter sogar unter die 30-Jahre-Grenze verschieben. Die Geschlechterproportionen fallen zugunsten männlicher Besucher aus. 60 % der Befragten waren männlich, 40 % weiblich, wobei die Geschlechterproportionen vom Alter abhängen, denn je älter die Besucher sind, desto höher ist der Männerüberschuss. In Bezug auf die Begleitsituation zeigte sich, dass etwa ein Fünftel (19 %) der Besucher alleine ins Deutsche Museum kamen, 30 % mit ihrem (Ehe-) Partner, 6 % mit Kindern, 12 % mit anderen Familienangehörigen, 14 % mit Freunden, Kollegen oder Bekannten, 20 % mit einer organisierten Gruppe. Im Vergleich zum Bevölkerungsdurchschnitt ist das Publikum des Deutschen Museums überdurchschnittlich gebildet und weist häufig einen naturwissenschaftlich-technischen Bildungshintergrund auf. Es befinden sich überproportional viele Akademiker in der Publikumsstichprobe, etwa 40 % der befragten Besucher sind Akademiker, wobei interessanterweise der Anteil der naturwissenschaftlich-technischen Akademiker denjenigen der geistes- oder sozialwissenschaftlichen mit zunehmendem Alter immer klarer überwiegt. Die wesentlichen Ergebnisse der Besucherstrukturanalyse von Klein (2000) konnten auch in anderen Besucherstudien des Deutschen Museums zur Ausstellungsevaluation repliziert werden (vgl. z. B. Geyer, Noschka-Roos & Lewalter, 2005; Lewalter, 2001; Lewalter & Geyer, 2003; Specht & Lewalter, 2011).

Nachfolgend werden für diese Arbeit relevante Befunde aus der Besucherforschung präsentiert, die in Evaluations- und experimentellen Studien gewonnen wurden.

Serrell (2002) untersuchte beispielsweise im Rahmen einer summativen Evaluationsstudie fünf Videos, die Bestandteil der Ausstellung „*The Changing Faces of Women's Health*“ waren und in denen verschiedene Frauen über die Themen Menopause, Herzerkrankung, Menstruation, Mammographie und Brustkrebs berichteten. Die Frauen in den Videos waren keine Schauspieler, sondern wirklich betroffene Personen, die in den Videos ihre eigene Lebensgeschichte zu dem jeweiligen Thema sehr emotional erzählten. Die *attracting power* der fünf Filme lag zwischen 12 % und 35 %, mit einem Gesamtdurchschnitt von 23 %. Die *holding power* variierte zwischen .03 und .50, die durchschnittliche *holding power* betrug .20. In dieser Studie wurde die *holding power* in Abhängigkeit von der Filmlänge operationalisiert: Die Besucher sahen zwischen 3 % und 50 % der gesamten Filmlänge, die durchschnittliche Betrachtungsdauer lag bei 20 % der Gesamtlänge des Films. In Befragungen, die zwei Monate später durchgeführt wurden, konnte festgestellt werden, dass die Videos sehr beliebt bei den Besuchern waren und gut erinnert wurden. Auch Allen (2004) konnte in einer Studie am Exploratorium in San Francisco belegen, dass narrative

Filme im Sinne eines personalisierten Storytellings in *follow up*-Interviews nach dem Museumsbesuch gut erinnert werden konnten.

In einer Metastudie analysierte Serrell (2002) zusätzlich mehrere Evaluationsstudien verschiedener amerikanischer Museen, Aquarien und Zoos und konnte feststellen, dass die durchschnittliche *attracting power* aller eingesetzten Filme bei 32 % lag, etwa ein Drittel aller Besucher sahen sich demzufolge Filme bei einem Besuch an. Die *holding power* betrug durchschnittlich 137 Sekunden, d. h. ein Film wurde im Mittel etwas über 2 Minuten geschaut. Ausgehend von einer durchschnittlichen Aufenthaltsdauer von etwa 20 Minuten in einer Ausstellung empfiehlt Serrell (2002) deshalb, dass die Dauer eines Videos nicht mehr als 3 Minuten überschreiten sollte. Noschka-Roos (2001) gibt hingegen eine großzügigere Empfehlung von 3 bis 5 Minuten, da die Betrachtungsdauer von Filmen in Ausstellungen je nach Museumskontext und Inhalt stark variieren kann; in kleinen Museen und Sonderausstellungen betrachten Besucher auch längere Filme vollständig.

Töpfer (2009) untersuchte im Rahmen einer Feld- und einer Laborstudie erstmals die Effekte der filmischen Personalisierung von Ausstellungsinhalten auf das Interesse und den Wissenserwerb. Als Untersuchungsgegenstand diente eine reale und virtuelle Version der Ausstellung „Leben mit Ersatzteilen“ des ZNT des Deutschen Museums zum Thema Medizintechnik, die jeweils mit einer neutralen Filmversion verglichen wurden.

In der Laborstudie wurde der Einfluss von personalisierten Filmen auf die Bewertung und den Wissenserwerb zu den Inhalten der virtuellen Version der Ausstellung „Leben mit Ersatzteilen“ im Rahmen einer zweifachen Postmessung (direkt vs. eine Woche später) in Abhängigkeit vom Filmtyp (personalisiert vs. neutral) untersucht. Allen 89 Probanden wurden vier Teilausstellungen zu den Themen „Auge und Sehhilfen“, „Ohr und Hörhilfen“, „Innere Organe und Herzunterstützungssysteme“ und „Gliedermaßen und Prothesen“ im Rahmen einer computerbasierten virtuellen Ausstellung präsentiert. Die personalisierten Filme stellten Betroffeneninterviews zu den Themen der Teilausstellungen dar, in denen eine von der Thematik betroffene Person in einem umgangssprachlichen und dialogorientierten Sprachstil (Ich-Erzählform) von ihrem Leben mit dem jeweiligen Ersatzteil (z. B. einer Beinprothese) erzählte. Die neutralen Filme wurden entgegen den Mayer'schen Prinzipien neutralisiert, indem statt eines Sprechers thematische Bilder gezeigt wurden (*Prinzip des Bildes*) und der Text in einer neutralen Erzählform („Betroffener“, „Patient“) von einer Stimme aus dem Off gesprochen wurde (*Personalisierungsprinzip*). Bei beiden Filmversionen wurden die identischen Informationsinhalte präsentiert.

Wie die Ergebnisse der Laborstudie zeigen konnten, wurden die personalisierten Videos im Vergleich zu den inhaltsäquivalenten, neutralen Videos besser bewertet und weckten mehr Interesse am Thema der Teilausstellung, motivierten jedoch entgegen der Vorannahmen

nicht zu einer positiveren Bewertung und längeren Verweildauer (*holding power*) in der zum jeweiligen Film begleitenden Teilausstellung. In Bezug auf den Wissenserwerb ergab sich bei der Analyse des Multiple-Choice-Wissenstests, dass die Ausstellungsinhalte, die von einem personalisierten Video präsentiert wurden, besser behalten wurden als bei den neutralen Versionen. Dieser Effekt zeigte sich sowohl direkt nach der Untersuchung als auch eine Woche später. Die vorgefundenen Behaltensunterschiede konnten jedoch hauptsächlich auf die Inhalte der Videos und nicht auf die begleitenden Teilausstellungen zurückgeführt werden. Mit Ausnahme der Teilausstellung „Auge und Sehhilfen“ konnte entgegen der Vorannahmen kein Einfluss des persönlichen Bezugs zur Thematik und des Vorwissens auf den Wissenserwerb zu den Teilausstellungen festgestellt werden.

Die Überlegenheit der personalisierten Videos gegenüber den neutralen Versionen konnte auch in der Feldstudie, die in der realen Ausstellung „Leben mit Ersatzteilen“ durchgeführt wurde, bestätigt werden. Variiert wurde analog zur Laborstudie der Filmtyp (personalisiert vs. neutral). Als abhängige Variable wurde bei 186 Ausstellungsbesuchern die Verweildauer (*holding power*) in der Gesamtausstellung und in den Teilausstellungen erhoben. Von den 186 beobachteten wurden 153 Personen zusätzlich befragt, wie interessant sie die jeweiligen Teilausstellungen fanden und wie vertieft sie in den einzelnen Teilausstellungen waren. Insgesamt konnte durch die verdeckte Beobachtung ermittelt werden, dass 46 Ausstellungsbesucher sich die Filme angeschaut hatten (neutral: $n = 12$; personalisiert: $n = 34$), von denen 30 befragt werden konnten. Zusätzlich wurde sowohl die *attracting power* als auch die *holding power* der beiden Filmtypen (personalisiert vs. neutral) erhoben.

Im Vergleich zu den neutralen Filmen konnte bei den personalisiert präsentierten Filmen eine signifikant höhere *attracting* und *holding power* nachgewiesen werden, d. h. es blieben mehr Besucher bei dem jeweiligen Video stehen und betrachteten es auch länger. Zudem verweilten Besucher, die sich einen der personalisierten oder neutralen Filme angeschaut hatten, signifikant länger in der jeweiligen Teilausstellung und der Gesamtausstellung als Besucher, die keine der Filme betrachteten. Im realen Ausstellungskontext konnte auch im Gegensatz zur Laborstudie ein Einfluss des persönlichen Bezugs zur Thematik auf das Besucherverhalten und die Bewertung (Interessanz) der Teilausstellungen festgestellt werden: Besucher mit einem hohen persönlichen Bezug zum Thema der einzelnen Teilausstellung verweilten länger in der jeweiligen Teilausstellung, bewerteten diese Teilausstellung gleichzeitig auch als interessanter und waren vertiefter als Besucher mit einem geringen Bezug zum Thema.

Zusammenfassend betrachtet konnte Töpfer (2009) den lernförderlichen Einfluss der filmischen Personalisierung von Ausstellungsinhalten erstmals im musealen Lernsetting nachweisen: Es konnte tatsächlich ein höheres Interesse am Thema geweckt und die

Behaltensleistung der vermittelten Inhalte gesteigert werden. Allerdings zeigten sich diese positiven Effekte auf den Wissenserwerb und das Interesse am Thema nur lokal in Bezug auf die Inhalte der Videos und übertrugen sich nicht auf die jeweiligen begleitenden Teilausstellungen.

Nachfolgend soll exemplarisch eine Computerstation vorgestellt werden, die für die Sonderausstellung „Leben mit Ersatzteilen“ zum Thema Medizintechnik im Deutschen Museum München eingesetzt wurde und einen Vorläufer der in dieser Arbeit untersuchten Dialogstation darstellt (vgl. Hauser, 2006, 2009). Abbildung 7 zeigt einen Screenshot des Diskussions- bzw. Meinungsterminals zum Thema Stammzellforschung, das in der Ausstellung „Leben mit Ersatzteilen“ des ZNT integriert war.

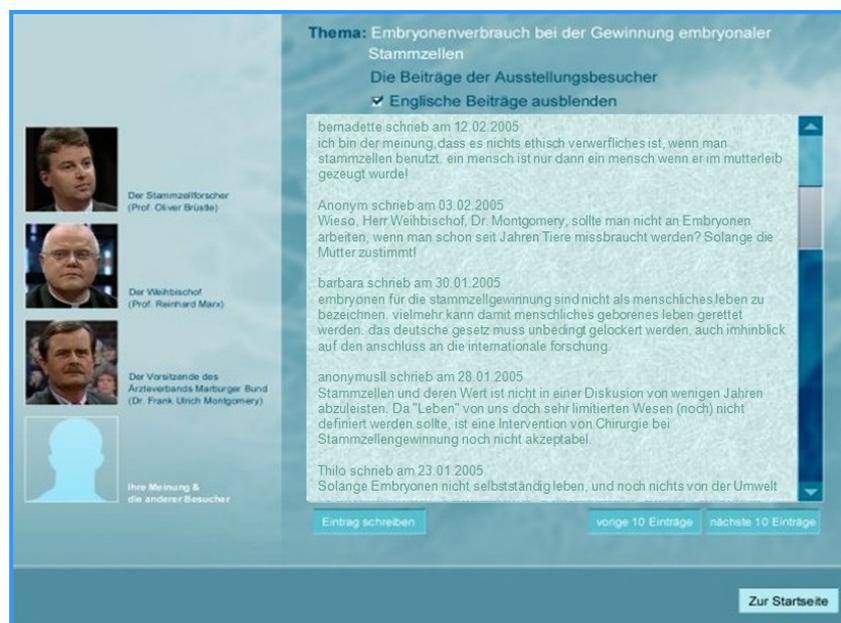


Abbildung 7: Meinungsterminal in der Ausstellung „Leben mit Ersatzteilen“ des ZNT zum Thema Stammzellforschung, Eingabe-PC (Screenshot Touchscreen-Interface).

In diesen Meinungsterminals, die in Form klassischer Touchscreen-Interfaces (Eingabe-PCs) realisiert wurden, konnten sich die Besucher über ethisch kontroverse Themen wie embryonale Stammzellforschung oder Organtransplantationen informieren sowie konträre Meinungen von Experten in Form von Videos abrufen. Darüber hinaus konnten die Besucher ihre eigene, persönliche Meinung zur Thematik über eine Tastatur in eine Meinungsdatenbank eintragen und die Meinungen anderer Besucher lesen.

Im Rahmen einer Besucherbefragung zur Evaluation der Sonderausstellung „Leben mit Ersatzteilen“ im ZNT des Deutschen Museums München konnte allerdings festgestellt werden, dass die Meinungsterminals von den Besuchern nur wenig und damit suboptimal genutzt wurden (Geyer et al., 2005). Von den 240 Befragten beschäftigte sich die Mehrheit von 145 (60.4 %) der befragten Besuchern überhaupt nicht, 29 (12.1 %) wenig, 43

(17.9 %) etwas, 20 ziemlich (8.3 %) und lediglich 1 Besucher (0.4 %) sehr intensiv mit den Eingabe-PCs. In diesem Zusammenhang stellt sich die Frage, ob die Eingabe-PCs möglicherweise ungünstig in der Ausstellung platziert waren, denn die Nutzung einer Multimedia-Station kann auch vom Standort abhängen (vgl. Hashagen, 2001). Eine andere mögliche Erklärung der vorgefundenen suboptimalen Nutzung könnte darin bestehen, dass die Meinungsterminals als unattraktiv von den Besuchern wahrgenommen wurden. Ferner zeigte sich laut Hauser (2009) „auch das Problem der teils zweifelhaften Qualität der Meinungsäußerungen von Besuchern in einem Setting, in dem die Inhalte von vielen nur oberflächlich rezipiert werden“ (S. 86).

Zusammenfassend lassen sich aus den vorgestellten Theorien und Ergebnissen aus der Besucherforschung für die inhaltliche Aufbereitung und didaktische Gestaltung multimedialer Lernumgebungen im Museum folgende Schlussfolgerungen ziehen: Aufgrund des selektiven und zeitlich eingeschränkten Nutzungsverhaltens der Besucher sollte eine multimediale Lernumgebung in kurzer, prägnanter und verständlicher Form die wesentlichen Lerninhalte vermitteln. Für den Einsatz von Videos in Ausstellungen empfiehlt Serrell (2002) eine Dauer von 2 bis höchstens 3 Minuten. Im Gegensatz zu langfristigen Interventionen wie z. B. dem Schulunterricht, stellen Computerstationen im Museum demzufolge sehr kurzfristige Interventionen dar.

Multimediale Lernumgebungen im Museum sollten als einzelne Ausstellungseinheit so gestaltet sein, dass sie unabhängig vom Gesamtkontext der Ausstellung funktionieren, Neugier und Interesse beim Besucher wecken und zu einer vertieften Auseinandersetzung mit dem dargestellten Inhalt, zur Kommunikation über das Ausstellungselement und zum nachhaltigen Lernen anregen (Prenzel, 2009; Schwan et al., 2008).

4.6 Empirische Studien zur Effektivität von multimedialen Lernumgebungen zur Vermittlung von kontroversen Wissenschaftsthemen

Zum Abschluss des Kapitels werden relevante empirische Befunde zur Lernwirksamkeit von multimedialen Lernumgebungen zur Vermittlung von kontroversen Wissenschaftsthemen präsentiert.

1) *Ergebnisse von zwei experimentellen Studien zum Lernpotential von Diskussionsterminals im Hinblick auf die Unterstützung von kritischem Denken und reflektiertem Urteilen im Museum im Bereich Nanotechnologie.*

Der aktuelle Forschungsstand hinsichtlich der Lernwirksamkeit von Dialogstationen, die speziell für den informellen Lernort Museum konzipiert wurden, ist noch recht dürftig, da

es sich um völlig neuartige Lernumgebungen handelt. Erste Befunde zur Lernwirksamkeit und zur Effektivität bestimmter Funktionalitäten dieser innovativen Lernumgebungen wurden von der Psychologin Kristin Knipfer (2009) im Rahmen eines Dissertationsprojektes geliefert. Knipfer (2009) untersuchte in zwei aufeinander aufbauenden experimentellen Studien am Beispiel eines prototypisch realisierten Diskussionsterminals zur Nanotechnologie, inwieweit der Wissenserwerb und eine fundierte Meinungsbildung zu komplexen und kontrovers diskutierten Wissenschaftshemen im Rahmen von Museumsausstellungen durch bestimmte Funktionalitäten explizit unterstützt und gefördert werden kann.

Meinungsbildung beschreibt und konzeptualisiert Knipfer (2009) dabei als kritisches Denken und reflektiertes Urteilen im Museumskontext. Wie der Name Diskussionsterminal bereits nahelegt, bildet die Diskussion und Debatte unter den Besuchern den zentralen Bestandteil dieser Lernumgebungen. Im Gegensatz zu der in dieser Arbeit untersuchten Dialogstation bieten Diskussionsterminals dem Besucher nicht nur die Möglichkeit, seine Meinung zu äußern und Feedback über das Meinungsbild der anderen Besucher zu erhalten, sondern darüber hinaus auch eigene Statements zum Thema schriftlich einzugeben und die Meinungen anderer Besucher zu lesen. Das von Knipfer (2009) prototypisch realisierte Diskussionsterminal lehnt sich an die Konzeption des Meinungsterminals Stammzellforschung in der Ausstellung „Leben mit Ersatzteilen“ an (siehe Kapitel 4.5, Abbildung 7).

Knipfer (2009) hält insbesondere drei zentrale Designprinzipien von Diskussionsterminals relevant, um Wissenserwerb und eine fundierte Meinungsbildung in Bezug auf kontrovers diskutierte und ambivalente Wissenschaftsthemen im Rahmen von Museumsausstellungen zu fördern: die Salienz (Auffälligkeit) der verfügbaren Argumente, die aktive Positionierung und die Information über andere Meinungen zum Thema. Ein Meinungsterminal bietet den Besuchern die Möglichkeit, kontroverse Expertenpositionen zur aktiven Auseinandersetzung vor der eigenen Meinungsäußerung evaluieren zu können, im Anschluss daran seine eigene Meinung eingeben zu können und im Gegenzug Feedback über das Meinungsbild der anderen Besucher in Form einer sozialen Vergleichsinformation zu erhalten.

Diese drei Faktoren wurden von Knipfer (2009) in zwei Experimenten gezielt experimentell manipuliert, um deren Einfluss auf den Wissenserwerb und die Meinungsbildung im Hinblick auf kontroverse Wissenschaftsthemen zu untersuchen.

Nachfolgend werden die wichtigsten Ziele, die Stichprobe und das Design sowie zentrale Befunde der zwei Studien präsentiert. In beiden experimentellen Studien haben die Versuchsteilnehmer zunächst etwa 20-30 Minuten eine virtuelle Ausstellung zum Thema Nanotechnologie frei exploriert. Diese Ausstellung präsentierte sowohl grundlagenwissenschaftliche Erkenntnisse aus den Nanowissenschaften als auch kontroverse Fragen zu

Chancen und Risiken zur Nanotechnologie und enthielt dazu zahlreiche divergierende Expertenpositionen zu diesem aktuellen kontroversen Wissenschaftsthema.

Ziele Studie 1. Primäres Ziel der ersten Studie bestand darin, den Einfluss der Salienz der Argumente in Form von Expertenstatements und der aktiven Positionierung durch die Äußerung seiner eigenen persönlichen Meinung auf den Wissenserwerb, die Meinungsqualität und die Einstellungen zur Nanotechnologie untersuchen. Es sollte untersucht werden, ob aktive Meinungsäußerung und die Salienz der verfügbaren Informationen zu einem höheren Wissenserwerb und einer deliberativeren Meinungsbildung zum Thema Nanotechnologie führen. Aus theoretischer Sicht wurde angenommen, dass die Salienz der Argumente die Kontroversität des Themas salient macht, die Integration kontroverser Informationen anregt und infolgedessen eine argumentbasierte Meinungsbildung induziert. Aktive Positionierung sollte wiederum das persönliche Involvement erhöhen, die Reflexion der kontroversen Informationen und die Abstraktion des erworbenen Wissens durch die schriftliche Explikation der eigenen Meinung fördern.

Stichprobe und Design Studie 1. Auf der Basis eines 2x2-faktoriellen experimentellen Designs (Faktor 1: keine aktive Positionierung vs. aktive Positionierung; Faktor 2: keine Salienz der Argumente vs. Salienz der Argumente in Form von Expertenstatements) wurden 60 Studierende zufällig einer der vier resultierenden Lernbedingungen zugewiesen.

Die Kontrollgruppe ($n = 14$; keine Salienz der Argumente/keine aktive Positionierung) bearbeitete ein Multiple-Choice-Quiz, indem lediglich Fakten über die virtuell besuchte Ausstellung abgefragt wurden. In der zweiten Lernbedingung ($n = 16$; Salienz der Argumente/keine aktive Positionierung) wurde der Faktor Salienz der Argumente durch die Bearbeitung eines Drag & Drop-Quiz realisiert. Aufgabe der Probanden bestand darin, acht kontroverse Expertenstatements den entsprechenden Fotos der Experten zuzuordnen, die in der besuchten Ausstellung zusammen präsentiert wurden. In der dritten experimentellen Bedingung ($n = 16$, keine Salienz der Argumente/aktive Positionierung) gaben die Versuchsteilnehmer per Schieberegler ihre persönliche Meinung zum Thema Nanotechnologie auf einer Skala von „-100“ (bin absolut gegen Nanotechnologie) bis „+100“ (bin absolut für Nanotechnologie) an und tippten ein Statement zu ihrer persönlichen Meinung ein. In der vierten Bedingung wurden beide Faktoren implementiert ($n = 14$; Salienz der Argumente/aktive Positionierung), indem die Studienteilnehmer acht Expertenstatements aus vier verschiedenen Anwendungsbereichen der Nanotechnologie, die jeweils ein Pro- und ein Kontra-Argument enthielten, im Hinblick auf die eigene Zustimmung und die persönliche Relevanz bewerten sollten. Aufgrund der expliziten Aufforderung zur aktiven Auseinandersetzung mit diesen bereits in der Ausstellung verfügbaren Expertenstatements wurde eine erhöhte Salienz der Argumente erzielt. Danach wurden die Versuchsteilnehmer dieser

Gruppe gebeten, ihr Gesamturteil zur Nanotechnologie abzugeben und diesbezüglich ein Statement einzugeben.

Als abhängige Variable wurden der Wissenserwerb, die Meinungsqualität und die Einstellungen zur Nanotechnologie erhoben. Der Wissenserwerb wurde über eine schriftliche Abfrage aller erinnerten Pro- und Contra Argumente (*recall*) zur Nanotechnologie (*Argument Repertoire*) operationalisiert. Zur Erfassung der Meinungsqualität wurden Essays verwendet, in denen die Probanden aufgefordert wurden, ihre Meinung zum Thema Nanotechnologie schriftlich zu begründen. Die Essays wurden anhand eines sechsstufigen Kodierschemas (von 0 niedrig bis 5 hohe Meinungsqualität) ausgewertet (*Reflective Judgment Score*). Die theoretische Grundlage für das verwendete Kategorienschema zur Erfassung der Meinungsqualität bildete das *Reflective Judgment Model* von King & Kit-chener (1994). Anhand 12 semantischer Differentiale wurden die Einstellungen zur Nanotechnologie erhoben (Prä-/Post-Messung). Bei gleichzeitiger Implementierung der Faktoren „Salienz der Argumente“ und „aktive Positionierung“ wurde angenommen, dass sich im Vergleich zu den anderen Lernbedingungen ein größeres Argument Repertoire, eine höhere Meinungsqualität im abschließenden Essay und von den Voreinstellungen unabhängige Einstellungen zeigen sollten.

Zentrale Befunde Studie 1. Die Ergebnisse zeigen zunächst, dass kein signifikanter Unterschied im Argument Repertoire zwischen den vier Lernbedingungen bestand: Alle vier Gruppen erinnerten gleich viel Pro- und Kontra-Argumente zur Nanotechnologie. Allerdings zeigte sich ein signifikanter Haupteffekt für die Bedingungen „Salienz der Argumente“ in Bezug auf die Breite des *Argument Repertoire*: Die Studierenden, die sich aktiv mit den Expertenstatements auseinandergesetzt haben, konnten signifikant mehr Argumente aus unterschiedlichen Anwendungsbereichen der Nanotechnologie erinnern. Die aktive Auseinandersetzung mit kontroversen Expertenstatements führte auch zu einer höheren Meinungsqualität im abschließenden Essay, in dem die persönliche Meinung ausführlich begründet werden sollte: Es konnte ein signifikanter Haupteffekt für beide Gruppen mit Salienz der Argumente festgestellt werden. In Bezug auf die Einstellungen zur Nanotechnologie konnte festgestellt werden, dass die aktive Positionierung ohne Salienz der Argumente zu einem sog. *confirmation bias* führte: Die Bewertung der Nanotechnologie nach dem Ausstellungsbesuch hing stark von den Voreinstellungen zur Nanotechnologie ab. Lediglich bei den Bedingungen mit Salienz der Argumente waren die Einstellungen zur Nanotechnologie nach dem virtuellen Museumsbesuch unabhängig von den Voreinstellungen. Insgesamt betrachtet, d. h. über alle Gruppen hinweg, konnten die Museumsbesucher mit Hilfe des prototypisch realisierten Meinungsterminals in Bezug auf eine fundierte Meinungsbildung effektiv unterstützt werden. Allerdings zeigte sich ein starker *myside*

bias im Argument Repertoire und in der Begründung der eigenen Meinung, d. h. die Besucher setzten sich nicht genügend mit möglichen Gegenpositionen auseinander und integrierten Gegenargumente nicht in ihre persönliche Meinung.

Resumee. Die Salienz der Argumente am Diskussionsterminal bildete den entscheidenden Faktor für eine argumentbasierte Meinungsbildung und ein reflektiertes Urteil. Die aktive Positionierung wiederum hatte keinen Einfluss auf die abhängigen Variablen, es zeigte sich auch keine Interaktion der beiden Faktoren. Trotz der positiven Ergebnisse im Hinblick auf den Wissenserwerb und die Meinungsqualität zeigte sich ein starker *myside bias* in der Argumentation bei den Versuchsteilnehmern. Ausgehend von diesem Befund wurde in einer zweiten Studie deshalb der Frage nachgegangen, inwieweit die Information darüber, zu welchem Gesamturteil die anderen Besucher im Durchschnitt gekommen sind, zur Bildung einer Meinung beiträgt, die auch mögliche Gegenpositionen berücksichtigt und integriert.

Ziele Studie 2. Die zweite Studie untersuchte konkret den Einfluss sozialer Vergleichsinformation auf den Wissenserwerb und die Meinungsbildung. Es wurde angenommen, dass die aktive Auseinandersetzung mit den (Gegen-)Positionen anderer Besucher den in der ersten Studie festgestellten *myside bias* in der Argumentation verringert und zu einer höheren Meinungsqualität führt. Insbesondere die Konfrontation mit einer zur eigenen Meinung im Konflikt stehenden Gegenposition sollte eine Reflexion des eigenen Verständnisses anstoßen, die wiederum die Elaboration und Integration von Gegenargumenten induziert und zu einer höheren Meinungsqualität führen sollte.

Stichprobe und Design Studie 2. Um den Einfluss sozialer Vergleichsinformation durch den Meinungs austausch auf die Meinungs- und Urteilsbildung zu untersuchen, wurde der Konfliktgrad zwischen der eigenen Position und der Meinung der anderen Besucher experimentell variiert. Der Studie lag ein 1x3 faktorielles Design (Stufen: ohne vs. kongruentes vs. konflikthafte Feedback) zugrunde. 61 Studierende wurden zufällig den drei resultierenden Lernbedingungen zugewiesen. Nach dem virtuellen Ausstellungsbesuch wurden alle drei Lernbedingungen von der Lernumgebung aufgefordert, ihre persönliche Meinung zur Nanotechnologie auf einer Ratingskala von „-100“ („Ich bin absolut gegen Nanotechnologie“) bis „+100“ („ich bin absolut für Nanotechnologie“) anzugeben sowie ihre Meinung ausführlich zu begründen. Die Kontrollgruppe „kein Feedback“ ($n = 22$) erhielt nach dieser Meinungsäußerung keine Information über das Meinungsbild der anderen Besucher. Die Experimentalgruppe „kongruentes Feedback“ ($n = 20$) dagegen erhielt die Information, dass das eigene Urteil mit dem durchschnittlichen Gesamturteil der anderen Besucher übereinstimmt. Zusätzlich wurden drei zur eigenen Meinung konsistente Statements von anderen Besuchern zur Verfügung gestellt. Die Probanden der Experimentalgruppe

„konflikthafte Feedback“ ($n = 19$) wurden hingegen mit einer Information konfrontiert, bei der die eigene, persönliche Meinung zur Nanotechnologie von der durchschnittlichen Gesamtmeinung der anderen Besucher abwich. Zusätzlich wurden drei zur eigenen Meinung inkonsistente Besucherstatements zur Verfügung gestellt. Das manipulierte Feedback wurde bei beiden Experimentalgruppen je nach Bedingung abhängig vom Gesamtrating zur Nanotechnologie auf der Basis eines Algorithmus berechnet (vgl. Knipfer, 2009).

Nach der Feedbackgabe über die durchschnittliche Meinung der anderen Besucher wurden die Probanden aufgefordert, zum zweiten Mal ihre persönliche Meinung zur Nanotechnologie anzugeben und ihre Meinung in einem Essay ausführlich zu begründen. In Studie 2 wurde neben der Meinungsqualität (*Reflective Judgment Score*) und dem Wissenserwerb (*Argument Repertoire*) als zusätzliche abhängige Variable die Änderung des Gesamturteils zur Nanotechnologie nach dem Feedback erhoben. Als weitere Variable in Bezug auf das Argument Repertoire wurde der *myside bias* Index erfasst, der das Verhältnis von erinnerten Argumenten, welche die eigene Position unterstützen, zu Gegenargumenten angibt. Je größer der Wert, desto stärker ist der *myside bias*. Zusätzlich wurden die Studienteilnehmer, nachdem sie ihre eigene Position im Essay begründeten, gebeten, drei Gegenargumente zu ihrer eigenen Position nennen und diskutieren (*Counterposition/Rebuttal Score*).

Zentrale Befunde Studie 2. Die Ergebnisse der zweiten Studie zeigen, dass Lernende, die mit einem zu ihrer eigenen Meinung konfligierenden Feedback konfrontiert wurden, signifikant häufiger ihr zuvor abgegebenes Gesamturteil zur Nanotechnologie revidierten. Hoher Konsens beeinflusste demzufolge die Urteilsbildung sehr deutlich, obwohl die Untersuchungsteilnehmer allein urteilten und kein normativer Druck ausgeübt wurde. Die Lernenden der Bedingung „inkongruentes Feedback“ konnten auch mehr Argumente sowohl für als auch gegen ihre Position erinnern als Lernende, die kongruentes oder kein Feedback erhielten. Zudem fiel der *myside bias* in der Argumentation für die Konfliktbedingung signifikant geringer aus als für die beiden anderen Lernbedingungen zusammen: Diese Lernbedingung konnte annähernd die gleiche Anzahl an meinungskonsistenten (*mysided*) und meinungsinkonsistenten Argumenten (*othersided*) erinnern. Die Lernbedingung „konfligierendes Feedback“ zeigte auch signifikant höhere Werte in der Meinungsqualität im Essay (*Reflective Judgment Score*) und in der Aufgabe, Gegenargumente zu nennen und zu diskutieren als die beiden anderen Bedingungen (*Counterposition/Rebuttal Construction Score*).

Knipfer (2009) konnte zusätzlich, so wie es die Theorien zum sozialen Einfluss postulieren, das Phänomen der *sozialen Projektion* (Allport, 1924) bzw. den *false consensus effect* (Ross et al., 1977) nachweisen: Die Versuchsteilnehmer erwarteten, selbst der Mehrheitsmeinung anzugehören und gingen davon aus, dass ihre eigene Meinung von der Mehrheit

geteilt wird. Dementsprechend wurde die Information, dass die anderen Besucher eine andere Meinung vertraten, auch als konflikthaft wahrgenommen und löste bei dieser Lernbedingung kognitive Dissonanz aus, die psychisch unangenehm erlebt wurde. Knipfer (2009) interpretiert vor diesem Hintergrund ihre Befunde folgendermaßen:

Die Auflösung der resultierenden kognitiven Dissonanz erfolgte nicht auf soziale Weise (normativer Druck, Assimilation der eigenen Meinung), sondern es wurden die am Meinungsterminal verfügbar gemachten Gegenargumente in die eigene Meinung integriert. Dies führte zu einem geringeren *myside bias* im *Argument Repertoire* und einer höheren Meinungsqualität im Essay. Außerdem konnten in dieser experimentellen Bedingung mehr Gegenargumente zur eigenen Meinung konstruiert und erfolgreich die eigene Meinung dagegen verteidigt werden. (S. 107)

Resumee. Konfligierendes Feedback führte im Vergleich zu den beiden anderen Lernbedingungen zu einer signifikant häufigeren Änderung des Gesamturteils nach dem Feedback, einem geringeren *myside bias* im *Argument Repertoire* und einer höheren Meinungsqualität im Essay, die sich vor allem durch eine bessere Integration von Gegenargumenten in der Begründung der eigenen Position äußerte. Außerdem konnten in dieser experimentellen Bedingung mehr Gegenargumente zur eigenen Meinung generiert und diskutiert werden. Demzufolge kommt dem Feedback über die Meinungen der anderen Besucher eine entscheidende Rolle bei der Meinungsbildung in Bezug auf kontroverse Wissenschaftsthemen bei. In beiden Experimenten konnte die Effektivität des Diskussionsterminals in Bezug auf kritisches Denken und reflektiertes Urteilen nachgewiesen werden. Insbesondere die Salienz relevanter Pro- und Kontra-Argumente zum Thema erwies sich als ausschlaggebend für eine Bewertung, die unabhängig ist von Voreinstellungen. Die Möglichkeit, die Meinung anderer Besucher und dabei auch Gegenpositionen kennenzulernen, trägt in großem Maße dazu bei, dass auch Gegenargumente in die eigene Bewertung integriert werden und somit die Qualität der eigenen Meinung steigt.

Aus beiden vorgestellten Studien lassen sich laut Knipfer (2009) konkrete Gestaltungsempfehlungen für innovative Medienanwendungen in Wissenschaftsmuseen ableiten. Diskussionsterminals unterstützen kritisches Denken und reflektiertes Urteilen in informellen Lernsettings wie Museumsausstellungen insbesondere dann, wenn Besuchern, die Möglichkeit geboten wird, sich mit kontroverser Information in Form von Pro- und Kontra-Argumenten aktiv auseinanderzusetzen *bevor* eine eigene Positionierung erfolgt und wenn Besuchern die Möglichkeit zum Meinungsaustausch mit anderen Besuchern geboten wird und *gezieltes Feedback* über das Meinungsbild der anderen Besucher gegeben wird.

2) *Summative Evaluationsstudie eines Computerlernprogramms zu ethischen Problemen in der Gentechnik.* Eine weitere, für diese Arbeit relevante Studie wurde von Götz (2001) durchgeführt, der für den schulischen Unterricht ein Computerlernprogramm zur Förde-

rung moralisch-ethischer Überzeugungen von Jugendlichen im Bereich Gentechnik entwickelte und evaluierte. Ziel der konzipierten Lernumgebung EPROG (Ethische PRObleme in der Gentechnik) bestand darin, Schülern zu ermöglichen, sich ein eigenständiges und gut begründetes Urteil über die ethischen Probleme der Gentechnik bilden zu können. Nachfolgend werden die Ziele, Stichprobe und Design sowie zentrale Befunde der Evaluationsstudie zum Lernprogramm EPROG präsentiert.

Ziele. Im Rahmen einer summativen Evaluation sollte die Effektivität des Lernprogramms EPROG nachgewiesen werden. Die zu überprüfenden Evaluationskriterien bildeten dabei vor allem das Wissen im Bereich Gentechnik, die Einstellungen zur Gentechnik im Sinne von Hoffnungen und Befürchtungen (vgl. Todt & Götz, 1997), das Entscheidungsverhalten sowie die Sicherheit der Entscheidung.

Stichprobe und Design. Die Evaluationsstudie wurde anhand eines einfachen Kontrollgruppen-Versuchsplans durchgeführt. Dabei bildeten 50 Studierende die Kontrollgruppe, die kein Treatment erhielt, die Experimentalgruppe setzte sich aus 54 Studierenden zusammen, die das Programm im Durchschnitt etwa 1 Stunde bearbeiteten. Bei beiden Lernbedingungen wurden anhand eines Prä-/Posttests das Wissen über Gentechnik und die Einstellungen zur Gentechnik erhoben. Im Rahmen des Lernprogramms wurde die Experimentalgruppe mit fünf verschiedenen ethisch/moralischen Dilemmata zum Thema Gentechnik in Textform konfrontiert: zur (1) Erbkrankheit Chorea Huntington, zur (2) Pränatal-Diagnostik, zum (3) gentechnisch hergestellten Medikament Insulin, zu (4) gentechnisch veränderten Lebensmitteln sowie zu (5) transgenen Tieren. Nach jedem präsentierten Dilemma wurden die Studenten von der Lernumgebung aufgefordert, eine Entscheidung über die Wahl einer Alternative zu treffen sowie ihre Entscheidung schriftlich zu begründen. Im Anschluss daran konnten bei Bedarf zusätzliche, vertiefende Informationen zu Grundbegriffen der Gentechnik abgerufen werden. Nach Bearbeitung dieser Informationen wurden die Probanden gebeten, ihre zweite und finale Entscheidung darüber zu treffen, wie sie sich in dieser Dilemmasituation verhalten und entscheiden würden. Anschließend musste wiederum die getroffene Entscheidung schriftlich begründet werden.

Zentrale Befunde. Die Ergebnisse der summativen Evaluationsstudie konnten zeigen, dass durch die Bearbeitung der computerunterstützten Lernumgebung der Wissenserwerb deutlich gesteigert werden konnte. Die Studierenden der Experimentalgruppe, die mit EPROG lernten, erzielten signifikant bessere Leistungen im Post-Test zum Wissen über Gentechnik als die Studierenden der Kontrollgruppe, die das Programm nicht bearbeiteten (im Vergleich zum zuvor durchgeführten Prä-Test). Außerdem konnte festgestellt werden, dass Probanden mit einem geringeren Vorwissen am meisten von der computerunterstützten Lernumgebung profitierten (ATI-Effekt). Ein weiteres wesentliches Studienergebnis be-

stand darin, dass sich durch die Bearbeitung des Programms die Einstellungen zur Gentechnik nicht wesentlich veränderten. Dieser Befund überraschte allerdings keineswegs, da es sich wie Götz (2001) bemerkt, bei diesem Lernprogramm lediglich um eine sehr kurzfristige Intervention handelte. Ferner wurden die anfänglich getroffenen Entscheidungen durch die Bearbeitung der zusätzlichen Informationen nur selten revidiert: Es ergaben sich deskriptiv keine großen Unterschiede in der Richtung der ersten Entscheidung vor Bearbeitung der Informationen und der zweiten, finalen Entscheidung nach Bearbeitung der zusätzlichen Informationen. Des Weiteren konnte die Hypothese, dass die Entscheidungen nach den Informationen im Mittel sicherer ausfallen als die Entscheidungen vor den Informationen nur in Bezug auf ein Dilemma empirisch bestätigt werden. Lediglich bei Dilemma 1, bei dem die Probanden sich entscheiden mussten, ob sie sich auf die Erbkrankheit Chorea Huntington untersuchen lassen würden, zeigte sich ein signifikanter Unterschied in der Sicherheit der Entscheidung vor und nach der Bearbeitung der Zusatzinformationen. Bei Dilemma 3 hingegen, das den Einsatz von gentechnisch erzeugtem Insulin thematisierte, zeigte sich ein konträrer Befund; in diesem Fall wurden die Probanden durch die vertiefenden Informationen tendenziell eher verunsichert. Bei allen anderen drei Dilemmata blieb die Entscheidungssicherheit vor und nach Bearbeitung der zusätzlichen Informationen relativ konstant. Demzufolge konnte durch die Bereitstellung zusätzlicher, vertiefender Informationen keine signifikante Steigerung der Sicherheit bei der Bewertung der einzelnen Dilemmata (mit Ausnahme von Dilemma 1) bewirkt werden. Insgesamt betrachtet, d. h. über alle fünf Dilemmata hinweg, konnte jedoch eine kontinuierliche Zunahme der Sicherheit der Entscheidung durch die Bearbeitung des Lernprogramms festgestellt werden. Zudem fielen die Bewertungen der Gentechnik nach der Bearbeitung des Programms differenzierter aus als vor dem Programm. Bei der Analyse der Begründungen der Entscheidungen konnte festgestellt werden, dass utilitaristische Argumente von den Schülern insgesamt betrachtet, d. h. über alle fünf Dilemmata hinweg, häufiger angeführt wurden als deontologische Argumente. Die Verwendung der beiden Argumentationstypen hing jedoch stark vom jeweiligen Dilemma, d. h. vom Thema bzw. Kontext ab.

Resumee. Zusammenfassend betrachtet konnten durch die Bearbeitung der Lernumgebung EPROG signifikante Lernfortschritte erzielt werden, insbesondere bei vorwissenschwachen Studierenden. Götz (2001) selbst empfiehlt einen Einsatz des Programms vor dem eigentlichen Schulunterricht, um auf diese Weise Interesse am Thema ethische Bewertung der Gentechnik bei den Schülern zu wecken. Als eine wichtige Voraussetzung dafür nennt Götz (2001), dass der Lehrer nicht nur auf die ethischen Probleme der Gentechnik eingehen und sie beschreiben sollte, sondern auch konkrete Ansätze liefern sollte, wie diese Dilemmata zu lösen sind.

5 Die multimediale Lernumgebung Dialogstation „Gentest“

Den Kontext der experimentellen Studie zur Analyse der Entscheidungsfindung und Argumentation im Bereich Medizinethik/PID bildet die multimediale Lernumgebung Dialogstation „Gentest“, die für den Einsatz in einer Museumsausstellung entwickelt wurde.

Als Untersuchungsgegenstand bzw. Experimentalumgebung wurde aus ökonomischen und untersuchungstechnischen Gründen ein Fallbeispiel aus den sieben existierenden ausgewählt und die Lernumgebung dementsprechend modifiziert. Die ausgewählte Dilemmageschichte zur PID-HLA lehnt sich an den in Kapitel 2.2.3 beschriebenen Fall Adam Nash an, der als erster Mensch im Labor zur Heilung seiner schwerkranken Schwester Molly gezeugt wurde. Die verwendete Experimentalumgebung wird im Methodenteil in Kapitel 8.3 ausführlich vorgestellt.

Zunächst werden wichtige Rahmenbedingungen (Kapitel 5.1) sowie die Ziele und Inhalte (Kapitel 5.2) der musealen Lernumgebung dargestellt. In den anschließenden Kapiteln werden die didaktische und mediale Gestaltung der Lernumgebung (Kapitel 5.3 und 5.4) sowie die konkrete Gestaltung der implementierten Feedbackinformation beschrieben (Kapitel 5.5). Zum Abschluss werden erste Evaluationsergebnisse präsentiert (Kapitel 5.6).

5.1 Rahmenbedingungen

Die Dialogstation „Gentest“ wurde als einzelnes Ausstellungselement von der Kuratorin für Life Sciences, Dr. Birte Hauser in Zusammenarbeit mit der Mediendesignerin Charlotte Kaiser für die Ausstellung zur Nano- und Biotechnologie des ZNT des Deutschen Museums konzipiert (Breitsameter et al., 2009).

Das Deutsche Museum in München ist das größte Museum zur Geschichte der Naturwissenschaften und Technik der Welt und wurde im Jahr 2009 von insgesamt rund 1,12 Millionen Menschen besucht (Deutsches Museum, 2009). Die im November 2009 neu eröffnete und 600 qm² große Dauerausstellung zur Nano- und Biotechnologie soll im Sinne eines PUR den Besuchern nicht nur die Ziele und Ergebnisse, sondern auch die Prozesse und Rahmenbedingungen aktueller nano- und biotechnologischer Forschung vermitteln (Breitsameter et al., 2009).

Die Dialogstation „Gentest“ ergänzt in der Ausstellung die Teilbereiche zum Thema Biotechnologie, in denen z. B. die molekularbiologischen Methoden der DNA-Analyse vermittelt werden und soll den Besuchern die ethischen Implikationen, die mit der Anwen-

dung und Durchführung einer genetischen Untersuchung verbunden sind, aufzeigen. Diese an die speziellen Bedingungen des informellen Lernorts Museum adaptierte multimediale Lernumgebung wurde ausgehend von den bisherigen Erfahrungen des in Kapitel 4.5 beschriebenen Meinungsterminals zum Thema Stammzellforschung entwickelt.

Auf eine ausführliche Darstellung der Gesamtausstellung und ihrer 17 Teilbereiche wird an dieser Stelle verzichtet, da eine Feldstudie im realen Lernsetting aus zeitlichen und organisatorischen Gründen leider nicht möglich war (Anmerkung: Aufgrund unvorhergesehener Schwierigkeiten bei den Baumaßnahmen wurde die Eröffnung der Ausstellung von Mai 2008 auf November 2009 verschoben). Eine detaillierte Beschreibung der ZNT-Ausstellung zur Nano- und Biotechnologie kann Gerber-Hirt & Noschka-Roos (2012) und der Besucherbefragung von Specht & Lewalter (2011) entnommen werden.

Nachfolgend wird die konkrete Umsetzung eines Prototyps der Dialogstation „Gentest“ vorgestellt.

5.2 Ziele und Inhalte

Zielgruppe. Mit der Lernumgebung und der Ausstellung zur Nano- und Biotechnologie sollen möglichst alle Besucher des Deutschen Museums, unabhängig von Alter, Bildungsstand oder Interesse an naturwissenschaftlich-technischen Themen erreicht und angesprochen werden (Gerber-Hirt & Noschka-Roos, 2012).

Zielsetzungen der Konzeption. Mit der Konzeption der Ausstellungseinheit Dialogstation „Gentest“ sind aus Sicht der Ausstellungsmacher verschiedene Annahmen und Erwartungen verbunden. Die Konzeption zielt darauf ab, dem Besucher die Vielschichtigkeit der ethischen Probleme von genetischen Untersuchungen deutlich zu machen und ihm zu ermöglichen, sich eine Meinung zu genetischen Untersuchungen zu bilden (Breitsameter et al., 2009; Hauser, 2009). Den Besuchern soll die Bandbreite der realen Anwendungssituationen aufgezeigt werden, in denen sich Personen für oder gegen eine genetische Untersuchung entscheiden müssen bzw. können (Gerber-Hirt & Noschka-Roos, 2012).

Im Rahmen der multimedialen Lernumgebung wurde das Thema der ethischen Aspekte der Gendiagnostik am Menschen dialogisch aufbereitet: Die Besucher werden mit divergierenden Meinungen und Standpunkten von Betroffenen und Experten konfrontiert und können ihre eigene Meinung zum jeweiligen Thema äußern und im Gegenzug wird ihnen das Meinungsbild der anderen Besucher zurückgespiegelt. Die Besucher sollen durch die Darstellung unterschiedlicher Meinungen und Argumente ohne explizite Lösungsvorgabe und der Spiegelung anderer Meinungen, dazu angeregt werden, „über ihre eigene Position nachzu-

denken oder einen Perspektivenwechsel vorzunehmen, nicht zuletzt durch die eigene aktive Positionierung“ (Gerber-Hirt & Noschka-Roos, 2012, S. 341).

Es ist *nicht* das Ziel der Lernumgebung, den Besucher zu einer Inanspruchnahme einer genetischen Untersuchung zu bewegen, sondern es soll ihm ermöglicht werden, sich mit der ethischen Problematik genetischer Untersuchungen kritisch-reflexiv auseinanderzusetzen, um am Diskurs zum gesellschaftlich relevanten Konfliktfeld genetischer Untersuchungen teilhaben zu können. Aus Sicht der Kuratoren sollte eine Lernumgebung entwickelt werden, die den Besucher nach seiner persönlichen Meinung zum Thema Gentest fragt und zur Auseinandersetzung mit den durch die Anwendung von genetischen Untersuchungen aufgeworfenen ethischen Fragen anregen soll (Hauser, 2006).

Des Weiteren soll der Besucher in die Lage versetzt werden, sich in die Rolle von Betroffenen hineinversetzen zu können, die Chancen und Risiken bezüglich der Durchführung eines genetischen Tests kritisch abwägen und ein reflektiertes Urteil fällen zu können.

Lernziele. Aus Sicht der Kuratorin und Mediendesignerin standen folgende zwei übergeordnete Lernziele bei der Konzeption der Lernumgebung im Vordergrund:

- 1) Aus *kognitiver Sicht* sollte den Besuchern zunächst grundlegendes Faktenwissen zu den verschiedenen prädiktiven, pränatalen und präimplantativen genetischen Untersuchungen vermittelt werden. Zentrales Lernziel der Lernumgebung bestand in der Förderung einer gut informierten und reflektierten Entscheidung bezüglich der Durchführung einer genetischen Untersuchung. Die Fallbeispiele und die zusätzlichen Informationen enthalten in kurzer und prägnanter Form die wichtigsten deskriptiven Informationen, die für oder gegen die Durchführung der jeweiligen genetischen Untersuchung sprechen und zur gesetzlichen Regelung in Deutschland und im Ausland.
- 2) Aus *motivationaler Sicht* sollte durch die direkte und personalisierte Ansprache einer fiktiven Fall-Person Neugier und Interesse geweckt und somit eine intensivere Auseinandersetzung mit der Thematik initiiert werden. Die Aufforderung zur aktiven Selbstpositionierung sollte diese Effekte in ähnlicher Weise verstärken, ebenso wie die technisch gesehen viel aufwändigere Einspielung über eine mediale Inszenierung (*talking heads*) statt der herkömmlichen Darbietung über ein simples Bildschirmterminal.

Inhalt. Der inhaltliche Schwerpunkt der Lernumgebung liegt auf den ethischen Aspekten von genetischen Untersuchungen. Dem Besucher werden authentische Fallgeschichten zu prädiktiven und pränatalen Gentests sowie zur PID präsentiert.

Inhaltlicher Aufbau. Die Lernumgebung bzw. der Prototyp der Dialogstation „Gentest“ besteht aus insgesamt sieben Fallbeispielen, die sich grob aus vier Komponenten zusammensetzen:

1. Einem personalisierten Film, der den Besucher via „*talking head*“ mit einer Dilemmageschichte einer betroffenen Person zum Thema Gentest konfrontiert.
2. Einer Aufforderung zur aktiven Positionierung, die dem Besucher ermöglicht, eine Entscheidung zu treffen, ob er anstelle der Person einen Gentest durchführen lassen würde oder nicht (dichotome Ja/Nein-Entscheidung).
3. Einer Feedbackinformation über die durchschnittliche Gesamtentscheidung der anderen Besucher (Soziale Vergleichsinformation/Konsensinformation in Form eines Umfrageergebnisses).
4. Zusätzlichen Informationen in Textform (z. B. sachliche Informationen, ethische Stellungnahmen von Experten und die Gesetzeslage in unterschiedlichen Ländern).

Abbildung 8 und 9 geben den Prototyp der Dialogstation „Gentest“ aus frontaler und seitlicher Perspektive wieder. In einem, aus weißen Spanplatten bestehenden, viereckigen Gehäuse sind 12 Halbrelieffmasken und ein Touchscreen-Monitor eingelassen. Auf vier von den 12 Halbreliiefs werden mit Hilfe von 2 Beamern, die sich im Inneren des Gehäuses befinden, Portrait-Videos projiziert, die einen 3-D-Effekt erzeugen. Integriert ist auch eine Loopfunktion, wenn kein Besucher die Dialogstation bedient. In diesem Ruhezustand blicken die bespielten Köpfe nach links, rechts oder geradeaus, so dass beim Besucher der Eindruck entstehen soll, „verfolgt“ zu werden.



Abbildung 8: Prototyp Dialogstation „Gentest“, Vorderansicht, Touchscreen-Monitor zur Bedienung (unten Mitte, Foto: Deutsches Museum).



Abbildung 9: Prototyp Dialogstation „Gentest“, seitliche Ansicht von links, 3-D-Effekt (Teilausschnitt, Foto: Deutsches Museum und Charlotte Kaiser, Berlin).

Wie Abbildung 8 zu entnehmen, erfolgt die Bedienung der Dialogstation über einen Touchscreen-Monitor. Hierbei handelt es sich um einen Computerbildschirm, den der Besucher durch Berührung mit dem Finger steuern und so zwischen einzelnen Seiten der Lernumgebung navigieren kann.

Folgende Abbildung 10 zeigt einen Screenshot der Startseite des Touchscreen-Monitors, aus dem ersichtlich wird, dass der Besucher zwischen sieben verfügbaren, von der Gentestproblematik betroffenen Personen bzw. Fallbeispielen frei auswählen kann.



Abbildung 10: Einstiegsseite bzw. Startseite der Dialogstation mit sieben frei auswählbaren Fallbeispielen zum Thema Gentest (Screenshot Touchscreen).

Nach Berühren der jeweiligen Person auf dem Touchscreen-Monitor wird das Video eines fiktiven, typischen Fallbeispiels zur Gentestproblematik abgespielt. Die Dauer der von Schauspielern gesprochenen Videos bzw. Fallbeispiele liegt zwischen 2 und 3 Minuten.

Im Fallbeispiel „Maik Schöller“ etwa steht der Protagonist vor der schwierigen Entscheidung, einen Gentest auf Chorea Huntington durchzuführen, nachdem die Erkrankung bei seinem Vater ausgebrochen ist und bei ihm und seiner Partnerin ein aktueller Kinderwunsch besteht. Die Wahrscheinlichkeit, dass er die spät manifestierende, tödlich verlaufende Krankheit an seine Kinder weitervererbt, beträgt 50 %.

Neben direkt betroffenen Personen werden auch Fälle indirekt-betroffener Personen präsentiert, beispielsweise eine Versicherungskauffrau, welche für die Verwendung von genetischen Untersuchungen zur Kalkulation des Risikos für den Abschluss von Lebensversicherungen plädiert. Jede der sieben „Dilemmageschichten“ endet mit einer Entscheidungsfrage an den Besucher, beispielsweise fragt „Maik Schöller“ am Ende seiner Geschichte: „Ob ich den Test doch an mir machen lasse? Würden Sie das an meiner Stelle tun?“

Anschließend wird eine kategoriale „Ja/Nein-Frage“ auf dem Touchscreen eingeblendet, die den Besucher auffordert, eine Entscheidung zu treffen, ob er an Stelle der Person einen Gentest durchführen lassen würde oder nicht (*aktive Positionierung*; siehe folgende Abbildung 11).



Abbildung 11: Aufforderung zur aktiven Positionierung über Ja/Nein-Abstimmung (Screenshot Touchscreen).

Nachdem der Besucher eine persönliche Entscheidung getroffen, d. h. abgestimmt hat, wird im Anschluss aggregiertes Feedback über alle bisher abgegebenen Besucherentscheidungen graphisch eingespielt (siehe Abbildung 12):

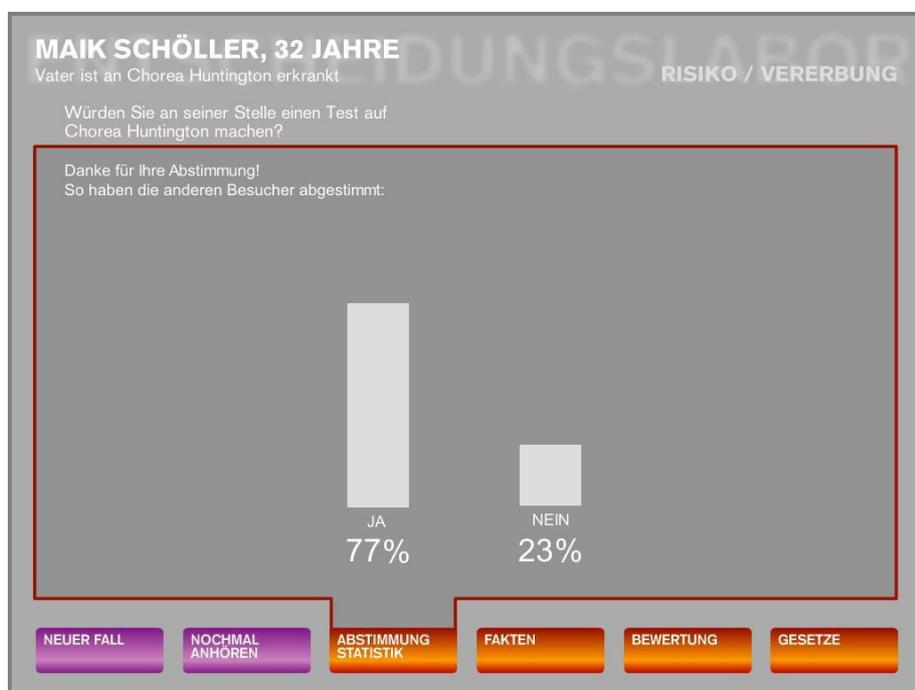


Abbildung 12: Feedback in Form einer sozialen Vergleichsinformation (Konsensinformation) über die Gesamtentscheidung der anderen Besucher (Screenshot Touchscreen).

Aus dieser Abstimmungsstatistik bzw. sozialen Vergleichsinformation wird ersichtlich, wie sich die vorherigen Besucher im Durchschnitt in diesem Fall entschieden haben. Wie Abbildung 12 zu entnehmen, würden z. B. beim Fallbeispiel Maik Schöller 77 % der Besucher, die insgesamt abgestimmt haben, an seiner Stelle einen Gentest auf Chorea Huntington durchführen lassen, 23 % dagegen nicht.

Nach seiner Entscheidung über einen vorgetragenen Fall erfährt der Besucher nicht nur, wie sich die anderen Besucher vor ihm entschieden haben, er kann auch auf dem Touchscreen-Monitor zusätzliche Informationen zum Fallbeispiel in Textform (On-Screen-Text) bei Bedarf frei abrufen, die jedoch nicht nach Pro und Contra geordnet sind, sondern nach:

1. *Fakten*. Unter dem Menüpunkt „Fakten“ befinden sich sachliche Informationen zum jeweils vorliegenden Fallbeispiel in Form von (Zahlen-)Angaben über verschiedene Tatsachen, denen eine positive bzw. negative Haltung gegenüber dem jeweiligen Gentest zu Grunde liegt.
2. *Bewertung*. Der Menüpunkt „Bewertung“ beinhaltet ein bis drei ausgewählte ethische Stellungnahmen bzw. Argumente von Experten wie Medizinern, Theologen oder Soziologen, aber auch von Institutionen wie z. B. dem Nationalen Ethikrat, die für oder gegen einen Gentest im jeweiligen Fallbeispiel sprechen.
3. *Gesetze*. Dieser Menüpunkt informiert über die unterschiedliche Gesetzeslage in verschiedenen Ländern, um aufzuzeigen, dass die Argumente für oder gegen einen Gentest in verschiedenen Ländern unterschiedlich gewichtet werden.

Diese zusätzlichen Informationen sollen aus Sicht der Kuratorin den Besucher dazu bewegen, über seine „aus dem Bauch“ getroffene Entscheidung nach dem gehörten Fall nochmals zu reflektieren. Der Aufruf der Zusatzinformationen nach Erhalt der Feedbackinformation über die Gesamtentscheidung der anderen Besucher ist optional, stattdessen kann der Besucher auch das Fallbeispiel nochmals hören oder einen neuen Fall aufrufen.

5.3 Didaktische Gestaltung

Bezüglich der didaktischen und medialen Gestaltung der Lernumgebung ist zunächst anzumerken, dass jene weniger lerntheoretisch fundiert, sondern vielmehr „intuitiv“ von der Kuratorin und Mediendesignerin konzipiert wurde. Dennoch können aus pädagogisch-psychologischer Sicht zahlreiche problemorientierte und multimediale Designprinzipien identifiziert werden, die bei der Konzeption der Lernumgebung umgesetzt wurden und nachfolgend skizziert werden.

Das Prinzip der *Authentizität* zeigt sich in der Dialogstation vor allem durch die integrierten Fallbeispiele von Personen, die von der Gentestproblematik betroffen sind. Die Ähnlichkeit zum Anchored-Instruction-Ansatz (CTGV, 1997) ist dabei offensichtlich: Im Mit-

telpunkt der Lernumgebung stehen authentische und komplexe Problemsituationen, die in Form einer (Dilemma-)Geschichte (*narratives Präsentationsformat*) und über Video multimedial präsentiert werden (*videobasiertes Präsentationsformat*). Durch die Verwendung von authentischen Fallbeispielen kann bei den Besuchern Interesse geweckt und eine hohe persönliche Relevanz erzeugt sowie die Identifikation mit dem Sprecher der Geschichte erleichtert werden. Zudem sind alle für die Problemlösung erforderlichen Informationen in die Lernumgebung eingebettet (*Prinzip der eingebetteten Daten*) und die verwendeten Problemsituationen entsprechen weitgehend der Komplexität realer Situationen (*Problemkomplexität*).

Multiple Kontexte lernen die Besucher durch die Darstellung von Fallbeispielen in unterschiedlichen Situationen bzw. Kontexten kennen. Da die Fallbeispiele aus der Perspektive betroffener Personen und zum Teil auch nicht-betroffener Personen (z. B. Versicherungskauffrau) bearbeitet werden und in den bereitgestellten Zusatzinformationen unterschiedliche Expertenmeinungen aufgezeigt werden, sind auch *multiple Perspektiven* gewährleistet. Konkret bedeutet dies, dass die Besucher die ethischen Aspekte genetischer Untersuchungen aus mehreren verschiedenen Zusammenhängen und aus verschiedenen Perspektiven aufgezeigt bekommen, damit das Wissen nicht auf bestimmte Situationen fixiert bleibt.

Ein *sozialer Kontext* kann in der Lernumgebung dadurch hergestellt werden, dass die Besucher die Fallbeispiele auch kooperativ bearbeiten können.

Die *instruktionale Unterstützung* erfolgte in der Lernumgebung durch die Möglichkeit, nach der Entscheidung zusätzliche Informationen in Form von Fakten, ethischen Stellungnahmen von Experten oder zur Gesetzeslage abrufen zu können. Zusätzlich wurde den Besuchern in Bezug auf ihre Entscheidung durch die Gabe von Feedback in Form einer Statistik über die durchschnittliche Entscheidung der anderen Besucher Rückmeldung anhand einer sozialen Vergleichsinformation gegeben (*Scaffolding*).

5.4 Medial-technologische Gestaltung

Innerhalb der Lernumgebung wurden einige Designprinzipien von Mayer (2005a) umgesetzt, die nachfolgend dargestellt werden. Im Rahmen der Lernumgebung werden sowohl verbale als auch bildliche Informationen verwendet, indem z. B. die Dilemmageschichten in Form von Videos präsentiert werden (*Multimediaprinzip*). Die Erläuterungen zu den Videos, den bewegten Bildern, werden in gesprochener Sprache und nicht durch geschriebenen Text präsentiert (*Modalitätsprinzip*). Zudem erfolgte keine Doppelung der gesprochenen Texte zum Video durch geschriebenen Text und geschriebener Text wurde nicht durch einen zusätzlichen Sprechertext vorgelesen (*Redundanzprinzip*). Bei der Konzeption

der Lernumgebung erfolgte eine Konzentration auf die zentralen und wesentlichen Lerninhalte, die geschriebenen Texte und die Sprechertexte der Videos wurden kurz und präzise formuliert, irrelevante Informationen wurden weggelassen (*Kohärenzprinzip*).

Zur Bedienung der Medienstation und für die Darstellung der zusätzlichen Informationen in geschriebener Textform, welche der Besucher zusätzlich zum Fallbeispiel bei Bedarf abrufen kann, wurde ein Touchscreen-Monitor verwendet. Der Aufbau der Menüführung auf dem Bildschirm wurde bewusst einfach gehalten. Alle Zusatzinformationen zum jeweiligen Fallbeispiel werden in gut lesbarer Größe und räumlich nah beieinander auf einer Bildschirmseite präsentiert, ohne dass über eine Bildlaufleiste nach unten gescrollt werden muss (*räumliches Kontiguitätsprinzip*). Mithilfe interaktiver Elemente in Form von Buttons wird dem Besucher ermöglicht, die Zusatzinformationen unmittelbar und zeitnah bei Bedarf zur Betrachtung heranzuziehen (*zeitliches Kontiguitätsprinzip*).

In den schriftlich auf dem Touchscreen dargebotenen Zusatzinformationen wurden außerdem fettgedruckte Überschriften verwendet, welche die Kernaussagen der jeweiligen Information repräsentieren und aus denen hervorgeht, ob die Zusatzinformation für oder gegen eine genetische Untersuchung spricht (*Signalisierungsprinzip*). Die Zusatzinformationen werden ebenso nicht als eine Gesamteinheit angeboten, sondern wurden in Teileinheiten (Fakten, Bewertung, Gesetze) aufgeteilt und die Lernenden können die Zusatzinformationen in ihrer eigenen Geschwindigkeit abrufen (*Segmentierungsprinzip*).

Im Rahmen der multimedialen Präsentation ist der jeweilige Sprecher der Dilemmageschichte in Form eines *talking head* zu sehen (*Prinzip des Bildes*), besitzt eine menschliche, akzent- und dialektfreie Stimme (*Prinzip der Stimme*) und verwendet einen personalisierten Sprachstil („Würden Sie das an meiner Stelle tun?“, „Unser Sohn...“) (*Personalisierungsprinzip*).

Aufgrund der Verwendung von *talking heads* wurde die „Illusion eines Face-to-Face-Kontaktes“ geschaffen. Demzufolge liegen günstige Bedingungen für das Auftreten einer High-Level-PSI vor, die zu einem starken interpersonalem Involvement mit der Medienfigur und einer intensiven Beschäftigung und Elaboration mit derselben und den von ihr vermittelten Inhalten führen sollte. Durch die direkte persönliche Ansprache des *talking head* (*direkte Adressierung, verbale Bezugnahme*) und die Hinwendung zum Besucher und dem Blick im Sinne eines „Augengrußes“ (*direkte Adressierung, non-verbale Bezugnahme*) wird dem Besucher Nähe und Unmittelbarkeit suggeriert. Die *talking heads* sind auf Gesichtshöhe eines normalwüchsigen Menschen positioniert, so dass eine niedrige räumliche Distanz realisiert wurde und die Wahrscheinlichkeit hoch ist, dass die Besucher das Gefühl haben, persönlich angesprochen zu werden (*niedrige räumliche Distanz*).

Außerdem handelt es sich um eine Videoaufnahme einer realen Person (*Artifizialität: natürlich; Anthropomorphismus: menschlich*) und die Personen weisen eine hohe Bildschirmpräsenz auf, da sie auffällig inszeniert wurden und dauerhaft zu sehen sind (*hohe Obtrusivität und hohe Persistenz*).

Bei der technischen Umsetzung wurde viel Wert auf eine hohe Benutzerfreundlichkeit gelegt, um auch Besuchern mit keinen oder nur basalen Computerkenntnissen eine effektive und einfache Nutzung der Medienstation zu ermöglichen. Die Lernumgebung ist in ihrer Nutzung selbsterklärend und erfordert keine große Expertise im Umgang mit dem Computer. Die Programmierung der Lernumgebung erfolgte mit dem Autorensystem Macromedia Director. Als Hardware wurden ein PC mit zwei Grafikkarten, ein Touchscreen-Monitor und zwei Beamer zur Projektion der Portrait-Videos auf die Halbreliëfmasken verwendet.

Zusammenfassend bleibt festzuhalten, dass die multimediale Lernumgebung Dialogstation „Gentest“ in hohem Maße einer gemäßigt konstruktivistischen Auffassung vom Lernen gerecht wird und die Anforderungen einer problemorientierten und multimedial adäquat gestalteten Lernumgebung erfüllt.

5.5 Gestaltung der Feedbackinformation

Ausgehend von den theoretischen Überlegungen in Kapitel 4.4 wird die konkrete Gestaltung der implementierten Feedbackinformation beschrieben. Damit Feedback einen nachweislichen Lernvorteil erbringen kann, muss es lernförderlich gestaltet sein und vom Empfänger in geeigneter Weise rezipiert werden (Krause, 2007).

Im Rahmen der Lernumgebung wird den Museumsbesuchern nach der individuell getroffenen Entscheidung zum jeweiligen Fallbeispiel Feedback in Form einer sozialen Vergleichsinformation (Konsensinformation) über die prozentuale Gesamtentscheidung der anderen Besucher gegeben (vgl. Abbildung 12, S. 156).

Im Hinblick auf die Gestaltung des Feedbacks wurde eine *einfache* Rückmeldung gegeben, also ein *niedriger Informationsgehalt* realisiert. Die Konsensinformation gibt dem Museumsbesucher lediglich Aufschluss darüber, wie sich die anderen Besucher vor ihm in Bezug auf das jeweilige Fallbeispiel entschieden haben, also ob die eigene, getroffene Entscheidung von einer Minderheit (Minorität) oder einer Mehrheit (Majorität) vertreten wird.

Mit Hilfe dieser Konsensinformation können die Lernenden ihre eigene Entscheidung mit der durchschnittlichen Gesamtentscheidung der anderen Besucher vergleichen und validieren. Das Feedback wird somit vor dem Hintergrund einer *sozialen Bezugsnorm* gegeben (Wie ist meine eigene Entscheidung im Vergleich zu den anderen Besuchern?).

Da hoher Konsens eine gewisse soziale Validierung sowie Unterstützung der eigenen Entscheidung impliziert, kann eine entscheidungskongruente Konsensinformation auch als ergänzender Hinweis auf die Korrektheit der getroffenen Entscheidung interpretiert werden. Dementsprechend wird den Lernenden im weitesten Sinne auch zurückgemeldet, inwieweit die eigene Entscheidung „richtig“ oder „falsch“ ist. Feedback wird implizit damit auch auf der Basis einer *sachlichen Bezugsnorm* gegeben, die sich konkret auf die im Mittelpunkt stehende Aufgabe (Entscheidung) bezieht. Angesichts dieser Eigenschaften ist das Feedback auch als *informierend* zu bezeichnen, da die Lernenden eine Information erhalten, die sie für ihren Entscheidungsprozess nutzen können.

Das Feedback wird *automatisch* in standardisierter Form und *sofort* nach dem Treffen der eigenen Entscheidung durch die Lernumgebung gegeben und hat somit eher einen *formativen* Charakter, da es auch für den nachfolgenden Lernprozess genutzt werden kann.

Das Feedback erfolgte *schriftlich* und wurde nicht auditiv dargeboten. Auf diese Weise kann jeder Lernende das Feedback in seinem individuellen Lerntempo rezipieren. Bei der Gestaltung des Feedbacks haben die Kuratorin und die Mediendesignerin besonders auf Verständlichkeit geachtet, z. B. einfache Wortwahl, kurze Sätze, eine anschauliche Grafik, die sich auf das Wesentliche beschränkt (vgl. Langer et al., 1999).

Erwartete Effekte der Feedbackinformation. Auf der Basis der theoretischen Ausführungen und der Befundlage wird angenommen, dass sich die Feedbackinformation im Vergleich zu keiner Rückmeldung positiv auf den Lernerfolg und den Lernfortschritt auswirkt. Zudem wird in Anlehnung an die Ergebnisse von Knipfer (2009; siehe Kapitel 4.6) erwartet, dass die Konfrontation mit einem zur eigenen Entscheidung inkongruenten Feedback in Form einer sozialen Vergleichsinformation über die Gesamtentscheidung anderer Personen, Dissonanz erzeugt, die nachfolgend durch eine Entscheidungsrevision reduziert wird. Ferner sollte ein im Konflikt zur eigenen Entscheidung stehendes Feedback auch zu einer informierteren und reflektierteren Entscheidung führen, die explizit auch Gegenpositionen berücksichtigt. Diese positiven Auswirkungen von konfligierendem Feedback von Knipfer (2009) auf Indikatoren der Meinungsqualität sollen anhand der vorliegenden Arbeit zur Entscheidungsfindung und Argumentation im Bereich Medizinethik repliziert werden. Die konkrete Operationalisierung der Variable Feedback (Stufen: kongruent vs. inkongruent vs. kein) wird in Kapitel 8.4.2 im Methodenteil beschrieben.

5.6 Erste Evaluation der Lernumgebung

Zum Abschluss dieses Kapitels werden Ergebnisse einer formativen Evaluation zum Prototypen der Dialogstation „Gentest“ berichtet sowie ausgewählte Befunde einer summativen

Evaluationsstudie der ZNT-Ausstellung zur Nano- und Biotechnologie im Deutschen Museum München berichtet, für welche die Medieninstallation als einzelnes Ausstellungselement konzipiert wurde.

Formative Evaluation des Prototyps der Dialogstation „Gentest“. Der Prototyp der Dialogstation „Gentest“ wurde bereits 2 Jahre vor der Eröffnung der ZNT-Ausstellung zur Nano- und Biotechnologie im Rahmen einer Felduntersuchung im Deutschen Museum formativ evaluiert (Hänle, 2008; Klein, Donecker, Hänle & Herrmann, 2009). Zum damaligen Zeitpunkt war in der ZNT-Ausstellung ein begehbare Kubus geplant, der aus der Dialogstation „Gentest“ und fünf Informationssäulen zu den Themen Risiko, Datenschutz, Vererbung, vorgeburtliche Tests und Therapie bestehen sollte. Jede dieser Säulen sollte ein authentisches Exponat zum jeweiligen Thema beinhalten (z. B. eine Punktionsnadel für das Thema vorgeburtliche Tests, mit der fetale Zellen zur DNA-Untersuchung aus dem Mutterleib entnommen werden oder ein Tresor für das Thema Datenschutz). Innerhalb des Prototyps wurden diese fünf Exponate und Themen in einer kleinen Vitrine (einer sog. *Stele*) zusammengefasst. Der Museumsbesucher erhielt auf diese Weise die Möglichkeit, sich an der Stele anhand von authentischen Objekten und zusätzlichen Texten zum jeweiligen Thema ausführlicher über die in den Fallbeispielen behandelten Themen/Aspekten zu genetischen Untersuchungen zu informieren. Der Prototyp „Entscheidungslabor Gentest“ bestand demzufolge aus der Dialogstation und einer Stele mit originalen Exponaten zum Thema Gentest. Als provisorischer Standort des Prototyps wurden aufgrund der thematischen Nähe die Arkaden vor dem Ausgang der Ausstellungen Pharmazie und Chemie im 1. OG des Deutschen Museums gewählt. Nachfolgend werden die wichtigsten Ziele, die Stichprobe und das Design sowie zentrale Befunde der Studie kurz dargestellt.

Ziele. Da sich die Dialogstation zum damaligen Zeitpunkt noch in der Entwicklungsphase befand, wurde eine formative Evaluationsstudie durchgeführt. Ziele der formativen Evaluation bestanden zum einen in der Aufdeckung von Schwachstellen oder Problemen, die bei der Bedienung der Dialogstation auftreten können, zum anderen in der Ermittlung der Akzeptanz und Nutzung der Medienstation und der Stele durch die Museumsbesucher, um anhand der Analyseergebnisse Optimierungsmöglichkeiten zu eruieren. Allerdings sollte bei der vorliegenden formativen Evaluation berücksichtigt werden, dass die Dialogstation - als isoliert untersuchtes Ausstellungselement - anders wirken könnte als im Gesamtkontext der zukünftigen Ausstellung (vgl. Graf & Treinen, 1983; Noschka-Roos, 1994).

Stichprobe und Design. Für die Feldstudie wurden im Erhebungszeitraum von November 2006 bis Februar 2007 die Methoden Beobachtung ($n = 300$), Logfile-Analyse ($n = 10792$) und Befragung ($n = 90$) verwendet. Die Logfile-Analyse und die Beobachtung wurden zur Untersuchung des Besucherverhaltens, insbesondere der Nutzung des Prototypen einge-

setzt, wohingegen die Befragung vornehmlich der Ermittlung der Akzeptanz und zur Optimierung der Dialogstation diene.

Bei der verdeckten Beobachtung wurden systematisch die *attracting* und *holding power* erhoben (vgl. Noschka-Roos, 1994; Shettel, 1968). In Bezug auf die *attracting power* wurden alle Besucher erfasst, die an der Medienstation vorbeigelaufen bzw. stehen geblieben sind, letztere unabhängig davon, wie lange sie sich aufhielten an der Station oder ob sie an der Station aktiv wurden. Bei der *holding power* wurden lediglich die Verweilzeiten derjenigen Besucher, die an der Dialogstation aktiv wurden, mit einer Stoppuhr gemessen. Mit der Logfile-Analyse wurden die Interaktionsweisen der Besucher mit der Dialogstation erfasst (vgl. Noschka-Roos, 1994; Heinecke, Bumann & Kerstan, 1995; Heinecke, 2002). Wenn der Prototyp benützt wird, erzeugt er Benutzungsprotokolle, sog. *Logfiles*, in denen alle Benutzeraktionen detailliert mit Datum und Uhrzeit aufgezeichnet werden. Durch diese automatische Protokollierung lassen sich zum einen Aufrufhäufigkeiten ermitteln, zum anderen lassen sich Rückschlüsse auf die Verweildauer bei bestimmten Informationen und die Navigationsmöglichkeiten ziehen. Die mündliche Befragung diene der Untersuchung der Akzeptanz der Besucher sowie der Optimierung der Medienstation. Die Besucher wurden während der Beschäftigung mit dem Prototypen beobachtet und anschließend in einem halbstrukturierten Interview befragt. Insgesamt wurden 90 Interviews durchgeführt. Der Evaluationsfragebogen umfasste offene und geschlossene Items und bezog sich z. B. auf positive und negative Aspekte der Dialogstation, motivationale Hintergründe zur Nutzung und Verbesserungsvorschläge. Die Interviews wurden mit einem Aufnahmegerät aufgezeichnet und anschließend transkribiert.

Zentrale Befunde. Der Prototyp der Dialogstation „Gentest“ wurde überwiegend gemeinschaftlich und von einem jungen Publikum mit einem relativ hohen Bildungsniveau genutzt. Dies entspricht weitgehend den in der Besucherforschung berichteten Befunden (vgl. z. B. Klein, 2000; Geyer et al., 2005; Lewalter, 2001; Lewalter & Noschka-Roos, 1993; Specht & Lewalter, 2011).

Die Evaluationsergebnisse zeigten eine hohe Akzeptanz der Museumsbesucher bezüglich der Dialogstation. Die Mehrheit von 53 % der Besucher blieb an der Station stehen, was für eine hohe *attracting power* spricht (z. B. Lewalter & Noschka-Roos, 1993; Serrel, 2002). Jedoch beachteten nur 16 % der Besucher die authentischen Exponate zum Thema Gentest in der Vitrine. Bezüglich der Frage, ob über die Dialogstation auch das Interesse an den authentischen Exponaten in der Stele geweckt werden konnte, lässt sich dies somit sehr klar verneinen. Die von den Ausstellungsmachern erwünschte Verknüpfung virtuell eingespielter Inhalte und authentischer Objekte gelang demzufolge nicht (Hauser, 2005, 2006). Die Dialogstation konnte die Attraktivität der Exponate, d. h. die Aufmerksamkeits-

zuwendung zum Exponat und somit die Interessensgenese nicht ausreichend unterstützen. Mögliche Erklärungen für die niedrige *attracting power* der Stele und der darin befindlichen Exponate könnte zum einen der Standort sein, da die Stele zum Rücken der Besucher stand (Hashagen, 1997). Zum anderen könnte die geringe Benutzungshäufigkeit auch darauf zurückzuführen zu sein, dass die Stele bzw. die darin befindlichen Exponate nicht attraktiv genug gestaltet waren.

Die durchschnittliche Verweilzeit (*holding power*) der Besucher an der Dialogstation lag bei 154 Sekunden, also etwa 2.5 Minuten, die verglichen mit den Befunden von Serrel (2002) etwas höher ausfiel. Innerhalb der Stichprobe zeigten sich jedoch extrem unterschiedliche Nutzungsdauern. Die kürzeste Aufenthaltsdauer lag bei 14 Sekunden, die Spitzenzeit betrug 1018 Sekunden, also etwa 17 Minuten. Abbildung 13 gibt die Verteilung der Aufenthaltsdauer (*holding power*) bei der beobachteten Stichprobe ($N = 300$) wieder:

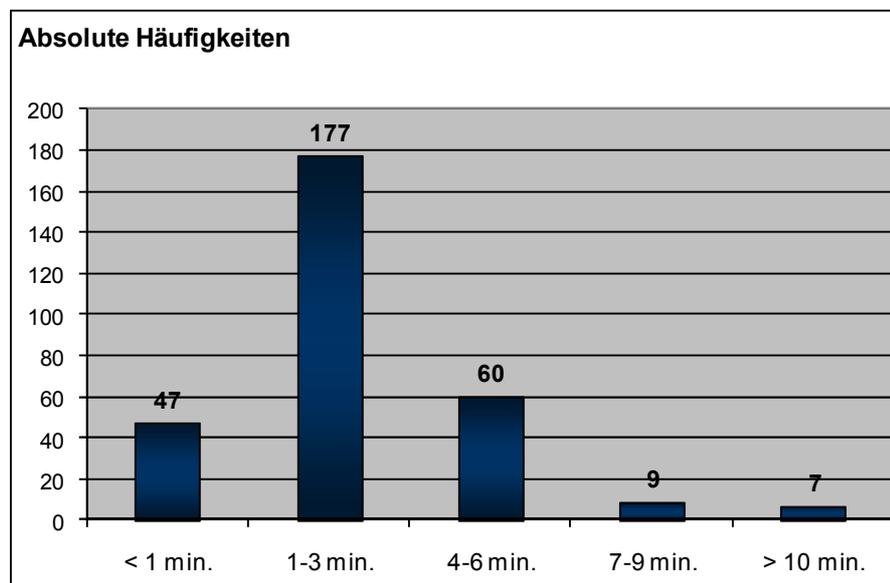


Abbildung 13: Aufenthaltsdauer (*holding power*) der Besucher an dem Prototypen der Dialogstation „Gentest“ in Minuten, absolute Häufigkeiten, Verteilung der Werte ($N=300$).

47 Museumsbesucher (15.7 %) hielten sich weniger als 1 Minute am Gerät auf, testeten und probierten das Touchscreen-System aus und gingen dann weiter, ohne sich wirklich mit der Thematik der Dialogstation auseinandergesetzt zu haben. Der Hauptanteil von 177 Benutzern (59 %) hielt sich 1-3 Minuten am Gerät auf und konnten einen kleinen Einblick in die Thematik der Dialogstation gewinnen. 20 % der Museumsbesucher (60) verweilten 4-6 Minuten. Lediglich 3 % der Benutzer (9) hielten sich 7-9 Minuten auf, 2.3 % (7) länger als 10 Minuten. Insgesamt lässt sich somit festhalten, dass die Dialogstation „Gentest“ von den Besuchern in extrem unterschiedlicher Weise genutzt wird. Immerhin über die Hälfte der Besucher verweilten über 1 bis 3 Minuten an der Dialogstation und konnten somit einen kleinen Einblick in die Thematik erhalten.

Hinsichtlich der motivationalen Hintergründe der Nutzung konnte festgestellt werden, dass es gelang, nicht nur das Fachpublikum, sondern auch Besucher mit eher niedrigem Vorwissen und Interesse zur Auseinandersetzung mit der Dialogstation zu animieren. Überraschenderweise hing der häufigste Grund, sich nicht länger mit der Station zu beschäftigen, nicht unmittelbar von der Dialogstation selbst, wie z. B. dem Thema ab, sondern von den Besuchsabsichten: Am häufigsten wurden in den Interviews zeitliche Aspekte genannt (z. B. *„Das Museum ist so groß, da kann man nicht alles anschauen“*), gefolgt von der fehlenden Einbettung der Dialogstation in einen größeren thematischen Kontext (z. B. *„wir haben nicht sofort verstanden, worum es überhaupt geht“*).

Absolutes „Highlight“ der Dialogstation „Gentest“ stellten die *talking heads* dar, die von über der Hälfte der Befragten aufgrund ihrer Dreidimensionalität/Plastizität und Authentizität positiv hervorgehoben wurden (z. B. *„Ich fand die Darstellung mit den Köpfen interessant, es sprach einen mehr an, als wenn ein Film ablaufen würde, der Persönlichkeitsfaktor, der da mitspielte, war schon ansprechend“*; *„Die Gesichter, das war sehr gut gemacht, das macht betroffen“*; *„Die Darstellung von den Gesichtern ist faszinierend, dieses plastische“*). Die Besucher lobten ferner das interessante und spannende Thema der Dialogstation, den ermöglichten Dialog durch die aktive Positionierung bzw. Abstimmungsstatistik (z. B. *„die Frage am Ende, dass man angeregt wird, darüber nachzudenken“* bzw. *„Fand es auch spannend, zu sehen, wo lieg ich mit meiner Einschätzung, also wie die anderen abgestimmt haben; hat beides meiner Einschätzung entsprochen, z. B. bei der Fanconi-Anämie, dass es fifty/fifty war... und bei der Chorea Huntington, dass mehr auf ja entschieden haben“*) und die zusätzlich abrufbaren Informationen (z. B. *„Mit den gesetzlichen Regelungen das fand ich gut, dass man sehen konnte, wo geht so was, wo geht so was nicht? Also, dass man auch ein bisschen tiefer einsteigen konnte“*).

Aus Besuchersicht konnte eine hohe Zufriedenheit mit der Bedienung des Touchscreen-Monitors festgestellt werden. Die Mehrheit der befragten Museumsbesucher war mit der Handhabung des Touchscreen-Systems zufrieden und lobte die sehr gute, einfache, übersichtliche und problemlose Bedienung der Dialogstation.

Als Schwachstellen des Prototyps wurden die zu langen Startzeiten und die als zu lang empfundene Dauer der Videos aufgedeckt. Diese Ergebnisse konnten zusätzlich durch die Befunde der Logfile-Analyse untermauert werden; beispielsweise zeigten sich hohe Abbruchquoten nach dem Aufruf der Fallbeispiele, die vermutlich auf die zu langen Startzeiten und auf die als zu lang empfundene Dauer der Videos zurückgeführt werden können. Zudem ergab die Logfile-Analyse, dass die zusätzlich verfügbaren Informationen nur sehr wenig von den Besuchern abgerufen und genutzt wurden. Außerdem wurden das provisorische Design des Prototyps (z. B. *„Gesamteindruck sehr weiß“*, *„freundlichere Gestaltung von dem Kasten“*) sowie die fehlende kontextuelle Einbettung in eine thematische Ausstellung

bemängelt. Die als zu lang empfundenen Startzeiten und Dauer der Videos sowie der fehlende Kontext dürften sich vermutlich auch auf die Verweilzeiten negativ ausgewirkt haben.

Erwartungsgemäß bezogen sich die meisten Verbesserungsvorschläge der befragten Besucher auf den zu langsamen Fallstart und die als zu lang empfundene Dauer der Videos; die Besucher wünschen sich schnellere Starts der Videos und eine verkürzte Dauer der Fallvideos. Eine Optimierung der Dialogstation sahen viele Befragte zudem in einem verbesserten Design und einer besseren kontextuellen Einbettung. Durch die Integration der Dialogstation in die zukünftige ZNT-Ausstellung werden diese beiden letztgenannten Schwachpunkte automatisch behoben, da die Dialogstation sich dort im Kontext der genetischen Analysemethoden befindet und die Reliefs zur Projektion der Portrait-Videos für die *talking heads* in eine Wand eingelassen sind (siehe folgende Abbildung 14).

Resumee. Die Dialogstation „Entscheidungslabor Gentest“ setzt eine innovative Form der musealen Präsentation des kontroversen Wissenschaftsthemas ethische Aspekte genetischer Untersuchungen um. Insgesamt betrachtet zeigen die Evaluationsergebnisse, dass die Dialogstation von den Besuchern des Deutschen Museums gut akzeptiert wurde. Dies spricht dafür, Dialogstationen wie diese im Museumsbereich weiterhin und verstärkt einzusetzen. Basierend auf den obig dargelegten Befunden wurden Konsequenzen für die Optimierung der Dialogstation gezogen. Aufgrund der suboptimalen Nutzung der Stele wurde im Rahmen der Ausstellung auf die realen Objekte (Exponate) verzichtet. Demzufolge fungiert die Dialogstation „Gentest“ in der Ausstellung selbst als (digitales) Exponat (Noschka-Roos, 2006). Um die Verweilzeiten (*holding power*) der Museumsbesucher vor der Dialogstation zu steigern, wurden kürzere Startzeiten der Fallbeispiele implementiert und die Dauer der Fallbeispiele deutlich gesenkt. Zu diesem Zweck wurden die Videos neu gedreht und die Schauspieler sprechen ihren Text wesentlich schneller und flüssiger. Da die Zusatzinformationen nur wenig genutzt wurden, können die Besucher diese nun bei Bedarf sowohl vor als auch nach der Positionierung frei abrufen. Ob und inwieweit durch diese Änderungen der Gestaltung der Medienstation die Verweilzeiten der Museumsbesucher gesteigert werden können und somit eine intensivere Auseinandersetzung mit den Inhalten und der Thematik gefördert werden kann, sollte durch weitere Forschung, insbesondere im Gesamtkontext der Ausstellung, noch geklärt werden.

Besucherbefragung in der Ausstellung zur Nano- und Biotechnologie im Deutschen Museum München. Die ZNT-Ausstellung zur Nano- und Biotechnologie des Deutschen Museums München wurde im Jahr 2010 im Rahmen einer Besucherbefragung ($N = 259$) im Hinblick auf die Besucherstruktur, die Nutzung und Bewertung der Ausstellung allgemein, der einzelnen Teilbereiche und der verwendeten Medien summativ evaluiert (Specht &

Lewalter, 2011). Bevor die zentralen Ergebnisse der Besucherbefragung skizziert werden, soll der Vollständigkeit halber noch die im konkreten Ausstellungskontext realisierte „Gentest-Station“ kurz präsentiert werden, da jene im Vergleich zur Prototyp-Version geringfügig modifiziert wurde.

Im konkreten Ausstellungskontext realisierte Lernumgebung. Dem Besucher werden in der ZNT-Ausstellung zur Nano- und Biotechnologie, die im November 2009 eröffnet wurde, sechs verschiedene Fallbeispiele in Form von *talking heads* präsentiert. Die für die Reliefprojektion erforderlichen Gesichtsmasken sind in eine Wand zwischen zwei Teilbereichen der Ausstellung eingelassen (Gerber-Hirt & Noschka-Roos, 2012; Specht & Lewalter, 2011). Hinter der Projektionswand befindet sich ein kleiner Raum mit dem benötigten technischen Equipment. Die Größe der durch die Reliefprojektion entstandenen sechs *talking heads* beträgt 15 x 25 x 10 cm (Breite x Höhe x Tiefe) (Breitsameter et al., 2009).

Abbildung 14 zeigt ein Bild der Dialogstation „Gentest“ in der ZNT-Ausstellung zur Nano- und Biotechnologie.



Abbildung 14: Finale Dialogstation „Gentest“ in der ZNT-Ausstellung zur Nano- und Biotechnologie im Deutschen Museum (Foto: Deutsches Museum, Anmerkung: Gesichter der beiden Ausstellungsbesucher wurden aus Datenschutzgründen unkenntlich gemacht).

Die Reduktion der Anzahl der Fallbeispiele gegenüber der Prototyp-Version von sieben auf sechs Fälle erfolgte aus technischen und räumlichen Gründen. Eine weitere minimale Änderung betraf die integrierten Zusatzinformationen, die im Gegensatz zur Version des Prototyps vom Besucher nicht nur nach, sondern auch vor der Entscheidung frei bei Bedarf abgerufen werden können. Der Menüpunkt Bewertung, der ethische Stellungnahmen von Experten zum jeweiligen Gentest enthielt, wurde in Ethik umbenannt. Zusätzlich werden

bei der Feedbackinformation über die Gesamtentscheidung der anderen Besucher die Fallzahlen eingeblendet, also wie viele Besucher insgesamt abgestimmt haben (z. B. „356 Besucher haben abgestimmt. Vielen Dank für Ihre Stimme!“). Eine weitere Änderung gegenüber der Prototyp-Version stellte die Integration einer englischen Sprachversion dar, um auch ausländischen Besuchern bei Aufruf eines English-Buttons auf dem Touchscreen die Nutzung der Medienstation zu ermöglichen.

Zentrale Befunde. Die Ergebnisse der Besucherstrukturanalyse zeigen zunächst, dass das befragte Publikum in Bezug auf soziodemographische Daten wie Alter, Geschlecht, Bildungsniveau und dem Berufszweig weitestgehend demjenigen früherer Besucherstrukturanalysen des Deutschen Museums entspricht (vgl. z. B. Klein, 2000; Geyer et al., 2005; Lewalter, 2001; Lewalter & Geyer, 2003).

Die Ausstellung zur Nano- und Biotechnologie wurde von den Besuchern nach dem Schulnoten-System im Durchschnitt mit der Gesamtnote „gut“ recht positiv bewertet. Auf die Frage, ob es etwas gibt, was den Besuchern besonders gut an der Ausstellung gefallen hat und warum, konnte kein spezielles Highlight in der gesamten Ausstellung identifiziert werden. Die *talking heads* im Speziellen wurden lediglich von einem Besucher genannt, weil sie einen „ethischen Konflikt“ thematisieren (Specht & Lewalter, 2011, S. 28). In Bezug auf die Frage, welche Medien besonders gut gefallen haben und warum, nannten lediglich drei Besucher die *talking heads*, wovon zwei dies folgendermaßen begründeten: weil man „dort zur Implantations-Diagnostik abstimmen konnte“ und weil sie „innovativ und interessant“ sind (Specht & Lewalter, 2011, S. 37). Das optische Design und die Gestaltung der Ausstellung haben insgesamt betrachtet den Besuchern gefallen. Anlass zur Kritik gaben bei den Ausstellungsbesuchern am häufigsten defekte und nichtfunktionierende Geräte bzw. Ausstellungselemente. Diese mangelnde Funktionalität dürfte sich vermutlich auch auf die Bewertung und Nutzung der *talking heads* negativ ausgewirkt haben, da jene aufgrund noch nicht behobener technischer Probleme nicht dauerhaft in Betrieb sind. Als ein weiterer Kritikpunkt der Ausstellung wurden inhaltspezifische Aspekte genannt, wie z. B. die schwere Verständlichkeit der Ausstellungsthemen für Laien oder Personen ohne entsprechende Vorkenntnisse. Zudem wurden von den Besuchern eine klare Struktur sowie ein Überblick über die gesamte Ausstellung vermisst.

Insgesamt betrachtet konnten die Ergebnisse der summativen Evaluationsstudie von Specht & Lewalter (2011) zeigen, dass die ZNT-Ausstellung zur Nano- und Biotechnologie beim Publikum des Deutschen Museums gut ankam und auf positive Resonanz stieß. Die Autoren der Studie empfehlen zur Optimierung der Ausstellung die Funktionstüchtigkeit der verwendeten Medien zu verbessern sowie Überlegungen anzustellen, mit welchen Maßnahmen Besucher sich innerhalb der Ausstellung besser orientieren können.

6 Zusammenfassung und Konsequenzen für die vorliegende Studie

In diesem Kapitel werden die zentralen Ausführungen im theoretischen Teil dieser Arbeit zusammengefasst und Konsequenzen für die vorliegende Studie abgeleitet.

Mit dieser Arbeit wird versucht, einen ersten Beitrag zum Nachweis der Effektivität einer innovativen musealen Lernumgebung im Hinblick auf die Vermittlung von kontroversen Wissenschaftsthemen zu leisten. Die Medieninstallation Dialogstation „Gentest“ wurde von einer Kuratorin in Zusammenarbeit mit einer Mediendesignerin konzipiert, um als einzelnes Element in einer Ausstellung zum Thema Nano- und Biotechnologie im Deutschen Museum eingesetzt zu werden.

Zentrales Ziel dieser Untersuchung besteht darin, unter möglichst kontrollierten, d. h. labor-experimentellen Bedingungen die Lernwirksamkeit dieser multimedialen Lernumgebung und die Wirkung einer Feedbackinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung anderer Personen im Hinblick auf die Entscheidungsfindung und Argumentation zu untersuchen. Als Experimentalumgebung wurde aus untersuchungstechnischen und ökonomischen Gründen lediglich ein Fallbeispiel zur PID-HLA von den sieben existierenden verwendet (siehe Kapitel 8.3).

Kapitel 2 befasste sich ausführlich mit dem maßgeblichen theoretischen Hintergrund zur Präimplantationsdiagnostik (PID), da dieses beispielhafte medizinethische und kontrovers diskutierte Thema als Untersuchungsgegenstand in dieser Arbeit fungierte. Hierbei konnte gezeigt werden, dass die Durchführung einer PID-HLA zum Zwecke der Rettung eines totkranken Kindes aus ethischer Sicht aufgrund des hohen Embryonenverbrauchs und der Instrumentalisierung des „Retterkindes“ und den damit befürchteten psycho-sozialen und gesellschaftlichen Folgen äußerst problematisch ist (vgl. Rehmann-Sutter, 2007). Ferner ist eine PID-HLA in Deutschland derzeit gesetzlich verboten. Die Analyse der im ethischen Diskurs von Befürwortern und Kritikern am häufigsten angeführten Pro- und Contra-Argumente bezüglich der Legitimität der PID hat zusätzlich offenbart, dass die Argumente für und wider PID-HLA für jeden Einzelnen unterschiedlich überzeugend sind und sehr kontrovers diskutiert werden (vgl. Schräer, 2009).

Darauf aufbauend erfolgte in *Kapitel 3* eine genaue Betrachtung der theoretischen Grundlagen zur Entscheidungsfindung und Argumentation im Bereich Medizinethik. Es konnte gezeigt werden, dass bio- und medizinethische Dilemmasituationen, wie sie z. B. im Zusammenhang mit der PID auftreten, dadurch gekennzeichnet sind, dass meist nur zwei

Handlungsoptionen („Ja vs. Nein“ bzw. „Pro vs. Contra“) und keine klaren Lösungskriterien im Sinne von „falsch“ oder „richtig“ existieren (Eggert, 2008; Heitkamp, et al., 2005; Pfeifer, 2003). Als Erfolgskriterium zur Lösung von diesen schlecht strukturierten und komplexen Problemstellungen ohne eindeutige Lösung nennt Jonassen (2000) die Artikulation einer Präferenz mit einer gewissen Rechtfertigung. Demnach verlangt eine persönlich getroffene Entscheidung stets nach einer ethischen Rechtfertigung bzw. Begründung. Moralische Entscheidungen können utilitaristisch oder deontologisch begründet werden (vgl. Höhle, 2001b; Höhle & Lude, 2004).

Die bei der Problemlösung ablaufenden Denk- und Entscheidungsprozesse können durch den Prozess des informellen Schlussfolgerns (*informal reasoning*) beschrieben werden, der wiederum durch Argumentation erfasst werden kann (Zohar & Nemet, 2002). Die Grundlage für gutes informelles Schlussfolgern bildet kritisches Denken (*critical thinking*), das eine bewusste selbstregulative Urteilsbildung erfordert, welche die Interpretation, Analyse und Bewertung von Informationen und Argumenten sowie das Ziehen einer begründeten Schlussfolgerung beinhaltet (Facione, 1990).

Wie die psychologische Forschung jedoch nachweisen konnte, neigen Personen bei Entscheidungen zu systematischen Urteils- und Entscheidungsverzerrungen: Typischerweise werden Informationen, welche die eigene Entscheidung bestätigen, bevorzugt gesucht und Informationen, die der eigenen Entscheidung widersprechen, eher vernachlässigt (*confirmation bias*) (vgl. z. B. D'Alessio & Allen, 2002; Frey, 1981, 1986; Jonas, Schulz-Hardt et al., 2001). Auch im Hinblick auf die Informationsbewertung konnten systematische Verzerrungen zugunsten präferenzkonsistenter gegenüber -inkonsistenten Informationen belegt werden (*prior belief effect*; Edwards & Smith, 1996). Darüber hinaus werden mögliche Gegenpositionen häufig nur selten in die eigene Entscheidung miteinbezogen (*myside bias*) (vgl. z. B. Baron, 1995; Toplak & Stanovich, 2003).

Ein geringer *confirmation* und *myside bias* werden aber als zentrale Indikatoren für kritisches Denken und das Treffen einer reflektierten Entscheidung angesehen (vgl. Baron, 1995; Knipfer, 2009; Toplak & Stanovich, 2003; West et al., 2008).

Kapitel 4 widmete sich den theoretischen Grundlagen zur Konzeption der in dieser Arbeit untersuchten multimedialen Lernumgebung. Es wurden eine Reihe von Ansätzen vorgestellt, aus denen sich didaktische und mediale Design-Prinzipien für multimediale Lernumgebungen ableiten lassen. Im Hinblick auf eine effektive Gestaltung von multimedialen Lernumgebungen sollte darauf geachtet werden, ein geeignetes und angemessenes didaktisches Design zu verwenden und die Lerninhalte in medial adäquater Weise aufzubereiten (vgl. Fischer et al., 2009; Kopp & Mandl, 2009; Mayer, 2005a; Niegemann, 2001; Reinmann & Mandl, 2006). Auf den konkreten Einsatz im informellen Lernfeld Museum

bezogen, stellen multimediale Lernumgebungen in der Regel sehr kurzfristige Interventionen dar und sollten als einzelne Ausstellungseinheit so gestaltet sein, dass sie unabhängig vom Gesamtkontext der Ausstellung funktionieren und möglichst positive kognitive und motivationale Lernprozesse bei den Besuchern anstoßen (vgl. Prenzel, 2009; Scheersoi, 2006; Schwan et al., 2008; siehe Kapitel 4.5).

In *Kapitel 5* wurde die multimediale Lernumgebung Dialogstation „Gentest“, die als Kontext der Untersuchung diente, ausführlich vorgestellt. Im Hinblick auf die didaktische und mediale Gestaltung konnte gezeigt werden, dass die museale Lernumgebung Dialogstation „Gentest“ in hohem Maße einer gemäßigt konstruktivistischen Auffassung vom Lernen gerecht wird (Gerstenmaier & Mandl, 1995) und die Anforderungen einer problemorientierten und multimedial adäquat gestalteten Lernumgebung erfüllt (Mayer, 2005a; Reinmann & Mandl, 2006).

Grundlegendes Lernziel der Dialogstation „Gentest“ besteht aus kognitiver Sicht darin, das Fällen einer informierten und reflektierten Entscheidung in Bezug auf medizinethische Dilemmata, die im Zusammenhang mit genetischen Untersuchungen auftreten, zu unterstützen und zu fördern. In einer ersten formativen Evaluationsstudie zur Lernumgebung im Feld zeigte sich in Übereinstimmung mit zahlreichen Befunden aus der Besucherforschung ein äußerst zeitlich eingeschränktes und selektives Nutzungsverhalten beim Museumspublikum (siehe Kapitel 5.6).

Angesichts dieser Befunde wird der Frage, inwieweit sich die Lernumgebung in kognitiver Hinsicht bewährt, in dieser Studie auf experimentellem Wege nachgegangen. Daneben war eine Feldstudie im realen Ausstellungskontext aus zeitlichen Gründen leider nicht möglich, da aufgrund unvorhergesehener Schwierigkeiten bei den Baumaßnahmen der ursprüngliche Eröffnungstermin der Ausstellung zur Nano- und Biotechnologie im Deutschen Museum von Mai 2008 auf November 2009 verschoben wurde. Aufgrund zeitlicher Rahmenbedingungen und Gründen der Ökonomie konnte eine Überprüfung der in Kapitel 5.2 beschriebenen motivationalen Lernziele der Dialogstation „Gentest“ nicht realisiert werden.

Ausgehend von den theoretischen Überlegungen in *Kapitel 3* wurde in dieser Untersuchung als Indikator des Lernerfolgs bzw. des Lernfortschritts die Argumentationsqualität im Rahmen einer Prä-Post-Messung herangezogen. Um die Argumentationsqualität zu erfassen, wurden Essays verwendet, in denen die Probanden ihre Entscheidung schriftlich begründeten. Zur Klassifikation der Qualität der Argumentationen wurde das in Kapitel 3.2.3 beschriebene Toulmin-Schema verwendet. Knipfer (2009) konnte bereits positive kognitive Auswirkungen von Diskussionsterminals auf Indikatoren der Argumentations- bzw. Meinungsqualität im Bereich Nanotechnologie feststellen. Diese ersten positiven Befunde zur Lernwirksamkeit von Dialogstationen im Hinblick auf die Unterstützung und

Förderung von kritischem Denken und reflektiertem Urteilen zu kontroversen Wissenschaftsthemen im Museum sollen mit dem hier dargestellten Experiment im Bereich Medizinethik repliziert werden.

In dieser Untersuchung wird auch analysiert, welche Argumente von den Lernenden am häufigsten zur Begründung der eigenen Entscheidung angeführt werden und ob die getroffene Entscheidung eher utilitaristisch oder deontologisch gerechtfertigt und begründet wird (vgl. Höhle, 2001b). Zudem soll in Anlehnung an Sadler & Zeidler (2005) überprüft werden, ob eher rational oder emotional argumentiert wird und welchen Einfluss Emotionen und Intuitionen bei der Lösung der Dilemmageschichte haben (vgl. Haidt, 2001).

Ein weiteres zentrales Ziel des Experiments besteht darin, die Wirkung bzw. den Einfluss der implementierten Feedbackinformation auf das Entscheidungsverhalten und die Argumentationsqualität zu untersuchen. In Kapitel 4.4 wurden deshalb die theoretischen Grundlagen zur Feedbackinformation beschrieben. Feedback muss, um einen nachweislichen Lernvorteil erbringen zu können, lernförderlich gestaltet sein und vom Empfänger in geeigneter Weise rezipiert werden (vgl. Krause, 2007). Knipfer (2009) konnte empirisch nachweisen, dass konfligierendes Feedback, das im Widerspruch zur eigenen Meinung stand, Dissonanz auslöste (Festinger, 1957; Frey & Gaska, 1993) und in einer signifikant häufigeren Änderung des Gesamturteils zur Nanotechnologie resultierte. Im Vergleich zu zwei Kontrollbedingungen, die kein bzw. ein meinungskonsistentes Feedback erhielten, führte konfligierendes Feedback auch zu einem geringeren *myside bias* im Argument Repertoire und einer besseren Integration von Gegenargumenten in der Begründung der eigenen Position sowie besseren Diskussion von Gegenargumenten. Diese positiven Befunde von konfligierendem Feedback auf Indikatoren der Meinungsqualität sollen in dieser Studie im Bereich der medizinethischen Entscheidungsfindung und Argumentation repliziert werden. Um den Einfluss der implementierten Feedbackinformation im Hinblick auf die Unterstützung und Förderung einer informierten und reflektierten Entscheidung zu untersuchen, wurde in Anlehnung an die Studie von Knipfer (2009) die soziale Vergleichsinformation in Form eines Umfrageergebnisses in Abhängigkeit von der eigenen Entscheidung gezielt manipuliert (unabhängige Variable: Feedback; Stufen: kongruent, inkongruent bzw. konfligierend, ohne Feedback).

Ferner sollen auch Erkenntnisse darüber gewonnen werden, inwieweit sich das in der Lernumgebung eingesetzte Präsentationsformat in Form eines *talking heads* bewährt und ob sich ebenfalls in Übereinstimmung mit den Befunden von Töpper (2009) ein lernförderlicher Einfluss der filmischen Personalisierung von Wissensinhalten zeigt (siehe Kapitel 4.5). Zusätzlich soll analysiert werden, welche Bedeutung den Lernvoraussetzungen (z. B. Vorwissen) der Lernenden im vorliegenden Kontext zukommt.

Für diese Untersuchung wurde ein Messwiederholungsdesign mit drei Messzeitpunkten verwendet: Die erste Messung erfolgte nach der Präsentation des personalisierten Films (Messzeitpunkt t1: Film), die zweite nach Feedbackgabe (Messzeitpunkt t2: Feedback) und die dritte nach Bearbeitung der zusätzlichen, vertiefenden Informationen (Messzeitpunkt t3: Informationen). Auf diese Weise war es möglich, den Einfluss des Präsentationsformates (*talking head*), der Feedbackinformation und der Zusatzinformationen auf das Entscheidungsverhalten und die Argumentation getrennt ermitteln zu können und mögliche Konfundierungseffekte zu vermeiden.

Da Ausstellungsgestalter in erster Linie an praktischen Empfehlungen für ihre Ausstellungsarbeit interessiert sind, soll zusätzlich geprüft werden, ob und inwieweit die didaktische und mediale Gestaltung der musealen Lernumgebung basierend auf den Untersuchungsergebnissen optimiert werden kann (Prenzel & Thoma, 2006).

Wie eine empirische Studie von Brähler & Stöbel-Richter (2004) zeigen konnte, besteht in der deutschen Allgemeinbevölkerung nur ein äußerst geringes Interesse und Wissen über die PID (siehe Kapitel 2.5). Ausgehend von diesen Befunden erschien eine Adaption der in dieser Studie verwendeten Ratingskalen zur Erhebung von Lernvoraussetzungen wie dem persönlichen Bezug, dem Interesse und den Einstellungen zum Inhaltsgebiet PID als nicht sinnvoll: Denn Einstellungen zu einem Gegenstand, zu dem nur wenig oder kein Wissen da ist, sind nur sehr schwer zu erheben. Stattdessen wurden diese Variablen auf die übergeordnete Domäne „genetische Untersuchungen“ hin operationalisiert, da hier mit einem höheren Vorwissen der Lernenden zu rechnen war und die PID einen mit genetischen Untersuchungen verknüpften Anwendungsbereich darstellt.

In Anlehnung an die Befunde von Götz (2001), der keine Veränderungen in den Einstellungen zur Gentechnik durch die Bearbeitung eines (kurzfristigen) Lernprogramms zu ethischen Problemen der Gentechnik feststellen konnte, wurde auf eine Prä-/Post-Messung der Einstellungen zu genetischen Untersuchungen in dieser Studie verzichtet (siehe Kapitel 4.6).

Zusammenfassend lässt sich festhalten, dass die vorgestellten theoretischen Überlegungen und empirischen Befunde dafür sprechen, dass Dialogstationen eine sinnvolle und geeignete Maßnahme im Hinblick auf die Unterstützung und Förderung von kritischem Denken und einer informierten und reflektierten Entscheidung zu kontroversen Wissenschaftsthemen im musealen Lernkontext darstellen. Es wird angenommen, dass die Auseinandersetzung mit der Lernumgebung Dialogstation „Gentest“ einen bedeutsamen Lernfortschritt im Sinne einer höheren Argumentationsqualität bewirkt. Angesichts der Befundlage zur Effektivität von konfligierenden Feedbacks wird erwartet, dass sich positive Auswirkungen auf Indikatoren der Argumentationsqualität zeigen.

7 Fragestellungen und Ableitung der Hypothesen

In diesem Kapitel werden die Fragestellungen der durchgeführten experimentellen Laborstudie aufgeführt und Hypothesen abgeleitet.

Zentrales Ziel dieser Untersuchung besteht darin, unter möglichst kontrollierten, d. h. labor-experimentellen Bedingungen die Lernwirksamkeit einer multimedialen Lernumgebung und die Wirkung einer Feedbackinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung anderer Personen zu untersuchen.

Ausgehend von einem 3x3-faktoriellen Design wurden die Faktoren "Feedback" (Stufen: kongruent, inkongruent, ohne) und „Messzeitpunkt“ (Stufen: t1: Film, t2: Feedback, t3: Informationen) in Hinblick auf das Entscheidungsverhalten, die Entscheidungssicherheit und Indikatoren der Argumentationsqualität untersucht. Zusätzlich sollen auch Erkenntnisse über die Rolle von Einflussfaktoren auf das Entscheidungsverhalten und die Argumentationsqualität gewonnen werden.

Die Fragestellungen lassen sich in fünf Themenkomplexe gliedern: Fragen zur parasozialen Interaktion mit der Medienfigur (Kapitel 7.1), Fragen zum Entscheidungsverhalten (Kapitel 7.2), Fragen zur Argumentationsqualität (Kapitel 7.3) und zu den Argumentationsmustern (Kapitel 7.4) sowie Fragen zu Einflussfaktoren auf das Entscheidungsverhalten und die Argumentationsqualität (Kapitel 7.5).

Um die Lesbarkeit dieses Kapitels nicht zu beeinträchtigen, wird auf eine inflationäre Auflistung von Hypothesen verzichtet und es werden nur acht zentrale Hypothesen formuliert.

7.1 Analyse der parasozialen Interaktion mit der Medienfigur

Fragestellung 1: In welchem Ausmaß setzen sich die Lernenden mit der Medienfigur auseinander?

Dem Zwei-Ebenen-Modell parasozialer Interaktionen von Hartmann et al. (2004a) zufolge liegen aufgrund der medialen Inszenierung in Form von *talking heads* günstige Bedingungen für eine High-Level-PSI bzw. ein starkes interpersonales Involvement mit den virtuellen Figuren vor. Dementsprechend wurde angenommen, dass bei den Lernenden ein starkes interpersonales Involvement mit der Medienfigur vorliegt.

Fragestellung 2: In welchem Zusammenhang steht die parasoziale Interaktion mit der wahrgenommenen Attraktivität der Medienfigur?

Wie Hartmann & Klimmt (2005) in einer Studie feststellen konnten, wird die Intensität der PSI durch die wahrgenommene Attraktivität der Medienfigur beeinflusst: Je attraktiver die Medienperson bewertet wird, desto intensiver fallen die resultierenden perzeptiv-kognitiven, affektiven und konativen PSI-Teilprozesse aus. Es wurde also erwartet, dass die empfundene Attraktivität der Medienfigur mit den PSI positiv korreliert.

7.2 Analyse des Entscheidungsverhaltens

Die Probanden wurden in dieser Untersuchung mit einem medizinethischen Dilemma zum Thema PID-HLA konfrontiert, zu welchem sie insgesamt drei stellvertretende Handlungsentscheidungen zwischen zwei Alternativen treffen mussten („Ja“ vs. „Nein“). Die Entscheidungsfrage lautete: „Würden Sie an ihrer Stelle für eine Präimplantationsdiagnostik nach Amerika gehen?“ (siehe Abbildung 20, Kapitel 8.3, S. 187).

Die erste, vorläufig zu treffende Entscheidung erfolgte unmittelbar nach der Präsentation der Dilemmageschichte durch den *talking head* (erster Messzeitpunkt t1), die zweite, vorläufige Entscheidung nach der Präsentation der Feedbackinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung der anderen Versuchsteilnehmer (zweiter Messzeitpunkt t2). Die dritte, d. h. finale und endgültige Entscheidung erfolgte nach der Bearbeitung der zusätzlichen und vertiefenden Informationen (dritter Messzeitpunkt t3).

Das Entscheidungsverhalten wurde im Hinblick auf die Richtung, die Sicherheit und die Änderung der Entscheidungen für alle drei Lernbedingungen im zeitlichen Verlauf (d. h. über alle drei Messzeitpunkte hinweg) analysiert.

Fragestellung 3: In welche Richtung gehen die Entscheidungen?

In dieser Fragestellung wird bei allen drei Lernbedingungen die Richtung der drei getroffenen Entscheidungen, also welche der beiden Handlungsalternativen im zeitlichen Verlauf präferiert wird, deskriptiv analysiert.

Fragestellung 4: Inwieweit kommt es durch die Bearbeitung der Lernumgebung zu einer Zunahme der Entscheidungssicherheit und inwieweit beeinflusst die Kongruenz der Feedbackinformation die Entscheidungssicherheit?

Menschen sind in der Regel bestrebt, zu einer objektiv richtigen und korrekten Entscheidung zu gelangen und ihre eigene Entscheidung mit derjenigen von anderen Personen zu vergleichen und zu validieren (vgl. Accuracy-Motiv: Chaiken et al., 1989; Festinger, 1954; Frey et al., 1993). In mehrdeutigen Situationen, wie z. B. dem Fällen einer Entscheidung zum kontroversen Wissenschaftsthema PID-HLA, herrscht in der Regel eine große Unsicherheit darüber, ob die eigene Entscheidung richtig bzw. korrekt ist.

Dementsprechend wird erwartet, dass die Sicherheit der Entscheidung zum ersten Messzeitpunkt (t1) relativ gering ausfällt. Da hoher Konsens eine valide Entscheidung impliziert, sollte die Sicherheit der eigenen Entscheidung zum zweiten Messzeitpunkt (t2) bei der Lernbedingung „kongruentes Feedback“ steigen, wenn die eigene Entscheidung von einer Mehrheit vertreten wird (Erb, 1998). In Analogie dazu wird erwartet, dass bei der Lernbedingung „inkongruentes Feedback“ die Konfrontation mit einer zur eigenen Entscheidung konfligierenden Konsensinformation zu einer Abnahme der Entscheidungssicherheit führen sollte.

In einer Evaluationsstudie zu einem Computerlernprogramm zum Thema ethische Aspekte der Gentechnik konnte Götz (2001) die Hypothese, dass die anfänglich getroffenen Entscheidungen zu bioethischen Dilemmata nach der Bearbeitung von zusätzlichen, vertiefenden Informationen im Mittel sicherer ausfallen als die Entscheidung vor den Informationen nur in Bezug auf ein Dilemma zur Erbkrankheit Chorea Huntington empirisch bestätigen. Allerdings konnte Götz (2001) eine kontinuierliche Zunahme der Sicherheit der Entscheidung durch die Bearbeitung der computerunterstützten Lernumgebung insgesamt, d. h. über alle 5 Dilemmata hinweg, feststellen. In Anlehnung an diese Befunde wird postuliert, dass durch die Bearbeitung der Lernumgebung die Sicherheit der Entscheidung im zeitlichen Verlauf zunimmt. Es wird also erwartet, dass die Sicherheit der Entscheidung durch die Bearbeitung der Zusatzinformationen steigt.

Bezüglich der Fragestellung 4 werden folgende drei Hypothesen aufgestellt:

Hypothese 1: Die Entscheidungssicherheit verläuft über die drei Messzeitpunkte nicht konstant. Es wird erwartet, dass die Entscheidungssicherheit im zeitlichen Verlauf durch die Bearbeitung der Lernumgebung zunimmt. Dementsprechend zeigt sich ein signifikanter Haupteffekt des Messwiederholungsfaktors „Entscheidungssicherheit“.

Hypothese 2: Die Entscheidungssicherheit verläuft im zeitlichen Durchschnitt nicht gleich bei den drei Gruppen. Dementsprechend zeigt sich ein signifikanter Haupteffekt des Gruppenfaktors „Lernbedingung“.

Hypothese 3: Es wird angenommen, dass bei inkongruentem Feedback die Sicherheit der Entscheidung nach Feedbackgabe zum zweiten Erhebungszeitpunkt (t2) abnimmt, bei kongruentem Feedback dagegen die Sicherheit der Entscheidung zunimmt. Zwischen dem zeitlichen Verlauf der Entscheidungssicherheit in Bezug auf die beiden Experimentalgruppen bestehen somit Wechselwirkungen. Dementsprechend sollte sich ein signifikanter Interaktionseffekt zwischen dem Gruppenfaktor „Lernbedingung“ und dem Messwiederholungsfaktor „Sicherheit der Entscheidung“ zeigen.

Fragestellung 5: Inwieweit führt inkongruentes Feedback zu einer Änderung der Entscheidung?

In dieser Fragestellung wird untersucht, wie die Probanden auf die Information reagieren, ob ihre eigene Entscheidung von einer Minderheit oder Mehrheit vertreten wird. Knipfer (2009) konnte empirisch nachweisen, dass konfligierendes Feedback, das im Widerspruch zur eigenen Meinung stand, kognitive Dissonanz auslöste, die psychisch unangenehm erlebt wurde und im Vergleich zu zwei Bedingungen mit kongruentem und ohne Feedback zu einer signifikant häufigeren Änderung des Gesamturteils über Nanotechnologie nach dem Feedback führte (vgl. Elliot & Devine, 1994; Festinger, 1957; Frey & Gaska, 1993).

Dementsprechend wird angenommen, dass bei den Probanden der Bedingung „inkongruentes Feedback“ durch die Konfrontation mit einer präferenzinkonsistenten Feedbackinformation eine dissonanzerzeugende Situation induziert wird, die in einer Revision der zuvor getroffenen Entscheidung resultieren sollte.

Bezüglich der Fragestellung 5 wurde folgende Hypothese formuliert:

Hypothese 4: Inkongruentes Feedback führt zu einer häufigeren Änderung der Entscheidung als kongruentes oder kein Feedback.

Fragestellung 6: Inwieweit führt die Bearbeitung der Zusatzinformationen zu einer Änderung der Entscheidung?

In dieser Fragestellung wurde analysiert, ob sich durch die Bearbeitung der Zusatzinformationen die zuvor getroffene Entscheidung ändert. Götz (2001) konnte beispielsweise feststellen, dass durch die Bearbeitung zusätzlicher, vertiefender Informationen die anfänglichen Entscheidungen zu ethisch-moralischen Dilemmata zum Thema Gentechnik nur sehr selten revidiert wurden.

7.3 Analyse der Argumentationsqualität

Knipfer (2009) konnte anhand einer experimentellen Studie feststellen, dass Lernende, die „inkongruentes“, d. h. „konfligierendes Feedback“ erhielten, eine signifikant höhere Argumentations- bzw. Meinungsqualität aufwiesen als Lernende der Bedingungen „kongruentes“ und „kein Feedback“.

Darüber hinaus konnte die Lernbedingung „konfligierendes Feedback“ mehr Argumente zur Unterstützung der eigenen Meinung generieren, mehr Gegenargumente nennen und diskutieren und wies einen geringeren *myside bias* in der Argumentation auf als die beiden anderen Gruppen.

Diese positiven Befunde von konfligierendem Feedback über das Meinungsbild der anderen Besucher auf Indikatoren der Meinungsqualität sollen auf den Kontext dieser vorliegenden Arbeit übertragen und repliziert werden. Dementsprechend ergeben sich folgende Fragestellungen und Hypothesen:

Fragestellung 7: Inwieweit kommt es durch die Bearbeitung der Lernumgebung zu einem Lernfortschritt und inwiefern wird der Lernfortschritt durch inkongruentes Feedback und durch die Bearbeitung der Zusatzinformationen gefördert?

Primäres Lernziel der Lernumgebung besteht in der Unterstützung und Förderung einer informierten und argumentativ gut begründeten Entscheidung im Hinblick auf ein medizinisches Dilemma zum Thema PID-HLA. Als Maß für den Lernfortschritt wird in dieser Studie die Argumentationsqualität vor und nach Bearbeitung der Zusatzinformationen, d. h. zu den beiden Messzeitpunkten t1 und t3, herangezogen.

Angesichts der Befunde von Knipfer (2009) zur Effektivität von Diskussionsterminals wird erwartet, dass die Auseinandersetzung mit der Lernumgebung insgesamt, d. h. über alle drei Gruppen hinweg, zu einem bedeutsamen Lernfortschritt, also einer höheren Argumentationsqualität nach Bearbeitung der Lernumgebung führt. Für die Lernbedingung „inkongruentes Feedback“ werden besonders positive Effekte im Hinblick auf einen Lernfortschritt postuliert.

Bezüglich der Fragestellung 7 werden folgende Hypothesen aufgestellt:

Hypothese 5: Die Bearbeitung der Lernumgebung führt zu einem bedeutsamen Lernfortschritt, der sich in einer höheren Argumentationsqualität nach Bearbeitung der Zusatzinformationen zeigt. Entsprechend ergibt sich ein Effekt des Messwiederholungsfaktors „Argumentationsqualität vor und nach Bearbeitung der Zusatzinformationen“.

Hypothese 6: Inkongruentes Feedback fördert den Lernfortschritt, der sich einer höheren Argumentationsqualität zeigt. Entsprechend ergibt sich ein Interaktionseffekt der Lernbedingung und des Messwiederholungsfaktors.

Fragestellung 8: Inwieweit wird die Generierung von Gegenargumenten durch inkongruentes Feedback gefördert?

In Anlehnung an die Befunde von Knipfer (2009) wird erwartet, dass Lernende der Bedingung „inkongruentes Feedback“ insgesamt mehr Gegenargumente generieren können als Lernende der Bedingung „kongruentes“ und „kein Feedback“. Bezüglich der Fragestellung 8 wird folgende Hypothese aufgestellt:

Hypothese 7: Lernende der Bedingung „inkongruentes Feedback“ generieren mehr Gegenargumente als Lernende der Bedingungen „kongruentes Feedback“ und „kein Feedback“.

Fragestellung 9: Inwieweit wird der myside bias durch inkongruentes Feedback verringert?

Den Befunden von Knipfer (2009) entsprechend wurde angenommen, dass Lernende der Bedingung „inkongruentes Feedback“ einen geringeren *myside bias* in der Argumentation aufweisen als die beiden anderen Gruppen. Bezüglich der Fragestellung 9 wird folgende Hypothese aufgestellt:

Hypothese 8: Lernende der Bedingung „inkongruentes Feedback“ weisen einen geringeren *myside bias* in der Argumentation auf als Lernende der Bedingung „kongruentes Feedback“ und „kein Feedback“.

7.4 Analyse der Argumentationsmuster

Fragestellung 10: Welche Argumentationsmuster (rational oder emotional-intuitiv) werden am häufigsten zur Begründung der eigenen Entscheidung verwendet?

Sadler & Zeidler (2005) konnten bei Schülern drei typische Begründungsmuster für die persönlich getroffene Entscheidung in Bezug auf gentechnologische Dilemmata rekonstruieren: rational, intuitiv und emotional. In dieser Fragestellung wird untersucht, ob eher emotional-intuitiv oder rational von den Probanden argumentiert wurde und welche Rolle Intuitionen und Emotionen bei der Lösung der Dilemmageschichte zur PID-HLA spielen (Sadler & Zeidler, 2005).

Eine informierte und gut begründete Entscheidung zu einem ethischen Dilemma sollte überwiegend auf rationalem Wege erfolgen (vgl. Haidt, 2001; Reitschert, 2009).

Fragestellung 11: Welche deontologischen und utilitaristischen Argumente werden von den Versuchspersonen am häufigsten formuliert?

Im Rahmen dieser Fragestellung wurde analysiert, welche Argumente am häufigsten bei der Begründung der eigenen Entscheidung und bei der Aufforderung von Gegenargumenten von den Probanden genannt wurden. Dabei wurden die formulierten Argumente in Anlehnung an Höhle (2001b) im Hinblick auf ihre ethische Tradition (deontologisch bzw. utilitaristisch) systematisiert.

Höhle (2003b) konnte beispielsweise in einer Studie an Schülern, die mit einem Dilemma zum Thema embryonale Stammzellforschung konfrontiert wurden, empirisch nachweisen, dass zur Begründung des eigenen moralischen Urteils deontologische Argumente häufiger angeführt wurden als utilitaristische. Konträr dazu konnte Götz (2001) bei der Evaluation eines Lernprogramms zu ethischen Aspekten der Gentechnik feststellen, dass die befragten

Studenten ihre getroffene Entscheidung häufiger utilitaristisch begründeten bzw. rechtfertigten als deontologisch. Die Verwendung der beiden Argumentationstypen war jedoch stark kontextabhängig.

7.5 Analyse von Einflussfaktoren auf das Entscheidungsverhalten und Argumentationsqualität

Fragestellung 12: In welchem Zusammenhang steht die Argumentationsqualität mit den erfassten Lernvoraussetzungen auf der einen und der parasozialen Interaktion mit der Medienfigur auf der anderen Seite?

Diese Analysen dienen der Aufdeckung von wichtigen Einflussgrößen beim Lernen mit der multimedialen Lernumgebung. Es sollten Erkenntnisse darüber gewonnen werden, welche Bedeutung verschiedenen Lernermerkmalen (z. B. Vorwissen) oder der Medienfigur (*talking head*) bei der Bearbeitung der Lernumgebung zukommt.

Nach Sadler & Fowler (2006) und Chang & Chiu (2008) kann das inhaltsbezogene Vorwissen als ein wichtiger Prädiktor im Hinblick auf die Argumentationsqualität zu kontroversen Wissenschaftsthemen angesehen werden. Dementsprechend wurde erwartet, dass das Vorwissen über Humangenetik und das subjektive Vorwissen zur PID mit der Argumentationsqualität positiv korreliert und sich bedeutsam auf die Argumentationsqualität auswirkt.

Ferner wurde in Anlehnung an Mayer (2005c) und Töpper (2009) angenommen, dass sich ein lernförderlicher Einfluss des eingesetzten Präsentationsformates in Form eines sichtbaren *talking heads* zeigt. Demzufolge sollte eine positive Assoziation der parasozialen Interaktion mit der Medienfigur und der Argumentationsqualität vorliegen.

Fragestellung 13: Welchen Einfluss haben die erfassten Lernvoraussetzungen und die Sicherheit der Entscheidung auf die Revision der Entscheidung?

In dieser Fragestellung wurde untersucht, inwieweit die Änderung der Entscheidung von den Lernvoraussetzungen und der Entscheidungssicherheit abhängt.

8 Methode

In diesem Kapitel werden Stichprobe (Kapitel 8.1) und Design (Kapitel 8.2), die verwendete Lernumgebung (Kapitel 8.3), die Operationalisierung der unabhängigen Variablen (Kapitel 8.4), der Versuchsablauf (Kapitel 8.5) sowie die Erhebungsinstrumente (Kapitel 8.6 und 8.7) der experimentellen Laborstudie vorgestellt.

8.1 Stichprobe

Am Experiment nahmen 72 Studierende der Ludwig-Maximilians-Universität München, der Technischen Universität München und der Hochschule München im Erhebungszeitraum von Ende Mai bis Juli 2009 teil. Die Studenten wurden über Aushänge im Kassenfoyer, in der Bibliothek und über einen Aufruf auf der Homepage des Deutschen Museums angeworben. Als weitere Rekrutierungsmaßnahmen kamen Mailinglisten sowie Aushänge bei verschiedenen universitären Fachschaften zum Einsatz. Die Versuchspersonen erhielten für die freiwillige Teilnahme an der Studie eine Aufwandsentschädigung in Höhe von 10 Euro. Nach dem Ende des Experiments wurde den Probanden von der Versuchsleiterin ein Codename zugewiesen; die Datenerhebung erfolgte folglich anonym.

Studierende wurden als Stichprobe aus zwei Gründen ausgewählt: Zum einen war es für diese Studie wichtig, eine relativ altershomogene Gruppe zu untersuchen, da die Einstellungen zu genetischen Untersuchungen vom Alter abhängen können (vgl. Berth et al., 2002). Zum anderen ist aus diversen Besucherstudien bekannt, dass das Publikum des Deutschen Museums relativ jung ist und ein überdurchschnittlich hohes Bildungsniveau aufweist (vgl. Geyer et al., 2005; Klein, 2000; Lewalter, 2001; Lewalter & Geyer, 2003; Specht & Lewalter, 2011). Verglichen mit dem Bevölkerungsdurchschnitt befinden sich in der Publikumsstichprobe des Deutschen Museums überproportional viele Akademiker. Da die naturwissenschaftlich-technischen Akademiker deutlich stärker vertreten sind als die geistes- oder sozialwissenschaftlichen, wurde bei der Probandenauswahl zusätzlich darauf geachtet, bevorzugt Studierende aus naturwissenschaftlich-technischen Studienfächern zu rekrutieren.

Dementsprechend waren die Mehrheit von 43 Versuchspersonen (59.7 %) Studierende naturwissenschaftlich-technischer Fachrichtungen, 29 Teilnehmer (40.3 %) studierten ein geistes- oder sozialwissenschaftliches Fach. Die Geschlechterverteilung war nahezu ausgewogen: Es nahmen 39 Frauen (54.2 %) und 33 Männer (45.8 %) teil. Die Anzahl der

Fachsemester variierte zwischen dem 2. und 13. Semester; der Mittelwert lag bei etwa 7 Semestern ($M = 6.63$, $SD = 3.50$). Die Versuchspersonen waren im Alter zwischen 19 und 40 Jahren, das Durchschnittsalter betrug knapp 25 Jahre ($M = 25.44$, $SD = 4.50$).

8.2 Design

Dem Experiment lag ein 3 (Feedback: kongruent, inkongruent, kein) x 3 (Messzeitpunkt: t1: Film, t2: Feedback, t3: Informationen) faktorielles, gemischtes Design mit Messwiederholung auf dem zweiten Faktor zugrunde. Das Feedback bildete hierbei den "between-subjects" - Faktor, der Messzeitpunkt den „within-subject“ - Faktor.

Tabelle 8 gibt einen Überblick über das Untersuchungsdesign.

Tabelle 8: Übersicht über das Untersuchungsdesign und Stichprobengröße ($N = 72$).

Feedback ("between")	Zeitpunkt ("within")		
	Film (t1)	Feedback (t2)	Informationen (t3)
1. Kongruentes Feedback (Experimentalgruppe 1)	$n=24$	$n=24$	$n=24$
2. Inkongruentes Feedback (Experimentalgruppe 2)	$n=24$	$n=24$	$n=24$
3. Kein Feedback (Kontrollgruppe)	$n=24$	$n=24$	$n=24$

Als abhängige Variablen wurden zu allen drei Messzeitpunkten die Richtung der getroffenen Entscheidung (dichotom: Ja vs. Nein) und die Sicherheit der Entscheidung erhoben. Die Argumentationsqualität wurde zu den zwei Messzeitpunkten t1 und t3 (vor und nach Bearbeitung der Zusatzinformationen) erfasst.

8.3 Verwendete Lernumgebung: Fallbeispiel PID-HLA

Als Lernumgebung zur Durchführung der Studie wurde von sieben existierenden ein Fallbeispiel der Dialogstation „Gentest“ zum Thema PID-HLA verwendet. Die Konzeption und Struktur der Lernumgebung wurden in Kapitel 5 sehr ausführlich dargestellt.

Zur Durchführung des Experiments wurde die Struktur der Lernumgebung gegenüber der Originalversion aus untersuchungstechnischen, ökonomischen und methodologischen Gründen leicht modifiziert. Leider konnte in dieser experimentellen Untersuchung kein ausschließlich postrezeptiver Messzugang gewählt werden, da nicht nur das Feedback (UV), sondern auch andere Faktoren wie beispielsweise die Medienfigur oder die Zusatzin-

formationen die Entscheidungsfindung und die Argumentationsqualität (AV's) potentiell beeinflussen können. Zur Vermeidung dieser Konfundierung, die eine kausale Interpretation der Untersuchungsergebnisse erschweren könnte, wurden zwei zusätzliche Messzeitpunkte in die Lernumgebung implementiert (siehe später).

Aus ökonomischen und untersuchungstechnischen Gründen wurde zunächst lediglich eine Fallgeschichte in diesem Experiment untersucht. Von den insgesamt sieben existierenden Fallbeispielen wurden vorab auf inhaltsanalytischem Wege die zwei mit dem höchsten Kontroversitäts- und Ambivalenzgrad ausgewählt: ein Fallbeispiel zum Thema Chorea Huntington und eine Fallgeschichte zum Thema PID-HLA. Im nächsten Schritt sollte mit Hilfe einer Voruntersuchung eine geeignete Auswahl zwischen diesen beiden Fallbeispielen getroffen werden. Als Auswahlkriterium wurde die Qualität der zusätzlichen Informationen im Hinblick auf deren Entscheidungsrelevanz und Verständlichkeit herangezogen.

11 Studierende erhielten beide Fallgeschichten mit allen Zusatzinformationen zur Bearbeitung. Jede einzelne Zusatzinformation wurde von den Studierenden hinsichtlich ihrer Entscheidungsrelevanz ("Wie wichtig ist diese Information für Ihre Entscheidung?") und ihrer Verständlichkeit ("Wie verständlich ist diese Information?") beurteilt. Die Entscheidungsrelevanz und Verständlichkeit der Zusatzinformationen wurden jeweils auf einer 10-stufigen Skala (1 = sehr unwichtig bis 10 = sehr wichtig bzw. 1 = sehr unverständlich bis 10 sehr verständlich) erhoben.

Tabelle 9 sind die gebildeten Gesamtmittelwerte aller Zusatzinformationen für beide Fallbeispiele hinsichtlich der Entscheidungsrelevanz und der Verständlichkeit zu entnehmen.

Tabelle 9: Pretest zur Auswahl des Fallbeispiels, Chorea Huntington vs. PID-HLA, Entscheidungsrelevanz und Verständlichkeit der Zusatzinformationen für beide Fallbeispiele (Gesamtmittelwerte und Standardabweichungen in Klammer).

Fallbeispiel	Chorea Huntington		PID-HLA	
	M	(SD)	M	(SD)
Entscheidungsrelevanz	5.24	(1.25)	6.70	(1.02)
Verständlichkeit	6.37	(1.87)	8.17	(1.20)

Anmerkung: Wertebereich von 1 niedrig bis 10 hoch.

Tabelle 9 verdeutlicht, dass die Zusatzinformationen des Fallbeispiels zur PID-HLA im Durchschnitt als deutlich entscheidungsrelevanter ($M = 6.70$, $SD = 1.02$) und verständlicher ($M = 8.17$, $SD = 1.20$) eingestuft wurden als beim Fallbeispiel zur Chorea Huntington ($M = 5.24$, $SD = 1.25$ bzw. $M = 6.37$, $SD = 1.87$).

Aufgrund der höheren Entscheidungsrelevanz und der besseren Verständlichkeit der Zusatzinformationen wurde für diese Untersuchung das Fallbeispiel zur PID ausgewählt.

Abbildung 15 gibt die für diese Untersuchung ausgewählte Dilemmageschichte „Margot Kreidler“ zum Thema PID-HLA in transkribierter Form wieder:

Margot Kreidler, 36 Jahre, möchte ein „Retterkind“

Unser Sohn Max ist fünf und hat die Fanconi-Anämie. Uns war aufgefallen, dass er irgendwie immer öfter krank wurde und schwächer wirkte als früher, und da sind wir zum Arzt gegangen. Und nun wissen wir sicher, er hat diese angeborene Bluterkrankung und seine Knochenmarkszellen bauen sich langsam immer weiter ab. Da kann man zwar Medikamente geben, aber richtig geheilt wird er dadurch nie.

Aber wenn er einen ganz genau passenden Knochenmarksspender bekäme, könnte er ganz gesund werden! Aber bis sich da mal einer findet! Mein Mann oder ich passen nicht gut genug, sonst hätten wir das natürlich gleich gemacht.

Jetzt haben wir uns gedacht, wo wir uns sowieso noch ein Kind wünschen, wäre es doch für alle am besten, wenn wir eine künstliche Befruchtung machen und am Embryo noch in der Schale testen lassen würden, welche Embryos am besten zu unserem Max passen würden.

In Deutschland darf man so was ja nicht machen, das hat mir der Arzt auch gleich gesagt, aber in Amerika gab es das schon, da wurden Kinder mit Fanconi-Anämie mit dem Nabelschnurblut von ihrem ausgesuchten Geschwisterchen völlig geheilt!

Sollte ich nicht auch nach Amerika gehen für so eine Präimplantationsdiagnostik? Würden Sie das an meiner Stelle tun?

Abbildung 15: Für die vorliegende Studie verwendetes Fallbeispiel zur PID-HLA, Transkript des Films.

In dem für diese Studie verwendeten Dilemma wurden die Versuchspersonen mit der Situation einer Mutter konfrontiert, die dringend einen Gewebespender für ihr an Fanconi-Anämie erkranktes Kind benötigt. Da kein lebender Familienteil als Spender in Frage kommt und sich keine passende Fremdspende finden lässt, besteht die Möglichkeit mittels einer PID-HLA einen passenden Gewebespender zur Heilung ihres Kindes erzeugen. Allerdings ist eine PID zur Erzeugung eines „Retterkindes“ in Deutschland gesetzlich verboten, könnte aber legal in Amerika durchgeführt werden.

In diesem Dilemma stehen lediglich zwei Handlungsoptionen zur Verfügung: Frau Kreidler kann entweder nach Amerika für eine PID-HLA gehen oder die Durchführung des Verfahrens ablehnen und versuchen, durch eine natürliche Schwangerschaft einen geeigneten Spender zu erzeugen bzw. auf konventionellem Wege weiterhin nach einem Spender suchen. Das in dieser Studie verwendete Beispiel lehnt sich im Großen und

Ganzen an den in Kapitel 2.2.3 beschriebenen Fall Adam Nash an, der als erstes „Retterkind“ in den USA geboren wurde. Die Länge des personalisierten Films beträgt 2 Minuten und 31 Sekunden. Die Fallgeschichte enthält in kurzer und prägnanter Form die wichtigsten deskriptiven Informationen, die für oder gegen die Durchführung einer PID-HLA sprechen und zur gesetzlichen Regelung in Deutschland.

Die Abbildungen 16 und 17 geben Bilder des *talking head* von dem für diese Studie verwendeten Fallbeispiel „Margot Kreidler“ aus frontaler und seitlicher Perspektive wieder. Abbildung 18 zeigt den Touchscreen-Monitor zur Bedienung der Lernumgebung.



Abbildung 16: Frontalansicht *talking head* „Margot Kreidler“ (Foto: Deutsches Museum).



Abbildung 17: Seitliche Ansicht *talking head* „Margot Kreidler“ (Foto: Deutsches Museum).



Abbildung 18: *talking head* „Margot Kreidler“, unten Touchscreen-Monitor zur Bedienung der Lernumgebung (Foto: Deutsches Museum).

Zunächst erforderte die Verwendung lediglich eines Fallbeispiels die Integration einer neuen Startseite, die den Aufruf des personalisierten Films durch die Versuchsperson ermöglicht. Die Aufforderung zum Start des Fallbeispiels wurde in der Instruktion gegeben („Bitte wenden Sie sich nun der Dialogstation zu und berühren Sie auf dem Touchscreen die Person Margot Kreidler!“ Siehe Kapitel 8.5 zum Versuchsablauf).

Folgende Abbildung 19 gibt einen Screenshot der Startseite zum Aufruf des Falls „Margot Kreidler“ wieder. Durch Berühren der Person bzw. des umgebenden grauen Feldes konnte der personalisierte Film von der Versuchsperson gestartet werden. Der Zeitpunkt des Aufrufs des Films wurde automatisch und sekundengenau durch die Lernumgebung per Logfile-Recording protokolliert.

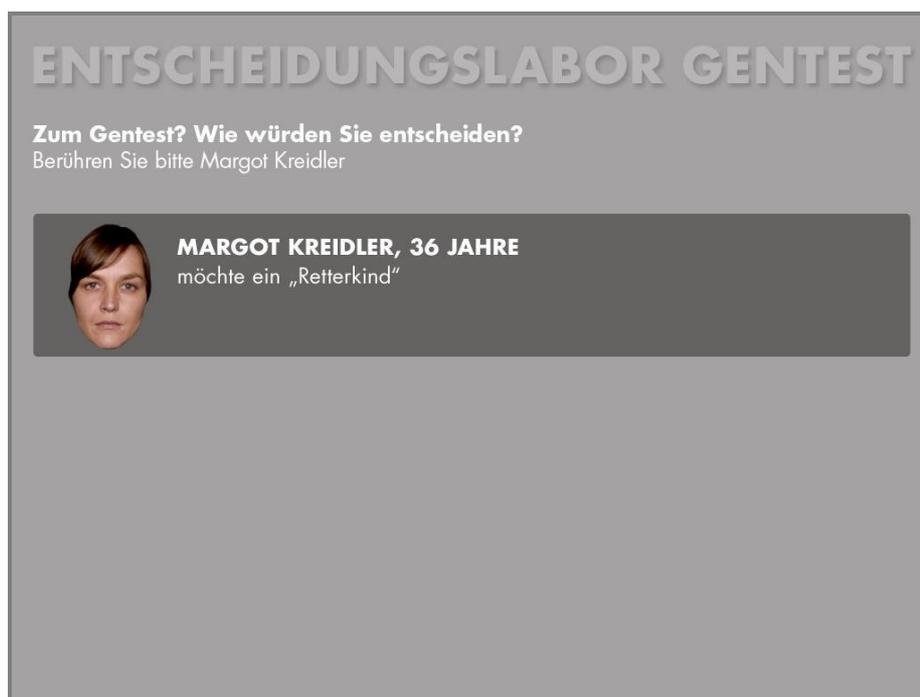


Abbildung 19: Startseite zum Aufruf des personalisierten Films (Screenshot Touchscreen).

Eine weitere, zusätzliche Modifikation der Lernumgebung stellte die Integration zusätzlicher Bildschirmseiten zur Erfassung von abhängigen Variablen während der Bearbeitung der Lernumgebung dar. In der ursprünglichen Fassung der Lernumgebung ist nach Präsentation des personalisierten Films bereits eine Aufforderung zur aktiven Positionierung vorhanden, die den ersten Messzeitpunkt (t_1) bildete (siehe folgende Abbildung 20). Jeweils eine weitere Aufforderung zur aktiven Positionierung wurde nach der Feedbackinformation (Messzeitpunkt t_2) und nach den Zusatzinformationen (Messzeitpunkt t_3) zusätzlich eingefügt. Demzufolge mussten die Probanden insgesamt drei Entscheidungen treffen: zwei vorläufige und eine finale.

Durch Betätigen des Ja- oder Nein-Buttons konnte der Proband seine persönliche, hypothetische Entscheidung treffen, ob er anstelle der Person für eine PID nach Amerika gehen würde oder nicht. Mit dem Aufruf des Ja- oder Nein-Buttons wurde das Entscheidungsverhalten als dichotome Variable automatisch durch die Lernumgebung per Logfile-Recording registriert.

Zur Erfassung weiterer Variablen zu den drei Messzeitpunkten wurde nach jeder der drei zu treffenden Entscheidungen eine Bildschirmseite mit einer Aufforderung zur Bearbeitung eines Frageblocks eingefügt. Insgesamt waren drei Kurzfragebögen zu bearbeiten. Alle drei Frageblöcke können im Anhang eingesehen werden und werden im Versuchsablauf in Kapitel 8.5 noch genauer beschrieben, die Kapitel 8.6 und 8.7 stellen die eingesetzten Instrumente vor. Zur Navigation zwischen den einzelnen Bildschirmseiten wurden innerhalb der Lernumgebung zusätzliche Weiter-Buttons integriert, deren Aufruf jeweils zur

nächsten Bildschirmseite führte. Abbildung 20 zeigt einen Screenshot der Bildschirmseite mit der Aufforderung zur aktiven Positionierung durch die Lernumgebung:

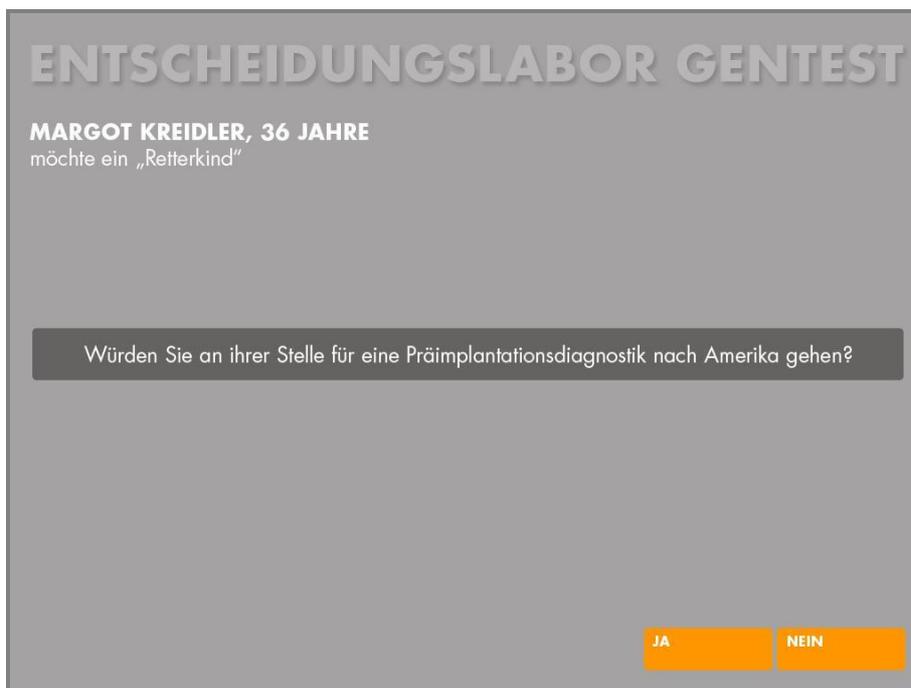


Abbildung 20: Aufforderung zur aktiven Positionierung durch die Lernumgebung zur Erhebung der getroffenen Entscheidung per Logfile-Recording (Screenshot Touchscreen).

Abbildung 21 gibt exemplarisch die Bildschirmseite der Aufforderung zur Bearbeitung des Frageblocks 1 wieder, die nach Betätigen des Ja- oder Nein-Buttons erschien:

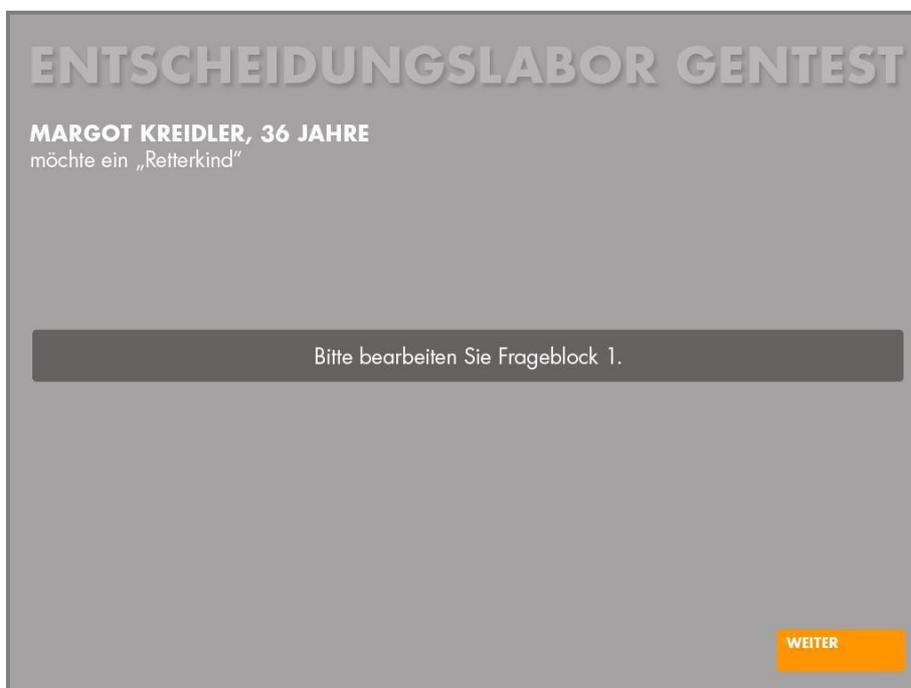


Abbildung 21: Aufforderung zur Bearbeitung von Frageblock 1 durch die Lernumgebung (Screenshot Touchscreen).

Am Ende jeden Frageblocks, nachdem alle Fragen beantwortet wurden, wurde der Proband aufgefordert, den Weiter-Button auf dem Touchscreen zu betätigen, um die Bearbeitung der Lernumgebung fortzusetzen („Wenn Sie alle Fragen beantwortet haben, drücken Sie bitte den orangen Weiter-Button auf dem Touchscreen der Dialogstation!“, vgl. Anhang).

Nach Präsentation des personalisierten Films und des Feedbacks der anderen Studienteilnehmer wurden alle Versuchsteilnehmer mit sechs zusätzlichen Informationen in Textform konfrontiert, die weitere Argumente für oder gegen die Durchführung einer PID-HLA umfassten (vgl. Anhang). Die Zusatzinformationen waren in drei Menüpunkte untergliedert:

1. *Fakten*: Unter diesem Menüpunkt befand sich eine sachliche Information zum sog. „PID-Tourismus“.
2. *Bewertung*: Der Menüpunkt Bewertung gab insgesamt vier ethische Stellungnahmen bzw. Argumente von Experten wie der Nationale Ethikkommission im Bereich der Humanmedizin der Schweiz oder dem Nationalen Ethikrat wieder.
3. *Gesetze*: Dieser Menüpunkt gab vergleichende Informationen zur Gesetzeslage zur PID in unterschiedlichen Ländern: Deutschland, Frankreich und Amerika.

Abbildung 22 gibt exemplarisch eine Zusatzinformation des Menüpunkts Bewertung wieder, die eine Contra-Position zur PID-HLA der Ethik-Kommission NEK-CNE (2007) unterbreitet (siehe Kapitel 2.4.3):

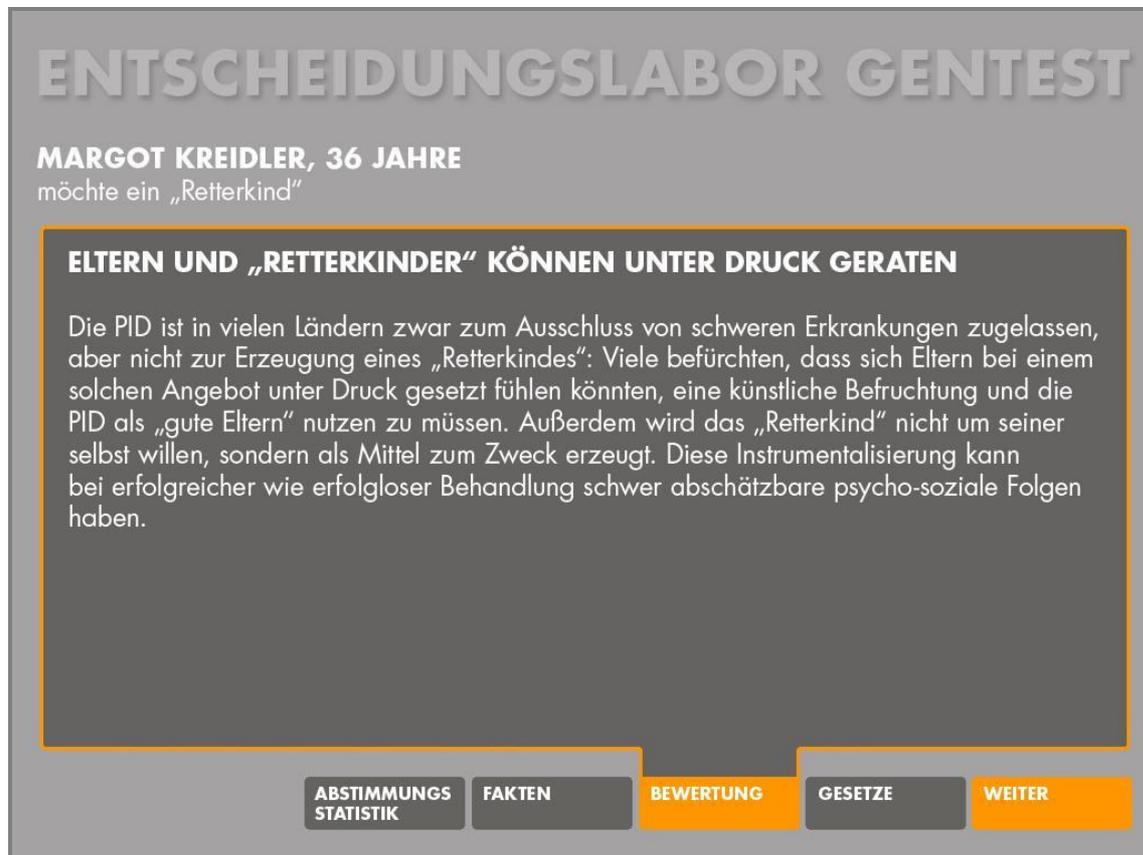


Abbildung 22: Beispiel Zusatzinformation Bewertung, Contra-Position der Nationalen Ethikkommission der Schweiz zur PID-HLA (Screenshot Touchscreen).

Wie Abbildung 22 zu entnehmen, wies jede Zusatzinformationen eine Überschrift auf, welche die Kernaussage der Information repräsentierte. Aus diesen Kernaussagen ging hervor, ob die jeweilige Zusatzinformation für oder gegen eine PID-HLA spricht.

Zum Abschluss der Bearbeitung der Lernumgebung, d. h. nach dem vollständigen Ausfüllen des letzten Frageblocks 3, war ein Ende-Button integriert, dessen Aufruf die Untersuchung mit dem Wortlaut: „Vielen Dank für Ihre Mitarbeit!“ beendete.

8.4 Operationalisierung der unabhängigen Variablen

8.4.1 Messzeitpunkt

Den “within-subject“ - Faktor bildete der Messzeitpunkt (t1: Film, t2: Feedback, t3: Informationen). Die Lernumgebung beinhaltete einen personalisierten Film, eine Feedbackinformation und zusätzliche Informationen in Textform. Um den getrennten Einfluss des personalisierten Films, der Feedbackinformation und der Zusatzinformationen auf die abhängigen Variablen zu erfassen, wurde eine dreimalige Wiederholungsmessung realisiert. Mit Hilfe dieser Messwiederholung sollten mögliche Konfundierungseffekte, die eine kausale Interpretation der Ergebnisse erschweren könnten, vermieden werden (vgl. Bortz & Döring, 2006). Abbildung 23 gibt zusammenfassend in einem Zeitstrahl einen Überblick über die drei im Rahmen des Messwiederholungsdesigns implementierten Erhebungszeitpunkte und Frageblöcke:

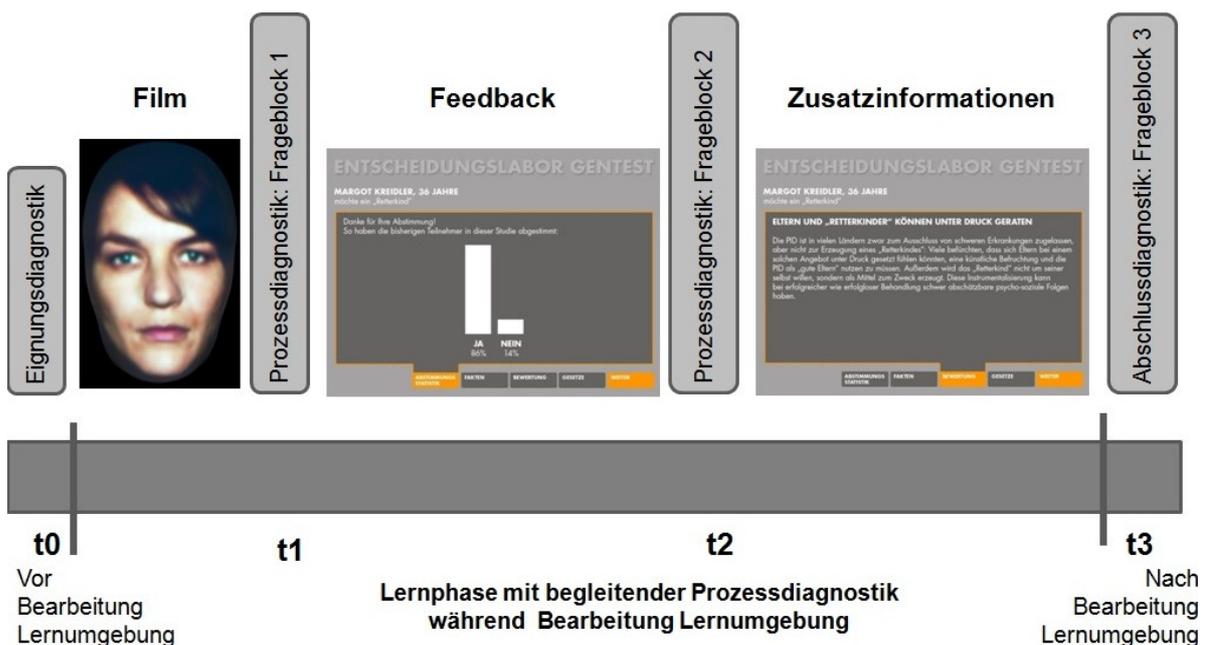


Abbildung 23: Überblick über die drei implementierten Messzeitpunkte und Frageblöcke (Zeitstrahl).

Vor der Bearbeitung der Lernumgebung wurden zunächst im Rahmen einer Eignungsdiagnostik (Messzeitpunkt t_0) individuelle Lernvoraussetzungen der Probanden erhoben.

Im Rahmen einer begleitenden Prozessdiagnostik während der Bearbeitung der Lernumgebung und einer Abschlussdiagnostik nach Bearbeitung der Lernumgebung wurden die Probanden aufgefordert, drei Kurzfragebögen (Frageblock 1 bis 3) zu bearbeiten.

Der erste Messzeitpunkt (t_1) erfolgte direkt nach der Präsentation des personalisierten Films via *talking head* und der ersten getroffenen Entscheidung. Als abhängige Variablen wurden die Richtung der Entscheidung (dichotom: Ja vs. Nein) und in Frageblock 1 die Entscheidungssicherheit, Variablen zur Medienfigur und zur Manipulationskontrolle sowie die Argumentationsqualität erhoben.

Unmittelbar nach Erhalt der Feedbackinformation und der zweiten, vorläufigen Entscheidung wurde der zweite Messzeitpunkt (t_2) eingefügt. Als abhängige Variablen wurden bei allen drei Gruppen die Richtung der Entscheidung (dichotom: Ja vs. Nein) und in Frageblock 2 die Entscheidungssicherheit sowie Variablen zur Manipulationskontrolle erfasst.

Die dritte und letzte Messung (t_3) erfolgte nach der vollständigen Bearbeitung der Lernumgebung bzw. der Zusatzinformationen und der final getroffenen, endgültigen Entscheidung. In dieser Abschlussdiagnostik wurden analog zum ersten Messzeitpunkt als abhängige Variablen die Richtung der Entscheidung (dichotom: Ja vs. Nein) und in Frageblock 3 die Entscheidungssicherheit und die Argumentationsqualität erhoben.

Alle drei Frageblöcke werden im Versuchsablauf in Kapitel 8.5 beschrieben, die eingesetzten Instrumente können Kapitel 8.6 und 8.7 sowie dem Anhang entnommen werden.

8.4.2 Feedback

Das Feedback wurde als “between-subjects“ - Faktor variiert (Stufen: kongruent vs. inkongruent vs. kein). Die 72 Probanden wurden zufällig einer der drei experimentellen Bedingungen zugewiesen. Das Feedback wurde nach der ersten, vorläufigen Entscheidung und der Bearbeitung des Frageblocks 1 gegeben. Es bestand aus einer sozialen Vergleichsinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung der anderen Versuchsteilnehmer. Aus dieser Konsensinformation wird ersichtlich, wie sich die anderen Studienteilnehmer insgesamt im Durchschnitt entschieden haben.

24 Probanden der Experimentalgruppe 1 (EG 1) erhielten ein kongruentes Feedback, in dem die eigene Entscheidung von der Mehrheit (86 %) der anderen Studienteilnehmer geteilt wurde, lediglich eine Minderheit (14 %) hatte sich anders entschieden. Bei der Lern-

bedingung „kongruentes Feedback“ stimmte folglich die Entscheidung der Mehrheit der anderen Studienteilnehmer mit der eigenen Entscheidung überein.

24 Probanden der Experimentalgruppe 2 (EG 2) wurden hingegen mit einem inkongruenten Feedback konfrontiert, bei der die eigene Entscheidung lediglich von einer Minderheit (14 %) vertreten wurde. Die Mehrheit von 86 % der anderen Studienteilnehmer traf eine andere Entscheidung. Demzufolge stand bei der Lernbedingung „inkongruentes Feedback“ die durchschnittliche Gesamtentscheidung der anderen Studienteilnehmer im Konflikt zur eigenen Entscheidung.

24 Probanden erhielten kein Feedback und bildeten die Kontrollgruppe. Sie erhielten lediglich eine Information über die stattgefundene Abstimmung der anderen Studienteilnehmer mit dem Hinweis, dass das Ergebnis dieser Umfrage leider noch nicht feststeht.

Bei beiden Experimentalgruppen wurde die Konsensinformation in Abhängigkeit von der eigenen Entscheidung manipuliert; variiert wurde die Kongruenz der Konsensinformation mit der eigenen Entscheidung. Experimentalgruppe 1 erhielt eine entscheidungskongruente Konsensinformation, Experimentalgruppe 2 hingegen eine entscheidungskongruente Konsensinformation. Bei Kongruenz wurde die eigene Entscheidung von einer Mehrheit, bei Inkongruenz dagegen von einer Minderheit vertreten.

Abbildung 24 zeigt exemplarisch einen Screenshot der Konsensinformation, die ein Proband der Bedingung „inkongruentes Feedback“ erhielt, nachdem er sich gegen eine PID („Nein“) entschieden hat:

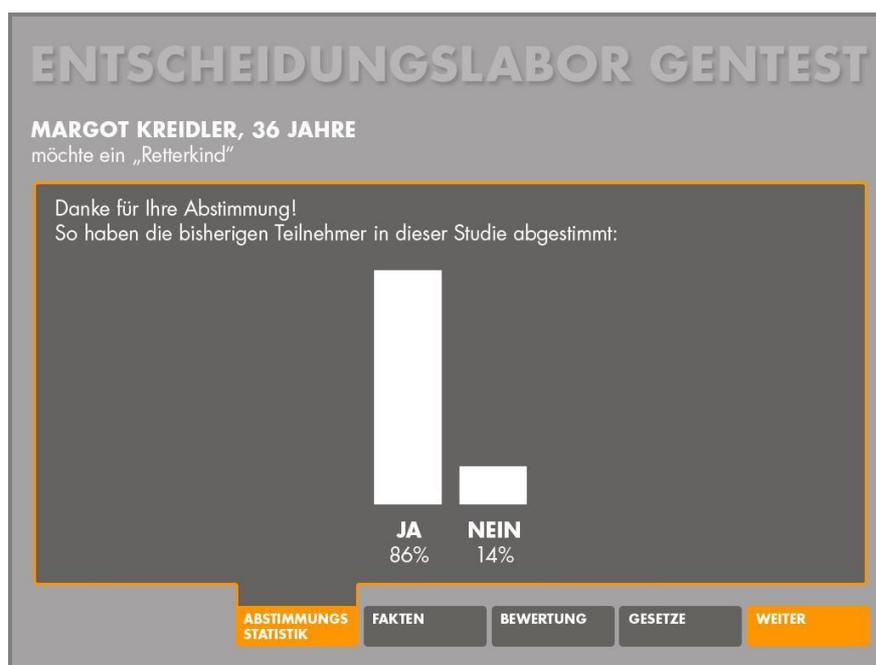


Abbildung 24: Beispiel Konsensinformation nach „Nein-Entscheidung“ für Bedingung „inkongruentes Feedback“ (Screenshot Touchscreen).

Abbildung 24 macht deutlich, dass bei inkongruentem Feedback der Proband, nachdem er sich gegen eine PID entschieden hat („Nein“) mit einer entscheidungsinkongruenten Mehrheitsposition von 86 % konfrontiert wurde.

Die verwendeten Prozentzahlen der Mehrheits- und Minderheitsposition (86 % vs. 14 %) waren fiktiv und wurden einer Studie von Erb et al. (2006) zur numerischen Stärke von Einflussgruppen entnommen. Als Vergleichsgruppe wurden der ursprünglichen Konzeption der Lernumgebung entsprechend und in Anlehnung an die Ähnlichkeitshypothese von Festinger (1954) die anderen Studienteilnehmer verwendet und somit eine niedrige soziale Distanz realisiert.

Im Rahmen einer formativen Evaluationsstudie zu einem Prototypen der Dialogstation „Gentest“ (siehe Kapitel 5.6) konnte anhand einer Logfile-Analyse festgestellt werden, dass sich 59 % (109) Besucher für eine PID („Ja“) und 41 % (76) gegen eine PID („Nein“) entschieden hatten (Hänle, 2008). Demzufolge fiel das in der Feldstudie ermittelte Ausmaß an Meinungs- bzw. Entscheidungsübereinstimmung (der Konsens) deutlich geringer aus als das in dieser Studie verwendete Abstimmungsergebnis, in der die Mehrheitsposition deutlich größer war als die Minderheitsposition.

8.5 Versuchsablauf

Das Experiment fand in Einzelsitzungen statt, die zwischen 37 und 74 Minuten dauerten. Zur Sicherung der ökologischen Validität der Untersuchung wurde für die Bearbeitung der Lernumgebung keine Lernzeit festgelegt. Die Untersuchung fand im Deutschen Museum im Raum des ehemaligen Besucherlabors zur Genforschung (DNA-Besucherlabor) statt.

Die Studie bestand aus zwei Untersuchungsabschnitten:

1. Der Eingangsdagnostik zur Erfassung der Lernvoraussetzungen der Probanden, die zwischen 10-15 Minuten dauerte.
2. Der Lernphase mit begleitender Prozessdiagnostik während der Bearbeitung der Lernumgebung sowie der Abschlussdiagnostik nach Bearbeitung der Lernumgebung. Die Dauer für den zweiten Untersuchungsteil lag zwischen 13 und 43 Minuten.

Zu Beginn des Experiments wurde den Probanden die Lernumgebung von der Versuchsleiterin kurz vorgestellt. Den Probanden wurde mitgeteilt, dass das Ziel der Untersuchung im Nachweis der Lernwirksamkeit der Lernumgebung besteht. Danach wurde der Ablauf der gesamten Untersuchung erklärt. Zusätzlich wurden die Probanden gefragt, ob andere, bisherige Versuchsteilnehmer etwas vom Sinn und Zweck dieser Studie erzählt haben und

wenn, ja, was diejenige Person erzählt hatte. Da keiner diese Kontrollfrage bejahte, konnten alle Probanden an der Studie teilnehmen.

Abbildung 25 gibt einen Überblick über den gesamten Ablauf der Erhebung:



Anmerkungen: * dreimalige Messwiederholung (t1-t3); ** zweimalige Messwiederholung (t1, t3).

Abbildung 25: Überblick über den gesamten Versuchsablauf des Experiments.

Eingangsdagnostik (Messzeitpunkt t0). Im ersten Untersuchungsteil, der etwa zwischen 10 und 15 Minuten dauerte, wurde mit Hilfe eines Pretests das themenspezifische Vorwissen zur Humangenetik erhoben. In einem Fragebogen wurden mittels Ratingskalen weitere kognitive Lernvoraussetzungen wie das subjektive Vorwissen über Gentests, zur PID und über Bioethik erfasst. Zusätzlich wurden motivationale und einstellungsbezogene Lernvoraussetzungen wie der persönliche Bezug, das thematische Interesse, die Einstellungen zu Gentests, die Präferenz für Deliberation und Intuition sowie die soziale Vergleichsorientierung und Orientierung an anderen erhoben. Der Fragebogen zur Eingangsdagnostik endete mit einem Teil zu soziodemographischen Angaben.

Lernphase und Prozessdiagnostik (Messzeitpunkte t1 und t2). Im Anschluss daran befassten sich die Probanden im zweiten Untersuchungsteil selbstständig mit der Lernumgebung. In einer Instruktion wurden den Probanden Informationen zur Aufgabenstellung und zum Umgang mit der Lernumgebung gegeben. Die Instruktion zur Bearbeitung des zweiten

Untersuchungsteils hing gut sichtbar rechts neben dem Touchscreen der Medienstation, so dass die Probanden jederzeit darauf zurückgreifen konnten und lautete wie folgt:

Liebe(r) Versuchsteilnehmer(in),

In der folgenden Dialogstation wird Ihnen ein Fallbeispiel zum Thema genetische Untersuchung/Präimplantationsdiagnostik präsentiert.

Ihre Aufgabe ist es, insgesamt **dreimal** eine Entscheidung darüber zu treffen, ob Sie anstelle der Person eine Präimplantationsdiagnostik (PID) machen würden oder nicht. Zudem sollen Sie immer so ausführlich wie möglich **Ihre Entscheidung begründen.**

Zusätzlich zum Fallbeispiel sollen Ihnen weitere Informationen bei der Entscheidungsfindung helfen. Bitte lesen Sie sich diese Zusatzinformationen sehr gründlich durch.

Zur Bedienung der Dialogstation drücken Sie bitte stets die orange unterlegten Buttons am rechten unteren Rand des Touchscreens, alle grau unterlegten Buttons sind inaktiv.

Bitte stellen Sie sich darauf ein, dass Sie während der Untersuchung direkt durch die Dialogstation aufgefordert werden, insgesamt **3 Frageblöcke** schriftlich auszufüllen. Diese Frageblöcke befinden sich rechts auf dem Tisch neben der Dialogstation in einer Mappe. Nachdem Sie alle Fragen eines Blockes beantwortet haben, drücken Sie bitte stets den orangen Weiter-Button auf dem Touchscreen der Dialogstation.

Falls Sie noch irgendwelche Fragen haben, melden Sie sich bitte bei der Versuchsleiterin!

Bitte wenden Sie sich nun der Dialogstation zu und berühren Sie auf dem Touchscreen die Person Margot Kreidler!

Abbildung 26: Instruktion zur Bearbeitung der Lernumgebung im zweiten Untersuchungsteil.

Die Instruktion macht deutlich, dass die Versuchspersonen mit einer fiktiven Entscheidungssituation konfrontiert wurden, in der sie im Verlauf der Untersuchung insgesamt drei Entscheidungen zwischen jeweils zwei Handlungsalternativen treffen sollen. Die Revision der Entscheidung wurde demzufolge angekündigt.

Während der Bearbeitung der Lernumgebung füllten die Probanden eine auf einem Tisch bereitliegende Mappe mit drei Frageblöcken aus. Es wurde darauf hingewiesen, dass erst nach expliziter Aufforderung durch die Lernumgebung der entsprechende Frageblock zu bearbeiten war (z. B. „Bitte bearbeiten Sie Frageblock 1“; siehe Kapitel 8.3, Abbildung 21).

Um ein versehentliches Weiterblättern der Probanden zu verhindern, wurden die drei Frageblöcke in der Mappe durch orangefarbene Zwischenseiten abgetrennt. Am Ende jeder der drei Frageblöcke wurde eine Aufforderung zur Fortsetzung der Bearbeitung durch die Lernumgebung gegeben („Wenn Sie alle Fragen beantwortet haben, drücken Sie bitte den orangen Weiter-Button auf dem Touchscreen der Dialogstation!“).

Frageblock 1 und Frageblock 3 waren dabei für alle drei Lernbedingungen identisch. Eine Ausnahme bildete Frageblock 2 zur Erfassung des Dissonanzerlebens bzw. des psychologischen Unbehagens und des wahrgenommenen Konflikts bei beiden Experimentalgruppen nach Erhalt des Feedbacks (vgl. Anhang).

Für diese Untersuchung war es sehr wichtig, das Dissonanzerleben bezüglich des Feedbacks unmittelbar vor der zweiten, vorläufigen Entscheidung zu erheben, da beispielsweise durch eine Revision der Entscheidung bereits eine Dissonanzreduktion stattfinden könnte. Aus diesem Grunde wurde bei beiden Experimentalgruppen das psychologische Unbehagen direkt nach Erhalt des Feedbacks erfasst. Durch die zeitnahe Messung des Dissonanzerlebens vor der Entscheidung konnte eine Verfälschung der Daten infolge einer Dissonanzreduktion ausgeschlossen werden.

Prozessdiagnostik (Messzeitpunkt t1, Frageblock 1). Im Anschluss an die erste getroffene Entscheidung gaben die Probanden auf einer Ratingskala ihre Sicherheit der Entscheidung an und hatten in einem Essay die Möglichkeit, ihre Entscheidung ausführlich zu begründen. Darauf folgend bearbeiteten die Probanden Ratingskalen, in denen die soziale Projektion der eigenen Entscheidung, der selbst erzeugte Konsens, das Bedürfnis nach sozialem Vergleich, die parasozialen Interaktionen mit der Medienfigur „Margot Kreidler“ sowie deren Attraktivität abgefragt wurden.

Prozessdiagnostik (Messzeitpunkt t2, Frageblock 2). Nach der zweiten getroffenen Entscheidung mussten alle drei Lernbedingungen erneut ihre Entscheidungssicherheit auf einer Ratingskala angeben. Um zu überprüfen, ob die Manipulation der unabhängigen Variable Feedback erfolgreich war, erhielten zur Manipulationskontrolle beide Experimentalgruppen unmittelbar nach Erhalt des Feedbacks, d. h. vor der zweiten, vorläufigen Entscheidung, zusätzliche Ratingskalen zum psychologischen Unbehagen und zum wahrgenommenen Konflikt zur Bearbeitung (Frageblock 2a). Um mögliche Ursachen einer Entscheidungsrevision eruieren zu können, wurden außerdem beide Experimentalgruppen gebeten, in einem Essay kurz ihre Entscheidung zu begründen (Frageblock 2b).

Abschlussdiagnostik (Messzeitpunkt t3, Frageblock 3). Den Abschluss des zweiten Untersuchungsteils bildete die Abschlussdiagnostik. Nach Bearbeitung der Zusatzinformationen trafen die Probanden ihre endgültige und letzte Entscheidung. Analog zu Frageblock 1 wurden die Entscheidungssicherheit mittels einer Ratingskala und die Argumentationsqualität bei allen drei Lernbedingungen in einem Essay erhoben. Zusätzlich wurden alle Probanden anhand einer offenen Frage aufgefordert, Gegenargumente zu ihrer endgültigen Position/Entscheidung zu nennen. Danach füllten alle Probanden einen Verdachtsfragebogen aus (*suspicion check*) aus.

Frageblock 3 endete mit der Aufforderung: “Wenn Sie alle Fragen beantwortet haben, ist die Untersuchung beendet. Drücken Sie bitte den orangen Ende-Button auf dem Touchscreen der Dialogstation. Geben Sie der Versuchsleiterin bitte ein Zeichen!“.

Probandenaufklärung. Nach Bearbeitung der Lernumgebung und der Abschlussdiagnostik erfolgte eine mündliche Aufklärung aller Probanden über das eigentliche Untersuchungsziel. Den Probanden wurde mitgeteilt, dass in diesem Experiment der Einfluss des Feedbacks anderer auf das Entscheidungsverhalten und die Argumentationsqualität untersucht wurde. Ferner wurde den Teilnehmern der beiden Experimentalgruppen eröffnet, dass das Feedback über die durchschnittliche Entscheidung der anderen Studienteilnehmer fiktiv war und in Abhängigkeit von der eigenen Entscheidung manipuliert wurde (kongruent vs. inkongruent). Aus ethischen Gründen kam der Probandenaufklärung in dieser Studie ein besonderer Stellenwert zu, da beide Experimentalgruppen mit einer fingierten Feedbackinformation konfrontiert und damit getäuscht wurden (vgl. Zimbardo & Gerrig, 2008).

Zusätzlich wurden alle Probanden gebeten, während des Zeitraums der Datenerhebung Stillschweigen über den Sinn und Zweck der Untersuchung, insbesondere über das manipulierte Feedback, gegenüber Bekannten und Kommilitonen zu bewahren. Den Abschluss des Experiments bildete die Auszahlung des Probandenhonorars von 10 Euro in bar; die Probanden wurden verabschiedet und mit Dank entlassen.

Versuchsleitung. Die Versuche wurden jeweils von der Autorin dieser Arbeit und einer Studentin der Geschichte der Naturwissenschaften und Technik geleitet. Die Studentin wurde vor Beginn der Untersuchung ausführlich in die objektive Durchführung des Experiments eingewiesen und durch die Autorin während des Untersuchungszeitraums supervidiert. Ferner lag ihr eine schriftliche Instruktion für die Versuchsdurchführung vor.

8.6 Instrumente

Alle Messungen vor, während und nach der Bearbeitung der Lernumgebung wurden in Form von papierbasierten Fragebögen realisiert. Eine Ausnahme bildeten die Lernzeit und das Entscheidungsverhalten. Für die Messung dieser Variablen kam das Verfahren der Logfile-Analyse zur Anwendung. Innerhalb der verwendeten Softwareumgebung wurde jeder Zugriff auf alle Buttons automatisch und sekundengenau mit protokolliert. Mit den so erstellten Logfiles konnte die Lernzeit berechnet und das Entscheidungsverhalten gemessen werden.

Während der Bearbeitung der Lernumgebung wurden mit Hilfe von drei Fragebögen bzw. Frageblöcken in einer begleitenden Befragung quantitative und qualitative Daten zu drei

Messzeitpunkten erhoben. Alle drei Frageblöcke waren in einer bereitliegenden Mappe enthalten (vgl. Anhang).

Die Fragebögen wurden als Papier-Bleistift-Version und nicht auf dem Touchscreen der Medienstation dargeboten, um eine mögliche Anwendung von Höflichkeitsregeln im Hinblick auf die Bewertung der Medienfigur zu vermeiden (siehe Kapitel 4.2.2; Media Equation Theory; Nass et al., 1999; Reeves & Nass, 1996).

Alle in dieser Untersuchung verwendeten Ratingskalen wurden auf der Basis bereits existierender und erprobter Instrumente konstruiert. Die folgenden Untersuchungsaspekte wurden im Regelfall auf einer 5-stufigen Likert-Skala von 1 (trifft nicht zu) bis 5 (trifft zu) erhoben. Abweichungen werden bei der anschließenden Darstellung der Skalen genannt.

8.6.1 Individuelle Lernvoraussetzungen

Vor der Bearbeitung der Lernumgebung wurden in Rahmen einer Eingangsdiagnostik (t_0) kognitive, motivationale und einstellungsbezogene Lernvoraussetzungen erfasst. Zusätzlich wurden demographische Daten wie das Alter, Geschlecht, Studienfach und die Semesterzahl erhoben.

Da in repräsentativen Studien festgestellt wurde, dass das Wissen über die PID in der Allgemeinbevölkerung nur sehr gering ist (vgl. Brähler & Stöbel-Richter, 2004; siehe Kapitel 2.5), erschien eine Adaption der in dieser Studie verwendeten Ratingskalen zur Erhebung des persönlichen Bezugs, zum Interesse und den Einstellungen zum Inhaltsgebiet der Lernumgebung, der PID, als nicht sinnvoll. Stattdessen wurden diese drei Variablen, da die PID einen mit genetischen Untersuchungen verknüpften Anwendungsbereich darstellt, auf das übergeordnete Inhaltsgebiet genetische Untersuchungen adaptiert.

8.6.1.1 Kognitive Lernvoraussetzungen

Im Hinblick auf kognitive Lernvoraussetzungen wurde das Vorwissen über Humangenetik und das subjektive Vorwissen zu genetischen Untersuchungen, zur PID und Bioethik erfasst.

Vorwissen über Humangenetik. Als Instrument zur Bestimmung des allgemeinen Wissens über Humangenetik wurde der Genetische Wissensindex (GeWi) von Berth et al. (2004) verwendet. Der GeWi setzt sich aus zwölf Aussagen zusammen, die ausschließlich humangenetische Inhalte thematisieren (z. B. „Einige genetische Erkrankungen kommen in bestimmten ethnischen Gruppen häufiger vor.“). Jedes dieser zwölf Items kann entweder

mit „richtig“ oder „falsch“ beantwortet werden. Aus der Summe der richtig beantworteten Aufgaben wird ein Gesamtscore gebildet (theoretisches Maximum: 12 Punkte). Je höher der erreichte Summenwert ist, umso ausgeprägter ist das Wissen über Genetik. Der GeWi stellt laut Autoren ein objektives, reliables, valides und aufgrund seiner kurzen Bearbeitungszeit von etwa 3 Minuten ein sehr ökonomisches Erhebungsinstrument dar.

Subjektive Einschätzung des Vorwissens zu genetischen Untersuchungen, zur PID und zur Bioethik. Um mögliche Unterschiede im subjektiven Vorwissen am Inhaltsgebiet der Lernumgebung zu erfassen, schätzten die Versuchspersonen ihren Wissensstand zu genetischen Untersuchungen, zur PID und zur Bioethik auf einer Skala von 1 (sehr wenig) bis 5 (sehr viel) ein.

8.6.1.2 Motivationale Lernvoraussetzungen

Hinsichtlich motivationaler Lernvoraussetzungen wurden der persönliche Bezug zu genetischen Untersuchungen, das Interesse am Thema Gentest, die Präferenz für Intuition und Deliberation, die soziale Vergleichsorientierung und die Orientierung an anderen erhoben.

Persönlicher Bezug zum Thema Gentest. In Anlehnung an Töpfer (2009) gaben die Versuchspersonen auf einer fünfstufigen Skala von 1 (keinen) bis 5 (sehr stark) ihren persönlichen Bezug zum Thema Gentest an (“Wie schätzen Sie Ihren persönlichen Bezug zum Thema Gentest ein? Zum Beispiel weil Sie privat oder sonst jemanden kennen, der von der Thematik betroffen ist?”).

Interesse am Thema Gentest. Das Interesse am Thema genetische Untersuchungen und den damit verbundenen ethischen Fragen wurde mit 5 Items erhoben. Diese Items stammen aus einer Studie von Todt & Götz (1998) bzw. Götz (2001) zum Interesse von Jugendlichen an Gentechnologie und wurden lediglich an das untersuchte Inhaltsgebiet adaptiert (z. B. “Ich möchte mehr darüber erfahren, welche Vorteile und welche Nachteile genetische Untersuchungen haben“, “Ich finde die Auseinandersetzung mit ethischen Problemen, die mit der Anwendung von genetischen Untersuchungen verbunden sind, spannend“). Die Reliabilität der Skala lag bei .70 (Cronbachs Alpha).

Präferenz für Deliberation und Intuition. Um individuelle Unterschiede in der Präferenz für kognitions- oder affektbasierte Entscheidungsstrategien zu erfassen, wurde das Inventar “Präferenz für Deliberation und Intuition“ von Betsch (2004) verwendet, das aus zwei unabhängigen Subskalen besteht. Die Präferenz für Deliberation, d. h. die Neigung, Entscheidungen aufgrund planvoller, bewusster Überlegungen zu treffen (z. B. “Bevor ich Entscheidungen treffe, denke ich meistens erst mal gründlich nach“), wurde mit 5 Items erfasst. Die interne Konsistenz der Gesamtskala war mit $\alpha = .82$ ausreichend reliabel. Die

Subskala zur Präferenz für Intuition umfasste ebenfalls aus 5 Items (Cronbachs Alpha = .66), die sich auf gefühlsbasierte Entscheidungsstrategien beziehen (z. B. “Bei meinen Entscheidungen spielen Gefühle eine große Rolle“).

Soziale Vergleichsorientierung. Eine weitere Skala erfasste die soziale Vergleichsorientierung der Versuchspersonen. Für diese Studie wurde die deutsche Übersetzung von Jonas & Mikula (2006) der *social comparison orientation* Skala von Gibbons & Buunck (1999) verwendet. Die Skala enthält 11 Items auf einer fünffach gestuften Likert-Skala, welche die Tendenz, sich mit anderen zu vergleichen, messen (z. B. “Ich bin nicht der Typ Mensch, der sich häufig mit anderen vergleicht“, „Es interessiert mich oft, was andere Leute in einer ähnlichen Situation wie meiner machen würden“). Die Reliabilität der Gesamtskala zur sozialen Vergleichsorientierung lag bei .84.

Selbstüberwachung (Orientierung an anderen). Orientierung an anderen stellt eine zentrale Dimension des Persönlichkeitsmerkmals Selbstüberwachung (*Self-Monitoring*, Snyder, 1979) dar. Die Skala setzte sich aus 5 Items zusammen, die aus der deutschen Übersetzung der Selbstüberwachungsskala von Collani & Stürmer (2007) übernommen wurden. Die Versuchspersonen mussten einschätzen, wie stark sie sich an anderen in sozialen Situationen orientieren (z. B. “Ich würde meine Meinungen (oder meine Art, Dinge zu tun) nicht ändern, nur um jemand anderen zufrieden zu stellen oder um ihm zu gefallen“, “Wenn ich unsicher bin, wie ich mich in einer sozialen Situation verhalten soll, suche ich im Verhalten anderer nach Hinweisen“). Die Reliabilität lag bei .60 (Cronbachs Alpha).

8.6.1.3 Einstellungen bezüglich genetischer Untersuchungen

Bei den Ratingskalen zu den *Einstellungen zu genetischen Untersuchungen* hatten die Lernenden zu den möglichen Vorteilen, Nachteilen und Befürchtungen in Bezug auf die molekulargenetische Diagnostik für erbliche Erkrankungen Stellung zu nehmen. Alle verwendeten Items zur Akzeptanz von genetischen Untersuchungen wurden aus einer deutschlandrepräsentativen Studie von Berth et al. (2002) übernommen.

Die Skala zu den *positiven Aspekten (Vorteilen)* von Gentests setzte sich aus 4 Items zusammen, die befürwortende Aussagen in Bezug auf genetische Untersuchungen enthielten (z. B. “Genetische Untersuchungen sind akzeptabel, weil alle das Recht haben, über ihre Gene Bescheid zu wissen und damit das eigene Leben und die Gesundheit beeinflussen können“), die Reliabilität war ausreichend ($\alpha = .60$).

Die Ratingskala zu den *negativen Aspekten (Nachteilen)* von genetischen Untersuchungen bestand aus 6 Items mit ablehnenden Aussagen (z. B. “Genetische Untersuchungen sind

nicht akzeptabel, weil die Ergebnisse zur Diskriminierung von Personen mit Krankheitsgenen führen könnten“), die Reliabilität der Skala war zufriedenstellend ($\alpha = .64$).

Die *Befürchtungen* in Bezug auf genetische Untersuchungen wurden mittels einer drei Items umfassenden Ratingskala erhoben, die den Missbrauch der Ergebnisse genetischer Untersuchungen für wissenschaftliche Zwecke, für Eugenik oder durch Dritte thematisieren (z. B. “Ich befürchte, dass genetische Untersuchungen zu Eugenik („Erbhygiene“) führen könnten“), die Reliabilität betrug .68 (Cronbachs Alpha).

8.6.2 Realisierte Lernzeit

Für die exakte Messung der individuellen Lernzeit jedes Probanden wurde das Verfahren der Logfile-Analyse verwendet. Innerhalb der Softwareumgebung wurde jeder Zugriff auf alle Buttons automatisch sekundengenau durch die Lernumgebung protokolliert. Mit den erstellten Logfiles konnte die Lernzeit aus der Differenz der Startzeit des Aufrufs des Fallbeispiels und dem Ende-Button, der nach dem Bearbeiten der Mappe auf dem Touchscreen Monitor der Lernumgebung zu betätigen war, berechnet werden. Bei der Berechnung der Lernzeit der beiden Experimentalgruppen wurde die Zeit für die Bearbeitung der zusätzlichen Ratingskalen zum psychologischen Unbehagen, zum wahrgenommenen Konflikt und des Kurzessays zur Begründung der zweiten, vorläufigen Entscheidung in Frageblock 2 berechnet und von der Gesamtlernzeit subtrahiert.

8.6.3 Manipulationskontrolle

Beim Manipulation Check wurde kontrolliert, ob die unabhängige Variable Feedback überhaupt in der vorgesehenen Weise in der Stichprobe realisiert wurde (Bortz & Döring, 2006; Perdue & Summers, 1986). Unmittelbar nach der Präsentation des Feedbacks (t_2) erhielten die Probanden der beiden Experimentalgruppen Ratingskalen zur Erfassung des psychologischen Unbehagens und zum wahrgenommenen Konflikt der Feedbackinformation zur Bearbeitung.

Psychologisches Unbehagen. Das psychische Unbehagen nach dem Feedback wurde als selbstberichteter, momentaner Gefühlszustand auf der Grundlage von 10 verschiedenen Adjektiven operationalisiert. In Anlehnung an das von Elliot & Devine (1994) entwickelte „Dissonanzthermometer“ und den Affektfragebogen PANAS-X (Watson & Clark, 1994) werden sowohl positive und negative Gefühle als auch der Zustand der Überraschung bezogen auf das Feedback thematisiert.

Die Ratingskala zum *positiven Affekt* bestand aus 4 Items (Cronbachs Alpha = .89), in denen die Versuchsteilnehmer ihre positiv erlebten Gefühle nach Erhalt des Feedbacks angeben sollten (z. B. „Als ich das Abstimmungsergebnis der anderen Studienteilnehmer erfahren habe, war ich erfreut“).

Der negative Affekt nach Feedbackgabe wurde mit drei Items erhoben (z. B. „Als ich das Abstimmungsergebnis der anderen Studienteilnehmer erfahren habe, war ich wie vor den Kopf geschlagen“). Die interne Konsistenz der Skala war mit $\alpha = .86$ hervorragend reliabel.

Der emotionale Zustand der *Überraschung* nach dem Feedback wurde mit drei Items erfasst (z. B. „Als ich das Abstimmungsergebnis der anderen Studienteilnehmer erfahren habe, war ich überrascht“). Die Reliabilität der Gesamtskala betrug .94.

Wahrgenommener Konflikt. Beide Experimentalgruppen, die kongruentes und inkongruentes Feedback erhielten, gaben auf einer Skala von 0 (überhaupt nicht) bis 10 (sehr) an, wie sehr sich ihre eigene Entscheidung und das Abstimmungsergebnis der anderen Studienteilnehmer unterschieden haben. Auf diese Weise konnte kontrolliert werden, ob die jeweiligen Experimentalgruppen das Feedback tatsächlich als entscheidungskongruent bzw. entscheidungskongruent interpretierten.

Zusätzlich wurde bei allen drei Lernbedingungen überprüft, ob soziale Projektion eintritt, also ob die Probanden annehmen, dass sich die anderen genauso entschieden haben wie sie selbst, und wie hoch das Bedürfnis nach sozialem Vergleich ist. Hierzu wurden allen drei Gruppen nach der ersten, vorläufigen Entscheidung (t1) Items zur sozialen Projektion der eigenen Entscheidung, dem selbst erzeugten Konsens und dem Bedürfnis nach sozialem Vergleich vorgelegt.

Bei der *sozialen Projektion der eigenen Entscheidung* schätzten die Probanden ein, wie sich die anderen Studienteilnehmer im Vergleich zu Ihnen entschieden haben (0 = haben sich anders entschieden vs. 1 = haben sich genauso entschieden).

Beim *selbst erzeugten Konsens* gaben die Probanden an, wie viel Prozent der anderen Studienteilnehmer sich Ihrer Meinung nach genauso wie sie entschieden haben (0 = 0 % haben sich genauso wie ich entschieden bis 5 = 100 % haben sich genauso entschieden).

Das *Bedürfnis nach sozialem Vergleich* wurde anhand einer Ratingskala von 1 (überhaupt nicht) bis 10 (sehr) abgefragt. Die Fragestellung lautete: „Wie interessant fänden Sie es, zu erfahren, wie die anderen Studienteilnehmer sich in diesem Fall entschieden haben?“

Suspicion check. Am Ende des Experiments (t3) wurde mit Hilfe der offenen Frage „Schildern Sie bitte in eigenen Worten, warum es in dieser Studie Ihrer Ansicht nach gegangen

ist“ überprüft, ob die drei Lernbedingungen das Untersuchungsziel (sozialer Einfluss durch Mehrheiten) erkannt haben. Teilnehmer, deren Angaben erkennen lassen, dass sie das Untersuchungsziel erkannt haben, sollten von der nachfolgenden Analyse ausgeschlossen werden.

Alle Items zur Manipulationskontrolle wurden mit Ausnahme des „*suspicion check*“ (Greitemeyer, 2000) in leicht abgewandelter Form der zweiten experimentellen Studie von Knipfer (2009) entnommen.

8.6.4 Parasoziale Interaktion und Attraktivität der Medienfigur

Die parasoziale Interaktion und Attraktivität der Medienfigur wurden zum ersten Messzeitpunkt (t1) erfasst. Um das interpersonale Involvement der Versuchspersonen mit der Medienfigur “Margot Kreidler“ zu operationalisieren, wurden 10 Items aus den Prozess-Skalen zur Messung von *parasozialen Interaktionen (PSI)* (Schramm & Hartmann, 2008) verwendet.

Bei der Skala zu den *kognitiven PSI-Teilprozessen* bewerteten die Probanden die Medienfigur und deren Aussagen anhand von vier Items (z. B. “Ich habe mir zu den Dingen, die Margot Kreidler gesagt hat, keine Meinung gebildet“). Des Weiteren wurde abgefragt, inwieweit ein Bezug zwischen Persona und Selbst hergestellt wurde (z. B. “Ich habe mich damit beschäftigt, was mich mit Margot Kreidler verbindet oder mich von ihr unterscheidet“). Die Reliabilität der Skala betrug .62 (Cronbachs Alpha) und war zufriedenstellend.

Bei den *affektiven PSI-Teilprozessen* schätzten die Untersuchungsteilnehmer in sechs Items ihre positiven wie negativen Gefühle gegenüber der Medienfigur ein (z. B. “Margot Kreidler war mir sympathisch“) und gaben ihre Emotionen an, die durch die Persona ausgelöst werden (z. B. “Was Margot Kreidler gesagt hat, löste keinerlei Emotionen bei mir aus“). Die interne Konsistenz der Skala zu den affektiven PSI-Prozessen war mit $\alpha = .74$ ausreichend reliabel. Konative PSI-Teilprozesse wurden in dieser Studie nicht erfasst, da sie eher selten in der alltäglichen Rezeption stattfinden.

Die Reliabilität der Gesamtskala zur parasozialen Interaktion, die sich aus den 10 Items der affektiven und kognitiven PSI-Prozesse zusammensetzte, lag bei .76 (Cronbachs Alpha).

Die *Attraktivität der Medienfigur* wurde mit 5 Items erhoben (Cronbachs Alpha = .83), die dem Speech Evaluation Instrument (SEI; Zahn & Hopper, 1985) bzw. einer Studie von Linek (2007) entnommen wurden. Die Versuchspersonen beurteilten die Attraktivität der Medienfigur anhand eines semantischen Differentials mit fünffachgestuften, bipolaren Adjektivpaaren (z. B. „kalt vs. warm“, „freundlich vs. unfreundlich“). Die Fragestellung wur-

de der Studie von Töpper (2009) entnommen und lautete: „Welchen Eindruck hat Margot Kreidler auf Dich gemacht?“.

8.6.5 Richtung, Sicherheit und Änderung der Entscheidungen

In Bezug auf das Entscheidungsverhalten wurden die Richtung und die Sicherheit der Entscheidung zu allen drei Messzeitpunkten erfasst sowie die Änderung der Entscheidung zu zwei Messzeitpunkten (t2 und t3) erhoben.

Das Entscheidungsverhalten bzw. die *Richtung der Entscheidung* wurde dichotom mit dem Verfahren des Logfile-Recording erfasst (0 = nein, 1 = ja). Das Entscheidungsverhalten bzw. alle drei Aufforderungen zur aktiven Positionierung waren Bestandteil der Lernumgebung. Jede der drei Aufforderungen zur aktiven Positionierung wurde über den Touchscreen-Monitor der Lernumgebung gegeben und lautete: „Würden Sie an ihrer Stelle für eine Präimplantationsdiagnostik nach Amerika gehen?“ (siehe Abbildung 20, Kapitel 8.3).

Entscheidungsänderung. Die Revision der ersten, vorläufigen Entscheidung und der zweiten, vorläufigen Entscheidung konnte aus den Logfile-Daten des Entscheidungsverhaltens berechnet werden und wurde dichotom codiert (0 = nein, keine Entscheidungsänderung, 1 = ja, Entscheidungsänderung).

Die *Sicherheit der Entscheidung* wurde mit der Frage „Wie sicher sind Sie sich in Ihrer Entscheidung?“ anhand einer 10-stufigen Ratingskala mit den Endpunkten 1 (überhaupt nicht) und 10 (sehr) erfasst. Dieses Item wurde einer Studie von Greitemeyer (2000) entnommen.

8.7 Erfassung der Argumentationsqualität und -muster

Zur Erfassung der Argumentationsqualität und der Argumentationsmuster wurden die schriftlich verfassten Essays, in denen die Probanden ihre persönlich getroffene Entscheidung zum präsentierten Dilemma zur PID-HLA begründeten, transkribiert und inhaltsanalytisch ausgewertet (vgl. Mayring, 2008). Begründungen für Entscheidungen liefern in der Regel Argumente für oder gegen eine Entscheidung.

Alle drei Lernbedingungen fertigten insgesamt zwei Essays zur Begründung ihrer eigenen Entscheidung im Rahmen eines Prä-/Posttests zum ersten und dritten Messzeitpunkt (t1 und t3) an. Der erste Essay (Prätest) wurde zum ersten Erhebungszeitpunkt nach der Darbietung des personalisierten Films (t1) verfasst, der zweite Essay (Posttest) zum dritten Erhebungszeitpunkt nach der Bearbeitung der Zusatzinformationen (t3). Die Anfertigung

der beiden Essays erfolgte jeweils nach der eingangs zu treffenden Entscheidung über die Zustimmung oder Ablehnung einer PID-HLA und nach Angabe der Entscheidungssicherheit anhand einer zehnstufigen Ratingskala. Die Instruktion zur Verfassung der Essays lautete: „Bitte begründen Sie Ihre Entscheidung so ausführlich wie möglich!“ (Erfassung der Argumentationsqualität, d. h. der Fähigkeit der Studenten, entscheidungsunterstützende (*mysided*) Argumente zu generieren).

Zum dritten Erhebungszeitpunkt wurden zusätzlich alle Teilnehmer der drei Lernbedingungen explizit dazu aufgefordert, in einem Essay Gegenargumente zu ihrer finalen und endgültigen Entscheidung zu formulieren. Die Fragestellung zur Generierung von Gegenargumenten lautete: „Es gab Studienteilnehmer, die sich anders als Sie entschieden haben. Mit welchen Argumenten könnten diese Personen ihre Entscheidung vertreten?“ (Erfassung der Fähigkeit der Studenten, Gegenargumente (*othersided*) zu konstruieren, d. h. mögliche Argumente gegen ihre Entscheidung zu generieren).

Die Probanden konnten dann ihre Entscheidungsbegründungen bzw. die generierten Gegenargumente in die entsprechenden Frageblöcke der bereitliegenden Mappe schreiben (vgl. Anhang, Frageblock 1 und 3). Ferner wurden beide Experimentalgruppen, die kongruentes bzw. inkongruentes Feedback erhielten, gebeten, in einem Kurzesay ihre Entscheidung nach Erhalt der Feedbackinformation zum zweiten Erhebungszeitpunkt (t2) kurz zu begründen. Die Instruktion lautete: „Bitte begründen Sie kurz, warum Sie sich so entschieden haben!“. Dieser zusätzliche Kurzesay bei beiden Experimentalgruppen diente der Eruierung von möglichen Gründen einer Entscheidungsrevision sowie der Überprüfung der Glaubwürdigkeit der verwendeten Konsensinformation. Bei der Bedingung „kein Feedback“ wurde aufgrund der fehlenden Konsensinformation auf einen derartigen Kurzesay verzichtet, da angenommen wurde, dass es bei dieser Lernbedingung zu keiner Entscheidungsrevision kommt.

Im Ergebnisteil werden aus den Essays entnommene exemplarische Zitate mit einem alphanumerischen Code gekennzeichnet, der jeweils den einzelnen Studienteilnehmer, die Lernbedingung, den Erhebungszeitpunkt und die zu diesem Messzeitpunkt getroffene Richtung der Entscheidung identifiziert:

[Probandencode_Erhebungszeitpunkt_Richtung der Entscheidung]

Die ersten vier Zeichen bilden stets den Probandencode. Die erste Nummer, die von 1 bis 72 reicht, identifiziert dabei jeden einzelnen Untersuchungsteilnehmer. Dieser Zahl folgt eine zweistellige alphanumerische Zeichenfolge, die sich aus einem Buchstaben und einer Zahl zusammensetzt und jeweils eine der drei Lernbedingungen repräsentiert: [b1] steht für die Lernbedingung „kongruentes Feedback“, [b2] für die Gruppe „inkongruentes Feed-

back“ und [b3] für „kein Feedback“. Die letzte alphanumerische Zeichenfolge gibt den Erhebungszeitpunkt des Zitats wieder: [t1] gibt den ersten Erhebungszeitpunkt wieder, [t2] den zweiten und [t3] den dritten und letzten Messzeitpunkt. Das folgende [Ja] oder [Nein] gibt zum jeweiligen Erhebungszeitpunkt die Richtung der getroffenen Entscheidung an, also ob dieser Studienteilnehmer für eine PID-HLA nach Amerika gehen würde oder nicht.

Beispielsweise steht [65b2_t3_Ja] für ein Zitat des 65. Probanden, der an der Studie teilgenommen hat und der Bedingung „inkongruentes Feedback“ zugewiesen wurde [65b2]. Das Zitat wurde dem Essay zum dritten Erhebungszeitpunkt nach Bearbeitung der Zusatzinformationen entnommen [t3]. Der Proband entschied sich zu diesem Messzeitpunkt für die Handlungsalternative [Ja], d. h. er würde eine PID in Amerika durchführen lassen.

Ziel der Analyse war es, die Qualität der Entscheidungsbegründungen, insbesondere in Bezug auf eine Veränderung durch die Bearbeitung der Lernumgebung, festzustellen und zu überprüfen, ob ein *myside bias* in der Argumentation bei den Untersuchungsteilnehmern vorliegt. Darüber hinaus sollten in Anlehnung an Sadler & Zeidler (2005) typische Begründungsmuster (rational, emotional und intuitiv) bei der Lösung der Dilemmageschichte rekonstruiert werden.

Nachfolgend werden daher die Kategorienschemata zur Erfassung der Argumentationsqualität (Kapitel 8.7.1), die Erfassung des *myside bias* Index (Kapitel 8.7.2) und der Argumentationsmuster (Kapitel 8.7.3) beschrieben.

8.7.1 Kategorienschema zur Erfassung der Argumentationsqualität

Um die Argumentationsqualität zu operationalisieren, wurden die beiden Essays mithilfe eines Kategoriensystems kodiert, welches in Orientierung an die Studien von Sadler & Fowler (2006), Foong & Daniel (2010) und Knipfer (2009) speziell für diese Untersuchung entwickelt wurde.

Den theoretischen Ausgangspunkt für die Analyse der Argumentationsqualität bildete das Modell von Toulmin (1958), das ausführlich in Kapitel 3.2.3 beschrieben wurde. Während keine oder eine Rechtfertigung der Entscheidung ohne eine valide Begründung die untersten Niveaus der Argumentationsqualität darstellten, setzen höhere Stufen das Formulieren von Begründungen (*data, warrants, backings*) sowie das Nennen und Widerlegen von Gegenpositionen voraus.

Als Kodierungseinheit zählte der gesamte geschriebene Text des Essays zum jeweiligen Erhebungszeitpunkt (erster Messzeitpunkt t1 und dritter Messzeitpunkt t3). Folgende Tabelle 10 gibt das in dieser Studie verwendete sechsstufige Kategorienschema zur

Operationalisierung der Argumentationsqualität wieder. Die sechs Kategorien zur Erfassung der Qualität der Argumentationen in Bezug auf die Lösung des medizinethischen Dilemmas zur PID-HLA werden nachfolgend näher beschrieben.

Die niedrigste Stufe des Kategorienschemas "0" wurde für Entscheidungsbegründungen vergeben, in denen die Probanden nicht in der Lage waren, ihre getroffene Entscheidung zu rechtfertigen, geschweige denn zu begründen. Wie dem Beispiel für Stufe "0" aus Tabelle 10 zu entnehmen, wurden auch Pseudo-Rechtfertigungen nicht als Argument akzeptiert (z. B. „Naja, vor 94 würde meine Oma auch nicht verstehen, warum Kontaktlinsen so praktisch sind“).

In der nachfolgenden Kategorie "1" ging es um Aussagen, die zwar eine Rechtfertigung, aber keine valide Begründung für die Entscheidung beinhalteten. Das Beispiel aus Tabelle 10 macht deutlich, dass der Proband seine Entscheidung zwar rechtfertigte, aber keine valide Begründung für seine Entscheidung gab.

In die Stufe "2" und "3" fallen Statements, die eine Rechtfertigung mit validen Begründungen (*grounds = data, warrants, backings*) enthielten. Der Unterschied zwischen den beiden Stufen bestand darin, dass auf Stufe "2" die Entscheidung relativ einfach durch eine oder zwei Begründungen gerechtfertigt wurde, wohingegen auf Stufe "3" drei oder mehr elaborierte und damit gut unterstützte Begründungen gegeben wurden. Wie dem Auszug des Zitats in Tabelle 10 zu entnehmen, gab der Teilnehmer der Stufe "2" lediglich eine einzige Begründung zur Rechtfertigung seiner Entscheidung an und zwar den bereits bestehenden Wunsch nach einem weiteren Kind. Im Gegensatz dazu offenbart die exemplarische Teilnehmeraussage auf Stufe 3 mehr Begründungen.

Folglich fallen zusammenfassend betrachtet unter die Kategorie "2" und "3" alle Äußerungen, die eine klar einseitige Argumentation beinhalten, da nur Argumente angeführt werden, welche die eigene Entscheidung unterstützen.

Niveau "4" und "5" wurden an Statements vergeben, die nicht nur begründete Rechtfertigungen, sondern auch (mindestens) eine Gegenposition zur eigenen Entscheidung enthielten. Äußerungen, in der die Entscheidung mit einer elaborierten Begründung gerechtfertigt wurde und (mindestens) eine Gegenposition formuliert wurde, wurden mit „4“ kodiert. Das Beispiel aus Tabelle 10 zeigt, dass diesem Versuchsteilnehmer ein Perspektivwechsel gelang und er neben Argumenten für seine Entscheidung auch Gegenpositionen nannte, diese jedoch nicht durch *rebuttals* widerlegte.

Die höchste Stufe "5" des Kategorienschemas wurde vergeben, wenn nicht nur Rechtfertigungen mit elabrierter Begründung gegeben wurden, sondern auch explizit (mindestens) eine Gegenposition zur eigenen Entscheidung genannt und widerlegt wurde.

Tabelle 10: Kategorienschema zur Operationalisierung der Argumentationsqualität (Stufen, Beschreibungen und Ankerbeispiele).

Stufe	Beschreibung	Ankerbeispiel (Originaltext)
0	Keine Rechtfertigung	„Naja, vor 94 würde meine Oma auch nicht verstehen, warum Kontaktlinsen so praktisch sind.“ [72b3_t1_Ja]
1	Rechtfertigung ohne valide Begründung	„Wenn es eine Lösung gibt, würde ich sie gern ausprobieren (keine valide Begründung) .“ [58b2_t1_Ja]
2	Rechtfertigung mit einfacher Begründung (eine oder zwei Begründungen)	„Wenn man berücksichtigt, dass die Familie sowieso noch ein Kind wollte (das zumindest behauptet), bietet es sich an (Begründung) . Es bekommen etliche Leute Kinder aus schlechteren Gründen bzw. ohne nachzudenken.“ [52b3_t1_Ja]
3	Rechtfertigung mit elaborierter Begründung (drei oder mehr Begründungen)	„Also ich habe mich für eine Behandlung in Amerika entschieden, weil ein zweites Kind sowieso der Wunsch ist (Begründung) . Durch eine Selektion der befruchteten Eizellen würden es für Max die richtigen Gene sein und primär für das neue Kind auch bedeuten, dass es die vererbliche Krankheit nicht bekommt (Begründung) . Es rettet ein Leben und gibt ein Leben (Begründung) .“ [27b3_t1_Ja]
4	Rechtfertigung mit elaborierter Begründung und Gegenposition (eine oder mehr Begründungen und mindestens eine Gegenposition)	„Grundsätzlich denke ich, dass eine Entscheidung für PID in diesem Fall allen Beteiligten zugute kommen könnte – das bereits 5-jährige Kind könnte Heilung erfahren, die es ohne ein Geschwisterkind zumindest nicht so schnell bekommen könnte (Gegenposition) . Das 2. Kind wird von den Eltern ohnehin gewünscht und wäre daher kein reiner „Ersatzteilspender“ (Gegenposition) . Allerdings denke ich auch, dass in Anbetracht der Risiken genetischer Selektion (-> Eugenik) PID zu Recht hierzulande derzeit illegal ist (Begründung) . Ein Umgehen des Verbots durch Auslandsbesuch halte ich daher nicht für sinnvoll (Begründung) . Es ist möglich, auch ohne PID einen passenden Spender zu finden (Begründung) .“ [37b3_t1_Nein]
5	Rechtfertigung mit elaborierter Begründung, Gegenposition und Widerlegung (eine oder mehr Begründungen und mindestens eine Gegenposition mit Widerlegung)	„Wenn Frau Kreidler nach Amerika geht, um durch eine PID das perfekte „Retterkind“ zu „erschaffen“, wird dadurch niemand geschädigt (Begründung) . Und ihrem Sohn Max kann geholfen werden (Begründung) . Es gibt also nur Gewinner bei der Sache (Begründung) . Das Gegenargument mit der Entscheidung von Embryos (Gegenposition) ist hinfällig, da bei Schwangerschaftsabbrüchen legal „Embryos“ in einem viel späteren Stadium „beseitigt“ werden (Widerlegung) . Warum also nicht bei der PID? Das Argument einer psychischen Belastung für Eltern und Retterkind (Gegenposition) kann ich nur sehr gering nachvollziehen, da sehr viele Menschen in belastete Verhältnisse hereingeboren werden und auch damit klar kommen müssen. Das Bewusstsein, ein „Retterkind“ zu sein, kann z. B. auch eine Quelle großen Stolzes für einen Menschen darstellen (Widerlegung) . Die Möglichkeit, dass Gentests ausgenutzt werden (höhere Versicherungsprämien für erblich belastete, Einstellungsverweigerung, Herabsehen auf Behinderte, Druck auf Eltern, PID durchzuführen) besteht durchaus (Gegenposition) . Allerdings gab es diese Probleme auch schon vor der PID. Die gesellschaftlichen Probleme, die dadurch aufgeworfen werden (Wert des Lebens, Stellung zu Behinderten, usw.) muss die Gesellschaft anders als durch das Verbot medizinisch sinnvoller Maßnahmen in den Griff bekommen (Widerlegung) .“ [57b3_t3_Ja]

Das Beispielzitat für die Stufe 5 aus Tabelle 10 macht deutlich, dass der Proband die in den Zusatzinformationen verfügbaren Gegenargumente in die eigene Entscheidung integrierte und durch Widerlegungen (*rebuttals*) entkräftete. Beispielsweise widerlegte er das Gegenargument der möglichen psycho-sozialen Risiken für das „Retterkind“ mit dem Hinweis, dass jene nicht zwangsläufig sind, sondern dass das Wissen, seinem Geschwister existenziell geholfen zu haben, es auch stolz machen kann.

Kodierprozedur und Interrater-Reliabilität. Die Kodierung der Essays wurde von der Verfasserin der Arbeit und einer Studentin vorgenommen. Zur Sicherung der Objektivität der Zuordnung der Bedeutungseinheiten wurden alle Essays von den beiden Raterinnen unabhängig voneinander analysiert. Anschließend wurden die beiden Kodierreihen auf Übereinstimmung überprüft. Dem obig dargelegten Kategorienschema zur Argumentationsqualität entsprechend wurden die qualitativen Daten aus den beiden Essays durch die Zuweisung einer Zahl von 0-5 in quantitative Daten überführt. Bezüglich der Transformation qualitativer in quantitative Daten bemerken Bortz & Döring (2006) auf S. 298 folgendes: „Mit Hilfe von Urteilern lassen sich aus Verbaldaten auch quantitative Daten auf höherem Skalenniveau (Ordinal-, Kardinalskala) erzeugen, indem die Texte in geordnete Kategorien sortiert oder auf Ratingskalen eingeschätzt werden“.

Da eine solche Messung der Argumentationsqualität auf Intervallskalenniveau sicherlich methodisch umstritten ist, wurden, um absolut sicher zu gehen, dass tatsächlich auch von einer Gleichabständigkeit der Skalenpunkte ausgegangen werden kann, Reliabilitätsmaße für intervall- und ordinalskalierte Daten berechnet. Denn, so bemerken Caspar & Wirtz (2002) auf S. 127, „indizieren sowohl die Maße für intervall- als auch für ordinalskalierte Daten, dass die Reliabilität zufriedenstellend ist, so ist die Frage des Skalenniveaus nicht kritisch für die Annahme der Reliabilität der Daten“. Dementsprechend wurden als Maß für die Interrater-Reliabilität sowohl der Intra-Klassen Korrelationskoeffizient (ICC) für intervallskalierte als auch der Rangkorrelationskoeffizient von Spearman für ordinalskalierte Daten berechnet.

Es zeigte sich eine hohe Übereinstimmung zwischen den Resultaten der beiden Raterinnen für die Argumentationsqualität zum ersten und dritten Messzeitpunkt. Der ICC-Wert für die Argumentationsqualität zum ersten Messzeitpunkt betrug .93 ($p = .000$) und es zeigte sich eine sehr hohe und substantielle Korrelation nach Spearman von $r_s = .81$ ($p = .000$). Auch der erreichte ICC-Wert von .95 und Spearman's Rho von .86 ($p = .000$) für die Interrater-Reliabilität in Bezug auf die Argumentationsqualität zum dritten Messzeitpunkt deutet auf eine sehr gute Urteilerkonkordanz hin. Die erreichte Interrater-Reliabilität zu beiden Messzeitpunkten war somit vor dem Hintergrund, dass ein Wert über .70 laut Wirtz & Caspar (2002) als zufrieden stellend gilt, hervorragend. Abschließend verglichen die bei-

den Raterinnen die nicht übereinstimmenden Kodierungen und beseitigten die bestehenden Differenzen in einem Konsensrating durch die gemeinsame Einigung auf einen Wert.

8.7.2 Myside bias Index

Von besonderem Interesse im Hinblick auf die Argumentationsqualität war außerdem der Vergleich der Anzahl der generierten entscheidungsunterstützenden Argumente (*mysided arguments*; Argumente, die mit der eigenen Entscheidung konsistent sind) mit der Anzahl der präferenzinkonsistenten Argumente (*othersided arguments*; (Gegen-)Argumente, die mit der eigenen Entscheidung inkonsistent sind).

Zur Berechnung des *myside bias* Index wurden über alle geschriebenen Essays hinweg die Häufigkeiten der generierten Argumente insgesamt, der entscheidungsunterstützenden Argumente und der Gegenargumente der Probanden deskriptiv ausgezählt. Aus der Differenz der entscheidungsunterstützenden Argumente und der Gegenargumente wurde der *myside bias* Index berechnet (vgl. Toplak & Stanovich, 2003).

Ein Wert < 0 im *myside bias* Index gibt an, dass die Versuchsteilnehmer mehr präferenzinkonsistente als -konsistente Argumente in den Essays angaben. Ein Wert $= 0$ zeigt, dass keine Verzerrung vorliegt und die Versuchsteilnehmer gleichviel Argumente für als auch gegen ihre Position generierten. Alle Werte > 0 geben an, dass mehr präferenzkonsistente als präferenzinkonsistente Argumente formuliert wurden und somit ein *myside bias* vorliegt.

8.7.3 Kategorienschema zur Erfassung der Argumentationsmuster

Die Argumentationen der Versuchspersonen wurden zusätzlich auf der Basis der bereits dargestellten theoretischen Überlegungen im Hinblick auf die Verwendung rationaler und emotional-intuitiver Argumentationsweisen analysiert. Zur Klassifikation der Argumentationsmuster wurde ein nominalskaliertes Kategoriensystem, das auf einer Studie von Sadler & Zeidler (2005) basierte, verwendet. Die beiden Autoren dieser Studie konnten drei typische Begründungsmuster bei der Lösung gentechnologischer Szenarien bei Schülern identifizieren: rationale, emotionale und intuitive Argumentationsmuster (siehe Kapitel 3.2.3 und 3.3). Da es sich auch in dieser Studie als äußerst schwierig erwies, emotionale und intuitive Argumentationsweisen eindeutig voneinander zu unterscheiden, werden diese beiden Begründungsmuster in dieser Studie zusammenfassend dargestellt (vgl. Dawson & Venville, 2009; Sadler & Zeidler, 2005). Eine Analyseeinheit bildeten die zwei anzuferti-

genden Essays. Die Statements aus beiden Essays konnten jeweils beiden Kategorien (rational vs. emotional-intuitiv) zugewiesen werden.

Bei einer *rationalen Argumentationsweise* lösten die Probanden die Dilemmageschichte auf der Basis von rationalen Kalkulationen. Als Beispiel kann die Versuchsperson [37b3_t1_Ja] (siehe folgende Tabelle 11) angeführt werden, die utilitaristische Prinzipien zur Begründung ihrer getroffenen Entscheidung anführte, indem sie eine Nutzen-Schaden-Kalkulation im Hinblick auf die Handlungsfolgen der von einer PID-HLA betroffenen Menschen anstellte. Eine *emotional-intuitive Argumentationsweise* bezog sich dagegen auf unverzügliche Reaktionen in Bezug auf den Kontext eines Szenarios oder auf emotionale Reaktionen in Bezug auf die in dem Dilemma betroffenen Personen. Dem Beispielzitat [70b1_t1_Nein] aus Tabelle 11 kann entnommen werden, dass dieser Versuchsteilnehmer eine Fürsorge-Perspektive einnahm; die Dilemmageschichte zur PID-HLA rief Empathie und Fürsorge gegenüber dem „Retterkind“ hervor.

Kodierprozedur und Interrater-Reliabilität. Die Kodierung der Argumentationsmuster wurde von der Autorin dieser Arbeit und einer Studentin unabhängig voneinander durchgeführt. Es konnten jeweils zwei Kategorien (rationales und emotional-intuitives Argumentationsmuster) über beide Essays hinweg gleichzeitig vergeben werden. Eine Mehrfachkodierung war somit möglich und kam bei der Mehrheit (51,4 %) der Begründungen vor. Die prozentuale Übereinstimmung in der Kategorisierung der beiden verwendeten Argumentationsmuster, die den prozentualen Anteil der Fälle angibt, in denen zwei Beurteiler das gleiche Urteil abgeben, lag bei über 95 %. Die erreichte prozentuale Beurteilerübereinstimmung war somit vor dem Hintergrund, dass für diese Studie in Anlehnung an Bortz & Döring (2006) ein Wert von 80 % zwischen den beiden Ratern als zufriedenstellend festgelegt wurde, sehr gut. Die nicht übereinstimmenden Kodierungen wurden anschließend von beiden Raterinnen durchgesprochen und die drei bestehenden Differenzen wurden durch ein Konsensrating behoben.

Um zu erfahren, welche deontologischen und utilitaristischen Argumente beim rationalen Argumentationstyp am häufigsten von den Lernenden zur Begründung der eigenen Entscheidung und bei der Aufforderung zur Nennung von Gegenargumenten formuliert wurden, wurden die transkribierten Essays zusätzlich mithilfe des Computerprogramms MAXQDA 10 inhaltsanalytisch ausgewertet. Die Auswertung der Daten erfolgte in Anlehnung an die Qualitative Inhaltsanalyse nach Mayring (2008). Hierfür wurden relevante Textstellen in ein Kategoriensystem eingeordnet, das aus deduktiv erstellten und induktiv ergänzten Kategorien bestand. Die genannten Argumente wurden im Hinblick auf ihre ethische Tradition (deontologisch: pflichten- und wertorientiert vs. utilitaristisch: folgen- und nutzenorientiert) systematisiert (vgl. Höhle, 2001b).

Tabelle 11: Kategorienschema zur Erfassung der Argumentationsmuster (Kategorien mit Beschreibung und Ankerbeispielen).

Kategorien mit Beschreibung

Rationales

Argumentationsmuster:

Rationale Kalkulationen wie Anwendung deontologischer und utilitaristischer Prinzipien, Abwägen von Chancen und Risiken sowie von Vor- und Nachteilen

Ankerbeispiel

„Aus Sicht des „Retterkinds“ wäre ich gegen eine PID, denn es läuft Gefahr „nur Mittel zum Zweck“ zu sein und somit als eigenständiger Mensch in den Hintergrund zu geraten. Welche psycho-sozialen Auswirkungen dies zur Folge hätte, ist schwer abzuschätzen“ [08b3_t3_Nein].

„Ich lehne Gentest und medizinischen Fortschritt nicht grundsätzlich ab, sehe es aber als kritisch an, wenn man sich ein Wunschkind aussuchen kann – auch wenn das Kind in diesem Fall eh gewollt ist und damit auch dem anderen geholfen werden kann. Würde es in Deutschland erlaubt sein, hätte ich mich vermutlich eher dafür entschieden. Ich halte nichts davon in Ausland zu gehen und dort Dinge tun zu lassen, die hier – aus i.d.R. guten Gründen verboten sind. Hierbei spielt für mich auch ein Sicherheitsaspekt eine Rolle. Ich finde eine Ausweitung der Knochenmarkspende zurzeit angebrachter als PID.“ [50b1_t3_Nein].

„Entscheidung gegen die „Fahrt nach Amerika“, weil: das zweite Kind würde evtl. nur als „Ersatzteillager“ fungieren, die Entscheidung für ein Kind sollte nicht nur daher kommen, da das andere sonst nicht leben kann, sonst müsste auch gestattet werden, dass sich Eltern aussuchen, wie ihr Kind aussehen soll – diese Gradwanderung ist in meinen Augen nicht weit (Aussehen bezüglich Geschlecht, Haarfarbe, Augenfarbe, alles was so möglich ist)“ [11b3_t1_Nein].

Emotional-intuitives

Argumentationsmuster

Emotionale Reaktionen in Bezug auf die in dem Dilemma betroffenen Personen, z. B. Fürsorge, Empathie, Sympathie und Beachten der Notlage

Intuitiv: “Bauchgefühl”, unverzügliche Reaktionen, häufig negativ und sehr stark, geht oft dem rationalen und emotionalen Argumentieren voraus

„Was muss das für ein Gefühl sein...später zu erfahren, dass man mehr oder weniger „gezüchtet“ wurde, um dem Bruder zu helfen“ [70b1_t1_Nein].

„Ich verstehe die Trauer und schwierige Lage der Frau wegen der Krankheit ihres Sohnes, aber es kann keine Lösung sein, selber mit Hilfe der Gentechnologie einen Knochenmarkspender zu erschaffen [...]. Ich finde den Gedanken schrecklich, dass dieses zweite Kind zur Rettung des ersten Kindes dienen muss, um leben zu dürfen. Dieses Verfahren ist mir einfach unheimlich und zuwider“ [35b2_t1_Nein].

„Die Vorstellung, helfen zu können, war stark in mir, aber die Funktionalisierung eines Menschen empfinde ich doch als sehr abstoßend“ [03b2_t1_Nein].

„Meine Entscheidung wurde stark beeinflusst durch ein Gefühl des Widerstands und des Unwohlseins, als Frau Kreidler davon sprach über die PID herausfinden zu können, welche Embryonen am besten zu ihrem Sohn passen könnten. Dabei kamen mir Assoziationen und Befürchtungen des Missbrauchs anderer „Wesen“ in den Sinn, die quasi für das eine Kind erhalten müssen“ [69b1_t1_Nein].

Anmerkungen: Kursiver Text stellt das rationale Argumentationsmuster dar, unterstrichener Text das emotional-intuitive Argumentationsmuster.

9 Ergebnisse

Die Darstellung der Ergebnisse erfolgt anhand der Reihenfolge der Fragestellungen und Hypothesen in Kapitel 7. Vorab werden zunächst die experimentellen Voraussetzungen der Studie überprüft (Kapitel 9.1). Im Anschluss daran folgt eine Analyse der parasozialen Interaktion mit der Medienfigur (Kapitel 9.2), des Entscheidungsverhaltens (Kapitel 9.3), der Argumentationsqualität (Kapitel 9.4) und der Argumentationsmuster (Kapitel 9.5). Zum Abschluss des Ergebnisteils werden Einflussfaktoren auf die Entscheidungsfindung und Argumentation analysiert (Kapitel 9.6).

Alle statistischen Analysen wurden mit dem Statistiktool SPSS 17.0 für Windows durchgeführt. Die qualitative Inhaltsanalyse der Essays erfolgte mithilfe des Programms MAXQDA 10.

9.1 Überprüfung der experimentellen Voraussetzungen

Trotz des Zufallsprinzips bei der Randomisierung besteht die Gefahr ungleicher Gruppen im Hinblick auf potentielle Störvariablen. Deshalb wurde zur Sicherung der internen Validität der Studie vorab überprüft, inwieweit die zwei experimentellen Gruppen und die Kontrollgruppe vergleichbar im Hinblick auf kognitive, motivationale und einstellungsbezogene Lernvoraussetzungen (Kapitel 9.1.1.) und die realisierte Lernzeit (Kapitel 9.1.2) sind. Zusätzlich wurde eine Manipulationskontrolle der unabhängigen Variablen Feedback durchgeführt und der Verdachtsfragebogen (*suspicion check*) daraufhin ausgewertet, ob das Untersuchungsziel (sozialer Einfluss von Mehrheiten) von den Probanden erkannt wurde (Kapitel 9.1.3).

9.1.1 Individuelle Lernvoraussetzungen

Insgesamt erzielten die Probanden im themenspezifischen Vorwissenstest zur Humangenetik (GeWi) durchschnittlich etwa 6 Punkte von 12 erreichbaren. Das Gesamtmittel betrug 6.19 Punkte ($SD = 1.63$). Die beste Leistung lag bei 10 Punkten, die schlechteste bei 3 Punkten. Am häufigsten wurden 5 Punkte (21; 29.2 %) erreicht; der Median betrug 6 Punkte. Es traten weder Boden- noch Deckeneffekte auf (siehe folgende Tabelle 12). Die Teilnehmer der Kontrollgruppe „ohne Feedback“ erzielten mit durchschnittlich 6.29 Punkten ($SD = 1.78$) den höchsten GeWi-Score. Deskriptiv etwas niedriger war das themenspezifische

sche Vorwissen bei den Teilnehmer der Experimentalgruppen „inkongruentes“ ($M = 6.25$, $SD = 1.68$) und „kongruentes Feedback“ ausgeprägt ($M = 6.04$, $SD = 1.49$). Die Standardabweichung war in allen drei Gruppen vergleichbar. Die deskriptiv minimal vorgefundene Punkte-Differenz im GeWi-Score zwischen den drei Lernbedingungen verfehlte erwartungsgemäß bei weitem die Signifikanzgrenze von 5 % ($F(2, 69) = 0.16$, *n.s.*).

Tabelle 12: Themenspezifisches Vorwissen zur Humangenetik für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Vorwissen Humangenetik	6.04	(1.49)	6.25	(1.68)	6.29	(1.78)	<i>n.s.</i>

Anmerkung: Wertebereich von 1 niedrig bis 12 hoch.

Beim selbsteingeschätzten Vorwissen bezüglich des Themas der Lernumgebung ergaben sich vor dem Hintergrund des theoretischen Maximums von 5.0 durchweg sehr niedrige Skalenmittelwerte (siehe folgende Tabelle 13). Am niedrigsten fiel deskriptiv das subjektive Vorwissen zur PID aus (Gesamtmittelwert: 1.49), etwas höher ausgeprägt war das Vorwissen bezüglich genetischer Untersuchungen (Gesamtmittelwert: 2.01) und zur Bioethik (Gesamtmittelwert: 2.13).

Tabelle 13: Subjektives Vorwissen zum Thema Gentest, PID und Bioethik für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Vorwissen Gentest	2.00	(0.83)	1.96	(0.81)	2.08	(0.78)	<i>n.s.</i>
Vorwissen PID	1.29	(0.62)	1.67	(1.01)	1.50	(0.83)	<i>n.s.</i>
Vorwissen Bioethik	2.38	(1.06)	2.17	(1.09)	1.83	(0.87)	<i>n.s.</i>

Anmerkung: Wertebereich von 1 niedrig bis 5 hoch.

Es bestanden sowohl im subjektiv eingeschätzten Vorwissen zum Thema Gentest als auch zur PID keine signifikanten Unterschiede zwischen den drei Gruppen ($F(2, 69) = 0.15$, *n.s.* bzw. $F(2, 69) = 1.21$, *n.s.*). Allerdings variierte das subjektive Vorwissen zur PID in der Bedingung „inkongruentes Feedback“ stärker als in den beiden anderen Gruppen. Dementsprechend wies der Levene-Test heterogene Varianzen auf ($F(2, 69) = 3.62$, $p < .05$). Bei

Verwendung des nonparametrischen Kruskal-Wallis-Test blieb der Unterschied ebenfalls nicht signifikant ($\chi^2 = 1.97$, *n.s.*). Auch hinsichtlich des subjektiv eingeschätzten Vorwissens zur Bioethik waren alle drei Gruppen vergleichbar ($F(2, 69) = 1.76$, *n.s.*). Es konnte somit davon ausgegangen werden, dass die Studierenden der drei Gruppen ein ähnliches Vorwissensniveau bezüglich des Inhaltsgebietes der Lernumgebung aufwiesen.

Für die interne Validität der Studie war es ebenfalls wichtig, zu untersuchen, ob sich die drei Lernbedingungen im Hinblick auf motivationale und einstellungsbezogene Lernvoraussetzungen unterscheiden.

Tabelle 14 verdeutlicht, dass der persönliche Bezug zum Thema Gentest bei allen drei Gruppen sehr deutlich unter dem theoretischen Durchschnittsbereich der Skala lag (Gesamtmittelwert: 1.92; theoretisches Maximum: 5). Zwischen den drei Gruppen zeigten sich im persönlichen Bezug zu genetischen Untersuchungen keine statistisch bedeutsamen Unterschiede ($F(2, 69) = 1.06$, *n.s.*).

Tabelle 14: Persönlicher Bezug zum Thema Gentest, Interesse am Thema Gentest, Präferenz für Deliberation und Intuition für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Persönlicher Bezug	1.92	(0.88)	1.71	(0.96)	2.13	(1.17)	<i>n.s.</i>
Interesse Gentest	3.92	(0.72)	3.91	(0.65)	3.93	(0.70)	<i>n.s.</i>
Präferenz f. Deliberation	3.74	(0.76)	3.68	(0.86)	3.61	(0.77)	<i>n.s.</i>
Präferenz f. Intuition	3.35	(0.65)	3.58	(0.81)	3.58	(0.57)	<i>n.s.</i>

Anmerkung: Wertebereich von 1 niedrig bis 5 hoch.

Das Interesse am Thema Gentest (Gesamtmittelwert: 3.92), die Präferenz für Deliberation (Gesamtmittelwert: 3.68) und Intuition (Gesamtmittelwert: 3.51) lagen im Gegensatz zum persönlichen Bezug deutlich über dem theoretischen Durchschnittsbereich von 3.0 der Skala. In Bezug auf diese drei Lernvoraussetzungsaspekte konnten ebenfalls keine signifikanten und substantiellen Unterschiede zwischen den drei Gruppen gefunden werden ($F(2, 69) = .01$, *n.s.* bzw. $F(2, 69) = 0.17$, *n.s.* bzw. $F(2, 69) = 0.93$, *n.s.*).

Wie folgender Tabelle 15 zu entnehmen, lag die soziale Vergleichsorientierung leicht über dem theoretischen Skalenmittel von 3.0 und war deskriptiv etwas höher ausgeprägt (Gesamtmittelwert: 3.32) als die Orientierung an anderen (Gesamtmittelwert: 2.82), die leicht unter dem theoretischen Skalenmittel lag.

Tabelle 15: Soziale Vergleichsorientierung und Orientierung an anderen für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Vergleichsorientierung	3.24	(0.78)	3.44	(0.76)	3.27	(0.50)	<i>n.s.</i>
Orientierung an anderen	2.65	(0.67)	2.99	(0.63)	2.83	(0.62)	<i>n.s.</i>

Anmerkung: Wertebereich von 1 niedrig bis 5 hoch.

Die drei Gruppen unterschieden sich statistisch nicht bedeutsam in diesen beiden Variablen ($F(2, 69) = 0.57, n.s.$ bzw. $(F(2, 69) = 1.72, n.s.)$).

Tabelle 16 gibt die Mittelwerte zu den Einstellungen zu genetischen Untersuchungen mit den Subskalen Positive Aspekte, Negative Aspekte und Befürchtungen wieder. Vor dem Hintergrund der fünfstufigen Skala fielen die positiven Aspekte (Vorteile) genetischer Untersuchungen im Durchschnitt bei allen drei Lernbedingungen hoch aus (Gesamtmittelwert: 3.55). Deskriptiv niedriger war der Skalenmittelwert zu den negativen Aspekten (Nachteilen) genetischer Untersuchungen ausgeprägt (Gesamtmittelwert: 3.06). Am höchsten fiel der Mittelwert zu den Befürchtungen in Bezug auf genetische Untersuchungen aus (Gesamtmittelwert: 3.78).

Tabelle 16: Einstellungen (Positive Aspekte, Negative Aspekte und Befürchtungen) zu genetischen Untersuchungen für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Positive Aspekte	3.53	(0.83)	3.48	(0.76)	3.63	(0.67)	<i>n.s.</i>
Negative Aspekte	3.11	(0.95)	3.11	(0.86)	2.94	(0.66)	<i>n.s.</i>
Befürchtungen	4.00	(0.79)	3.88	(0.79)	3.47	(1.10)	<i>n.s.</i>

Anmerkung: Wertebereich von 1 niedrig bis 5 hoch.

Insgesamt betrachtet konnte somit in Bezug auf das Meinungsbild zu genetischen Untersuchungen bei den Versuchsteilnehmern eher eine positive als eine negative Haltung gegenüber Gentests konstatiert werden: Es werden mehr Vorteile als Nachteile genetischer Untersuchungen gesehen. Allerdings werden auch Befürchtungen bezüglich genetischer Untersuchungen gehegt. Es bestanden keine statistisch bedeutsamen Unterschiede in den drei

Skalen (Positive Aspekte, Negative Aspekte, Befürchtungen) zwischen den Gruppen ($F(2, 69) = 0.22, n.s.$ bzw. $F(2, 69) = 0.36, n.s.$ bzw. $F(2, 69) = 2.23, n.s.$).

Zur Charakterisierung der Stichprobe kann zusammenfassend festgehalten werden, dass zwar ein hohes Interesse, simultan aber auch ein niedriges subjektives Vorwissen am Thema der Lernumgebung bei den Lernenden bestand. Zudem konnten ein geringer persönlicher Bezug und eher positive als negative Einstellungen zum Thema genetische Untersuchungen festgestellt werden.

Insgesamt betrachtet war die angestrebte Homogenität der drei experimentellen Bedingungen hinsichtlich der individuellen Lernvoraussetzungen in dieser Studie gewährleistet, alle drei Gruppen waren in ihren kognitiven, motivationalen und einstellungsbezogenen Lernvoraussetzungen vergleichbar. Die interne Validität der Studie konnte somit in Bezug auf diese erfassten Lernvoraussetzungsaspekte als gesichert gelten.

9.1.2 Realisierte Lernzeit

Zusätzlich wurde noch überprüft, ob sich die drei Gruppen in der realisierten individuellen Lern- bzw. Bearbeitungszeit der Lernumgebung unterscheiden. Da eine längere Lern- und Bearbeitungszeit auf eine intensivere Auseinandersetzung mit den Inhalten der Lernumgebung schließen lässt, könnte dies auch zu einer höheren Argumentationsqualität führen. Es sollte ausgeschlossen werden, dass sich ein Effekt der Lernumgebung im Hinblick auf die Argumentationsqualität lediglich durch eine unterschiedliche Bearbeitungszeit der Lernumgebung ergab.

Die maximale Zeit, in der sich die Studierenden mit der Lernumgebung auseinandersetzen konnten, war nicht begrenzt. Da beide Experimentalbedingungen (EG 1 und EG 2) einen Frageblock mehr zu bearbeiten hatten, wurde für die Signifikanzprüfung die Lernzeit bei beiden Gruppen um die Bearbeitungszeit dieses Frageblockes bereinigt.

Die Bearbeitungszeit der Lernumgebung variierte zwischen 14 und 44 Minuten, der Gesamtmittelwert der Lernzeit betrug 25.87 ($SD = 6.76$) Minuten.

Folgende Tabelle 17 macht deutlich, dass die Bedingung „kein Feedback“ sich am längsten mit der Lernumgebung und der Bearbeitung der Frageblöcke befasste ($M = 28.36, SD = 7.05$), gefolgt von der Bedingung „kongruentes Feedback“ ($M = 25.55, SD = 6.72$). Am kürzesten bearbeitete die Gruppe „inkongruentes Feedback“ die Lernumgebung ($M = 23.69, SD = 5.88$). Die Standardabweichungen waren in allen drei Gruppen vergleichbar.

Tabelle 17: Individuelle Lernzeit bzw. Bearbeitungszeit (in Minuten) der Lernumgebung für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Lernzeit	25.55	(6.72)	23.69	(5.88)	28.36	(7.05)	$p > .05$

Anmerkungen: Bei Bedingung 1 und 2 wurde die Bearbeitungszeit der zusätzlichen Ratingskalen und des Kurzsessays von der gesamten Lernzeit subtrahiert (bereinigte Lernzeit).

Die deskriptiv vorgefundenen Unterschiede in der Lernzeit zwischen den drei Gruppen verfehlten denkbar knapp die Signifikanzgrenze von 5 % ($F(2, 69) = 3.08, p > .05$).

Demzufolge befassten sich tendenziell nicht alle drei Bedingungen gleich lang mit der Lernumgebung und der Bearbeitung der Frageblöcke. Bei der Analyse der Argumentationsqualität in Kapitel 9.4 sollte daher zusätzlich untersucht werden, ob und inwieweit die Lernzeit als potentielle Störvariable die Argumentationsqualität beeinflusst (vgl. Bortz, 1999).

9.1.3 Manipulationskontrolle

Im Rahmen der Manipulationskontrolle wurde zunächst bei allen drei Lernbedingungen überprüft, ob eine soziale Projektion der eigenen Entscheidung im Sinne eines *false consensus effect* auftritt und wie hoch das Bedürfnis nach sozialem Vergleich ist. Im Anschluss wurde überprüft, ob die Manipulation der unabhängigen Variable Feedback bei beiden Experimentalgruppen erfolgreich war. Es wurde angenommen, dass die drei Lernbedingungen in der sozialen Projektion, dem selbst erzeugten Konsens und dem Bedürfnis nach sozialem Vergleich vergleichbar sind. Ferner wurde erwartet, dass zwischen beiden Experimentalgruppen signifikante Unterschiede im psychologischen Unbehagen nach dem Feedback und im wahrgenommenen Konflikt der Feedbackinformation bestehen.

Soziale Projektion, selbst erzeugter Konsens und Bedürfnis nach sozialem Vergleich. Folgende Tabelle 18 gibt die Mittelwerte der sozialen Projektion, dem selbst erzeugten Konsens und dem Bedürfnis nach sozialem Vergleich für alle drei Lernbedingungen wieder. Der Mittelwert der sozialen Projektion, bei der die Teilnehmer nach der ersten, vorläufigen Entscheidung einschätzen mussten, wie sich die anderen Studienteilnehmer im Vergleich zu ihnen entschieden haben, lag im Durchschnitt bei 0.71 ($SD = 0.46$). Bei der Variable Soziale Projektion ist zu beachten, dass sie dichotom erhoben wurde und so kodiert wurde,

dass der Wert 0 = „haben sich anders als ich entschieden“ und der Wert 1 = „die anderen Studienteilnehmer haben sich genauso entschieden“ entspricht. Der Mittelwert gibt daher an, dass 71 % der Probanden (51) schätzen, dass sich die anderen Studienteilnehmer genauso wie sie selbst entschieden haben, lediglich 29 % (21) dagegen glauben, dass die Versuchsteilnehmer sich anders entschieden haben.

Tabelle 18: Soziale Projektion, selbst erzeugter Konsens und Bedürfnis nach sozialem Vergleich für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Soziale Projektion ^a	0.79	(0.42)	0.67	(0.48)	0.67	(0.48)	<i>n.s.</i>
Selbst erzeugter Konsens ^b	3.58	(0.78)	3.21	(0.83)	3.13	(0.90)	<i>n.s.</i>
Bedürfnis nach sozialem Vergleich ^c	7.96	(2.69)	7.79	(2.60)	8.29	(2.07)	<i>n.s.</i>

Anmerkung: Wertebereich von ^a 0 niedrig bis 1 hoch; ^b 1 niedrig bis 5 hoch; ^c 1 niedrig bis 10 hoch.

Die soziale Projektion der eigenen Entscheidung fiel am höchsten bei der Lernbedingung „kongruentes Feedback“ aus ($M = 0.79$, $SD = 0.42$). Etwas weniger trat soziale Projektion bei den Bedingungen „inkongruentes Feedback“ und „kein Feedback“ auf, der Mittelwert war deskriptiv identisch ausgeprägt ($M = 0.67$, $SD = 0.48$). Der Unterschied in der sozialen Projektion zwischen den drei Gruppen war nicht signifikant ($F(2, 69) = 0.59$, *n.s.*).

Bezüglich der Signifikanztestung ist zu anmerken, dass eine varianzanalytische Verrechnung dichotomer Daten statt eines auf Häufigkeitsdaten basierenden Verfahrens wie dem Chi-Quadrat-Test legitim ist, wenn die Mittelwerte, so wie im vorliegenden Fall, ungefähr zwischen .20 und .90 liegen (Cox, 1970, p. 16; Jonas, Frey et al., 2001).

Beim selbst erzeugten Konsens sollten die Vpn einschätzen, wie viel Prozent der anderen Studienteilnehmer sich genauso wie sie selbst entschieden haben (von (1) = 0 % haben sich genauso wie ich entschieden bis (5) = 100 % haben sich genauso wie ich entschieden). Der Gesamtmittelwert bei allen drei Gruppen lag bei 3.31 ($SD = 0.85$), d. h. die Probanden schätzten, dass im Durchschnitt etwa 58 % der anderen Studienteilnehmer sich genauso entschieden haben wie sie selbst. Der selbst erzeugte Konsens fiel bei der Gruppe „kongruentes Feedback“ ($M = 3.58$, $SD = 0.78$) etwas höher aus, als bei den Gruppen „inkongruentes“ ($M = 3.21$, $SD = 0.83$) und „kein Feedback“ ($M = 3.13$, $SD = 0.90$). Die Unterschiede zwischen den drei Gruppen waren jedoch nicht signifikant ($F(2, 69) = 2.04$, *n.s.*).

Das Interesse, zu erfahren, wie sich die anderen Studienteilnehmer entschieden haben, fiel vor dem Hintergrund der 10-stufigen Ratingskala recht hoch aus (Gesamtmittelwert: 8.01, $SD = 2.45$). Die Unterschiede im Bedürfnis nach sozialem Vergleich zwischen den drei Gruppen waren aber nur marginal und nicht signifikant ($F(2, 69) = 0.25, n.s.$).

Unmittelbar nach Erhalt der Feedbackinformation zum zweiten Messzeitpunkt (t_2) wurde überprüft, ob die experimentelle Manipulation in Form von kongruentem bzw. inkongruentem Feedback von den Probanden der beiden Experimentalgruppen tatsächlich als entscheidungskongruent bzw. entscheidungsinkongruent interpretiert wurde. Hierzu schätzten beide Experimentalgruppen zunächst ihr psychologisches Unbehagen nach dem Feedback und anschließend ihren wahrgenommenen Konflikt ein.

Psychologisches Unbehagen nach Feedback. Deskriptiv betrachtet erlebten die Teilnehmer der Bedingung “inkongruentes Feedback“, welche die Information bekamen, dass 86 % der anderen Studienteilnehmer sich anders entschieden hatten als sie selbst, deutlich weniger positive Gefühle ($M = 1.67, SD = 0.83$) als die Teilnehmer der Bedingung “kongruentes Feedback“ ($M = 2.91, SD = 0.87$), bei denen sich 86 % genauso entschieden hatten. Die Standardabweichung war bei beiden Gruppen vergleichbar (siehe Tabelle 19).

Tabelle 19: Psychologisches Unbehagen nach dem Feedback für beide Experimentalgruppen: Positiver und negativer Affekt sowie Überraschung. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.

	1) Kongruentes Feedback		2) Inkongruentes Feedback		P
	M	(SD)	M	(SD)	
Positiver Affekt	2.91	(0.87)	1.67	(0.83)	$p = .000$
Negativer Affekt	1.26	(0.54)	2.69	(1.24)	$p = .000$
Überraschung	2.81	(1.18)	4.04	(1.20)	$p \leq .001$

Anmerkung: Wertebereich von 1 niedrig bis 5 hoch.

Erwartungsgemäß war der deskriptiv vorgefundene Unterschied im positiven Affekt zwischen den beiden Feedbackbedingungen hochsignifikant und substantiell ($t(46) = 5.05, p = .000$). Hier lag mit einer Effektstärke von $d = 1.46$ ein sehr großer und praktisch bedeutsamer Effekt vor.

Zudem berichteten Versuchspersonen, die ein zu ihrer eigenen Entscheidung inkongruentes Feedback erhielten, deutlich mehr negative Emotionen ($M = 2.69, SD = 1.24$) als Versuchspersonen, die eine zu ihrer Entscheidung kongruente Konsensinformation bekamen ($M = 1.26, SD = 0.54$). Die Mittelwerte der beiden Gruppen unterschieden sich hochsigni-

fikant voneinander ($t(31.33) = -5.17, p = .000$), die Größe des Effektes war mit $d = 1.50$ etwas höher als beim positiven Affekt und ebenfalls praktisch sehr bedeutsam.

Allerdings variierten die Standardabweichungen in der Bedingung „inkongruentes Feedback“ stärker als in der Bedingung „kongruentes Feedback“, der Levene-Test war signifikant ($F(1, 46) = 20.89, p = .000$). Aufgrund der heterogenen Varianzen wurde zusätzlich noch ein U-Test für unabhängige Stichproben durchgeführt (vgl. Bortz, 1999). Der Unterschied zwischen den beiden Gruppen blieb auch bei Verwendung eines nonparametrischen Verfahrens hochsignifikant ($U = 87.00, p = .000$).

Ferner waren die Lernenden der Bedingung „inkongruentes Feedback“ deutlich überraschter über das Abstimmungsergebnis der anderen Studienteilnehmer ($M = 4.04, SD = 1.20$) als Lernende der Bedingung „kongruentes Feedback“ ($M = 2.81, SD = 1.18$). Inferenzstatistisch konnte ein substantieller und praktisch bedeutsamer Effekt ($t(46) = -3.60, p \leq .001; d = 1.03$) nachgewiesen werden.

Abbildung 27 gibt die oben beschriebenen Befunde zum psychologischen Unbehagen bzw. Dissonanzerleben nach dem Feedback bei beiden Experimentalgruppen nochmals zusammenfassend wieder:

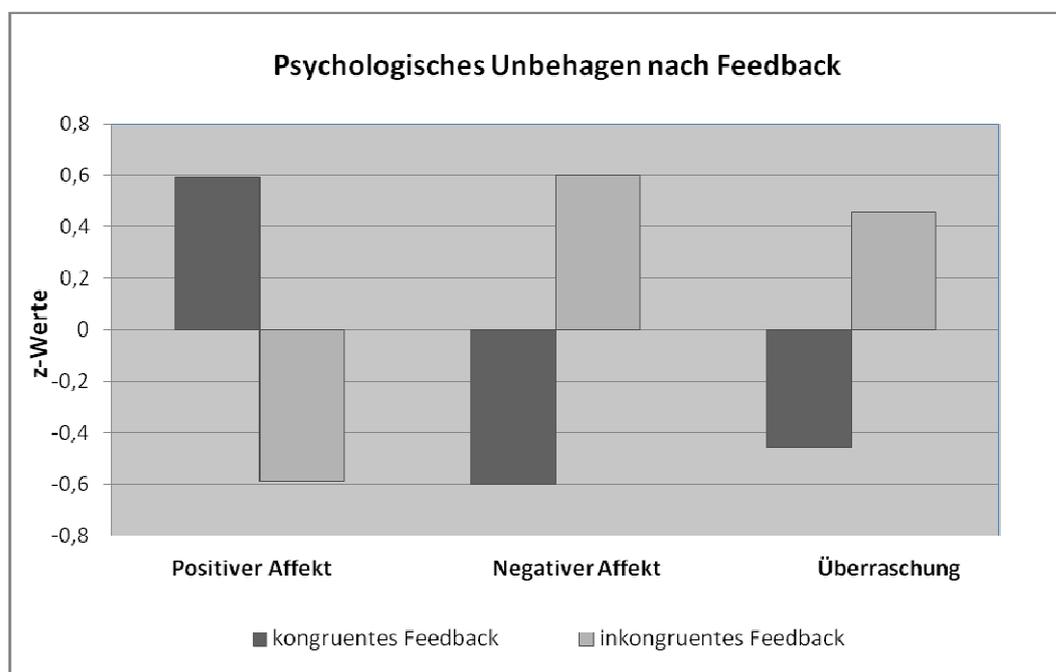


Abbildung 27: Erlebtes psychologisches Unbehagen bzw. Dissonanzerleben bei beiden Experimentalgruppen ($n = 48$) (z-Werte).

Es kann somit festgehalten werden, dass Untersuchungsteilnehmer, die ein zu ihrer eigenen Entscheidung inkongruentes, d. h. konfligierendes Feedback erhielten, signifikant und substantiell weniger positive Gefühle, aber mehr negative Gefühle erlebten und überraschter waren, als Teilnehmer, die kongruentes Feedback bekamen. Feedback, das im Konflikt mit

der eigenen Entscheidung steht, wird psychisch unangenehm erlebt und erzeugt demzufolge kognitive Dissonanz. Eine Möglichkeit, die entstandene Dissonanz zu reduzieren, besteht in der Revision der anfänglichen Entscheidung. Ob das festgestellte Dissonanzerleben bei der Experimentalgruppe „inkongruentes Feedback“ auch zu einer signifikanten Entscheidungsänderung führte, wird in Kapitel 9.3.3 untersucht.

Wahrgenommener Konflikt. Zusätzlich wurde kontrolliert, ob die Versuchspersonen die Kongruenz bzw. die Inkongruenz (den Konflikt) der Konsensinformation mit der eigenen Entscheidung richtig erkannten. Der wahrgenommene Konflikt bezüglich der Feedbackinformation wurde mit Hilfe der Frage „Wie sehr haben sich Ihre eigene Entscheidung und das Abstimmungsergebnis der anderen Studienteilnehmer unterschieden?“ erhoben.

Tabelle 20 veranschaulicht, dass der Mittelwert zum wahrgenommenen Konflikt bei der Feedbackbedingung „kongruent“ deskriptiv eindeutig geringer ($M = 2.50$, $SD = 0.78$) ausfiel als bei der Bedingung „inkongruent“ ($M = 9.08$, $SD = 1.02$).

Tabelle 20: Wahrgenommener Konflikt für beide Experimentalgruppen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.

	1) Kongruentes Feedback		2) Inkongruentes Feedback		P
	M	(SD)	M	(SD)	
Wahrgenommener Konflikt	2.50	(0.78)	9.08	(1.02)	$p = .000$

Anmerkung: Wertebereich von 1 niedrig bis 10 hoch.

Inferenzstatistisch konnte ein hochsignifikanter und zugleich in hohem Grade substanzieller Effekt beim wahrgenommenen Konflikt ($t(46) = -25.15$, $p = .000$; $d = 7.25$) nachgewiesen werden: Die Einschätzungen in den beiden Experimentalgruppen bezüglich des wahrgenommenen Konflikts unterschieden sich tatsächlich statistisch bedeutsam voneinander.

Zudem lagen wie erwartet bei der kongruenten Lernbedingung die Einschätzungen zum wahrgenommenen Konflikt im unteren Bereich des theoretischen Maximums von 10 (Minimum: 1, Maximum: 4), bei der inkongruenten Lernbedingung dagegen im oberen Bereich (Minimum: 7, Maximum: 10).

Alle Probanden der jeweiligen Feedbackbedingung hatten somit eine gute Erinnerung an die Mehrheitsverhältnisse und gaben folglich den Unterschied zwischen der eigenen Entscheidung und dem Abstimmungsergebnis der anderen Studienteilnehmer korrekt wieder.

Die Manipulationskontrolle konnte zeigen, dass die Operationalisierung der unabhängigen Variablen „Feedback“ erfolgreich war: Das Feedback konnte in dieser Untersuchung in der beabsichtigten Weise manipuliert werden. Darüber hinaus ergab die qualitative Auswertung der Antworten zum Verdachtsfragebogen (*suspicion check*), dass keiner der Probanden das eigentliche Untersuchungsziel der Studie erkannte (sozialer Einfluss durch Mehrheiten). Somit musste kein Versuchsteilnehmer von der Datenauswertung ausgeschlossen werden.

Insgesamt betrachtet konnten keine überzufälligen Unterschiede in den kognitiven, motivationalen sowie einstellungsbezogenen Lernvoraussetzungen und in der Lernzeit bei allen drei Lernbedingungen festgestellt werden. Die Randomisierung der Teilnehmer zu den drei Lernbedingungen war somit erfolgreich. Ferner gelang die experimentelle Manipulation der Kongruenz der Feedbackinformation. Die interne Validität der Studie konnte somit als gesichert gelten.

9.2 Analyse der parasozialen Interaktion mit der Medienfigur

Fragestellung 1: In welchem Ausmaß setzen sich die Lernenden mit der Medienfigur auseinander?

Nach der ersten getroffenen Entscheidung, unmittelbar nach der Präsentation der Dilemmageschichte durch den *talking head*, wurden zum ersten Messzeitpunkt (t1) die parasozialen Interaktionen und die wahrgenommene Attraktivität der Medienfigur bei allen drei Lernbedingungen erhoben. Es wurde angenommen, dass die empfundene Attraktivität der Medienfigur einen positiven Einfluss auf die Intensität der Auseinandersetzung mit der medialen Person hat. Demzufolge sollte sich eine signifikante und positive Korrelation der Attraktivität und der parasozialen Interaktion mit der Medienfigur zeigen.

Vorab wurde deskriptiv überprüft, wie intensiv sich die Lernenden mit der Medienfigur auseinandersetzen und wie attraktiv die Medienfigur eingeschätzt wurde. Folgende Tabelle 21 zeigt die Mittelwerte der beiden Variablen an.

Die Attraktivität der Medienfigur wurde vor dem theoretischen Hintergrund von 5.0 mittel bis hoch eingeschätzt. Der Skalenmittelwert betrug 3.38 ($SD = 0.70$), der Median betrug 3.40, der Modus 3.20, die Werte variierten zwischen 2.00 und 4.80. Die Lernenden empfanden die Medienfigur demnach als attraktiv.

Der Skalenmittelwert der parasozialen Interaktion mit der Medienfigur lag bei 3.25 ($SD = 0.69$) und variierte von 1.70 bis 4.90. Der Median und der Modus lagen jeweils bei 3.20. Deskriptiv betrachtet lag der Skalenmittelwert zur parasozialen Interaktion im Durch-

schnitt leicht über dem theoretischen Skalenmittel von 3.0. Es konnte somit eine mittlere bis hohe Motivation der Lernenden, sich intensiv mit der Medienfigur auseinanderzusetzen, festgestellt werden. Dabei fielen die kognitiven PSI-Prozesse ($M = 3.44$, $SD = 0.79$) etwas höher aus als die affektiven ($M = 3.09$, $SD = 0.89$).

Tabelle 21: Attraktivität der Medienfigur und parasoziale Interaktion mit der Medienfigur, insgesamt, d. h. über alle drei Gruppen hinweg. Mittelwerte (Standardabweichungen in Klammern), Minima, Maxima.

	M	(SD)	Min.	Max.
Attraktivität der Medienfigur	3.38	(0.70)	2.00	4.80
Parasoziale Interaktion (PSI)	3.25	(0.69)	1.70	4.90
Kognitive PSI-Prozesse	3.44	(0.79)	1.50	5.00
Affektive PSI-Prozesse	3.09	(0.89)	1.20	5.00

Anmerkung: Wertebereich von 1 niedrig bis 5 hoch.

Fragestellung 2: In welchem Zusammenhang steht die parasoziale Interaktion mit der wahrgenommenen Attraktivität der Medienfigur?

Um festzustellen, inwieweit die wahrgenommene Attraktivität der Medienfigur die Intensität der parasozialen Interaktionen (PSI) beeinflusst, wurden diese beiden Variablen miteinander korreliert.

Wie Tabelle 22 zu entnehmen, korrelierte die Attraktivität der Medienfigur positiv in mittlerer Höhe mit den parasozialen Interaktionen, der Zusammenhang war signifikant und substantiell ($r = .26$, $p < .05$): Je attraktiver die Medienfigur bewertet wurde, desto intensiver fielen die PSI-Prozesse aus (und umgekehrt).

Tabelle 22: Korrelation nach Pearson zwischen der Attraktivität der Medienfigur und der parasozialen Interaktion mit der Medienfigur.

	Attraktivität der Medienfigur
Parasoziale Interaktion mit Medienfigur	.26*

Anmerkungen: ** $p < .01$; * $p < .05$ (zweiseitige Signifikanzprüfung).

Die wahrgenommene Attraktivität der Medienfigur erhöhte somit die Motivation der Versuchspersonen, sich mit der medial dargestellten Person dezidierter auseinanderzusetzen und stellt einen wichtigen Einflussfaktor auf die parasozialen Interaktionen dar.

9.3 Analyse des Entscheidungsverhaltens

Alle drei Lernbedingungen mussten insgesamt drei Entscheidungen treffen, wie sie sich bezüglich der vorgeschlagenen Handlungsalternative, eine PID-HLA in Amerika durchführen zu lassen, verhalten würden („Ja“ vs. „Nein“). Die Entscheidungsfrage lautete: „Würden Sie an ihrer Stelle für eine Präimplantationsdiagnostik nach Amerika gehen?“.

Das Entscheidungsverhalten wurde in mehreren Schritten analysiert. Zunächst werden der zeitliche Verlauf der Richtung der Entscheidung (Kapitel 9.3.1) und der Sicherheit der Entscheidung (Kapitel 9.3.2) über alle drei Messzeitpunkte hinweg betrachtet (t1: Film, t2: Feedback, t3: Informationen). Im Anschluss daran befasst sich Kapitel 9.3.3 mit der Änderung der Entscheidung infolge des Feedbacks anderer zum zweiten Messzeitpunkt (t2) und der Bearbeitung der Zusatzinformationen zum dritten Messzeitpunkt (t3).

9.3.1 Richtung der Entscheidungen

Fragestellung 3: In welche Richtung gehen die Entscheidungen?

Zunächst wurde die Richtung der Entscheidungen zu allen drei Messzeitpunkten analysiert. Mit Richtung ist gemeint, ob sich die Probanden für („Ja“) oder gegen („Nein“) die Durchführung einer PID-HLA in Amerika entschieden haben.

Tabelle 23 gibt die absoluten Häufigkeiten der Richtung der Entscheidung für alle drei Lernbedingungen zu allen drei Messzeitpunkten wieder.

Tabelle 23: Richtung der Entscheidung (Ja vs. Nein) für alle drei Lernbedingungen zu allen drei Messzeitpunkten (zeitlicher Verlauf) (absolute Häufigkeiten).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback	
	Ja	Nein	Ja	Nein	Ja	Nein
Film (t1)	12	12	12	12	14	10
Feedback (t2)	10	14	11	13	14	10
Informationen (t3)	9	15	9	15	12	12

Die Richtung der ersten, vorläufigen Entscheidung zum ersten Erhebungszeitpunkt (t1) war sowohl bei der Bedingung „kongruentes“ als auch „inkongruentes Feedback“ deskriptiv identisch ausgeprägt. Es konnte eine neutrale Position konstatiert werden: Je 12 (50 %) Probanden würden für eine PID nach Amerika gehen, 12 (50 %) hingegen nicht. Die Teil-

nehmer der Gruppe „kein Feedback“ wiesen hingegen eine leichte Tendenz zur Pro-Position auf, 14 Teilnehmer (58.3 %) befürworteten eine PID, 10 (41.7 %) lehnten PID ab.

Nach Feedbackgabe über die Gesamtentscheidung der anderen Teilnehmer konnte zum zweiten Messzeitpunkt (t2) hinsichtlich der zweiten, vorläufigen Entscheidung bei beiden Experimentalgruppen ein Richtungswechsel der Gesamtentscheidungen von der Pro- zur Contra-Position festgestellt werden. Die Mehrheit von 14 bzw. 13 Teilnehmern (58.3 % bzw. 54.2 %) würden keine PID durchführen lassen, 10 bzw. 11 (41.7 % bzw. 45.8 %) dagegen schon. Wiederum 14 Teilnehmer der Gruppe „kein Feedback“ (58.3 %) entschieden sich für, 10 (41.7 %) gegen eine PID.

Bezüglich der dritten und finalen Entscheidung (t3) nahm bei beiden Experimentalgruppen die Ablehnung der PID weiter zu. Die Entscheidungen waren bei beiden Gruppen identisch ausgeprägt: Jeweils 15 (62.5 %) entschieden sich gegen eine PID, lediglich 9 (37.5 %) für eine PID. Bei der Gruppe „kein Feedback“ fiel die endgültige Entscheidung neutral aus, jeweils 12 (50 %) entschieden sich gegen bzw. für die Alternative einer PID zur Erzeugung eines „Retterkindes“.

Abbildung 28 gibt die Richtung der Entscheidungen insgesamt, d. h. über alle Gruppen hinweg, im zeitlichen Verlauf wieder:

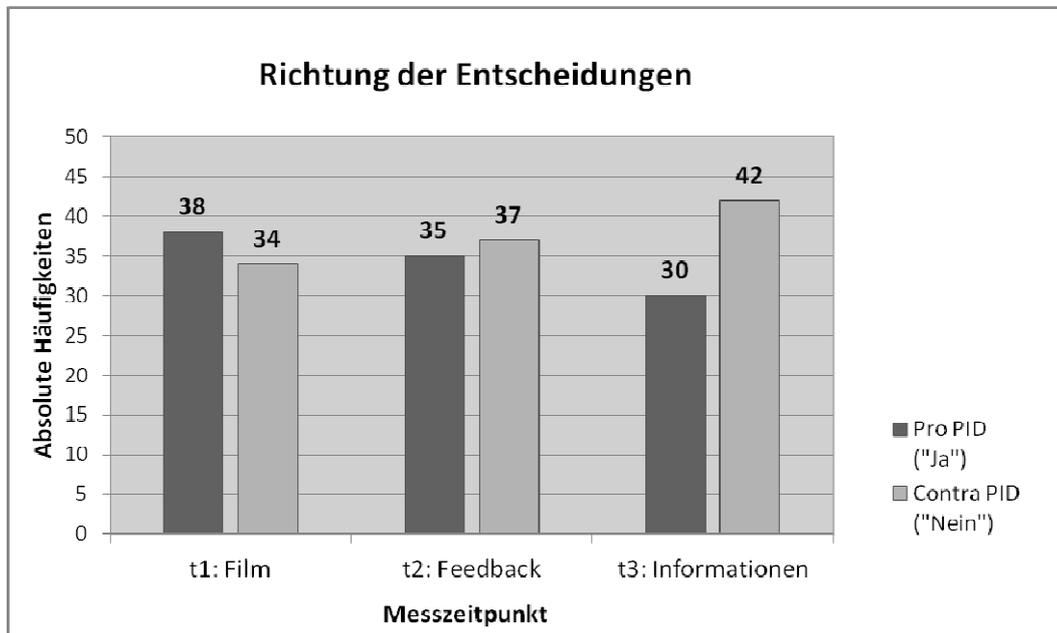


Abbildung 28: Richtung der Entscheidungen (Pro PID = „Ja“ vs. Contra PID = „Nein“) im zeitlichen Verlauf (Messzeitpunkte t1 bis t3), absolute Häufigkeiten (Gesamtgruppe).

Nach Präsentation des personalisierten Films zu ersten Messung (t1) konnte insgesamt betrachtet eine leichte Tendenz zur Pro-Position festgestellt werden. Die Mehrzahl von 38 Versuchspersonen (52.8 %) entschieden sich für die Alternative der PID, dagegen lehnten 34 Probanden (47.2 %) diese Alternative ab. Nach der Feedbackgabe zum zweiten Mess-

zeitpunkt (t2) konnte gegenüber der ersten Messung (t1) ein Richtungswechsel der Gesamtentscheidungen von der Pro- zur Contra-Position festgestellt werden. 37 Teilnehmer (51.4 %) entschieden sich nun gegen die Alternative einer PID, 35 (48.6 %) für eine PID. Nach Bearbeitung der Zusatzinformationen zur dritten Messung (t3) verstärkte sich der Richtungswechsel von der Pro- zur Contra-Position noch. Bei der finalen, endgültigen Gesamtentscheidung entschieden sich 42 Probanden (58.3 %) gegen und lediglich 30 (41.7 %) für eine PID.

Insgesamt betrachtet kann festgehalten werden, dass sich die Richtung der Entscheidung durch die Bearbeitung der Lernumgebung veränderte: Von einer anfänglichen Befürwortung der PID wird das gendiagnostische Verfahren zur Erzeugung eines „Retterkindes“ im Durchschnitt zum Schluss von der Mehrheit der Untersuchungsteilnehmer abgelehnt.

9.3.2 Sicherheit der Entscheidungen

Fragestellung 4: Inwieweit kommt es durch die Bearbeitung der Lernumgebung zu einer Zunahme der Entscheidungssicherheit und inwieweit beeinflusst die Kongruenz der Feedbackinformation die Entscheidungssicherheit?

Hypothesen. Es wird erwartet, dass die Entscheidungssicherheit durch die Bearbeitung der Lernumgebung im zeitlichen Verlauf zunimmt (Hypothese 1). Es wird zudem angenommen, dass die Entscheidungssicherheit im zeitlichen Durchschnitt nicht konstant bei den drei Lernbedingungen verläuft (Hypothese 2). Ferner wird erwartet, dass bei inkongruentem Feedback die Sicherheit der Entscheidung abnimmt, bei kongruentem Feedback dagegen die Sicherheit der Entscheidung zunimmt (Hypothese 3).

Ergebnisse. Zunächst werden die deskriptiven Werte der Sicherheit der drei getroffenen Entscheidungen für alle drei Lernbedingungen und insgesamt, d. h. über alle drei Gruppen hinweg, im zeitlichen Verlauf betrachtet (siehe folgende Tabelle 24).

Vor dem Hintergrund des theoretischen Maximums von 10 war die Sicherheit der Entscheidung nach der Präsentation des personalisierten Films (t1) über alle drei Lernbedingungen hinweg recht hoch ausgeprägt (Gesamtmittelwert: 5.94, $SD = 3.03$).

Am sichersten in der ersten, vorläufigen Entscheidung waren sich die Lernenden der Bedingung „inkongruentes Feedback“ ($M = 6.58$, $SD = 2.92$), etwas unsicherer war die Bedingung „kein Feedback“ ($M = 6.13$, $SD = 2.91$). Am unsichersten in ihrer Entscheidung war die Bedingung „kongruentes Feedback“ ($M = 5.12$, $SD = 3.21$). Bezüglich der zweiten, vorläufigen Entscheidung nach Erhalt des Feedbacks über die durchschnittliche Entscheidung der anderen Studienteilnehmer (t2) konnte deskriptiv insgesamt bei allen drei

Gruppen eine etwas höhere Sicherheit festgestellt werden (Gesamtmittelwert: 6.25, $SD = 3.03$). Erneut war die Gruppe „inkongruentes Feedback“ am sichersten in ihrer Entscheidung ($M = 6.92$, $SD = 3.22$), gefolgt von der Gruppe „kein Feedback“ ($M = 6.21$, $SD = 2.92$) und „kongruentes Feedback“ ($M = 5.63$, $SD = 2.93$). Nach Bearbeitung der Zusatzinformationen (t3) fiel die Sicherheit nach der dritten und finalen Entscheidung über alle drei Gruppen hinweg deutlich höher aus (Gesamtmittelwert: 6.83, $SD = 2.68$). Die sicherste Entscheidung gab nach Bearbeitung der Zusatzinformationen nun die Bedingung „kein Feedback“ ab ($M = 6.96$, $SD = 2.10$). Etwas entscheidungsunsicherer waren die Bedingung „inkongruentes“ ($M = 6.88$, $SD = 3.30$) und „kongruentes Feedback“ ($M = 6.67$, $SD = 2.62$).

Bei genauerer Betrachtung der deskriptiven Werte wird deutlich, dass nur bei den Bedingungen „kongruentes“ und „kein Feedback“ eine deutliche Zunahme der Entscheidungssicherheit durch die Bearbeitung der Zusatzinformationen verzeichnet werden konnte. Bei Teilnehmern, die „inkongruentes Feedback“ erhielten, nahm die Entscheidungssicherheit nach Bearbeitung der Informationen deskriptiv sogar geringfügig ab.

Tabelle 24: Entscheidungssicherheit im zeitlichen Verlauf (Messzeitpunkte t1 bis t3) für alle drei Lernbedingungen und gesamt. Mittelwerte (Standardabweichungen in Klammern).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		Gesamt	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Film (t1)	5.12	(3.21)	6.58	(2.92)	6.13	(2.91)	5.94	(3.03)
Feedback (t2)	5.63	(2.93)	6.92	(3.22)	6.21	(2.92)	6.25	(3.03)
Information (t3)	6.67	(2.62)	6.88	(3.30)	6.96	(2.10)	6.83	(2.68)

Anmerkung: Wertebereich von 1 niedrig bis 10 hoch.

Um den Einfluss des Gruppenfaktors „Lernbedingung“ sowie den Einfluss des Messwiederholungsfaktors auf die Entscheidungssicherheit zu testen, wurde eine zweifaktorielle Varianzanalyse mit Messwiederholung durchgeführt. Es ergab sich kein signifikanter Haupteffekt des Gruppenfaktors „Lernbedingung“ für die Entscheidungssicherheit ($F(2, 69) = 0.91$, *n.s.*). Im zeitlichen Durchschnitt bestanden keine signifikanten Unterschiede in den Messwerten zur Entscheidungssicherheit zwischen den drei Gruppen. Es kann folglich von einem konstanten zeitlichen Verlauf der Entscheidungssicherheit bei allen drei Gruppen ausgegangen werden.

Die Sicherheit der Entscheidung wird nicht bedeutsam durch die Lernbedingung beeinflusst. Hypothese 2 konnte nicht bestätigt werden.

Bezüglich des Messwiederholungsfaktors konnte zunächst festgestellt werden, dass die Annahme der Sphärizität verletzt wurde, da der Mauchly-Test signifikant war ($\chi^2(2) = 13.31, p < 0.1$). Infolgedessen wurden die Freiheitsgrade mit Hilfe von Greenhouse-Geisser ($\varepsilon = .85$) korrigiert. Es zeigte sich ein signifikanter Haupteffekt des Messwiederholungsfaktors ($F(1.70, 117.17) = 5.00, p < .05; \eta^2p = .07$). Somit ergab sich ein signifikant nicht konstanter Verlauf der Entscheidungssicherheit zu den drei Messzeitpunkten insgesamt, d. h. über alle Gruppen hinweg. Mit Hilfe eines Post-Hoc-Tests wurde untersucht, welche der drei Messzeitpunkte sich voneinander unterscheiden. Die Ergebnisse der paarweisen Vergleiche mit Bonferroni-Korrektur zeigten signifikante Unterschiede zwischen dem ersten und dritten Messzeitpunkt ($d_i = -0.89, p < .01$).

Die Entscheidungen nach Bearbeitung der Zusatzinformationen fallen substantiell sicherer aus als vor der Bearbeitung der Informationen. Hypothese 1 konnte bestätigt werden: Die Entscheidungssicherheit nahm durch die Bearbeitung der Lernumgebung kontinuierlich zu.

Zwischen den drei Gruppen und den drei Messzeitpunkten lagen keine Interaktionen vor ($F(3.40, 117.17) = 1.02, n.s.$). Die Sicherheit der Entscheidung der drei Lernbedingungen hat sich zu keinem der drei Erhebungszeitpunkte signifikant und substantiell unterschieden, Hypothese 3 konnte damit nicht bestätigt werden.

Abbildung 29 gibt zusammenfassend die Ergebnisse zum zeitlichen Verlauf der Sicherheit der Entscheidung für alle drei Gruppen wieder und verdeutlicht, dass die Messverläufe annähernd parallel verlaufen und keine Wechselwirkungen aufweisen.

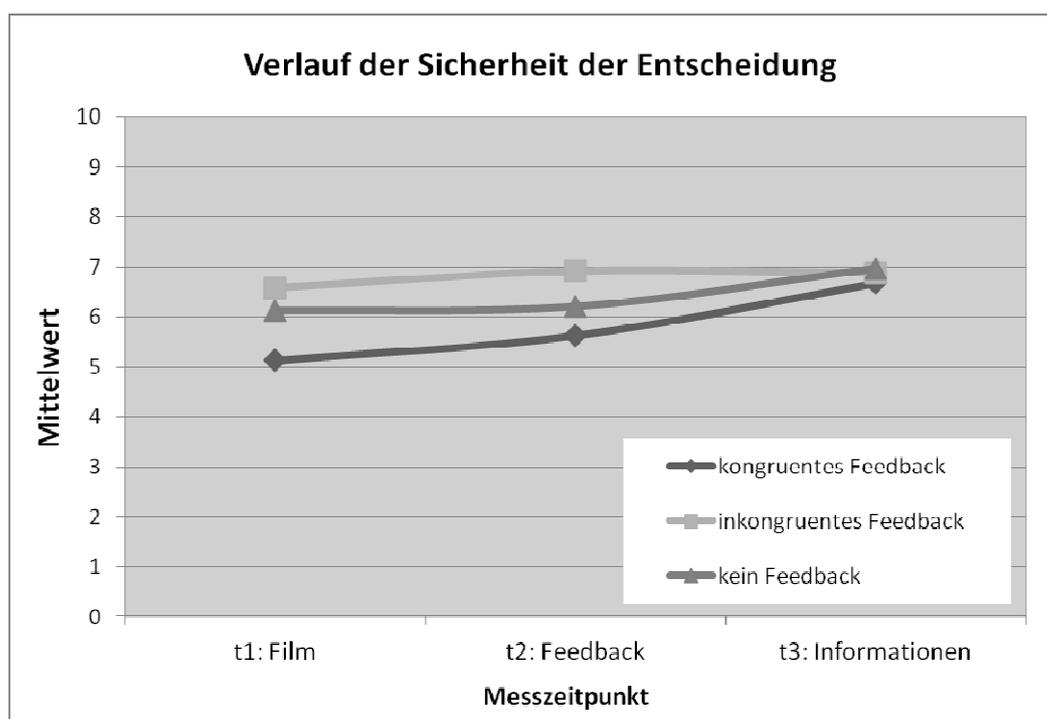


Abbildung 29: Zeitlicher Verlauf der Sicherheit der Entscheidung für alle drei Gruppen (Messzeitpunkte t1-t3), Mittelwerte, Liniendiagramm.

Die Sicherheit der Entscheidung fiel insgesamt und über alle drei Messzeitpunkte hinweg durchgängig hoch aus und nahm bei allen drei Lernbedingungen im zeitlichen Verlauf kontinuierlich zu. Es bestanden signifikante und substantielle Unterschiede hinsichtlich des zeitlichen Verlaufs der Sicherheit der Entscheidung: Durch die Bearbeitung der Zusatzinformationen fällt die Entscheidung signifikant sicherer aus als nach Erhalt des Feedbacks.

Die Entscheidungssicherheit wird demzufolge durch die Bearbeitung der zusätzlichen Informationen positiv beeinflusst. Das Feedback der anderen Studienteilnehmer hatte dagegen keinen Einfluss auf die Sicherheit der Entscheidung.

Zusammenfassend betrachtet lässt sich über alle drei Gruppen hinweg im zeitlichen Verlauf durch die Auseinandersetzung mit der Lernumgebung eine Zunahme der Sicherheit der getroffenen Entscheidungen feststellen. Die Erhöhung der Sicherheit der Entscheidung konnte auch in den Essays, die zur ergänzenden Auswertung herangezogen wurden, bestätigt werden. Insgesamt gaben 5 Probanden eine Zunahme der Sicherheit durch die Bearbeitung der Zusatzinformationen an, wie diese 2 Teilnehmerantworten belegen:

„Wie ein paar zusätzliche Informationen und einige Gedanken mehr zur PID die Meinung schon verändern. Die Gründe sind immer noch die aus Frageblock 2, nur dass die psycho-sozialen Bedenken größer geworden sind. Die Entscheidung fiel mir vor allem deshalb nun schon wesentlich leichter.“ [02b2_t3_Nein]

„Durch die mehrmaligen Wiederholungen des Ausdrucks „Retter-Kind“ und den Verweis auf die psycho-sozialen Probleme, die entstehen könnten, fiel mir die Entscheidung schon viel leichter als die beiden Male davor.“ [69b1_t3_Nein]

Ein Untersuchungsteilnehmer gab an, dass die Zusatzinformationen hilfreich bei der Entscheidungsfindung waren:

„All diese Informationen helfen mir, meine Entscheidung zu treffen.“ [25b3]

Allerdings äußerten auch 2 Probanden eine Abnahme der Entscheidungssicherheit bzw. eine Zunahme der Schwierigkeit der Entscheidung infolge der Bearbeitung der Zusatzinformationen, z. B.:

„Nachdem ich alles durchgelesen habe, habe ich es schwieriger gefunden, die Entscheidung zu treffen.“ [18b1_t3_Ja]

9.3.3 Änderung der Entscheidungen

Im nächsten Analyseschritt wurde untersucht, ob das Feedback über die Entscheidung der anderen Untersuchungsteilnehmer sowie die Bearbeitung der Zusatzinformationen einen Einfluss auf die Entscheidungsfindung haben, d. h. zu einer Revision der zuvor getroffenen Entscheidung führen.

Fragestellung 5: Inwieweit führt inkongruentes Feedback zu einer Änderung der Entscheidung?

Hypothese. Es wird angenommen, dass inkongruentes Feedback zu einer häufigeren Änderung der Entscheidung führt als kongruentes oder kein Feedback (Hypothese 4).

Tabelle 25 veranschaulicht die absoluten Häufigkeiten der Revision der ersten, vorläufigen Entscheidung nach Erhalt des Feedbacks über die durchschnittliche Gesamtentscheidung der anderen Studienteilnehmer zum zweiten Messzeitpunkt (t2).

Tabelle 25: Entscheidungsänderung nach Feedback (t2) für alle drei Lernbedingungen und insgesamt. Absolute Häufigkeiten.

		1) Kongruentes Feedback	2) Inkongruentes Feedback	3) Kein Feedback	Gesamt
Revision	ja	2	3	0	5
	nein	22	21	24	67

Lediglich 3 Teilnehmer (13 %) der Bedingung „inkongruentes Feedback“ revidierten ihre Entscheidung nach Erhalt der Information, dass die Mehrheit von 86 % der Studienteilnehmer sich anders entschieden hat. Die überwältigende Mehrheit von 21 Teilnehmern (87 %) blieb bei ihrer anfänglich getroffenen Entscheidung.

Bei der Bedingung „kongruentes Feedback“ kam es, trotz der Information, dass die Mehrheit von 86 % der Studienteilnehmer sich genauso entschieden hat, bei 2 Personen (8 %) zu einer Revision der zuerst getroffenen Entscheidung. Beide Personen gaben im Essay leider keine Begründung bezüglich ihrer Entscheidungsrevision an. 22 Personen (92 %) behielten ihre Entscheidung bei. In der Bedingung „kein Feedback“ kam es erwartungsgemäß zu keiner Entscheidungsänderung, alle 24 Probanden entschieden sich genauso wie beim ersten Mal.

Insgesamt betrachtet, d. h. über alle drei Lernbedingungen hinweg, behielt die überwältigende Mehrheit von 67 Teilnehmern (93.1 %) ihre anfänglich getroffene Entscheidung bei, lediglich 5 (6.9 %) Versuchspersonen revidierten ihre Entscheidung nach Erhalt der Feedbackinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung der anderen Versuchsteilnehmer. Deskriptiv betrachtet lässt sich somit keine wesentliche Änderung der Entscheidung von der ersten Messung nach Präsentation des Films zur zweiten Messung nach Erhalt des Feedbacks bei allen drei Gruppen verzeichnen.

Zur Signifikanzprüfung wurde der Freeman-Halton-Test verwendet, da die Voraussetzungen des Chi-Quadrat-Test, dass alle erwarteten Häufigkeiten mindestens 5 betragen sollen, verletzt war. Die Häufigkeiten der Entscheidungsrevision in den drei Lernbedingungen

unterschieden sich erwartungsgemäß nicht signifikant voneinander, die Signifikanzgrenze von 5 % wurde bei weitem verfehlt ($\chi^2 = 2.99, n.s.$).

Es kann somit festgehalten werden, dass Hypothese 4 nicht bestätigt werden konnte: Inkongruentes Feedback führte zu keiner signifikant häufigeren Entscheidungsrevision als kongruentes oder kein Feedback. Die Feedbackinformation hat demzufolge keinen Einfluss auf das Entscheidungsverhalten.

Zur Eruierung möglicher Ursachen der ausbleibenden Entscheidungsrevision nach entscheidungsinkongruenten Feedback wurden zusätzlich die Essays des zweiten Erhebungszeitpunkts herangezogen.

Die Auswertung der Essays ergab zunächst, dass sowohl die entscheidungskongruente als auch die entscheidungsinkongruente Feedbackinformation glaubwürdig gewesen zu sein schien. In keinem Essay wurde von den Probanden der Bedingungen „kongruentes“ oder „inkongruentes Feedback“ die Glaubwürdigkeit der Konsensinformation angezweifelt. Folgendes Teilnehmerzitat aus einem Kurzesay belegt, dass die Konsensinformation für glaubwürdig gehalten wurde:

„Ich finde es sehr erschreckend, dass das kollektive Bewusstsein schon so mit der Thematik „bearbeitet“ wurde und die Leute so einen Eingriff größtenteils gut heißen.“ [35b2_t2_Nein]

Von den insgesamt 5 Probanden, die ihre Entscheidung nach der Feedbackgabe änderten, gaben 3 Gründe für ihre Entscheidungsrevision an, z. B.:

„Ich habe mich nun umentschieden, ich würde zwar das Leben meines Sohns retten wollen, aber es gibt zum einen ja auch Methoden in Deutschland und zum anderen finde ich es nicht verantwortbar, wenn meinem zweiten Kind dabei etwas passiert. Ich denke, dadurch dass diese Technik noch so neu ist und die Auswirkungen unbekannt sind, würde ich mich in Verbindung mit meinen moralischen Bedenken aus Frageblock 1 dagegen entscheiden.“ [22b2_t2_Nein]

Dieser Studienteilnehmer bemerkte bereits im vorangegangenen ersten Essay, dass ihm die Entscheidung sehr schwer gefallen sei und er die PID nur anwenden würde, wenn das „Retterkind“ dabei keinen Schaden nimmt. Die Konfrontation mit einer entscheidungsinkongruenten Mehrheitsposition führte demnach zu einer Revision der anfänglichen, sehr unsicheren Entscheidung, in der die PID noch unter Vorbehalt befürwortet wurde.

Ein anderer Proband begründete seine Entscheidungsänderung folgendermaßen:

„Bei genauerem Überdenken kamen mir Zweifel bezüglich der „Richtigkeit“ meiner vorherigen Entscheidung. Das Problem ist natürlich, dass zwar einem Kind geholfen werden kann, ein anderes dafür aber bewusst ausgewählt und benützt wird, was man auch sehr kritisch sehen kann (wg. späteren evtl. psychischen Folgen für das Kind). Beim ersten Fragenblock habe ich mich sehr von meinem Gefühl für das kranke Kind und der Verzweiflung der Mutter leiten lassen.“ [65b2_t2_Nein]

Dieser Versuchsteilnehmer gab im ersten Essay (in Übereinstimmung mit 5 weiteren Probanden) auch an, dass im Hinblick auf seine endgültige Entscheidung noch weiterer Informationsbedarf besteht:

„Allerdings müssten für meine endgültige Entscheidung noch mehr Informationen vorhanden sein, z. B. wie sicher es ist, dass der optimale Spender auf diese Art gefunden werden kann und welche Risiken damit verbunden sind.“ [65b2_t1_Ja].

Wie diese beiden Äußerungen verdeutlichen, waren sich die Probanden, die ihre Entscheidung revidierten, generell in ihrer Entscheidung sehr unsicher, was zur Folge hatte, dass infolge der Konfrontation mit einer zur eigenen Entscheidung konfligierenden Mehrheitsposition die zuvor getroffene Entscheidung revidiert wurde (siehe auch Kapitel 9.6).

Bei der überwiegend ausbleibenden Änderung der Entscheidung nach inkongruentem Feedback dürften ethisch-moralische Grundwerte eine nicht zu unterschätzende Rolle gespielt haben, wie dieses Zitat illustriert:

„Die Tatsache, dass die anderen Studienteilnehmer sich für eine andere Möglichkeit entschieden haben, bringt mich nicht von meiner Entscheidung so leicht ab, die auf ethisch- und moralischen Grundwerten basiert.“ [60b2_t2_Nein]

Dies schließt jedoch nicht aus, dass der Erhalt einer inkongruenten Konsensinformation bei einigen Teilnehmern durchaus auch zu einer Abnahme der Entscheidungssicherheit führte, so äußerte eine Versuchsperson beispielsweise:

„Obwohl es mich verunsichert hat, dass die Mehrheit anders entschieden hat, denke ich noch immer, dass Frau Kreidler nach Amerika fahren sollte.“ [55b2_t2_Ja]

Ein anderer Teilnehmer bemerkte, dass die Konsensinformation für ihn keinen zusätzlichen Informationsgehalt in Bezug auf das Treffen seiner persönlichen Entscheidung ergab:

„Ich habe mich genauso wie beim ersten Mal entschieden, da ich denselben Informationsgehalt über die beiden Hauptpersonen (die beiden betroffenen Kinder) habe wie vorher. Wie andere Menschen in dieser Situation urteilen würden, bleibt ihnen überlassen, im Grunde ist es ja mein Gewissen dem gegenüber ich mich verantworten muss.“ [40b1_t2_Ja]

Dass hoher Konsens auch eine gewisse soziale Validierung und soziale Unterstützung der eigenen Entscheidung impliziert, verdeutlicht dieses Zitat eines Probanden der Lernbedingung „kongruent“:

„Ich wurde in meiner ersten Entscheidung ein bisschen mehr bestätigt, da sich andere auch so entschieden hätten.“ [53b1_t2_Ja]

Die präferenzkonsistente Feedbackinformation wurde von diesem Probanden als ergänzender Hinweis auf die Richtigkeit der eigenen getroffenen Entscheidung interpretiert.

Zusammenfassend betrachtet kann festgehalten werden, dass entgegen der Vorannahmen die Konfrontation mit einer zur eigenen Entscheidung inkongruenten Mehrheitsposition

nicht zu einer signifikanten und statistisch bedeutsamen Entscheidungsrevision führte, die anfänglich getroffene Entscheidung wurde beibehalten. Hypothese 4 konnte somit nicht bestätigt werden.

Eine präferenzinkonsistente Mehrheitsmeinung wird nicht ungeprüft in die eigene Entscheidung übernommen, wie dieses Zitat einer Versuchsperson der Lernbedingung „inkongruentes Feedback“ nochmals belegt:

„Trotz der Tatsache, dass die Mehrheit sich für einen solchen Eingriff entscheiden würde, bleibe ich bei meiner Meinung, da ich nach wie vor der Meinung bin, dass man keinem Kind eine solche Grausamkeit antun darf, um ein anderes zu retten, abgesehen davon, dass das kranke Kind dem gesunden gegenüber vermutlich ein Leben lang Schuldgefühle hätte, da die Eltern es nicht um seiner selbst willen, sondern um des kranken Kindes willen.“ [56b2_t2_Nein]

Fragestellung 6: Inwieweit führt die Bearbeitung der Zusatzinformationen zu einer Änderung der Entscheidung?

In einem nächsten Schritt wurde überprüft, ob die Bearbeitung der Zusatzinformationen (t3) zu einer signifikanten Änderung der zweiten, vorläufigen Entscheidung führt. Tabelle 26 gibt die absoluten Häufigkeiten der Änderung der Entscheidung nach Bearbeitung der Zusatzinformationen für alle drei Lernbedingungen wieder.

Tabelle 26: Entscheidungsänderung nach Bearbeitung der Zusatzinformation (t3) für alle drei Lernbedingungen und insgesamt. Absolute Häufigkeiten.

		1) Kongruentes Feedback	2) Inkongruentes Feedback	3) Kein Feedback	Gesamt
Revision	ja	3	2	2	7
	nein	21	22	22	65

Die überwiegende Mehrheit von 65 Probanden (90.3 %) behielt auch nach der Bearbeitung der Zusatzinformationen ihre Entscheidung bei, lediglich bei 7 Probanden (9.7 %) kam es zu einer Revision der zweiten, vorläufigen Entscheidung. Davon entfielen 3 Entscheidungsrevisionen (12.5 %) auf die Teilnehmer der Bedingung „kongruentes Feedback“, jeweils 2 (8.3 %) auf die Bedingungen „inkongruentes“ und „kein Feedback“.

Erwartungsgemäß ergab der Freeman-Halton-Test, dass der Unterschied zwischen den drei Lernbedingungen hinsichtlich der Änderung von der zweiten, vorläufigen zur dritten, finalen Entscheidung nicht signifikant war ($\chi^2 = 0.45$, *n.s.*). Auffällig ist, dass anhand einer Einzelfallanalyse festgestellt werden konnte, dass 6 von 7 Probanden ihre Entscheidung von „ja“ auf „nein“ revidierten. Von den insgesamt 7 Probanden, bei denen die Bearbei-

tung der Zusatzinformationen zu einer Entscheidungsänderung führte, gaben 4 Begründungen für ihre Entscheidungsrevision an, z. B.:

„Mit dem Hintergrundwissen, das ich über diesen Fall erhalten habe, wird mir bewusst, dass mit dieser Behandlung viele Risiken/Spätfolgen für das Geschwisterkind etc. entstehen können. Ich bin nicht der Meinung, dass diese Behandlung jedem frei zur Verfügung stehen sollte. Es sollten klare Grenzen für so ein Verfahren geschaffen werden.“ [63b1_t3_Nein]

Die Konfrontation mit möglichen Gegenpositionen durch die Bearbeitung der Zusatzinformationen führte offensichtlich bei diesem Versuchsteilnehmer zu einer Revision der zweiten, vorläufigen Entscheidung.

Zusammenfassend bleibt festzuhalten, dass über alle drei Gruppen und Messzeitpunkte hinweg, die Mehrheit von 60 Teilnehmern (83.3 %) ihre anfänglich getroffene Entscheidung beibehielt, lediglich eine Minderheit von 12 (16.7 %) revidierte ihre Entscheidung infolge des Feedbacks anderer oder der Bearbeitung der Zusatzinformationen. Alle 12 Probanden änderten ihre Entscheidung dabei jeweils nur einmal, d. h. eine hypothetisch mögliche zweimalige Revision der Entscheidung sowohl nach Feedbackgabe als auch nach Bearbeitung der Zusatzinformationen trat in dieser Untersuchung nicht auf.

Sowohl die Feedbackinformation als auch die Bearbeitung der Zusatzinformationen führten nicht zu einer signifikanten häufigeren Revision der zuvor getroffenen Entscheidung, vielmehr wurde die anfänglich getroffene Entscheidung beibehalten. Folgendes Teilnehmerzitat macht noch einmal deutlich, dass sowohl die Feedbackinformation als auch die Zusatzinformationen keinen bedeutsamen Einfluss auf die Entscheidungsfindung bezüglich des medizinethischen Dilemmas zur PID-HLA hatten:

„Die Informationen, die ich eben erhalten habe, haben meine Entscheidung, die sehr klar ist, nur bestärkt oder waren irrelevant. Es ist eine Grundsatzentscheidung, da spielt es keine Rolle, was technisch möglich wäre und wie viele Leute diese Technik schon nutzen oder gut heißen.“ [35b2_t3_Nein]

Weitere 5 Versuchspersonen gaben ebenfalls an, dass die Zusatzinformationen ihre Entscheidung bestätigt und bestärkt haben.

9.4 Analyse der Argumentationsqualität

Als Indikator des Lernfortschritts wurde die Argumentationsqualität zum ersten und dritten Messzeitpunkt (d. h. vor und nach Bearbeitung der Zusatzinformationen) herangezogen (Prä-/Post-Messung). Dazu wurden die Essays, in denen die Probanden zu beiden Messzeitpunkten ihre Entscheidung schriftlich begründeten, mithilfe des in Kapitel 8.7.1 beschriebenen Kategorienschemas kodiert und ausgewertet. Es wurde erwartet, dass sich die

zusätzlichen, vertiefenden Informationen positiv auf die Qualität der Argumentation zum dritten Messzeitpunkt auswirken. Insgesamt konnten für die 72 untersuchten Probanden 144 gültige Datensätze ausgewertet werden.

Da die interne Validität der experimentellen Studie aufgrund der tendenziell überzufälligen Unterschiede in der realisierten Lernzeit gefährdet war (siehe Kapitel 9.1.2), wurde vorab überprüft, inwieweit die Dauer der Auseinandersetzung mit der Lernumgebung mit der Argumentationsqualität zum ersten und dritten Messzeitpunkt assoziiert war. Zwischen der Lernzeit und der Argumentationsqualität zu beiden Messungen bestanden nur äußerst schwache Zusammenhänge, die bei weitem statistisch nicht bedeutsam waren ($r = .01$, *n.s.* bzw. $r = .10$, *n.s.*). Die Argumentationsqualität war somit unabhängig von der individuellen Bearbeitungs- bzw. Lernzeit der Probanden. Demnach handelte es sich bei der realisierten Lernzeit um keine tatsächliche Störvariable, die interne Validität der Studie war nicht bedroht. Auf eine kovarianzanalytische Kontrolle der Lernzeit wurde daher verzichtet.

Deskriptive Befunde zur Argumentationsqualität zum ersten und dritten Messzeitpunkt für die Gesamtgruppe. In Bezug auf die Argumentationsqualität konnten theoretisch 5 Punkte zu beiden Messzeitpunkten erzielt werden. Die Argumentationsqualität der Probanden aller drei Lernbedingungen lag zum ersten Messzeitpunkt leicht unter dem mittleren Bereich der Punkteskala. Der Mittelwert betrug 2.50 ($SD = 0.93$). Die beste Leistung lag bei 5 Punkten, die schlechteste bei 0 Punkten. Am häufigsten erreichten die Versuchsteilnehmer 2 Punkte, der Median betrug ebenfalls 2.

Folgende Abbildung 30 gibt die Verteilung der Punktwerte für die Argumentationsqualität zum ersten Messzeitpunkt für die Probanden aller drei experimentellen Gruppen wieder. Lediglich 6 Teilnehmer (8.3 %) konnten ihre Entscheidung nicht angemessen rechtfertigen oder valide begründen (*Stufe 0: Keine Rechtfertigung* bzw. *Stufe 1: Rechtfertigung ohne valide Begründung*). Die Hälfte der Versuchsteilnehmer (36; 50 %) erreichten die Stufe 2 (*Rechtfertigung mit einfacher Begründung*) und waren damit in der Lage, ein bis zwei Argumente zur Begründung ihrer Entscheidung zu formulieren. Ein Viertel (18; 25 %) der Probanden konnten drei oder mehr entscheidungsunterstützende Argumente generieren (*Stufe 3: Rechtfertigung mit elaborierter Begründung*). 11 Probanden (15.3 %) konnten ihre Entscheidung nicht nur elaboriert begründen, sondern berücksichtigten auch (mindestens) eine Gegenposition zu ihrem eigenen Standpunkt (*Stufe 4: Rechtfertigung mit elaborierter Begründung und Gegenposition*). Das höchste Niveau der Argumentationsqualität (*Stufe 5: Rechtfertigung mit elaborierter Begründung, Gegenposition und Widerlegung*) erreichte lediglich ein Versuchsteilnehmer (1.4 %). Dieser Proband konnte seine Entscheidung elaboriert begründen, (mindestens) eine Gegenposition nennen und durch Widerlegungen entkräften.

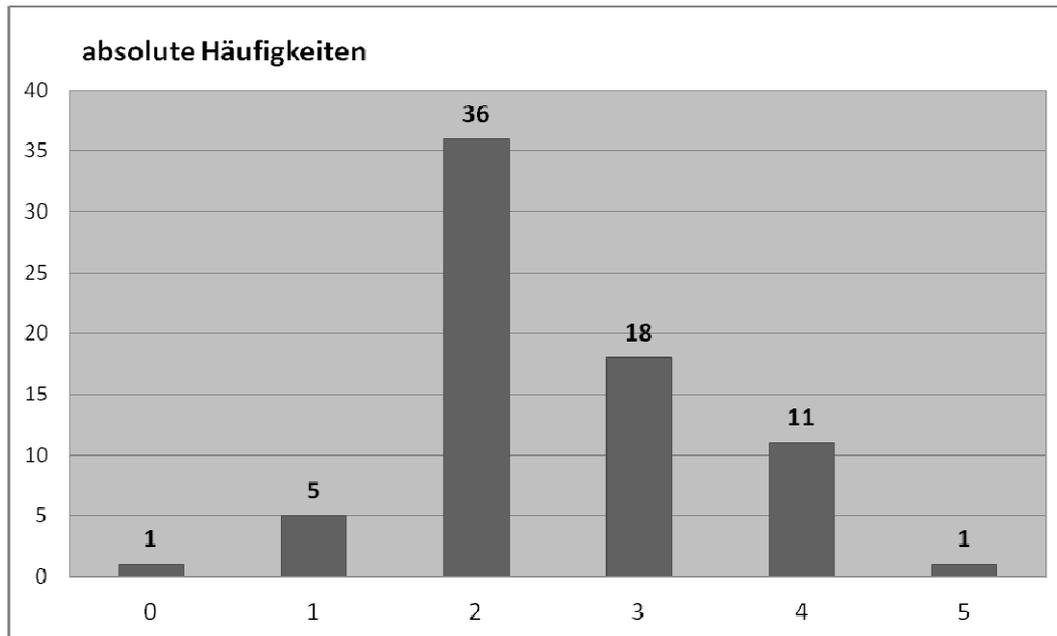


Abbildung 30: Argumentationsqualität zum ersten Messzeitpunkt: Verteilung der Punktwerte, absolute Häufigkeiten (Gesamtgruppe).

Insgesamt betrachtet konnte somit die Mehrzahl der Probanden ihre erste, anfängliche Entscheidung in Bezug auf die Dilemmageschichte zur PID-HLA zum ersten Messzeitpunkt im Durchschnitt mit etwa zwei bis drei Begründungen rechtfertigen.

Der Mittelwert der Argumentationsqualität zum dritten Messzeitpunkt, d. h. nach Bearbeitung der zusätzlichen, vertiefenden Informationen, fiel für die Gesamtgruppe deskriptiv höher aus als zum ersten Messzeitpunkt und lag nur noch sehr leicht unter dem mittleren Bereich der Skala ($M = 2.85$, $SD = 0.91$). Der Modus und der Median betragen jeweils 3. Das Minimum lag bei 0, das Maximum bei 5 Punkten. Bei knapp einem Drittel der Probanden (23; 31.9 %) konnte deskriptiv eine Änderung der Argumentationsqualität festgestellt werden. Bei den restlichen 49 Probanden (68.1 %) blieb die Argumentationsqualität dagegen auf einem konstanten Niveau.

Folgende Abbildung 31 veranschaulicht die Verteilung der Punktwerte für die Argumentationsqualität zum dritten Messzeitpunkt für die Gesamtgruppe. Gegenüber der ersten Messung lässt Abbildung 31 erkennen, dass die Mehrzahl der Probanden (31; 43.1 %) nach Bearbeitung der Lernumgebung nun auf Stufe 3 (*Rechtfertigung mit elaborierter Begründung*) argumentierten, d. h. drei oder mehr Begründungen zur Rechtfertigung ihrer finalen Entscheidung anführen konnten. Das Niveau 2 der Argumentationsqualität (*Rechtfertigung mit einfacher Begründung*) erreichten fast ein Drittel der Versuchspersonen (22; 30.6 %). In rund ein Fünftel der Entscheidungsbegründungen (14; 19.4 %) konnten neben entscheidungsunterstützenden Argumenten auch eine oder mehr Gegenpositionen genannt werden (*Stufe 4: Rechtfertigung mit elaborierter Begründung und Gegenposition*). Analog zum

ersten Messzeitpunkt wurden auch zur dritten Messung mögliche Gegenpositionen nur äußerst selten überzeugend widerlegt: Lediglich 2 (2.8 %) Probanden waren dazu in der Lage. Drei Entscheidungsbegründungen (3; 4.2 %) konnten als qualitativ sehr niedrig eingestuft werden (*Stufe 0: Keine Rechtfertigung* bzw. *Stufe 1: Rechtfertigung ohne valide Begründung*).

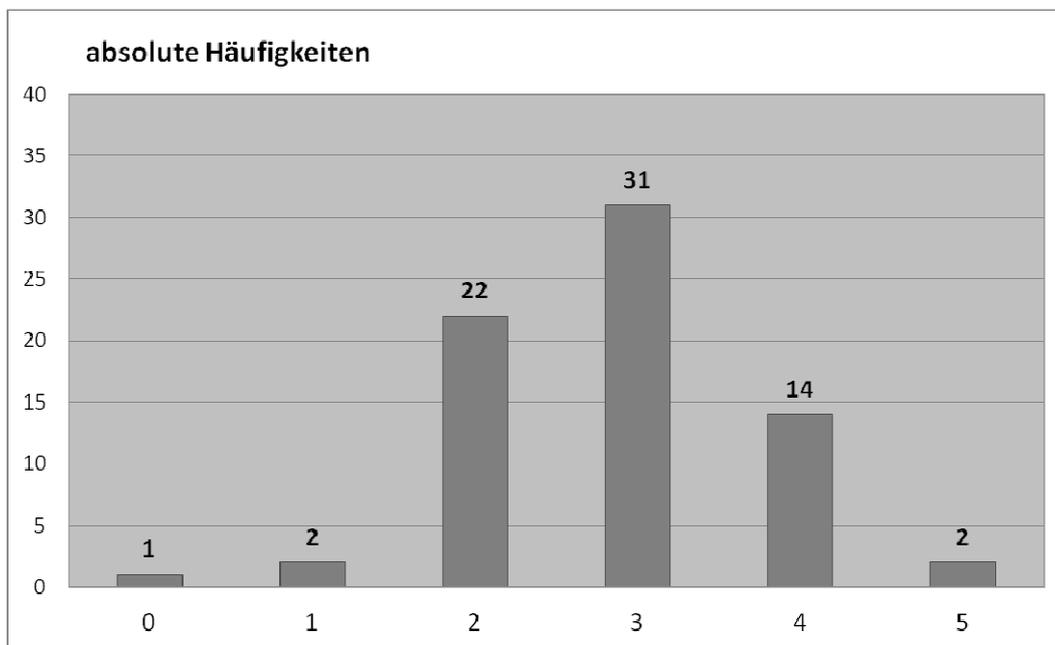


Abbildung 31: Argumentationsqualität zum dritten Messzeitpunkt: Verteilung der Punktwerte, absolute Häufigkeiten (Gesamtgruppe).

Nach der Bearbeitung der vertiefenden Informationen konnte die Mehrzahl der Versuchspersonen ihre Entscheidung elaborierter begründen. Allerdings offenbaren die Verteilungen der Punktwerte der Argumentationsqualität zu beiden Messungen auch, dass die überwiegende Mehrheit der Probanden nur Argumente zur Unterstützung der eigenen Entscheidung darlegten, eine Evaluation von Gegenargumenten fand nicht statt.

Fragestellung 7: Inwieweit kommt es durch die Bearbeitung der Lernumgebung zu einem Lernfortschritt und inwiefern wird der Lernfortschritt durch inkongruentes Feedback und durch die Bearbeitung der Zusatzinformationen gefördert?

Hypothesen. Es wurde postuliert, dass durch die Bearbeitung der Lernumgebung ein deutlicher Lernfortschritt erzielt wird. Es sollte sich ein signifikanter Effekt des Messwiederholungsfaktors ergeben (Hypothese 5). Zudem wurde angenommen, dass sich inkongruentes Feedback besonders günstig auf den Lernfortschritt auswirkt, demzufolge wurde eine Interaktion dieses Faktors mit dem Messwiederholungsfaktor erwartet (Hypothese 6).

Die Mittelwerte der Prä-Post-Messung zur Argumentationsqualität wurden mithilfe einer zweifaktoriellen Varianzanalyse mit Messwiederholung auf Signifikanz getestet, wobei die

Argumentationsqualität zum ersten und dritten Messzeitpunkt (vor und nach Bearbeitung der Zusatzinformationen) den Messwiederholungsfaktor mit 2 Stufen bildete. Den Gruppierungsfaktor stellte die Lernbedingung (kongruentes vs. inkongruentes vs. kein Feedback) mit 3 Stufen dar.

Tabelle 27 gibt die Mittelwerte zur Argumentationsqualität zum ersten und dritten Messzeitpunkt in Abhängigkeit von der Lernbedingung und insgesamt wieder und verdeutlicht, dass weder Boden- noch Deckeneffekte auftraten.

Tabelle 27: Argumentationsqualität zum ersten und dritten Messzeitpunkt (vor und nach Bearbeitung der Zusatzinformationen) (theoretisches Maximum: 5) für alle drei Lernbedingungen und insgesamt. Mittelwerte (Standardabweichungen in Klammern).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Kein Feedback		Gesamt	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Argumentationsqualität (t1)	2.38	(0.71)	2.67	(1.05)	2.46	(1.02)	2.50	(0.94)
Argumentationsqualität (t3)	2.75	(0.68)	2.96	(0.99)	2.83	(1.05)	2.85	(0.91)

Anmerkung: Wertebereich von 0 niedrig bis 5 hoch.

Wie bereits erwähnt, konnte für die Gesamtgruppe deskriptiv eine Zunahme der Argumentationsqualität vom ersten ($M = 2.50$, $SD = 0.94$) zum dritten Messzeitpunkt ($M = 2.85$, $SD = 0.91$) verzeichnet werden.

Am höchsten fiel die Argumentationsqualität zum ersten Messzeitpunkt in der Experimentalgruppe „inkongruentes Feedback“ ($M = 2.67$, $SD = 1.05$) aus, gefolgt von den Kontrollgruppen „ohne Feedback“ ($M = 2.46$, $SD = 1.02$) und „kongruentes Feedback“ ($M = 2.38$, $SD = 0.71$). Auch zum dritten Messzeitpunkt schnitt die Lernbedingung, die ein zur eigenen Entscheidung inkongruentes Feedback erhielt, deskriptiv geringfügig besser in der Argumentationsqualität ab ($M = 2.96$, $SD = 0.99$) als die beiden anderen Bedingungen „ohne“ ($M = 2.83$, $SD = 1.05$) und mit „kongruentem Feedback“ ($M = 2.75$, $SD = 0.68$).

Die Unterschiede zwischen den drei Lernbedingungen zu beiden Messzeitpunkten waren jedoch nur marginal; es ergab sich weder ein signifikanter Effekt des Gruppenfaktors „Lernbedingung“ für die Argumentationsqualität ($F(2, 69) = 0.49$, *n.s.*), noch lag eine statistisch bedeutsame Wechselwirkung des Gruppierungsfaktors mit dem Messwiederholungsfaktor vor ($F(2, 69) = 0.19$, *n.s.*). Somit bestanden keine substantiellen Unterschiede in den beiden Messwerten zur Argumentationsqualität zwischen den drei Gruppen, demzufolge kann von einem konstanten Lernfortschritt bei allen drei Gruppen ausgegangen wer-

den. Hypothese 6 konnte damit nicht bestätigt werden: Inkongruentes Feedback wirkte sich nicht nachweislich auf die Argumentationsqualität bzw. den Lernfortschritt aus.

Während der Gruppierungsfaktor und die Interaktion keine Signifikanz aufwiesen, konnte den deskriptiven Ergebnissen entsprechend inferenzstatistisch ein Effekt für den Messwiederholungsfaktor „Argumentationsqualität vor und nach Bearbeitung der Zusatzinformationen“ nachgewiesen werden. Die Punkte-Differenz zwischen den beiden Messungen der Argumentationsqualität für die Gesamtgruppe war hochsignifikant ($F(1, 69) = 29.64$, $p = .000$; $\eta^2_p = .30$). Da Cohen (1973) bei Varianzanalysen mit Messwiederholung vor einer Fehl- oder Überinterpretation des partiellen η^2 warnt, wurde hier als Alternative Cohens d berechnet (vgl. Dunlap, Cortina, Vaslow & Burke, 1996). Es lag ein mittlerer und praktisch bedeutsamer Effekt vor ($d = 0.61$).

Hypothese 5 konnte empirisch bestätigt werden: Durch die Bearbeitung der Lernumgebung bzw. Zusatzinformationen wurde wie erwartet ein deutlicher Lernfortschritt bewirkt.

Abbildung 32 veranschaulicht nochmals zusammenfassend die Veränderungen in der Argumentationsqualität im Lernverlauf für die Gesamtgruppe.

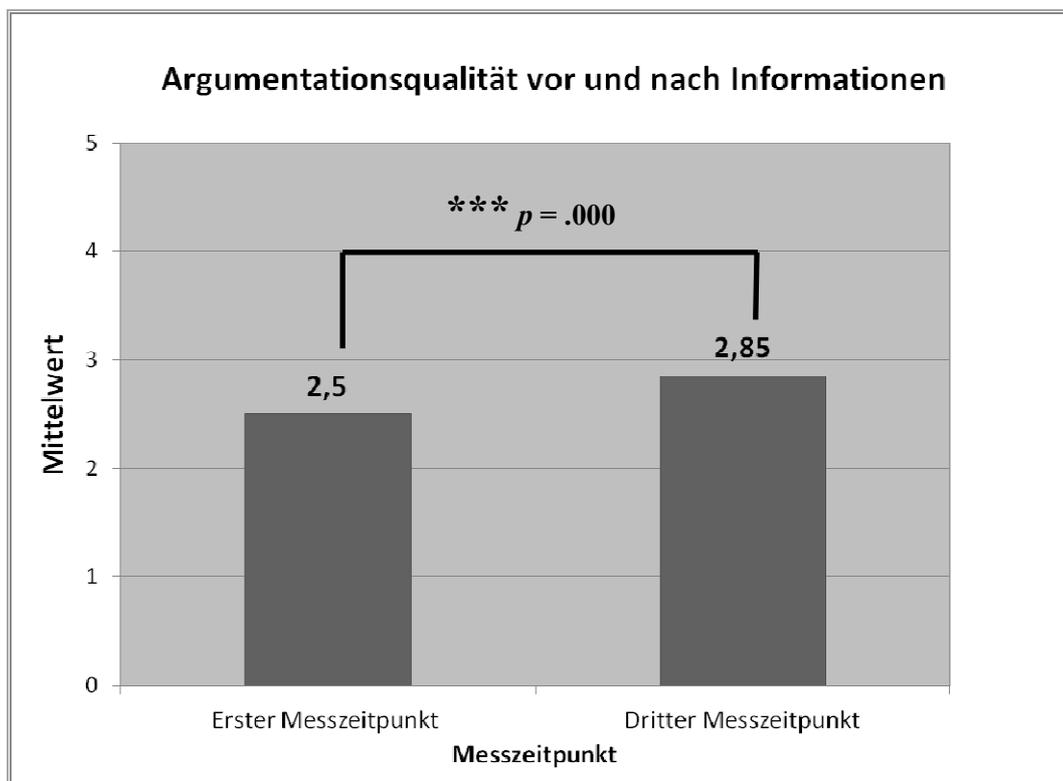


Abbildung 32: Argumentationsqualität zum ersten und dritten Messzeitpunkt (vor und nach Bearbeitung der Zusatzinformationen) für die Gesamtgruppe (Mittelwerte und Ergebnisse der Signifikanzprüfung, Balkendiagramm).

Wie Abbildung 32 zu entnehmen, stieg die Argumentationsqualität vom ersten bis zum dritten Messzeitpunkt deutlich an. Der Unterschied zwischen der ersten und dritten Mes-

sung war substantiell und praktisch bedeutsam. Insgesamt betrachtet konnte ein deutlicher Lernfortschritt durch die Bearbeitung der Lernumgebung bzw. der zusätzlichen Informationen erzielt werden. Anders als angenommen zeigte sich jedoch kein Effekt des Feedbacks anderer auf den Lernfortschritt.

Fragestellung 8: Inwieweit wird die Generierung von Gegenargumenten durch inkongruentes Feedback gefördert?

*Fragestellung 9: Inwieweit wird der *myside bias* durch inkongruentes Feedback verringert?*

Hypothesen. Es wird erwartet, dass Lernende der Bedingung „inkongruentes Feedback“ insgesamt mehr Gegenargumente generieren als Lernende der Bedingung „kongruentes“ und „kein Feedback“ (Hypothese 7). Außerdem wird angenommen, dass Lernende der Bedingung „inkongruentes Feedback“ einen geringeren *myside bias* in der Argumentation aufweisen als Lernende der beiden anderen Bedingungen (Hypothese 8).

Die 72 Versuchsteilnehmer führten im Ganzen betrachtet 312 Argumente, davon 185 präferenzkonsistente Argumente und 127 Gegenargumente zu ihrer eigenen Entscheidung an. Die deskriptiven Befunde für alle drei Gruppen zur durchschnittlichen Gesamtanzahl der generierten Argumente, der Anzahl der präferenzkonsistenten und -inkonsistenten Argumente sowie dem *myside bias* Index können folgender Tabelle 28 entnommen werden.

Tabelle 28: Gesamtanzahl der generierten Argumente, der entscheidungsunterstützenden Argumente, der Gegenargumente und *myside bias* Index für alle drei Lernbedingungen (Mittelwerte, Standardabweichungen und Ergebnisse der Signifikanzprüfung).

	1) Kongruentes Feedback		2) Inkongruentes Feedback		3) Ohne Feedback		P
	M	(SD)	M	(SD)	M	(SD)	
Gesamtanzahl Argumente	4.33	(1.86)	4.42	(2.52)	4.25	(1.96)	<i>n.s.</i>
Entscheidungsunterstützende Argumente	2.63	(0.97)	2.50	(1.29)	2.58	(1.10)	<i>n.s.</i>
Gegenargumente	1.71	(1.08)	1.92	(1.44)	1.67	(1.20)	<i>n.s.</i>
<i>Myside bias</i> Index	0.92	(0.88)	0.58	(1.06)	0.92	(1.21)	<i>n.s.</i>

Anmerkungen: *Myside bias* Index = Differenz entscheidungsunterstützende Argumente und Gegenargumente.

Im Mittel generierten die Versuchsteilnehmer insgesamt etwa vier Argumente ($M = 4.33$, $SD = 2.10$; Minimum: 0, Maximum: 10), davon $M = 2.57$ ($SD = 1.11$; Minimum 0, Maximum 5) entscheidungsunterstützende (*mysided*) Argumente und $M = 1.76$ ($SD = 1.24$;

Minimum 0, Maximum 5) Gegenargumente (*othersided*). Der *myside bias* Index betrug für die Gesamtgruppe im Durchschnitt 0.81 ($SD = 1.06$), das Minimum betrug -2, das Maximum 3. Da Werte > 0 angeben, dass mehr präferenzkonsistente als -inkonsistente Argumente generiert wurden, lag für die Gesamtgruppe ein *myside bias* und damit eine einseitige Argumentation und Verzerrung vor.

Zwischen den drei Lernbedingungen bestanden deskriptiv lediglich minimale Mittelwertsdifferenzen im Hinblick auf die Gesamtanzahl der generierten Argumente, der entscheidungsunterstützenden (*mysided*) Argumente, der Gegenargumente (*othersided*) sowie im *myside bias* Index. Dementsprechend konnte inferenzstatistisch auch kein bedeutsamer Unterschied zwischen den drei Gruppen in der Gesamtanzahl aller Argumente ($F(2, 69) = 0.04, n.s.$), der Anzahl der entscheidungskonsistenten Argumente ($F(2, 69) = 0.08, n.s.$), in der Anzahl der entscheidungsinkonsistenten Argumente ($F(2, 69) = 0.28, n.s.$) und im *myside bias* Index ($F(2, 69) = 0.79, n.s.$) nachgewiesen werden.

Die Hypothesen 7 und 8, dass Lernende der Bedingung „inkongruentes Feedback“ mehr Gegenargumente generieren und einen geringeren *myside bias* aufweisen als Lernende der Bedingungen „kongruentes“ und „kein Feedback“ konnten empirisch nicht bestätigt werden.

Abbildung 33 gibt zusammenfassend die Mittelwerte zur Anzahl der generierten Gegenargumente für alle drei Lernbedingungen wieder:

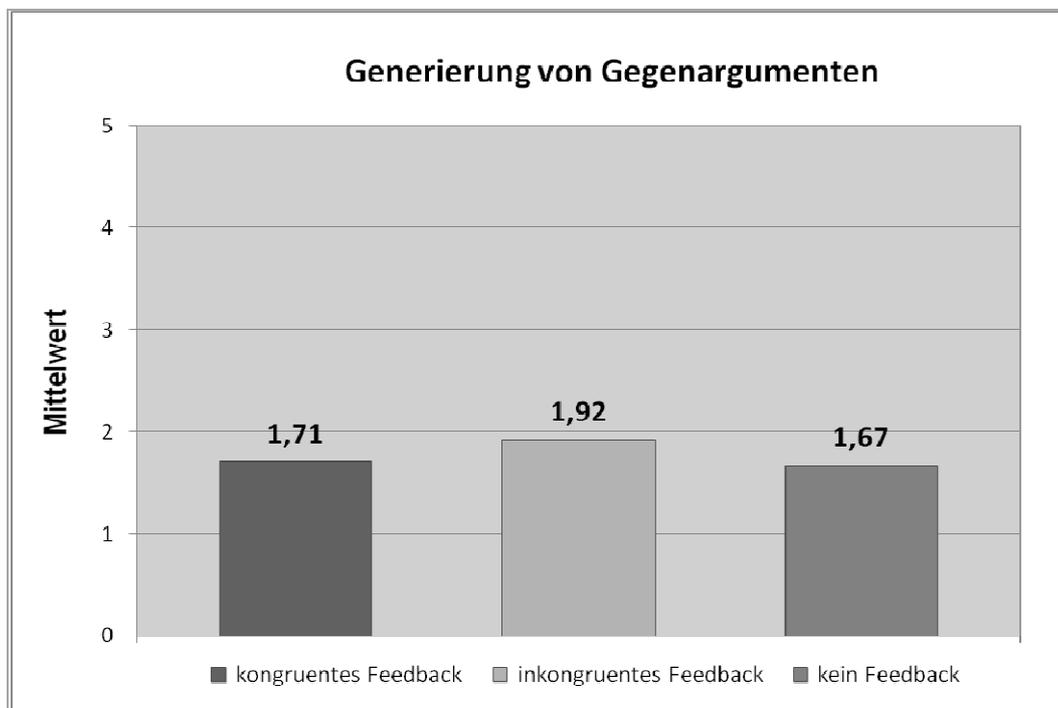


Abbildung 33: Anzahl der generierten Gegenargumente für alle drei Lernbedingungen, (Mittelwerte, Säulendiagramm).

Lernende, die mit einem zu ihrer eigenen Entscheidung inkongruenten Feedback konfrontiert wurden, konstruierten nicht substantiell mehr Gegenargumente zu ihrer eigenen Position und wiesen auch keinen geringeren *myside bias* auf als die Lernenden der beiden anderen Gruppen, die kein oder ein kongruentes Feedback erhielten. Anders als angenommen, hatte die Konfrontation mit einer Feedbackinformation, welche im Konflikt zur eigenen Entscheidung stand, keinen positiven Einfluss auf die Generierung von möglichen Gegenargumenten zur eigenen Entscheidung.

9.5 Analyse der Argumentationsmuster

Fragestellung 10: Welche Argumentationsmuster (rational oder emotional-intuitiv) werden am häufigsten zur Begründung der eigenen Entscheidung verwendet?

In dieser Fragestellung wird untersucht, auf welche Art und Weise die Versuchsteilnehmer ihre Entscheidungen in Bezug auf die Dilemmageschichte zur PID-HLA begründeten. Zur Klassifikation der verwendeten Argumentationsmuster zur Lösung der Dilemmageschichte wurden beide Essays (Prä-/Posttest), in denen die Probanden ihre eigene, persönlich getroffene Entscheidung zur PID-HLA begründeten, mit dem in Kapitel 8.7.3 beschriebenen Kategorienschema kodiert.

Analog zu den Studien von Sadler & Zeidler (2005) und Dawson & Venville (2009) konnten auch in dieser Arbeit die emotionalen und intuitiven Argumentationsweisen nur schwer voneinander unterschieden werden und werden daher in dieser Fragestellung zusammenfassend dargestellt.

Dementsprechend wurde analysiert, welche der beiden typischen Argumentationsmuster (rational oder emotional-intuitiv) am häufigsten zur Begründung der eigenen Entscheidung von den Versuchsteilnehmern verwendet werden, also ob überwiegend rational oder eher emotional-intuitiv bei der Lösung der Dilemmageschichte argumentiert wurde. Des Weiteren wurde geprüft, ob sich geschlechtsspezifische Unterschiede bei der Verwendung dieser zwei typischen Begründungsmuster zeigen.

Folgende Tabelle 29 gibt die Anzahl der Versuchsteilnehmer wieder, die das entsprechende Argumentationsmuster zur Lösung der Dilemmageschichte verwendeten. Dabei ist zu beachten, dass die Antworten der Probanden gemäß der Studie von Sadler & Zeidler (2005) jeweils beiden Argumentationsmustern gleichzeitig zugeordnet werden konnten.

Abgesehen von einer Ausnahme zeigte die überwältigende Mehrheit von 98.6 % der Versuchsteilnehmer (71) eine rationale Argumentationsweise in Bezug auf das Szenario zur PID-HLA, wohingegen lediglich etwas über die Hälfte der Vpn (37; 51.4 %) emotional-

intuitiv argumentierten. Bei der gezielten qualitativen Einzelfallanalyse konnte jedoch festgestellt werden, dass die emotionalen und intuitiven Reaktionen einen starken Einfluss auf das Entscheidungsverhalten hatten und recht häufig die endgültige individuelle Entscheidung determinierten.

Tabelle 29: Verwendete Argumentationsmuster (rational und emotional-intuitiv) zur Lösung der Dilemmageschichte. Absolute Häufigkeiten, Gesamt und nach Geschlecht sowie Ergebnisse der Signifikanztestung des Geschlechtervergleichs.

	Gesamt (<i>n</i> =72)	Männlich (<i>n</i> =33)	Weiblich (<i>n</i> =39)	Geschlechter- vergleich (χ^2)
Rationales Argumentationsmuster	71	33	38	0.86 (<i>n.s.</i>)
Emotional-intuitives Argumentationsmuster	37	12	25	5.51 ($p < 0.2$)

Es konnte kein statistisch signifikanter Häufigkeitsunterschied in der Verwendung des rationalen Argumentationstyps zwischen männlichen und weiblichen Versuchsteilnehmern nachgewiesen werden ($\chi^2(1) = 0.86$, *n.s.*).

Beim emotional-intuitiven Argumentationstyp konnte hingegen ein geschlechtsspezifischer Unterschied aufgedeckt werden. Beinahe zwei Drittel (67.6 %) aller weiblichen Versuchsteilnehmer (25) verwendeten eine emotional-intuitive Argumentationsweise, wohingegen nur etwa ein Drittel (32.4 %; 12) der männlichen Probanden diesen Argumentationstyp bei der Lösung der Dilemmageschichte äußerten. Der deskriptiv vorgefundene Häufigkeitsunterschied zwischen den weiblichen und männlichen Studenten in der Verwendung des emotional-intuitiven Argumentationstyps war signifikant ($\chi^2(1) = 5.51$, $p < 0.2$).

Im Vergleich zu den männlichen Studenten zeigten die Studentinnen demnach signifikant häufiger eine affektive Argumentationsweise und lösten das Dilemma zur PID-HLA vermehrt auf der Basis von Emotionen und Intuitionen. Allem Anschein nach wurden durch das Thema PID-HLA bei den weiblichen Studienteilnehmern mehr emotionale und intuitive Reaktionen ausgelöst als bei den männlichen.

Im Rahmen einer korrelationsstatistischen Analyse konnte ein positiver und signifikanter Zusammenhang in mittlerer Höhe zwischen der Präferenz für Intuition und einer emotional-intuitiven Argumentationsweise festgestellt werden ($r = .26$, $p < .05$). Die Verwendung des emotional-intuitiven Argumentationstyps hing folglich von der individuellen Präferenz für Intuition der Lernenden ab: Je höher die individuelle Präferenz für affektbasierte Entscheidungsstrategien ausgeprägt war, desto häufiger wurde eine emotional-intuitive Argumentationsweise verwendet (und umgekehrt).

Fragestellung 11: Welche deontologischen und utilitaristischen Argumente werden von den Versuchspersonen am häufigsten genannt?

In diesem Zusammenhang wurde analysiert, welche der 312 Argumente zur Rechtfertigung bzw. Begründung der persönlichen Entscheidung und bei der Nennung von Gegenpositionen am häufigsten von den Probanden formuliert wurden. Zusätzlich wurden die in den Essays angeführten Argumente im Hinblick auf ihre ethische Tradition (deontologisch: pflichten- und werteorientiert vs. utilitaristisch: folgen- und nutzenorientiert) systematisiert (vgl. Höhle, 2001b).

Um die Lesbarkeit dieses Kapitels nicht zu beeinträchtigen, werden in den folgenden Tabellen 30 und 31 lediglich die fünf häufigsten deontologischen und utilitaristischen Argumente wiedergegeben. Tabelle 30 sind die am häufigsten angeführten deontologischen Argumente zu entnehmen.

Tabelle 30: Verwendete deontologische Argumente (Anzahl und Beispiel).

Deontologische Argumente (Prinzipien/Werte/Pflichten)	Anzahl	Beispiel
Instrumentalisierung „Retterkind“ (Mittel zum Zweck)	44 (31/13)	„Ein Kind zu bekommen und schon im Vorhinein zu planen, dass es dem anderen Sohn ein Spender sein soll, finde ich moralisch sehr fragwürdig! Dieses Kind hat schließlich das Recht, um seiner Person wegen zu leben und nicht auf der Welt zu sein <u>für</u> jemand anderen“ [70b3_t1_Nein].
Tötung Menschenleben (Verwerfung Embryonen)	28 (13/15)	„Da für mich menschliches Leben schon bei der Befruchtung beginnt, halte ich dieses Vorgehen für menschenunwürdig. Auch ein Embryo ist „Leben“ [5b1_t3_Nein]
Gesetzliches Verbot in Deutschland	14 (11/3)	„Solange der rechtliche Rahmen in der BRD noch nicht gegeben ist ..., ist das Vorgehen nicht akzeptabel“ [71b1_t3_Nein].
Eingriff in die naturgegebene Ordnung	5 (2/3)	„Es ist verständlich, dass die Frau ihren Sohn retten will, dennoch sind die Konsequenzen einer solchen Handlung im Großen und deshalb auch im Kleinen nicht tragbar: nicht alles, was machbar ist, ist gut. Der Mensch greift bei diesem Beispiel zu weit in die Natur ein“ [60b2_t1_Nein].
Pflicht der Eltern zur Rettung ihres kranken Kindes	4 (2/2)	„Als Mutter hat man die Pflicht bzw. möchte man sein Kind über alles retten, wenn es keine zweite verwirklichtbare Möglichkeit gibt, muss man alles tun, damit das Kind weiter leben kann“ [18b1_t3_Ja].

Anmerkungen: Die Zahl hinter jeder Kategorie gibt die Anzahl der Personen wieder, die dieses Argument nannten (Anzahl entscheidungsunterstützende Argumente/Gegenargumente).

An erster Stelle rangierte die Instrumentalisierung des „Retterkindes“ (44), gefolgt von der Tötung von Menschenleben (28) infolge der Verwerfung von Embryonen. An dritter Stelle wurde das gesetzliche Verbot der PID-HLA in Deutschland (14) genannt, an vierter und fünfter Stelle der Eingriff in die naturgegebene Ordnung durch die PID (5) und die Pflicht der Eltern zur Lebensrettung ihres Kindes (4).

Als utilitaristischer Sicht wurden folgende fünf Argumente von den Probanden am häufigsten formuliert (siehe Tabelle 31).

Tabelle 31: Verwendete utilitaristische Argumente (Anzahl und Beispiel).

Utilitaristische Argumente (Folgen/Nutzen)	Anzahl	Beispiel
Lebensrettung/Heilung des kranken Geschwisterkindes	64 (36/28)	„Ich habe mich für ja entschieden, weil es mir einleuchtete, dass hier ein Kind von einer Krankheit geheilt werden kann und somit die Methode einen positiven Zweck hat“ [65b2_t1_Ja].
Psycho-soziale Folgen für das „Retterkind“	40 (29/11)	„Risiken bezogen auf soziale und psychologische Aspekte. Ich könnte mir vorstellen, dass das Kind, wenn es erfährt, dass es nur zu lebensrettenden Maßnahmen des anderen „erstellt“ wurde, es deutliche Schäden hinterlassen könnte!“ [11b3_t3_Nein].
Denkbare Alternativen	30 (20/10)	„Frau Kreidler sollte lieber ein „Zufallskind“ statt ein „Retterkind“ bekommen und ihre Energie in die Beziehung zu Mann und Kind stecken, statt in Amerika zur Untersuchung zu gehen“ [35b2_t1_Nein].
„Slippery slope“ (Missbrauchsgefahr: Designerbaby, Eugenik)	23 (12/11)	„Es stellt einen Eingriff in das Leben dar. Heilen durch Genforschung ist etwas anderes als „schaffen“. Bald würde sie dann auch Größe, Augenfarbe und Charaktereigenschaften des Kindes beeinflussen wollen“ [38b2_t3_Nein].
Diskriminierung/Stigmatisierung behinderter Menschen	11 (5/6)	„Ich habe mich unentschieden, weil ich die Stigmatisierung von behinderten Personen oder Eltern, die nicht das perfekte Kind wollen, für gar nicht unwahrscheinlich und sehr problematisch halte“ [55b2_t3_Nein].
Wunsch nach einem weiteren Kind	11 (7/4)	„Ausschlaggebend für meine Entscheidung war aber der Hinweis von Fr. Kreidler, dass sie und ihr Partner sowieso noch ein Kind geplant hatten und es deswegen eben nicht nur auf ein „Retterkind“ reduziert werden kann. Es ist ein eigenständiges, gewolltes Kind mit zusätzlichen Eigenschaften“ [40b1_t3_Ja].

Anmerkungen: Die Zahl hinter jeder Kategorie gibt die Anzahl der Personen wieder, die dieses Argument nannten (Anzahl entscheidungsunterstützende Argumente/Gegenargumente).

An erster Stelle stand die Lebensrettung bzw. Heilung des schwer erkrankten Geschwisterkinds (64), an zweiter Stelle die möglichen psycho-sozialen Folgen für das „Retterkind“ (40) und an dritter die denkbaren Alternativen zur PID, wie z. B. eine natürliche Schwangerschaft oder Knochenmarkspendenetze bzw. Nabelschnurblutbanken (30). An vierter Stelle wurden „slippery slope“-Argumente (23) genannt, wie z. B. die Gefahr einer Tendenz zu Designerbabys (Geschlecht, usw.) oder einer neuen Eugenik, an fünfter Stelle die mögliche Diskriminierung und Stigmatisierung von behinderten Menschen (11) sowie der Wunsch nach einem weiteren (gesunden) Kind (11).

Die qualitative Analyse der Argumente ergab zusätzlich, dass zwei Argumente aus den bereitgestellten Zusatzinformationen nur wenig oder gar nicht in der eigenen Entscheidung bzw. bei der Nennung von Gegenargumenten berücksichtigt wurden. Dies betraf zum einen das Argument zum PID-Tourismus (1), das lediglich einmal genannt wurde sowie das Argument zu den möglichen negativen Auswirkungen auf das Menschenbild und das Selbstverständnis des Menschen (0), das kein einziges Mal angeführt wurde.

Insgesamt betrachtet ergab die Frequenzanalyse, dass von den Probanden mehr utilitaristische (204; 65 %) als deontologische Argumente (108; 35 %) zur Rechtfertigung der persönlichen Entscheidung und bei der Nennung von Gegenpositionen formuliert wurden: Demnach wurden häufiger Argumente angeführt, welche die Richtigkeit einer moralischen Handlung nach der Qualität der absehbaren Handlungsfolgen und dem kollektiven Gesamtnutzen beurteilen und nicht nach moralischen Prinzipien, Werten oder Pflichten.

9.6 Analyse von Einflussfaktoren auf das Entscheidungsverhalten und die Argumentationsqualität

Da sich die Probanden der drei Lernbedingungen nur marginal im Entscheidungsverhalten und in der Argumentationsqualität unterschieden, wurden die in dieser Fragestellung untersuchten Einflussfaktoren auf diese Variablen nicht gruppenspezifisch, sondern für die Gesamtgruppe analysiert.

Fragestellung 12: In welchem Zusammenhang steht die Argumentationsqualität mit den erfassten Lernvoraussetzungen auf der einen und der parasozialen Interaktion mit der Medienfigur auf der anderen Seite?

Um Anhaltspunkte für das Zustandekommen von interindividuellen Unterschieden in Bezug auf die Argumentationsqualität zu bekommen, wurden zunächst die erfassten Lernvoraussetzungen mit der Argumentationsqualität beider Messungen in Beziehung gesetzt. Folgende Tabelle 32 gibt die resultierenden bivariaten Korrelationen wieder.

Die Mehrzahl der Korrelationen zwischen den erhobenen Lernvoraussetzungsaspekten und der Argumentationsqualität zum ersten und dritten Messzeitpunkt fielen durchweg sehr niedrig aus und verfehlten die Signifikanzgrenze.

Tabelle 32: Korrelationen nach Pearson zwischen Lernvoraussetzungen und der Argumentationsqualität zum ersten und dritten Messzeitpunkt ($N =$ zwischen 71 und 72).

	Argumentationsqualität (t1)	Argumentationsqualität (t3)
Vorwissen über Humangenetik	.32**	.33**
Subjektives Vorwissen Gentest	.07	.10
Subjektives Vorwissen PID	.30*	.26*
Subjektives Vorwissen Bioethik	.14	.11
Persönlicher Bezug Gentest	.18	.22
Interesse am Thema Gentest	.04	.06
Präferenz für Deliberation	-.06	-.02
Präferenz für Intuition	-.01	-.04
Vorteile von Gentests	-.09	-.01
Nachteile von Gentests	-.11	-.01
Befürchtungen bezgl. Gentests	-.06	-.02
Soziale Vergleichsorientierung	.11	.12
Orientierung an anderen	-.09	-.02

Anmerkungen: ** $p < .01$; * $p < .05$ (zweiseitige Signifikanzprüfung).

Es gab jedoch zwei Ausnahmen: Das Vorwissen über Humangenetik korrelierte signifikant und positiv in mittlerer Höhe mit der Argumentationsqualität zum ersten und dritten Messzeitpunkt (t1 und t3) ($r = .32$, $p < .01$ bzw. $r = .33$, $p < .01$). Ebenfalls positiv, aber etwas geringer fiel der Zusammenhang zwischen dem subjektiven Vorwissen zur PID und der Argumentationsqualität bei beiden Messungen aus, die Korrelationen waren substantiell ($r = .30$, $p < .05$ bzw. $r = .26$, $p < .05$).

Sowohl das Vorwissen über Humangenetik als auch das subjektive Vorwissen zur PID waren statistisch bedeutsam assoziiert mit der Argumentationsqualität zum ersten und dritten Messzeitpunkt: Je höher das Vorwissen über Humangenetik und das subjektive Vor-

wissen zur PID ausgeprägt war, desto besser fiel die Argumentationsqualität zum ersten und dritten Messzeitpunkt aus (und umgekehrt).

Die Argumentationsqualität erwies sich -das Vorwissen ausgenommen- zu beiden Messungen als nahezu unabhängig von den berücksichtigten Lernvoraussetzungsaspekten.

Zusätzlich wurde überprüft, ob die parasoziale Interaktion mit der Medienfigur mit der Argumentationsqualität zu beiden Messzeitpunkten assoziiert war. Wie Tabelle 33 zu entnehmen, korrelierten sowohl die parasoziale Interaktion insgesamt ($r = .19$, *n.s.* bzw. $r = .16$, *n.s.*) als auch die affektiven PSI-Prozesse nur schwach positiv und statistisch nicht bedeutsam mit der Argumentationsqualität zu beiden Messungen ($r = .11$, *n.s.* bzw. $r = .06$, *n.s.*). Die Argumentationsqualität war somit unabhängig von den PSI insgesamt und den affektiven PSI.

Tabelle 33: Korrelation nach Pearson zwischen der parasozialen Interaktion (PSI) mit der Medienfigur, den affektiven und kognitiven PSI-Teilprozessen und der Argumentationsqualität zum ersten und dritten Messzeitpunkt ($N = 72$).

	Argumentationsqualität (t1)	Argumentationsqualität (t3)
Parasoziale Interaktion	.19	.16
Affektive PSI	.11	.06
Kognitive PSI	.28*	.31**

Anmerkungen: ** $p < .01$; * $p < .05$ (zweiseitige Signifikanzprüfung).

Es zeigte sich jedoch eine Assoziation der kognitiven PSI-Prozesse mit der Argumentationsqualität, es lag eine positive und substantielle Korrelation in mittlerer Höhe vor ($r = .28$, $p < .05$ bzw. $r = .31$, $p < .01$): Je stärker sich die Lernenden kognitiv mit der Medienfigur auseinandersetzten, desto besser fiel die Argumentationsqualität zu beiden Messzeitpunkten aus (und umgekehrt). Die Argumentationsqualität wurde demnach positiv beeinflusst von der kognitiven PSI.

Fragestellung 13: Welchen Einfluss haben die erfassten Lernvoraussetzungen und die Sicherheit der Entscheidung auf die Revision der Entscheidung?

Um den Einfluss der erfassten Lernvoraussetzungen und der Entscheidungssicherheit auf die Revision der Entscheidung zu ermitteln, wurde ein t-Test für unabhängige Stichproben durchgeführt.

Die unabhängige Variable bildete die Entscheidungsrevision (0 = nein, keine Entscheidungsänderung, 1 = ja, Entscheidungsänderung). Als abhängige Variablen wurden die im Rahmen der Eignungsdiagnostik erhobenen Lernvoraussetzungsaspekte und die Entschei-

derungssicherheit untersucht. Durch eine Zusammenfassung der Entscheidungsrevision und der Sicherheit der Entscheidung, die nach jeder der drei zu treffenden Entscheidungen erfasst wurde, konnte die Anzahl der durchzuführenden Signifikanztests reduziert werden.

Zwischen den Lernenden, die ihre Entscheidung revidierten und den Lernenden, die ihre Entscheidung beibehielten, bestanden in allen erhobenen und untersuchten Lernvoraussetzungsaspekten keine signifikanten Unterschiede ($p > .1$).

Eine Ausnahme bildete lediglich das subjektive Vorwissen zur PID (siehe Tabelle 34). Das subjektiv eingeschätzte Vorwissen war deskriptiv bei den Lernenden, die ihre Entscheidung nicht änderten ($M = 1.55$, $SD = 0.89$) deutlich höher ausgeprägt als bei den Lernenden, die ihre Entscheidung revidierten ($M = 1.17$, $SD = 0.39$). Im Hinblick auf das subjektive Vorwissen zur PID bestand ein signifikanter Unterschied zwischen den beiden Gruppen ($t(38.30) = 2.38$, $p < .05$).

Tabelle 34: Subjektives Vorwissen zur PID und Sicherheit der Entscheidung für die beiden Gruppen Entscheidungsrevision ($n = 12$) und keine Entscheidungsrevision ($n = 60$) (Mittelwerte, Standardabweichungen in Klammer und Ergebnisse der Signifikanzprüfung).

	Entscheidungs- revision ($n = 12$)		Keine Entscheidungs- revision ($n = 60$)		p
	M	(SD)	M	(SD)	
Subjektives Vorwissen PID ^a	1.17	(0.39)	1.55	(0.89)	$p < .05$
Sicherheit der Entscheidung ^b	4.19	(2.71)	6.77	(2.32)	$p < .001$

Anmerkung: Wertebereich von ^a 1 niedrig bis 5 hoch, ^b 1 niedrig bis 10 hoch.

Allerdings variierte das subjektive Vorwissen zur PID in der Gruppe „keine Entscheidungsrevision“ stärker als in der Bedingung „Entscheidungsrevision“: Der Levene-Test wies heterogene Varianzen auf ($F(38.30) = 9.48$, $p < .01$). Der Unterschied blieb bei Verwendung des U-Tests für unabhängige Stichproben jedoch statistisch nicht bedeutsam, da die Signifikanzgrenze von 5 % knapp verfehlt wurde (MR Gruppe „Entscheidungsrevision“, 30.67, MR Gruppe „keine Entscheidungsrevision“: 37.67, $U = 290.00$, $p > .1$). Lernende mit niedrigem Vorwissen zur PID revidierten demnach ihre Entscheidung zumindest tendenziell häufiger als Lernende mit hohem Vorwissen.

Wie Tabelle 34 zu entnehmen, fiel die globale Sicherheit der Entscheidung bei der Gruppe „Entscheidungsrevision“ ($M = 4.19$, $SD = 2.71$) deutlich geringer aus als bei der Gruppe „keine Entscheidungsrevision“ ($M = 6.77$, $SD = 2.32$). Die Standardabweichung war in beiden Gruppen vergleichbar. Der Mittelwertsunterschied zwischen den beiden Gruppen war signifikant und statistisch bedeutsam ($t(70) = 3.41$, $p \leq .001$). Zudem lag hier mit einer

Effektgröße von $d = 1.02$ ein sehr großer Effekt vor. Lernende, die ihre Entscheidung änderten, waren substantiell unsicherer in ihrer Entscheidung, als Lernende, die ihre Entscheidung nicht revidierten.

Die Befunde zur Bedeutung potenzieller Einflussfaktoren auf das Entscheidungsverhalten und die Argumentation lassen sich wie folgt zusammenfassen: Im Hinblick auf die Argumentationsqualität zeigte sich ein mittlerer Effekt des inhaltsbezogenen Vorwissens und der kognitiven parasozialen Interaktion mit der Medienfigur. Höheres Vorwissen über Humangenetik und PID sowie eine stärkere kognitive Auseinandersetzung mit der Medienfigur gingen jeweils mit einer höheren Argumentationsqualität zu beiden Messzeitpunkten einher (und umgekehrt).

In Bezug auf das Entscheidungsverhalten konnte nachgewiesen werden, dass die Sicherheit der Entscheidung einen entscheidenden Einflussfaktor auf das Entscheidungsverhalten bzw. die Änderung der Entscheidung darstellte. Lernende, die ihre Entscheidung im Verlauf der Bearbeitung der Lernumgebung revidierten, waren unsicherer in ihrer Entscheidung als Lernende, die ihre Entscheidung beibehielten.

10 Diskussion

Der Schwerpunkt der vorliegenden Arbeit lag auf der Untersuchung, inwieweit eine für den musealen Kontext konzipierte multimediale Lernumgebung eine informierte und gut begründete Entscheidung im Bereich Medizinethik/PID ermöglicht. Ein weiterer Untersuchungsschwerpunkt lag auf der Frage, welchen Einfluss eine Feedbackinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung anderer Personen auf die Entscheidungsfindung und die Argumentationsqualität zu einem medizinethischen Dilemma hat.

Im Gesamtbild zeigte sich, dass eine informierte und argumentativ gut begründete Entscheidung zum kontroversen Wissenschaftsthema PID-HLA durch die multimediale Lernumgebung angemessen unterstützt und gefördert werden konnte. Allerdings konnte entgegen der Vorannahmen kein statistisch bedeutsamer Effekt der Feedbackinformation auf das Entscheidungsverhalten und die Argumentationsqualität nachgewiesen werden. Diese unerwarteten Ergebnisse werden in diesem Kapitel schrittweise im Verlauf der Diskussion der Befunde erklärt.

Nachfolgend werden die einzelnen Ergebnisse der Studie in der Reihenfolge der Untersuchungsfragen zusammengefasst und im Rückgriff auf die Theorie diskutiert.

10.1 Überprüfung der experimentellen Voraussetzungen

Lernvoraussetzungen. Im Vergleich zur Allgemeinbevölkerung verfügten die Probanden über ein durchschnittliches allgemeines Wissen über Humangenetik (vgl. Berth et al., 2004). Im Hinblick auf die weiteren untersuchten kognitiven Lernvoraussetzungsaspekte ergaben sich durchweg sehr niedrige Mittelwerte. Die mittleren Ausprägungen zum selbst eingeschätzten Vorwissen zu genetischen Untersuchungen, zur PID und zum Thema Bioethik lagen sehr deutlich im unteren Bereich der Skala. Ein besonders niedriger Skalenmittelwert zeigte sich im subjektiven Vorwissen zur PID. Dieser Befund ist allerdings nicht verwunderlich, da es sich in einer repräsentativen Studie herausgestellt hat, dass das Wissen über die PID in der deutschen Allgemeinbevölkerung generell recht gering ist (vgl. Brähler & Stöbel-Richter, 2004).

Die erhobenen Lernvoraussetzungsaspekte in Bezug auf das Vorwissen zum Inhaltsgebiet der Lernumgebung können somit aus empirischer Sicht zwar als repräsentativ, aus pädagogisch-psychologischer Sicht aber eher als ungünstig angesehen werden.

Dem niedrigen Vorwissensniveau entsprechend konnte auch nur ein sehr geringer persönlicher Bezug bei den Versuchsteilnehmern zum Thema genetische Untersuchungen festgestellt werden. Auffällig ist, dass dem eher ungünstigen subjektiven Vorwissen und geringen persönlichen Bezug ein stark ausgeprägtes Interesse am Thema Gentest gegenüberstand. Die mittlere Ausprägung zum Interesse an genetischen Untersuchungen lag sehr deutlich im oberen Bereich der Skala. Als mögliche Ursache für diese vorgefundene Merkmalskombination der untersuchten Stichprobe können vermutlich Selbstselektionsprozesse bei der Rekrutierung der Teilnehmer herangezogen werden. Die Teilnahme an der Studie erfolgte auf freiwilliger Basis, wahrscheinlich nahmen nur Studierende mit einem hohen Interesse am Thema der Lernumgebung teil.

In Bezug auf die Einstellungen zu genetischen Untersuchungen lagen die mittleren Ausprägungen zu den eingeschätzten positiven Aspekten im oberen Bereich, zu den negativen Aspekten dagegen im mittleren Bereich der Skala. Demzufolge zeigte sich bei den Probanden eher eine Zustimmung in Bezug auf die molekulardiagnostische Diagnostik für erbliche Erkrankungen. Von der Mehrheit der Befragten werden somit mehr mögliche Vor- als Nachteile von Gentests gesehen. Dieser positiven Grundhaltung standen allerdings auch ausgeprägte Befürchtungen im Hinblick auf den Missbrauch der Ergebnisse genetischer Untersuchungen für wissenschaftliche Zwecke, für Eugenik oder durch Dritte gegenüber. Der Skalenmittelwert zu den Befürchtungen fiel deskriptiv am höchsten aus und lag deutlich im oberen Skalenbereich. Im Hinblick auf das Meinungsbild der untersuchten Stichprobe ist insgesamt betrachtet zu konstatieren, dass trotz der positiven Grundhaltung, auch Nachteile und insbesondere Befürchtungen in Bezug auf genetische Untersuchungen gesehen werden. Das vorgefundene Ergebnismuster in Bezug auf die Akzeptanz genetischer Untersuchungen - positive Grundhaltung, aber auch kritische Artikulation von Nachteilen und Befürchtungen - entspricht weitestgehend den Ergebnissen der deutschlandrepräsentativen Studie von Berth et al. (2002), deren Items in leicht adaptierter Form in dieser Untersuchung verwendet wurden.

Im Hinblick auf die Tendenz, sich mit anderen Personen zu vergleichen bzw. sich in sozialen Situationen an anderen Personen zu orientieren, ergaben sich mittlere Ausprägungen bei den Untersuchungsteilnehmern. Die soziale Vergleichsorientierung lag leicht über dem Skalenmittel, wohingegen die Orientierung an anderen leicht unter dem mittleren Bereich der Skala lag. Bei den Probanden dieser Untersuchung lag demzufolge weder eine besonders starke noch eine besonders niedrig ausgeprägte soziale Vergleichsorientierung und Orientierung an anderen vor.

Sicherung der internen Validität. Die Lernenden der drei Bedingungen unterschieden sich weder im themenspezifischen Vorwissen über Humangenetik, noch im subjektiven Vor-

wissen über genetische Untersuchungen, zur PID und über Bioethik. Auch in Bezug auf die erfassten motivationalen und einstellungsbezogenen Lernvoraussetzungen bestanden keine überzufälligen Unterschiede zwischen den drei Lernbedingungen. Insgesamt betrachtet zeigten sich somit vergleichbare Ausgangsbedingungen zwischen den drei Gruppen im Hinblick auf die vor der Untersuchung im Rahmen der Eignungsdiagnostik erhobenen Kontrollvariablen. Die interne Validität der experimentellen Studie konnte somit im Hinblick auf die erfassten individuellen Lernvoraussetzungen gesichert werden.

Allerdings bestanden tendenziell Unterschiede in der realisierten Lernzeit zwischen den drei Lernbedingungen: Die Lernenden der Bedingung „kein Feedback“ befassten sich etwas länger mit der Lernumgebung und der Bearbeitung der Frageblöcke als die beiden Experimentalgruppen. Die Lern- bzw. Bearbeitungszeit stellte jedoch keine potentielle Störvariable dar, da sie nicht bedeutsam mit der Argumentationsqualität zu beiden Messungen assoziiert war.

Manipulationskontrolle. Im Rahmen der Manipulationskontrolle konnte in Übereinstimmung mit den Theorien zum sozialen Einfluss und den Untersuchungsergebnissen von Knipfer (2009) nachgewiesen werden, dass *soziale Projektion* (Allport, 1924) bei den Versuchsteilnehmern eintrat. Die Studienteilnehmer nahmen an, dass ihre eigene Entscheidung von der Mehrheit geteilt wird, also dass sich die anderen genauso entschieden haben wie sie selbst. Der *false consensus-Effekt* konnte damit repliziert werden: Die Versuchspersonen erwarteten, selbst der Meinungsmehrheit anzugehören (Ross et al., 1977).

Zudem bestand ein hohes Bedürfnis nach sozialem Vergleich bei den Probanden, d. h. ein hohes Interesse daran, zu erfahren, wie sich die anderen Studienteilnehmer in Bezug auf die Dilemmageschichte zur PID-HLA entschieden haben. Die Mittelwerte zur sozialen Projektion, dem selbst erzeugten Konsens und dem Bedürfnis nach sozialem Vergleich entsprechen in ihrer deskriptiven Ausprägung weitestgehend den in der Studie von Knipfer (2009) berichteten Werten. Es zeigten sich keine überzufälligen Unterschiede zwischen den drei Lernbedingungen in diesen drei Variablen.

Folglich waren die Voraussetzungen zum Entstehen bzw. Auslösen von kognitiver Dissonanz infolge der Konfrontation mit einer präferenzinkonsistenten Konsensinformation bei den Teilnehmern der Lernbedingung „inkongruentes Feedback“ gegeben. Dementsprechend wurde auch die Information, dass die anderen Versuchsteilnehmer eine andere Entscheidung trafen, von der Bedingung „inkongruentes Feedback“ auch als konflikthaft wahrgenommen und löste kognitive Dissonanz aus, die psychisch unangenehm erlebt wurde (Elliot & Devine, 1994; Festinger, 1957): Teilnehmer, die ein zur eigenen Entscheidung inkongruentes Feedback erhielten, berichteten mehr negative, weniger positive Gefühle

und waren deutlich überraschter als Teilnehmer, deren Entscheidung durch kongruentes Feedback bestätigt wurde.

Die Befunde zur Manipulationskontrolle konnten zeigen, dass die Voraussetzungen zur Testung der Hypothesen in dieser experimentellen Studie erfüllt wurden. Die im Rahmen dieser Untersuchung aufgestellten Hypothesen zum Einfluss des Feedbacks in Form einer Konsensinformation auf das Entscheidungsverhalten und die Argumentationsqualität konnten also bestätigt oder widerlegt werden.

Darüber hinaus erkannte kein Versuchsteilnehmer das eigentliche Ziel der Studie (sozialer Einfluss durch Mehrheiten) oder dass die präsentierte Feedbackinformation in Form einer Umfrageergebnisse über die Entscheidung der anderen Versuchsteilnehmer manipuliert war. Die experimentelle Manipulation der unabhängigen Variable „Feedback“ war somit erfolgreich.

10.2 Parasoziale Interaktion mit der Medienfigur

Fragestellung 1: In welchem Ausmaß setzen sich die Lernenden mit der Medienfigur auseinander?

Auch wenn die Interpretierbarkeit deskriptiver Daten auf Intervallskalenniveau generell begrenzt ist, lässt die Höhe der Mittelwerte in Anlehnung an das Zwei-Ebenen-Modell parasozialer Interaktionen (Hartmann et al., 2004a) eher auf eine intensiv-starke Beschäftigung (*High-Level-PSI*) mit der virtuellen Figur schließen. Dabei fiel die Auseinandersetzung auf kognitiver Ebene etwas höher aus als auf der affektiven. Die Versuchsteilnehmer empfanden die Medienfigur als attraktiv.

Fragestellung 2: In welchem Zusammenhang steht die parasoziale Interaktion mit der wahrgenommenen Attraktivität der Medienfigur?

Die wahrgenommene Attraktivität der Medienfigur bildete eine wichtige Determinante auf die Intensität der PSI-Prozesse. Es konnte ein positiver Einfluss der Attraktivität der Medienfigur auf die Motivation, sich mit der Persona intensiv auseinanderzusetzen, festgestellt werden. Der Zusammenhang zwischen der Attraktivität und den PSI war signifikant und substantiell, fiel jedoch etwas niedriger aus als die in der Studie von Hartmann & Klimmt (2005) berichteten Regressionskoeffizienten. Versuchsteilnehmer, welche die Medienfigur attraktiver bewerteten, beschäftigten sich intensiver mit derselben und den von ihr vermittelten Inhalten (und umgekehrt).

Die Befunde von Hartmann & Klimmt (2005) konnten repliziert werden: Die Attraktivität der Medienfigur stellte einen entscheidenden Einflussfaktor auf die Intensität der PSI-

Prozesse dar. Die PSI zur Medienperson fielen umso intensiver aus, je attraktiver die Medienfigur von den Lernenden eingeschätzt wurde.

Zusammenfassend lässt sich festhalten, dass die Vorannahmen, dass bei den Lernenden ein starkes interpersonales Involvement mit der virtuellen Figur vorliegt und die empfundene Attraktivität der Medienfigur mit den PSI positiv korreliert, bestätigt werden konnten. Die Darstellung der Medienpersona (z. B. hohe Bildschirmpräsenz, direkte Adressierung und dialogorientierter Sprachstil) begünstigte somit tatsächlich das Entstehen von parasozialen Interaktionen.

Das in der multimedialen Lernumgebung eingesetzte Präsentationsformat eines *talking heads* hat sich somit im Hinblick auf das Zustandekommen von sozialen Reaktionen bei den Lernenden bewährt. Der Social Agency Theory (Mayer, 2005c) zufolge sollte die Aktivierung einer sozialen Reaktion zu einer Steigerung der aktiven, kognitiven Informationsverarbeitung und zu besseren Lernergebnissen im Sinne einer höheren Argumentationsqualität führen (siehe Kapitel 10.6).

10.3 Entscheidungsverhalten

Die Probanden wurden im Rahmen dieser Untersuchung aufgefordert, insgesamt dreimal eine hypothetische Entscheidung für oder gegen die Durchführung einer PID-HLA in Amerika zu treffen, wobei sie sich in die Rolle der von dem Dilemma betroffenen Person hineinversetzen sollten.

Fragestellung 3: In welche Richtung gehen die Entscheidungen?

Die deskriptive Analyse der Häufigkeitsverteilungen der drei getroffenen Entscheidungen im Hinblick auf die Zustimmung oder Ablehnung einer PID-HLA ergab, dass sich die Richtung der Entscheidungen bei den Untersuchungsteilnehmern im zeitlichen Verlauf änderte. Zum ersten Messzeitpunkt (t1), unmittelbar nach der Präsentation der Dilemmageschichte, entschied sich die knappe Mehrheit der Untersuchungsteilnehmer für die Durchführung einer PID-HLA in Amerika, d. h. stimmte der durch die Medienfigur implizit vertretenen Position zu.

Das in dieser experimentellen Studie ermittelte Abstimmungsergebnis der Probanden zum ersten Messzeitpunkt (t1) entspricht weitestgehend demjenigen das im Rahmen einer formativen Evaluationsstudie zu einem Prototyp der Dialogstation „Gentest“ im Deutschen Museum anhand einer Logfile-Analyse ermittelt werden konnte. Auch im realen Feldsetting hatte sich eine knappe Mehrheit für die Durchführung der PID entschieden (Hänle, 2008).

Bei der zweiten Messung konnte gegenüber der ersten ein Richtungswechsel der Entscheidungen festgestellt werden, der zum dritten Messzeitpunkt noch zunahm: Die Mehrheit der Probanden entschied sich nun gegen eine PID-HLA in Amerika. Zwischen den drei Lernbedingungen konnten deskriptiv betrachtet keine wesentlichen Unterschiede in Bezug auf die Richtung der drei getroffenen Entscheidungen festgestellt werden.

Der Befund, dass sich durch die Bearbeitung der Lernumgebung insgesamt die Richtung der Gesamtentscheidungen änderte, lässt sich aus persuasionstheoretischer Sicht folgendermaßen erklären. Bei Versuchsteilnehmer, die ihre Entscheidung revidierten, kann von einer heuristischen Verarbeitung ausgegangen werden, in der die präsentierten Informationen und Argumente nicht detailliert und kritisch geprüft, sondern wahrscheinlich einfache Entscheidungsregeln (Heuristiken) angewendet wurden. Entscheidungen, die auf heuristischer Verarbeitung beruhen, sind in der Regel sehr instabil.

Der im Lernverlauf konstatierte Richtungswechsel von einer anfänglichen Befürwortung der PID-HLA zu einer späteren Ablehnung kann vermutlich auf Merkmale des Kommunikators bzw. der Kommunikation zurückgeführt werden. In den bereitgestellten Zusatzinformationen wurden mehr Contra- als Pro-Argumente von Experten zur ethischen Bewertung der PID-HLA präsentiert (vgl. Anhang). Es ist denkbar, dass in diesem Zusammenhang die Contra-Position unter der Anwendung von einfachen Entscheidungsmechanismen übernommen wurde (z. B. *the more arguments the better*: Je höher die Anzahl der Argumente, desto wahrscheinlicher ist Richtigkeit der Schlussfolgerung; Petty & Cacioppo, 1984) oder *expert's statements can be trusted*: Expertenheuristik; Chaiken et al., 1989).

Die Ergebnisse der Einzelfallanalyse, in der festgestellt wurde, dass 6 von 7 Entscheidungsrevidierern nach der Bearbeitung der Zusatzinformationen ihre Entscheidung von „ja“ auf „nein“ änderten, sprechen für die Plausibilität dieser Annahme. In diesem Zusammenhang ist es auch denkbar, dass in Bezug auf die erste Entscheidung zunächst der von der Medienfigur implizit vertretenen Position („Ja“ bzw. „Pro PID-HLA“) zugestimmt wurde, da jene als direkt betroffene Person, z. B. für glaubwürdig und vertrauenswürdig gehalten und/oder als attraktiv und sympathisch empfunden wurde (z. B. Sympathieheuristik: *people agree with people they like*; Chaiken, 1980).

Fragestellung 4: Inwieweit kommt es durch die Bearbeitung der Lernumgebung zu einer Zunahme der Entscheidungssicherheit und inwieweit beeinflusst die Kongruenz der Feedbackinformation die Entscheidungssicherheit?

Überrascht hat die relativ hohe Sicherheit der Entscheidung bei den Versuchsteilnehmern, denn für diese Untersuchung wurde ein hochgradig ambivalentes und kontroverses medizinisches Thema ausgewählt. Die mittleren Ausprägungen zur Sicherheit der getroffene-

nen Entscheidung zum ersten Messzeitpunkt (t1) lagen bei allen drei Lernbedingungen im oberen Bereich der Skala. Damit hatten die Probanden entgegen der Vorannahmen keine größeren Schwierigkeiten eine erste, vorläufige Entscheidung in Bezug auf die Dilemmageschichte zur PID-HLA zu fällen. Trotz der hohen Komplexität des zu beurteilenden Gegenstandes fiel die Entscheidung bereits zu Beginn relativ sicher aus.

Möglicherweise kann die unerwartet hohe Entscheidungssicherheit bei den Untersuchungsteilnehmern auf das Erleben von Dissonanz unmittelbar nach der getroffenen Entscheidung zurückgeführt werden. Die aus einer Entscheidung resultierende Dissonanz kann beispielsweise dadurch aufgelöst werden, dass die Sicherheit bezüglich der Richtigkeit der Entscheidung geändert wird und letztendlich zu einer erhöhten Entscheidungssicherheit führt (vgl. Frey & Gaska, 1993). Ob allerdings diese Annahme tatsächlich zutrifft, darüber kann auf Grundlage der Befunde nur gemutmaßt werden.

Durch die Bearbeitung der Lernumgebung konnte wie erwartet ein bedeutsamer Anstieg der Sicherheit der Entscheidung festgestellt werden. Es hat sich im Einklang mit den Ergebnissen von Götz (2001) herausgestellt, dass die Entscheidungen nach Bearbeitung der zusätzlichen, vertiefenden Informationen im Mittel signifikant und substantiell sicherer ausfallen als die Entscheidungen vor der Bearbeitung der Informationen. Bei der gezielten qualitativen Einzelfallanalyse konnten zahlreiche Kommentare der Probanden identifiziert werden, die den Grundtenor, dass die Zusatzinformationen hilfreich und nützlich bei der Lösung der Dilemmageschichte waren, bestätigten.

Zwischen den drei Lernbedingungen bestanden keine substantiellen Unterschiede in der Sicherheit der Entscheidung zu allen drei Messzeitpunkten. Die Kongruenz der Konsensinformation mit der eigenen Entscheidung hatte keinen Einfluss auf die Entscheidungssicherheit. Weder die Information, dass die eigene Entscheidung von einer Mehrheit vertreten wurde, noch die Information, dass die eigene Entscheidung von einer Minderheit geteilt wurde, führte zu einer substantiellen Änderung der Entscheidungssicherheit. Die Entscheidungssicherheit verlief demnach bei allen drei Gruppen annähernd konstant.

Fragestellung 5: Inwieweit führt inkongruentes Feedback zu einer Änderung der Entscheidung?

Im Rahmen der Manipulationskontrolle konnte zunächst nachgewiesen werden, dass die Konfrontation mit einer zur eigenen Entscheidung inkongruenten Feedbackinformation als konflikthaft wahrgenommen wurde und bei dieser Lernbedingung kognitive Dissonanz auslöste, die psychisch unangenehm erlebt wurde.

Das im *manipulation check* festgestellte Erleben von kognitiver Dissonanz ging allerdings entgegen der Vorannahmen nicht mit einer signifikant häufigeren Entscheidungsrevision

bei der Lernbedingung „inkongruentes Feedback“ einher. Lernende, die eine im Konflikt zur eigenen Entscheidung stehende Konsensinformation erhielten, revidierten nicht häufiger ihre zuvor abgegebene Entscheidung und übernahmen nicht die in der Konsensinformation vertretene Mehrheitsposition. Hoher Konsens beeinflusste somit die individuelle Entscheidungsfindung in Bezug auf die Dilemmageschichte zur PID-HLA nicht: Lediglich bei 3 Probanden konnte eine Revision der ersten, vorläufigen Entscheidung nach inkongruentem Feedback beobachtet werden.

Die Ergebnisse von Mackie (1987) und insbesondere Knipfer (2009), die feststellen konnten, dass Lernende, die mit einem zu ihrer eigenen Meinung konfligierenden Feedback konfrontiert wurden, signifikant und substantiell häufiger ihr zuvor abgegebenes Gesamturteil zur Nanotechnologie revidierten, konnten damit im Rahmen dieser Untersuchung im Bereich Medizinethik nicht repliziert werden. Vielmehr wurde die am Anfang getroffene Entscheidung überwiegend beibehalten und das von Haidt (2001) postulierte rechtfertigende Entscheidungsverhalten bei moralischen Sachverhalten konnte bestätigt werden.

Da keine überzufälligen Unterschiede in den individuellen Lernvoraussetzungen zwischen den drei Gruppen bestanden, kann das Ausbleiben eines Effekts der inkongruenten Feedbackinformation nicht auf vorab vorhandene Unterschiede in den erfassten Lernvoraussetzungsaspekten zurückgeführt werden. Ferner konnten auch keine Anhaltspunkte dafür identifiziert werden, dass die Probanden die Glaubwürdigkeit der Feedbackinformation bezweifelten. Darüber hinaus konnte im *manipulation check* festgestellt werden, dass sowohl die Lernbedingung „kongruentes“ als auch „inkongruentes Feedback“ eine gute Erinnerung an die Mehrheitsverhältnisse der präsentierten Konsensinformation hatten und die Übereinstimmung bzw. den Unterschied zwischen der eigenen Entscheidung und dem Abstimmungsergebnis der anderen Studienteilnehmer korrekt wiedergaben.

Der Befund zum wahrgenommenen Konflikt spricht dafür, dass das präsentierte Feedback von den Probanden bewusst rezipiert und aktiv verarbeitet wurde (*mindful reception*; vgl. Bangert-Drowns et al., 1991).

Für die überraschenderweise ausbleibende Entscheidungsrevision nach inkongruentem Feedback können mehrere mögliche Ursachen herangezogen werden. Zum einen ist die Revision einer bereits getroffenen Entscheidung oft mit einem hohen Aufwand verbunden (und führt u. U. erneut zu Dissonanz), so dass die entstandene Dissonanz vermutlich auf eine andere Weise reduziert wurde (vgl. Frey & Gaska, 1993): z. B. durch Abwertung der Informationsquelle hinsichtlich Glaubwürdigkeit, Kompetenz oder Relevanz (siehe später: Defense-Motiv) oder die Suche nach bestätigenden Informationen für die eigene Entscheidung (siehe Kapitel 10.4). Zum anderen kann der informationale soziale Einfluss von Mehrheiten von verschiedenen Faktoren moderiert werden (vgl. Werth & Mayer, 2008):

der Art des Urteilsgegenstandes, der Entscheidungssicherheit, der persönlichen Bedeutsamkeit der Entscheidung, der wahrgenommenen Glaubwürdigkeit/Expertise der anderen Versuchsteilnehmer und der Gruppengröße (Samplegröße), die nachfolgend diskutiert werden.

Eine mögliche Ursache des ausbleibenden Effekts der Feedbackinformation auf das Entscheidungsverhalten liegt zunächst in der *Art des Urteilsgegenstands* begründet. In ethischen Dilemmasituationen existieren keine klaren Lösungskriterien im Sinne von „richtig“ oder „falsch“. Dementsprechend bildet das Entscheidungsverhalten der anderen in diesen Situationen keine adäquate Informationsquelle: Konformität macht in diesem Kontext „keinen Sinn“ (Werth & Mayer, 2008, S. 292), da das Bedürfnis nach Korrektheit nicht befriedigt werden kann. Die Tatsache, dass ein Versuchsteilnehmer im Kurzessay angab, dass die Entscheidungen der anderen keinen zusätzlichen Informationswert für ihn in Bezug auf die Lösung der Dilemmageschichte ergaben, ist kompatibel mit dieser Interpretation.

Des Weiteren kommt die sehr *hohe Sicherheit der Entscheidung* der Probanden in Betracht, denn je unsicherer in der Regel Personen sind, desto stärker fällt die Vergleichsmotivation aus und desto mehr verlassen sie sich auf die Entscheidungen der anderen (vgl. Frey et al., 1993; Werth & Mayer, 2008). Die über den gesamten Lernverlauf festgestellte unerwartet hohe Sicherheit der Entscheidung führte vermutlich dazu, dass das Entscheidungsverhalten resistenter gegenüber dem Beeinflussungsversuch war. Für die Plausibilität dieser Annahme spricht zum einen, dass die Entscheidungssicherheit nach Erhalt des inkongruenten Feedback nicht bedeutsam sank und zum anderen, dass sich die Sicherheit der Entscheidung als ein wichtiger Einflussfaktor auf die Revision der Entscheidung erwies (siehe Kapitel 10.6). Außerdem fielen die soziale Vergleichsorientierung (Gibbons & Buunck, 1999) und die Orientierung an anderen (Snyder, 1979) bei den Teilnehmern dieser Untersuchung nicht sehr hoch aus.

Als eine weitere wichtige Moderatorvariable kommt die *persönliche Bedeutsamkeit einer korrekten Entscheidung* in Betracht. Mit steigender Motivation, eine korrekte Entscheidung zu fällen, z. B. bei hoher persönlicher Bedeutsamkeit der Entscheidung, steigt der informative Einfluss. Da die Probanden im Rahmen dieser Untersuchung lediglich eine hypothetische Entscheidung treffen mussten, die wahrscheinlich auch nur wenig Bezug zu ihrem alltäglichen Leben aufwies, ist zu vermuten, dass die persönliche Bedeutsamkeit der Entscheidung bei den Probanden eher gering war.

Eine weitere naheliegende Erklärung für die Wirkungslosigkeit der Feedbackinformation konnte auch darin bestehen, dass die als Einflussgruppe verwendeten anderen Versuchsteilnehmer für die Lernenden keine verlässliche Quelle in Bezug auf das Fällen der indivi-

duellen Entscheidung darstellten. Denn die wahrgenommene *Glaubwürdigkeit/Expertise anderer Personen* hat einen starken Einfluss: Je mehr Expertise anderen Menschen zugeschrieben wird, umso wertvoller werden jene als Informationsquelle. Dementsprechend erhöht sich mit steigender Glaubwürdigkeit bzw. Expertise der informationale Einfluss, bei geringer sinkt er. Es ist daher nicht auszuschließen, dass sich bei Verwendung einer anderen Einflussgruppe, wie z. B. Ethikexperten, ein Effekt der Konsensinformation zeigen könnte.

Der informationale Einfluss nimmt aber auch mit steigender *Gruppengröße (Samplegröße)* zu. Die in der Lernumgebung implementierte Feedbackinformation enthielt keine Angabe über die Samplegröße, d. h. darüber, wie viele Versuchsteilnehmer insgesamt abgestimmt hatten. Vermutlich war die Beeinflussungsstärke der Konsensinformation aufgrund der fehlenden Samplegröße einfach zu schwach, das Gesamtergebnis über die Entscheidung der anderen Personen hatte keinen zusätzlichen Informationswert und erhöhte damit auch nicht den informativen Einfluss.

Zur Erklärung der hohen Resistenz gegenüber dem Beeinflussungsversuch bieten auch die Annahmen des HSM von Chaiken et al. (1989) einen angemessenen Rahmen. Die Wirkungslosigkeit der inkongruenten Feedbackinformation kann auch als Indikator für eine systematische Informationsverarbeitung der Probanden angesehen werden, denn über den systematischen Weg gebildete Entscheidungen sind relativ stabil. Offenbar prüften die Lernenden detailliert und kritisch die Reliabilität der präsentierten Konsensinformation, so dass deren Wirkung auf die individuelle Entscheidung nicht zuletzt auch wegen der fehlenden Samplegröße dramatisch abgeschwächt wurde.

Denkbar ist auch, dass die durch das inkongruente Feedback erzeugte kognitive Dissonanz ein Bestreben nach Verteidigung der bereits bestehenden Entscheidung auslöste, die zu einer verzerrten Verarbeitung und Abwertung der inkonsistenten Konsensinformation führte (*Defense-Motiv*, Verteidigungsmotivation; vgl. Chen & Chaiken, 1999; Giner-Sorolla & Chaiken, 1997).

Das Ausbleiben eines Effekts der inkongruenten Feedbackinformation auf die Revision der Entscheidung dürfte vermutlich auch methodisch bedingt sein. Alle Probanden schrieben Essays, in denen sie ihre Entscheidungen so ausführlich wie möglich begründeten. In der Regel geht das Abgeben einer schriftlichen Begründung für eine Entscheidung mit einer erhöhten Überzeugtheit von der Richtigkeit gegenüber der vertretenen Position und einer erhöhten Bindung (*Commitment*) gegenüber der gewählten Entscheidungsalternative einher, welche wiederum zu einem Anstieg des *confirmation bias* führt (vgl. Jonas, Frey et al., 2001; Schwarz, Frey & Kumpf, 1980).

Zudem waren die *zeitlichen Abstände* zwischen den drei zu treffenden Entscheidungen äußerst kurz. Hier dürften auch Gewöhnungseffekte, die dem Messwiederholungsdesign geschuldet sind, eine nicht zu unterschätzende Rolle gespielt haben.

Eine weitere mögliche Ursache besteht in der *Art der Operationalisierung der Variable*. Im Gegensatz zur Studie von Knipfer (2009), die die Meinung zur Nanotechnologie sehr differenziert mithilfe einer 200-stufigen Intervallskala (von -100 bis +100) erfasste (!), wurde in der Untersuchung das Entscheidungsverhalten - der ursprünglichen Konzeption der Lernumgebung entsprechend - lediglich nominalskaliert gemessen. Möglicherweise war dieses methodische Vorgehen nicht sensitiv genug, um Veränderungen innerhalb des Untersuchungszeitraums festzustellen.

Darüber hinaus gab Knipfer (2009) ein *elaborierteres Feedback*, indem sie zusätzlich zur konfligierenden Konsensinformation gezielt drei meinungsinkonsistente Besucherstatements zur Verfügung stellte, d. h. die Probanden erhielten auch Auskunft darüber, *warum* die anderen Studienteilnehmer so denken und auf welchen Argumenten ihre Meinung basierte.

In diesem Zusammenhang stellt sich auch die Frage nach dem Einfluss des *Inhaltsgebietes der Lernumgebung*. Die Nanotechnologie stellt im Gegensatz zur Biotechnologie eine relativ junge Wissenschaft dar, zu der vermutlich noch wenig starke (Vor-) Einstellungen bestehen. In dem in dieser Studie verwendeten Inhaltsgebiet der Medizinethik dürfte dagegen mit stärkeren und robusteren Einstellungen und Wertvorstellungen zu rechnen sein.

Nicht auszuschließen ist auch, dass in dieser Untersuchung ein sog. Bumerang-Effekt auftrat (vgl. Frey & Gaska, 1993): Werden relativ änderungsresistente Kognitionen, wie z. B. stark emotional gefärbte Werte angegriffen, kann die bestehende Dissonanz auch durch eine Verfestigung/Verstärkung der bereits getroffenen Entscheidung reduziert werden. Diese Annahme lässt sich allerdings nicht anhand der Daten überprüfen, da das Entscheidungsverhalten nicht intervallskaliert, sondern lediglich dichotom gemessen wurde.

Überraschenderweise änderten in dieser experimentellen Studie auch 2 Teilnehmer der Bedingung „kongruentes Feedback“ ihre Entscheidung nach Erhalt des Feedbacks und übernahmen die in der Konsensinformation vertretene Minderheitsposition. Leider konnten den Kurzesays keine detaillierten Gründe für diese unerwartete Entscheidungsrevision entnommen werden. Die Übernahme einer Minderheitsposition in die eigene Entscheidung kann bei diesen Personen jedoch vermutlich auf das Bedürfnis nach Einzigartigkeit zurückgeführt werden (Snyder & Fromkin, 1980): Neben dem Wunsch, „dazuzugehören“ können Personen unter bestimmten Bedingungen auch ein Bestreben danach entwickeln, eben nicht zu sein, wie „die anderen“ und unterscheidbar zu sein, um ihre Individualität

und Besonderheit zu betonen. Individuen tendieren insbesondere zur Übernahme einer Minderheitsposition, wenn sie sich von der anonymen „grauen Masse“ der Mehrheit positiv differenzieren können (Imhoff, 2005; Imhoff & Erb, 2009). In diesem Fall erscheint die Minderheitsposition insgesamt positiver als die der Mehrheit, weil sie der Person ermöglicht, die eigene Individualität zu demonstrieren.

Fragestellung 6: Inwieweit führt die Bearbeitung der Zusatzinformationen zu einer Änderung der Entscheidung?

Die Bearbeitung der Zusatzinformationen führte nicht zu einer substantiell häufigeren Revision der zweiten, vorläufigen Entscheidung. Insgesamt betrachtet behielt die überwiegende Mehrheit der Probanden nach Bearbeitung der Zusatzinformationen ihre Entscheidung bei, lediglich 7 revidierten ihre Entscheidung. Damit konnten die Ergebnisse von Götz (2001), der in einer Evaluationsstudie zu einem Lernprogramm zu ethischen Problemen der Gentechnik feststellen konnte, dass die anfänglich getroffenen Entscheidungen nur sehr selten durch die Bearbeitung von Zusatzinformationen geändert wurden, in dieser Studie empirisch bestätigt werden.

Zusammenfassend kann festgehalten werden, dass das Entscheidungsverhalten der Probanden relativ robust war. Weder die Information darüber, wie andere sich entscheiden noch die zusätzlichen, vertiefenden Informationen führten zu einer überzufällig häufigeren Revision der eingangs gefällten Entscheidungen.

10.4 Argumentationsqualität

Fragestellung 7: Inwieweit kommt es durch die Bearbeitung der Lernumgebung zu einem Lernfortschritt und inwiefern wird der Lernfortschritt durch inkongruentes Feedback und durch die Bearbeitung der Zusatzinformationen gefördert?

Als Indikator für den Lernfortschritt wurde in dieser Untersuchung die Argumentationsqualität zu den Messzeitpunkten t1 und t3 (vor und nach Bearbeitung der Zusatzinformationen) herangezogen. Die Qualität der Argumentationen konnte durch das in dieser Studie verwendete Kategorienschema reliabel abgebildet werden. Sowohl die Reliabilitätskoeffizienten für ordinal- als auch für intervallskalierte Daten erwiesen sich als hervorragend reliabel. Einwände gegen eine Berechnung der Daten auf Intervallskalenniveau konnten damit entkräftet werden (vgl. Wirtz & Caspar, 2002). Verwandte Kategoriensysteme zur Klassifikation der Argumentationsqualität haben sich bereits in anderen Studien im Bereich SSI bewährt (vgl. z. B. Sadler & Fowler, 2006; Knipfer, 2009).

Effektivität der Lernumgebung im Hinblick auf den Lernerfolg bzw. Lernfortschritt. Die deskriptiven Befunde ergaben, dass die Argumentationsqualität insgesamt, d. h. über alle drei Lernbedingungen und beide Messzeitpunkte hinweg, auf einem unerwartet hohen Niveau rangierte. Die Mittelwerte zur Argumentationsqualität zu beiden Messzeitpunkten befanden sich nur leicht unter dem mittleren Bereich des sechsstufigen Kategorienschemas. Die Mehrzahl der Versuchspersonen war folglich in der Lage, ihre in Bezug auf die Dilemmageschichte getroffene Entscheidung logisch argumentativ und elaboriert zu begründen. Im Durchschnitt konnten etwa zwei bis drei Begründungen zur Rechtfertigung der eigenen Entscheidung formuliert werden. Eine überraschend große Anzahl der Entscheidungsbegründungen wies demnach eine hohe Argumentationsqualität auf.

Die Leistungen der Probanden können angesichts der anspruchsvollen Dilemmageschichte und vor dem Hintergrund der Befunde von Jiménez-Aleixandre et al. (2000), Osborne et al. (2004), Sadler et al. (2004) sowie Sadler & Fowler (2006), die bei Schülern feststellen konnten, dass Entscheidungen im Bereich SSI häufig nur sehr einfach begründet wurden, somit als gutes Resultat gewertet werden. Es hätten sich durchaus auch Bodeneffekte zeigen können.

Bei näherer Betrachtung der Verteilung der Punktwerte zur Argumentationsqualität zu beiden Messzeitpunkten war aber auch zu erkennen, dass Gegenpositionen nur sehr selten in der eigenen Entscheidung berücksichtigt und überzeugend widerlegt wurden.

Wie erwartet führte die Bearbeitung der Lernumgebung zu einem deutlichen und praktisch bedeutsamen Lernfortschritt, der sich in einem Anstieg der Argumentationsqualität zum dritten Messzeitpunkt manifestierte. Für jede der drei Gruppen ließ sich ein deutlicher Lernfortschritt feststellen. Die Verteilungen der Punktwerte zu beiden Messzeitpunkten für die Gesamtgruppe unterstreichen zusätzlich diesen Befund. Ein Großteil der Lernenden konnte damit einen Lernzuwachs verzeichnen und war in der Lage, nach Bearbeitung der Zusatzinformationen eine informiertere und argumentativ besser begründete Entscheidung in Bezug auf die Dilemmageschichte zur PID-HLA zu fällen.

Für die Lernenden hat sich folglich die Auseinandersetzung mit der Lernumgebung gelohnt. Die Bearbeitung der Zusatzinformationen wirkte sich positiv auf die Argumentationsqualität aus: Viele Probanden nutzten die in den bereitgestellten Zusatzinformationen enthaltenen Argumente zur Begründung ihrer eigenen Entscheidung. Durch die Arbeit mit der Lernumgebung konnte eine informierte und gut begründete Entscheidung im Bereich Medizinethik in hohem Maße gefördert werden. Das grundlegende Ziel der experimentellen Laborstudie wurde damit erreicht und die ersten positiven Befunde hinsichtlich der Lernwirksamkeit von Dialogstationen konnten bestätigt werden (Knipfer, 2009).

Berücksichtigt man die Komplexität der zu lösenden Dilemmageschichte und den Umstand, dass die untersuchte museale Lernumgebung lediglich eine sehr kurzfristige Intervention darstellt, wird deutlich, dass für die Lernenden aller drei Bedingungen erhebliche Lernfortschritte zu verzeichnen waren. Für so eine kurze Intervention ist dies schon eine bemerkenswerte Änderung und spricht für die Lernwirksamkeit der multimedialen Lernumgebung Dialogstation „Gentest“.

Einfluss der Feedbackinformation auf den Lernerfolg bzw. Lernfortschritt. Wider Erwarten ergab sich kein statistisch bedeutsamer Effekt der Feedbackinformation auf den Lernfortschritt. Der Lernfortschritt im Sinne einer höheren Argumentationsqualität und einem geringeren *myside bias* in der Argumentation wurde weder durch die Präsentation einer zur eigenen Entscheidung inkongruenten Konsensinformation noch durch kongruentes Feedback zusätzlich gesteigert. Die positiven Resultate konfligierender Feedbacks im Hinblick auf die Argumentationsqualität, die von Knipfer (2009) anhand eines prototypisch realisierten Diskussionsterminals auf dem Gebiet der Nanotechnologie erzielt wurden, konnten in dieser Untersuchung nicht repliziert werden. Inkongruentes Feedback war also nicht so lernförderlich wie angenommen: Durch die Konfrontation mit einer zur eigenen Entscheidung inkongruenten Feedbackinformation konnte kritisches Denken und eine reflektierte Entscheidung, die auch Gegenpositionen berücksichtigt, im Bereich Medizinethik nicht substantiell unterstützt und gefördert werden.

Fragestellung 8: Inwieweit wird die Generierung von Gegenargumenten durch inkongruentes Feedback gefördert?

*Fragestellung 9: Inwieweit wird der *myside bias* durch inkongruentes Feedback verringert?*

Die Probanden generierten im Mittel etwa 4 Argumente ($M = 4.33$) zur ethischen Bewertung der PID-HLA. Dabei konnten die Probanden durchschnittlich etwa zwei bis drei Argumente zur Unterstützung der eigenen Entscheidung ($M = 2.57$) und etwas weniger als zwei Gegenargumente zu ihrer eigenen Entscheidung formulieren ($M = 1.63$). Folglich konnten mehr präferenzkonsistente als präferenzinkonsistente Argumente in allen drei Gruppen konstruiert werden. Dementsprechend offenbarte der Gesamtmittelwert zum *myside bias* Index ($M = 0.81$) auch, dass eine einseitige Argumentation bei den Untersuchungsteilnehmern vorlag.

In Übereinstimmung mit zahlreichen Studien konnte auch in dieser Untersuchung das Phänomen des *myside bias* empirisch nachgewiesen werden, obwohl explizit zur Nennung von Gegenargumenten aufgefordert wurde (vgl. z. B. Baron, 1995; Toplak & Stanovich, 2003).

Entgegen der Vorannahmen waren die Lernenden der Bedingung „inkongruentes Feedback“ nicht in der Lage, mehr entscheidungsunterstützende Argumente und Gegenargumente zu generieren als die beiden anderen Kontrollbedingungen und wiesen auch keinen geringeren *myside bias* in der Argumentation auf. Die Konfrontation mit einer zur eigenen Entscheidung konfligierenden Feedbackinformation führte demnach nicht dazu, dass mehr Gegenargumente konstruiert werden konnten und verringerte auch nicht den *myside bias* in der Argumentation.

Obwohl sich in dieser Untersuchung nicht alle Erwartungen insbesondere in Hinblick auf die Effektivität der implementierten Feedbackinformation bestätigten, wurden insgesamt betrachtet, d. h. über alle drei Lernbedingungen hinweg, bedeutsame Effekte erzielt, die für die Lernwirksamkeit der von einer Kuratorin und Mediendesignerin realisierten multimedialen Lernumgebung sprechen. Die Mehrheit der Probanden war nach der Lernphase in der Lage, die komplexe und anspruchsvolle Dilemmageschichte zur PID-HLA erfolgreich zu lösen und eine informierte und gut begründete Entscheidung zu treffen. Ähnlich positive Resultate zur Effektivität und Effizienz von Diskussionsterminals wurden von Knipfer (2009) erzielt. Die Lernumgebung scheint demnach bestens geeignet zu sein, die Entscheidungsfindung und Argumentation zu einem kontroversen Wissenschaftsthema zu unterstützen und zwar unabhängig von der zusätzlich integrierten Feedbackinformation.

Einschränkend muss allerdings bemerkt werden, dass durch die Auseinandersetzung mit der Lernumgebung die Berücksichtigung und Widerlegung von möglichen Gegenpositionen in der eigenen Entscheidung nicht ausreichend gestärkt und gefördert werden konnte. Dem häufig auftretenden Problem des *myside bias* in der Argumentation konnte auch mit der Konfrontation einer inkongruenten Feedbackinformation und durch die Bereitstellung von zusätzlichen, vertiefenden Informationen nicht erfolgreich begegnet werden. Sowohl die Feedbackinformation als auch die Zusatzinformationen zahlten sich in dieser Hinsicht nicht aus.

Vielmehr lässt sich an dieser Stelle vermuten, dass die von den Probanden erlebte Dissonanz auf einem anderen Weg reduziert wurde (vgl. Frey & Gaska, 1993) und ein *confirmation bias* in Richtung der getroffenen Entscheidung auftrat und die in den Zusatzinformationen präsentierten Argumente nicht unvoreingenommen, sondern in Abhängigkeit von der bereits bestehenden Entscheidungspräferenz bewertet wurden (vgl. Edwards & Smith, 1996; Frey, 1981, 1986; Greitemeyer et al., 2003; Jonas et al., 2003; Jonas, Schulz-Hardt et al., 2001; Lord et al., 1979; Traut-Mattausch et al., 2004). Denn Argumente, die der eigenen Entscheidung widersprechen, werden typischerweise abgeschwächt oder ignoriert, da präferenzkonsistente Informationen in der Regel für glaubwürdiger, entscheidungsrelevanter und überzeugender gehalten werden als präferenzinkonsistente Argumen-

te. Das Zitat eines Untersuchungsteilnehmers, der angab, dass die Zusatzinformationen für seine Entscheidung irrelevant waren oder diese nur bestärkt haben sowie 5 weitere Zitate von Probanden, die ebenfalls auf eine solche Bestätigungstendenz hinweisen, sprechen für die Plausibilität der Annahme einer verzerrten Informationsbewertung.

Da ein geringer *confirmation* und *myside bias* als zentrale Indikatoren für kritisches Denken und reflektiertes Urteilen angesehen werden, legen die Befunde nahe, dass kritisches Denken und das Treffen einer reflektierten Entscheidung sowohl durch die Auseinandersetzung mit der Lernumgebung als auch durch die Feedbackinformation nur unzureichend unterstützt und gefördert werden konnten (vgl. Baron, 1995; Knipfer, 2009; Toplak & Stanovich, 2003; West et al., 2008).

Auf die Frage, warum die Feedbackinformation keinen bedeutsamen Einfluss auf die Indikatoren der Argumentationsqualität hatte, gibt es eine plausible Antwort: die Gestaltung des Feedbacks. Vermutlich war die alleinige Präsentation einer Konsensinformation in Form eines Umfrageergebnisses zu schwach, um einen lernförderlichen Effekt auf Indikatoren der Argumentationsqualität zu bewirken. Im Rahmen der Lernumgebung wurde lediglich eine *einfache* Rückmeldung gegeben und damit ein *niedriger Informationsgehalt* realisiert. Die Konsensinformation gibt lediglich Aufschluss darüber, wie sich andere Personen in Bezug auf das Fallbeispiel entschieden haben. Knipfer (2009) realisierte in ihrer Untersuchung dagegen einen *hohen Informationsgehalt* und gab *elaborierteres* Feedback, indem sie zusätzlich zur Konsensinformation gezielt drei zur eigenen Meinung inkonsistente bzw. konsistente Statements von anderen Besuchern bereitstellte, die Auskunft geben, *warum* die anderen Studienteilnehmer eine abweichende bzw. die gleiche Meinung vertreten und auf welchen Argumenten möglicherweise ihre Meinung basierte.

Eine Feedbackinformation über das Meinungsbild von anderen Personen mag sich somit in einem anderen Inhaltsbereich wie der Nanotechnologie und in einer elaborierteren Form bewähren, bei den hier untersuchten Probanden im Bereich Medizinethik konnte mit einer einfachen Rückmeldung kein lernförderlicher Effekt nachgewiesen werden.

10.5 Argumentationsmuster

Fragestellung 10: Welche Argumentationsmuster (rational oder emotional-intuitiv) werden am häufigsten zur Begründung der eigenen Entscheidung verwendet?

In Bezug auf die Argumentationsmuster konnte bei der Analyse der Argumentationen aufgezeigt werden, dass die Versuchsteilnehmer ihre Entscheidung bevorzugt aufgrund von rationalen Überlegungen (*reflektiv*) als auf der Basis von Emotionen und Intuitionen (*intui-*

tiv) trafen. Bis auf eine Ausnahme stellten alle Versuchsteilnehmer rationale Kalkulationen an, d. h. sie trafen eine wissens- und wertebasierte Verhaltensentscheidung. In Übereinstimmung mit den Ergebnissen von Sadler & Zeidler (2005) konnte in dieser Studie ebenfalls eine Prädominanz des rationalen Argumentationstyps festgestellt werden, die meisten Probanden zeigten auch eine Kombination der beiden Argumentationsmuster. Die intuitiven und emotionalen Reaktionen gingen häufig dem rationalen Argumentieren voraus und determinierten oft die Entscheidung.

Somit konnte auch in dieser Untersuchung in Übereinstimmung mit den Annahmen des *social intuitionist model* von Haidt (2001) und den Befunden von Sadler & Zeidler (2005) nachgewiesen werden, dass Intuitionen und Emotionen eine wichtige Rolle in moralischen Entscheidungsprozessen einnehmen und das moralische Urteilsverhalten häufig beeinflussen. Im Einklang dazu steht auch, dass neuere Befunde aus der Hirnforschung und der Neurobiologie zeigen, dass Emotionen für das Treffen von (rationalen und vernünftigen) Entscheidungen nicht nur wichtig und hilfreich, sondern sogar unerlässlich sind (vgl. Damásio, 1997, 2011).

Im Hinblick auf die Verwendung des emotional-intuitiven Argumentationsmusters zeigten sich geschlechtsspezifische Unterschiede: Frauen lösten die Dilemmageschichte zur PID-HLA häufiger auf der Basis von Emotionen und Intuitionen als Männer. Dieser Befund lässt sich wahrscheinlich auf die verwendete Dilemmageschichte zurückführen. Beispielsweise konnten Döbert & Nunner-Winkler (1986, zitiert nach Höbke, 2001b) zeigen, dass nicht das Geschlecht per se, sondern die Nähe zu einem Dilemma, d. h. die erzeugte persönliche Betroffenheit, über die Qualität des moralischen Urteils entscheidet. In Analogie dazu lässt sich vermuten, dass bei Frauen durch das Dilemma zur PID-HLA, das zudem von einer weiblichen Person präsentiert wurde, eine stärkere persönliche Betroffenheit erzeugt wurde, die dazu führte, dass mehr emotionale und intuitive Reaktionsweisen hervorgerufen wurden als bei den männlichen Teilnehmern.

Ferner zeigte sich, dass die Verwendung einer emotional-intuitiven Argumentationsweise nicht unabhängig war von der Präferenz für Intuition: Es zeigte sich eine positive und statistisch bedeutsame Korrelation in mittlerer Höhe. Daraus kann gefolgert werden, dass Lernende, die eine höhere Präferenz für Intuition aufwiesen, häufiger bei der Lösung der Dilemmageschichte emotional-intuitiv argumentierten als Lernende, die eine geringere Neigung für affektbasierte Entscheidungsstrategien haben (und umgekehrt). Dieser Befund war aus theoretischer Sicht plausibel und spricht für die Validität der eingesetzten Skala zur Präferenz für Intuition von Betsch (2004).

Bei der Analyse der rationalen Argumentationsmuster fiel auf, dass bei den Studierenden eine utilitaristische Argumentationsweise gegenüber der deontologischen überwog. Die

von den Lernenden getroffenen Entscheidungen in Bezug auf das Dilemma zur PID-HLA wurden eher utilitaristisch als deontologisch begründet und gerechtfertigt. Die meisten Probanden führten häufiger Argumente an, welche die Richtigkeit einer moralischen Handlung nach der Qualität der absehbaren Handlungsfolgen und dem kollektiven Gesamtnutzen beurteilen und nicht nach moralischen Prinzipien, Werten oder Pflichten. Dies steht im Einklang mit den Ergebnissen von Götz (2001) und widerspricht den Befunden von Hößle (2003b), die feststellen konnte, dass bei der Begründung des eigenen moralischen Urteils zum Thema embryonale Stammzellforschung von Schülern deontologische Argumente deutlich häufiger angeführt wurden als utilitaristische Argumente.

Aus pädagogisch-psychologischer Sicht war es sehr bedauerlich, dass nicht alle Argumente aus den bereitgestellten Informationen zur Begründung der eigenen Entscheidung genutzt wurden. Es bietet sich daher an, im Rahmen der Lernumgebung eine andere Darbietungsform der zusätzlichen, vertiefenden Informationen zu erproben (siehe Kapitel 11.2.2).

10.6 Einflussfaktoren auf das Entscheidungsverhalten und die Argumentationsqualität

Fragestellung 12: In welchem Zusammenhang steht die Argumentationsqualität mit den erfassten Lernvoraussetzungen auf der einen und der parasozialen Interaktion mit der Medienfigur auf der anderen Seite?

Die Analyse potenzieller Einflussfaktoren auf das Entscheidungsverhalten und die Argumentationsqualität war ergiebig. In Bezug auf die Argumentationsqualität konnte ein Vorwissenseffekt identifiziert werden: Bei Lernenden mit hohem Vorwissen war die Argumentationsqualität zu beiden Messzeitpunkten höher ausgeprägt als bei Lernenden mit niedrigem Vorwissen.

Damit konnte die allgemeine Bedeutung, die dem Vorwissen in der pädagogisch-psychologischen Forschung im Hinblick auf den Lernerfolg (Ausubel, 1968; Bereiter, 1995; Dochy, 1992; Duit, 1999; Gagné, 1965) und auf eine höherwertige Argumentation (Means & Voss, 1996; Perkins & Salomon, 1989) zukommt, bestätigt werden. Das Vorwissen erwies sich als ein bedeutsamer Einflussfaktor für die Qualität der Argumentation.

Vorwissenstärkere profitierten in höherem Maße von der Auseinandersetzung mit der Lernumgebung und konnten ihre Entscheidung elaborierter begründen als Vorwissenschwächere. Somit konnten die positiven Effekte von inhaltsbezogenem Fachwissen auf die Argumentationsqualität im Bereich SSI auch in dieser Studie empirisch nachgewiesen

werden (Chang & Chiu, 2008; Dawson & Venville, 2009; Sadler & Fowler, 2006; Sadler & Zeidler, 2004).

Allerdings fielen die ermittelten Korrelationskoeffizienten in dieser Studie nicht sehr hoch aus. Wie die deskriptiven Befunde zum Vorwissen über das Inhaltsgebiet der Lernumgebung ergaben, verfügten die Lernenden insgesamt nur über ein sehr geringes subjektives Vorwissen über die PID und über ein durchschnittliches allgemeines Wissen über Human-genetik. Gemäß dem Threshold Model of Content Knowledge Transfer (TMCKT) von Sadler & Donnelly (2006), dass von einem nicht-linearen Zusammenhang zwischen dem Vorwissen und der Argumentationsqualität ausgeht, ist es nicht auszuschließen, dass die korrelationsstatistischen Effekte höher ausgefallen wären, wenn die Lernenden insgesamt über ein höheres Vorwissensniveau zum Thema der Lernumgebung verfügt hätten.

Niedriges Vorwissen ging jedoch mit einer geringeren Argumentationsqualität zu beiden Messzeitpunkten einher. Demzufolge konnten die bestehenden Vorwissensunterschiede zwischen den Lernenden auch durch die Bearbeitung der zusätzlichen, vertiefenden Informationen nicht kompensiert werden. Es ist in diesem Zusammenhang durchaus denkbar, dass für Lernende mit einem geringen Vorwissen das Fallbeispiel und die zusätzlichen Informationen zu komplex waren und zu kognitiver Überlastung geführt haben (Chandler & Sweller, 1991) und sie dementsprechend nicht in der Lage waren, die präsentierten Informationen und Argumente auch effektiv in die Begründung der eigenen Entscheidung miteinzubeziehen. Die Lernenden mit hohem Vorwissen konnten offenbar die Zusatzinformationen besser in ihre schon vorhandenen Wissensstrukturen einbinden und relevante Informationen und Argumente aus den Lerninhalten auch effektiver zur Begründung der eigenen Entscheidung nutzen.

Bei allen anderen untersuchten Lernvoraussetzungsaspekten zeigten sich durchweg niedrige und sehr niedrige Korrelationen mit der Qualität der Argumentation zu beiden Messzeitpunkten. Die Argumentationsqualität war folglich - mit Ausnahme des Vorwissens - weitgehend unabhängig von den erfassten Lernvoraussetzungen.

Im Hinblick auf den Einfluss der PSI mit der Medienfigur auf die Argumentationsqualität zu beiden Messzeitpunkten zeigte sich lediglich ein Effekt auf kognitiver Ebene. Je stärker sich die Lernenden kognitiv mit der Medienfigur und den von ihr vermittelten Inhalten beschäftigten, desto höher fiel die Qualität der Argumentation aus (und umgekehrt). Eine höhere kognitive PSI ging demnach mit einer besseren Begründung der Entscheidung und besseren Lernergebnissen einher. Die narrative und personalisierte Präsentation einer Dilemmageschichte durch einen menschlichen sprechenden Kopf (*talking head*) förderte die Motivation der Lernenden, sich stärker kognitiv mit der Medienperson und den in der Lernumgebung präsentierten Informationen und Argumenten auseinanderzusetzen.

Dieser Befund ist keineswegs trivial, denn es hätten sich durchaus in Anlehnung an die Cognitive Load Theory (Chandler & Sweller, 1991; Sweller, 1999; Sweller et al., 1998) auch negative Effekte des eingesetzten Präsentationsformats zeigen können. Das Aufzeigen einer sichtbaren Person hätte die Lernenden auch ablenken und zu einer extrinsischen kognitiven Belastung (*Extraneous Load*) und Beeinträchtigung der Verarbeitung der eigentlich relevanten Informationen und Argumente führen können (*seductive detail effect*) (Mayer, 2005c). Die obig dargelegten Befunde sprechen dafür, dass dem nicht so war.

Damit konnte in Übereinstimmung mit der Social Agency Theory (Mayer, 2005c) und den Ergebnissen von Töpper (2009) der lernförderliche Einfluss personalisierter Filme auch in dieser Untersuchung nachgewiesen werden. Das im Rahmen der Lernumgebung eingesetzte Präsentationsformat eines sichtbaren *talking heads*, der den Besucher direkt in einem dialogorientierten Sprachstil adressiert, hat sich folglich bewährt.

Fragestellung 13: Welchen Einfluss haben die erfassten Lernvoraussetzungen und die Sicherheit der Entscheidung auf die Revision der Entscheidung?

In Bezug das Entscheidungsverhalten konnte festgestellt werden, dass lediglich die Sicherheit der Entscheidung einen Einfluss auf die Änderung der Entscheidung hatte: Wer sich in seiner Entscheidung sehr unsicher war, revidierte signifikant häufiger seine zuvor getroffene Entscheidung. Dieser Befund war inhaltlich plausibel und spricht in Übereinstimmung mit den Ergebnissen zur Änderung der Entscheidung für sich (siehe Kapitel 10.3.).

11 Gesamtdiskussion und Konsequenzen

Zum Abschluss dieser Arbeit werden zunächst die zentralen Ergebnisse der experimentellen Laborstudie kurz zusammengefasst und diskutiert (Kapitel 11.1), bevor Konsequenzen für die weitere Forschung und Praxis aufgezeigt werden (Kapitel 11.2).

11.1 Zusammenfassung und Diskussion der zentralen Ergebnisse

Zentrales Ziel dieser Arbeit war es, die Lernwirksamkeit einer von einer Kuratorin und Mediendesignerin konzipierten multimedialen Lernumgebung im Bereich Medizinethik im Hinblick auf die Förderung der Entscheidungsfindung und Argumentation empirisch zu untersuchen und nachzuweisen. Ein weiteres Studienziel bestand darin, die Wirkung bzw. den Einfluss einer implementierten Feedbackinformation in Form eines Umfrageergebnisses über die durchschnittliche Gesamtentscheidung anderer Personen auf die Entscheidung und die Argumentationsqualität zu untersuchen. Ferner sollten auch Erkenntnisse über die Bedeutung von verschiedenen Lernvoraussetzungen (z. B. Vorwissen) und der Medienfigur (*talking head*) im vorliegenden Kontext gewonnen werden.

Um die Effektivität der multimedialen Lernumgebung Dialogstation „Gentest“ und der Feedbackinformation im Hinblick auf die Förderung einer informierten und reflektierten Entscheidung im Bereich Medizinethik/PID zu überprüfen, wurde eine experimentelle Laborstudie durchgeführt. Der Studie lag ein 3 (Feedback: kongruent, inkongruent, kein) x 3 (Messzeitpunkt: t1: Film, t2: Feedback, t3: Informationen) faktorielles, gemischtes Design mit Messwiederholung auf dem zweiten Faktor zugrunde. Der „between-subject“-Faktor Feedback wurde in Abhängigkeit von der getroffenen Entscheidung gezielt manipuliert: Variiert wurde die Kongruenz (Übereinstimmung) zwischen der eigenen Entscheidung und der Konsensinformation über die durchschnittliche Gesamtentscheidung der anderen Versuchsteilnehmer. Im Rahmen dieser Untersuchung wurde eine Analyse der parasozialen Interaktion mit der Medienfigur, des Entscheidungsverhaltens, der Argumentationsqualität, der Argumentationsmuster und von Einflussfaktoren auf die Entscheidung und die Argumentationsqualität durchgeführt.

Nachfolgend werden bezogen auf diese fünf Hauptanalyseschwerpunkte die zentralen Ergebnisse im Hinblick auf die grundlegenden Ziele der Studie nochmals kurz zusammengefasst und diskutiert: die Lernwirksamkeit der multimedialen Lernumgebung (Kapitel 11.2.1), der Einfluss der Feedbackinformation auf die Entscheidung und die Argumentati-

onsqualität (Kapitel 11.1.2) und die Bedeutung der Medienfigur und des Vorwissens im vorliegenden Lernkontext (Kapitel 11.1.3).

11.1.1 Lernwirksamkeit der multimedialen Lernumgebung

Mit der vorliegenden Arbeit wurde versucht, einen ersten Beitrag zur Lernwirksamkeit einer innovativen musealen Lernumgebung zu leisten.

Das primäre Lernziel der multimedialen Lernumgebung besteht aus kognitiver Sicht darin, das Fällen einer informierten, gut begründeten und reflektierten Entscheidung im Bereich Medizinethik angemessen zu unterstützen und zu fördern. Der Lernerfolg bzw. Lernfortschritt wurde mit Hilfe von zwei Essays anhand einer Prä-/Post-Messung der Argumentationsqualität erhoben. Im Rahmen der Lernumgebung wurden als instruktionale Unterstützungsmaßnahme nach dem Treffen der eigenen Entscheidung zusätzliche, vertiefende Informationen bereitgestellt, die mit einer Verbesserung der Argumentationsqualität einhergehen sollten. Es wurde angenommen, dass sich die Auseinandersetzung mit der multimedialen Lernumgebung bzw. mit den Zusatzinformationen positiv auf die Argumentationsqualität auswirkt und sich ein deutlicher Lernfortschritt zeigt.

Insgesamt betrachtet, d. h. über alle Gruppen hinweg, konnte gezeigt werden, dass die Bereitstellung von zusätzlichen, vertiefenden Informationen mit einer höheren Argumentationsqualität und einem praktisch bedeutsamen Lernfortschritt einherging.

Das grundlegende Ziel der Studie wurde damit erreicht: Mit den zur Erfassung der Lernwirksamkeit eingesetzten Methoden konnten positive Effekte im Hinblick auf die Unterstützung und Förderung einer informierten und gut begründeten Entscheidung festgestellt werden. Dieser Befund entspricht den Ergebnissen von Knipfer (2009) und bestätigt die angenommene Lernwirksamkeit von Dialogstationen.

Allerdings zeigte sich trotz dieser positiven Ergebnisse ein starker *myside bias* in der Begründung der eigenen Entscheidung und bei der Generierung von Gegenargumenten, d. h. die Lernenden setzten sich nicht genügend mit möglichen Gegenpositionen zu ihrer eigenen Entscheidung auseinander und integrierten Gegenargumente nicht in ihre Entscheidung. Die Berücksichtigung und Widerlegung von möglichen Gegenpositionen in der eigenen Entscheidung konnten demnach durch die Auseinandersetzung mit der Lernumgebung nicht angemessen unterstützt und gefördert werden. Einige Zitate von Untersuchungsteilnehmern deuten auch daraufhin, dass ein *confirmation bias* beim Umgang mit den Zusatzinformationen auftrat. Ein geringer *confirmation* und *myside bias* gelten jedoch

als zentrale Indikatoren für kritisches Denken und reflektiertes Urteilen (vgl. Baron, 1995; Knipfer, 2009; Toplak & Stanovich, 2003; West et al., 2008).

Im Rahmen der Analyse der Argumentationsmuster konnte zusätzlich gezeigt werden, dass bei der Lösung der Dilemmageschichte zur PID häufiger rational argumentiert wurde als emotional-intuitiv. Die Entscheidungsfindung erfolgte somit vorwiegend auf rationalem Wege als auf der Basis von Emotionen und Intuitionen. Allerdings spielten die emotionalen und intuitiven Reaktionen eine sehr wichtige Rolle bei der Entscheidungsfindung, da sie oft dem rationalen Argumentationstyp vorausgingen und das Fällen der Entscheidung recht häufig determinierten.

Zusammenfassend betrachtet konnte durch die Auseinandersetzung mit der multimedialen Lernumgebung zwar das Treffen einer informierten und argumentativ gut begründeten Entscheidung zu einem medizinethischen Dilemma gefördert werden, kritisches Denken und eine reflektierte Entscheidung, die auch explizit mögliche Gegenpositionen berücksichtigt, konnten dagegen nicht adäquat unterstützt werden. In dieser Hinsicht konnten die positiven Befunde von Knipfer (2009) zum Lernpotential von Diskussionsterminals leider nicht repliziert werden.

11.1.2 Einfluss der Feedbackinformation auf die Entscheidung und Argumentationsqualität

Im Rahmen der Arbeit wurde auch der Fragestellung nachgegangen, inwiefern durch eine Feedbackinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung anderer Personen kritisches Denken sowie eine informierte und reflektierte Entscheidung im Bereich Medizinethik angemessen gefördert werden kann. Es wurde angenommen, dass vor allem die Konfrontation mit einer zur eigenen Entscheidung inkongruenten Feedbackinformation positive Auswirkungen auf Indikatoren der Argumentationsqualität hat. Die Annahmen eines lernförderlichen Einfluss der inkongruenten Feedbackinformation beruhen auf einer Studie von Knipfer (2009) (siehe Kapitel 4.6).

Wider Erwarten konnte empirisch jedoch kein Einfluss der Feedbackinformation auf Indikatoren der Argumentationsqualität nachgewiesen werden. Inkongruentes Feedback führte im Vergleich zu den beiden anderen Lernbedingungen weder zu einer höheren Argumentationsqualität im abschließenden Essay noch zu einem geringeren *myside bias* in der Argumentation. Lernende, die ein inkongruentes Feedback nach dem Treffen ihrer eigenen Entscheidung erhielten, konnten ihren Standpunkt nicht elaborierter begründen, nicht mehr Argumente gegen ihre Entscheidung anführen und wiesen keinen geringeren *myside bias*

in der Argumentation auf als Lernende, die „kongruentes“ oder „kein Feedback“ erhielten. Die Befunde von Knipfer (2009) zur positiven Wirkung konfligierenden Feedbacks auf Indikatoren der Argumentationsqualität konnten damit im Bereich Medizinethik nicht repliziert werden. Obwohl die Konfrontation mit einer zur eigenen Entscheidung im Konflikt stehenden Feedbackinformation nachweislich kognitive Dissonanz auslöste und psychisch unangenehm erlebt wurde, revidierten Lernende der Bedingung „inkongruentes Feedback“ nicht häufiger ihre zuvor abgegebene Entscheidung und übernahmen nicht die in der Konsensinformation vertretene Mehrheitsposition. Auch in dieser Hinsicht konnten die von Knipfer (2009) gewonnenen Ergebnisse im Bereich Nanotechnologie nicht bestätigt werden.

Feedback in Form einer sozialen Vergleichsinformation über die Gesamtentscheidung anderer Personen hatte somit entgegen der Vorannahmen keinen nachweislichen Effekt auf die Entscheidung und die Argumentationsqualität: Durch die Konfrontation mit einer inkongruenten Feedbackinformation wurden weder das Entscheidungsverhalten noch die Argumentationsqualität beeinflusst.

11.1.3 Bedeutung der Medienfigur und des Vorwissens im vorliegenden

Lernkontext

Wie die Analyse von Einflussfaktoren auf die Argumentationsqualität zeigen konnte, kam der Medienfigur (*talking head*) und dem Vorwissen zum Inhaltsgebiet beim Lernen mit der multimedialen Lernumgebung Dialogstation „Gentest“ eine bedeutsame Rolle zu.

Es konnte im Einklang mit den Annahmen der Social Agency Theory (Mayer, 2005c) und den Ergebnissen von Töpper (2009) nachgewiesen werden, dass die Präsentation einer Dilemmageschichte durch eine Medienfigur (*talking head*) einen lernförderlichen Effekt hatte. Die Qualität der Argumentation zu beiden Messungen war positiv assoziiert mit der kognitiven parasozialen Interaktion mit der medialen Person. Lernende, die sich kognitiv intensiv mit der Medienfigur und den von ihr vermittelten Inhalten beschäftigten, konnten ihre Entscheidung besser begründen als Lernende, die sich nur oberflächlich-schwach mit der Persona auf kognitiver Ebene auseinandersetzen.

Einen weiteren wichtigen Einflussfaktor auf die Argumentationsqualität bildete im vorliegenden Lernkontext das Vorwissen zum Inhaltsgebiet der Lernumgebung. Es ergaben sich positive und substantielle Korrelationen des Vorwissens über Humangenetik und zur PID mit der Argumentationsqualität zu beiden Messungen. Vorwissensstärkere profitierten demnach stärker von der Auseinandersetzung mit der Lernumgebung als Vorwissens-

schwächere. Folglich kompensierten die zusätzlich bereitgestellten Informationen nicht ein geringes Vorwissen.

11.2 Konsequenzen für die Forschung und Praxis

Ausgehend von der zusammenfassenden Darstellung der zentralen empirischen Ergebnisse der experimentellen Untersuchung werden nun Konsequenzen für die weitere Forschung (Kapitel 11.2.1) und pädagogische Praxis gezogen (Kapitel 11.2.2).

11.2.1 Konsequenzen für die Forschung

Aus den Ergebnissen der vorliegenden Arbeit lassen sich folgende Implikationen für die weitere Forschung ableiten.

Das untersuchungsmethodische Vorgehen hat sich in dieser Arbeit als erfolgreich erwiesen und bewährt: Die experimentelle Laborstudie war hinreichend intern valide, da der Lernfortschritt in Form einer höheren Argumentationsqualität nach Bearbeitung der Zusatzinformationen eindeutig auf die multimediale Lernumgebung zurückgeführt werden konnte. Ferner konnte die Argumentationsqualität bzw. der Lernfortschritt mit dem in dieser Studie eingesetzten Kategorienschema reliabel gemessen werden. Es war im Hinblick auf die Lernwirksamkeit der musealen Lernumgebung sehr erstaunlich, dass trotz dieser sehr kurzfristigen Intervention überhaupt ein statistischer Nachweis eines Effektes gelang. Die Studie gibt weiterhin Aufschluss darüber, dass im Bereich Medizinethik eine Feedbackinformation in Form eines Umfrageergebnisses über die Gesamtentscheidung anderer Personen keinen bedeutsamen Einfluss auf das individuelle Entscheidungsverhalten und die Argumentationsqualität hatte.

Allerdings geht die strikte Kontrolle und Ausschaltung untersuchungsbedingter Störvariablen, die eine hohe interne Validität garantieren, häufig zu Lasten der externen (ökologischen) Validität (vgl. Bortz & Döring, 2006). Die Studie fand in einem klar definierten und künstlichen Setting statt, das nur wenig mit dem informellen und selbstgesteuerten Lernen im realen Ausstellungskontext gemein hat. Als Untersuchungsobjekte wurden nur Studierende ausgewählt und die Studie bezog sich lediglich auf einen klar abzugrenzenden Inhaltsbereich, einem Fallbeispiel zur PID-HLA. Es ist daher anzunehmen, dass die experimentell gewonnenen Ergebnisse nicht vollständig auf andere Personengruppen wie das Museumspublikum, die informelle Lernsituation im Museum oder die anderen Fallbeispiele der Lernumgebung bzw. andere Inhaltsgebiete generalisiert werden können.

Eine weitere wesentliche Einschränkung der Generalisierbarkeit der Untersuchungsergebnisse besteht darin, dass ein Messwiederholungsdesign im Rahmen des Experiments verwendet wurde. Durch dieses Vorgehen konnte zwar eine detaillierte Analyse der Entscheidungsprozesse und des Lernfortschritts zu den einzelnen Messzeitpunkten unter kontrollierten Bedingungen gewährleistet werden, allerdings bringt eine Messwiederholung auch einige Nachteile mit sich, die sich unter dem Begriff der Sequenzeffekte subsumieren lassen (Hussy, Schreier & Echterhoff, 2009): Durch die wiederholte Messung kann es zu kurzfristigen Übungs- und Gedächtniseffekten, aber auch Ermüdungs- und Sensibilisierungseffekten kommen, die während und nach der Bearbeitung der Lernumgebung wirksam werden und das Ergebnis der Post-Messung (positiv oder negativ) beeinflussen können.

Zur Kontrolle von Übungs- und Gedächtniseffekten und zur Überprüfung der Nachhaltigkeit der erzielten kognitiven Effekte der Lernumgebung würde sich daher die Durchführung einer Follow-Up-Messung empfehlen. Auf diese Weise könnte auch überprüft werden, ob die Lernumgebung zum nachhaltigen Lernen anregt (vgl. Prenzel, 2009) und zu einer späteren Auseinandersetzung mit dem Thema führt (z. B. Lesen von vertiefender Literatur und Rezipieren von Medienbeiträgen). Leider ließ sich in dieser Studie eine Follow-Up-Messung aus organisatorischen und ökonomischen Gründen nicht realisieren.

Eine weitere Restriktion der Studie liegt darin, dass die Probanden in dieser Untersuchung ein Honorar für ihre Teilnahme erhielten. Dies mag zu größerer Persistenz und mentaler Anstrengung bei den Versuchsteilnehmern geführt haben als dies möglicherweise beim selbstgesteuerten, informellen Lernen im musealen Setting mit der Lernumgebung der Fall wäre.

Die experimentell gewonnenen Befunde zur Lernwirksamkeit der Lernumgebung bedürfen demzufolge der Replikation unter ökologisch validen Bedingungen. In weiteren Studien im Feld sollte geklärt werden, ob die in diesem Experiment vorgefundenen Ergebnisse auch auf das museale Setting übertragbar und unter "realen" Lernbedingungen zu beobachten sind. Da die multimediale Lernumgebung bereits in einer Museumsausstellung zur Nano- und Biotechnologie zum Einsatz kommt, bietet sich insbesondere ein Feldexperiment an. Zudem könnten auch Untersuchungen mit den anderen Fallbeispielen der Lernumgebung und an einer größeren Population die Allgemeingültigkeit der Befunde vergrößern.

Um eine Brücke zwischen Theorie und Praxis zu schlagen (vgl. Fischer, Waibel & Wecker, 2005; Hauser, Noschka-Roos, Reussner & Zahn, 2009), empfiehlt sich besonders die Anwendung des sog. integrativen Forschungsparadigmas (Stark, 2001; Stark & Mandl, 2001b), das versucht, wissenschaftliches Wissen auf eine anwendungsorientierte Weise zu generieren und dadurch die Kluft zwischen Theorie und Praxis zu reduzieren. Grundprin-

zip des integrativen Forschungsansatzes bildet die explizit anwendungsbezogene Generierung wissenschaftlicher Erkenntnisse, weitere zentrale Prinzipien sind sukzessive Optimierung, Replikation und Evaluation in der Praxis, d. h. die Verifizierung des experimentell gewonnenen Wissens im Feld. Der integrative Forschungsansatz ermöglicht es, nicht nur theoretische, sondern auch praxisrelevante Erkenntnisse zu generieren, die zur Optimierung der Lernumgebung genutzt werden können. Die Evaluation und Weiterentwicklung der multimedialen Lernumgebung sollte dementsprechend anhand systematisch aufeinander aufbauenden Feld- und Laborstudien erfolgen.

Eine angemessene und sinnvolle Herangehensweise für Feldstudien im Ausstellungssetting stellt in Anbetracht des oftmals selektiven und zeitlich eingeschränkten Nutzungsverhaltens der Museumsbesucher eine Kombination des sog. *cued* und *non-cued testing* dar (vgl. Noschka-Roos, 1994). Zum einen sollten die Ausstellungsbesucher indirekt, d. h. versteckt, während der Beschäftigung mit der Lernumgebung beobachtet und nach der Nutzung mithilfe eines Fragebogens und/oder halbstrukturierten Interviews befragt werden (*con-cued testing*). Zum anderen sollten auch Museumsbesucher direkt vor der Nutzung der Lernumgebung rekrutiert werden, um eine Prä-Post-Studie im Feld durchführen zu können (*cued testing*).

Um Veränderungen im Entscheidungsverhalten und der Argumentationsqualität differenzierter nachweisen zu können, wäre es für weitere Studien auch wünschenswert, längere Untersuchungszeiträume realisieren zu können, z. B. indem innerhalb der Lernumgebung mehr Informationen bereitgestellt werden.

Im Fokus der präsentierten Studie stand lediglich die Überprüfung der kognitiven Lernziele der musealen Lernumgebung. Aus kognitiver Sicht konnten mit der multimedialen Lernumgebung positive Effekte erzielt werden. Eine erfolgreiche Lernumgebung sollte jedoch neben positiven kognitiven Aspekten auch möglichst günstige motivationale Effekte hervorrufen, insbesondere im Hinblick auf einen Einsatz in der Praxis (vgl. Stark, Gruber, Renkl & Mandl, 1998). In dieser Arbeit konnte eine zusätzliche Überprüfung der motivationalen Lernziele aufgrund zeitlicher Rahmenbedingungen und aus ökonomischen Gründen nicht realisiert werden. Ob die positiven kognitiven Effekte der Lernumgebung auch mit positiven motivationalen Effekten einhergehen, sollte daher noch in weiteren Studien, insbesondere im Feld, geklärt werden. Angesichts der Befunde zum Einfluss von Intuitionen und Emotionen auf das moralische Urteilsverhalten sollten in Folgestudien generell auch verstärkt emotionale Aspekte erfasst werden.

Dazu kommt, dass auf der Basis der Befunde über die genauen Gründe der ausbleibenden Wirksamkeit der Feedbackinformation sowie der mangelnden Integration von möglichen Gegenpositionen in der Begründung der eigenen Entscheidung bei den Probanden nur spe-

kuliert werden kann. Um diesbezüglich weitere Erkenntnisse zu erhalten, erscheint es sinnvoll, im Rahmen von Replikationsstudien das Lerngeschehen differenzierter und umfassender (unter besonderer Berücksichtigung von emotionalen und motivationalen Aspekten) zu erfassen, z. B. mit Hilfe von entsprechenden Fragebögen, nachträglichem lauten Denken oder halbstrukturierten Interviews.

In Anbetracht der Befunde zum psychologischen Unbehagen nach dem Feedback sollten auch die emotionalen Konsequenzen von Feedback näher untersucht werden (Heckhausen, 1989; Musch, 1999), beispielsweise ob eine negative Abweichung des Feedbacks mit der eigenen Entscheidung für den Lernerfolg ungünstige Emotionen wie Scham, Peinlichkeit, Ärger oder Angst auslöst und eine positive Rückmeldung günstige, z. B. Stolz, nach sich zieht.

Darüber hinaus würde es sich auch anbieten, konkret zu untersuchen, ob und inwieweit beim Umgang mit den zusätzlich bereitgestellten Informationen systematische Verzerrungen zugunsten präferenzkonsistenter gegenüber -inkonsistenten Informationen bei den Lernenden zu beobachten sind (vgl. z. B. Greitemeyer et al., 2003; Jonas, Frey et al., 2001; Jonas et al., 2003).

Ein methodischer Schwachpunkt der Studie besteht in der Berechnung des *myside bias* Index. In den Studien von Toplak & Stanovich (2003) und Knipfer (2009) wurde der *myside bias* Index über eine schriftliche Abfrage (*recall*) aller erinnerten Argumente für (*mysided arguments*) und gegen die eigene Position (*othersided arguments*) erhoben. In dieser Untersuchung wurden die Probanden jedoch lediglich explizit dazu aufgefordert, Gegenargumente zu nennen, auf einen *recall* aller erinnerten entscheidungsunterstützenden Argumente wurde verzichtet, um die Probanden nicht durch eine zu umfangreiche Datenerhebung zu ermüden und zu demotivieren. Stattdessen wurden zur Bestimmung der Anzahl der entscheidungsunterstützenden Argumente die beiden Essays zur Begründung der eigenen Entscheidung verwendet. In diesem Zusammenhang ist es durchaus denkbar, dass die Probanden vielleicht mehr entscheidungsunterstützende (*mysided*) Argumente hätten generieren können als sie tatsächlich zur Begründung ihrer eigenen Entscheidung herangezogen haben. Möglicherweise wäre damit der *myside bias* in der Argumentation bei den Probanden in dieser Untersuchung sogar noch viel höher ausgefallen. Diese vorliegende methodische Verzerrung galt jedoch für *alle* drei Lernbedingungen und ändert nichts am wesentlichen Ergebnis dieser Studie, dass ein *myside bias* in der Argumentation bei den Untersuchungsteilnehmern vorlag und sowohl durch inkongruentes Feedback als auch durch die Bearbeitung der Lernumgebung nicht verringert werden konnte.

Eine weitere methodische Schwierigkeit ergab sich bei der Erfassung der individuellen Lernvoraussetzungen zum Inhaltsgebiet der Lernumgebung. Die in diesem Experiment

eingesetzten Ratingskalen zum Interesse und den Einstellungen zum Lernthema wurden auf die übergeordnete Domäne „genetische Untersuchungen“ adaptiert, da ausgehend von Befunden einer deutschlandrepräsentativen Studie zum damaligen Erhebungszeitraum im Jahr 2009 nur mit einem sehr geringen Vorwissen zur PID zu rechnen war (vgl. Brähler & Stöbel-Richter, 2004). Dieses methodische Vorgehen war sicherlich nicht optimal, insbesondere im Hinblick auf die Analyse von Einflussfaktoren auf die Entscheidung und Argumentationsqualität. In Anbetracht der im Jahr 2010 und 2011 wieder intensiv entfachten medialen Diskussion zur ethischen Legitimität der PID kann jedoch zum gegenwärtigen Zeitpunkt davon ausgegangen werden, dass sich der Wissensstand der Allgemeinbevölkerung zur PID erhöht hat und entsprechende Einstellungen zu diesem Thema gebildet wurden (vgl. Singer et al., 1999; Stockdale, 1999). In Folgestudien könnten daher das Interesse und die Einstellungen der Lernenden zum Thema PID erfasst und deren Einfluss auf das Entscheidungsverhalten und die Argumentationsqualität ermittelt werden.

Zusammenfassend lässt sich festhalten, dass trotz der genannten Limitationen der in dieser Arbeit berichteten Untersuchungsergebnisse Dialogstationen zur Unterstützung und Förderung einer informierten und gut begründeten Entscheidung in Bezug auf kontroverse Wissenschaftsthemen im Museumskontext vielversprechend erscheinen. Durch die Ergebnisse des Experiments wurden jedoch auch viele neue Fragen aufgeworfen, die verdeutlichen, dass noch erheblicher Forschungsbedarf im Hinblick auf die Lernwirksamkeit dieser innovativen Lernumgebungen und insbesondere der Feedbackinformation besteht. Die Ergebnisse dieser Pilotstudie können vor allem zur Hypothesengenerierung für künftige Folgestudien genutzt werden.

11.2.2 Konsequenzen für die pädagogische Praxis

Die vorliegende Arbeit konnte einen ersten wertvollen Beitrag zum Nachweis der Effektivität einer innovativen und an die speziellen Bedingungen des informellen Lernorts Museum adaptierten Lernumgebung leisten. Ausgehend von den positiven Befunden zur Lernwirksamkeit kann für die museumspädagogische Praxis die Schlussfolgerung gezogen werden, dass sich der Einsatz von multimedialen Lernumgebungen wie der Dialogstation „Gentest“ im Rahmen von Ausstellungskonzepten im Museum empfiehlt.

Es spricht für die Effektivität und Effizienz der Lernumgebung, dass angesichts dieser sehr kurzfristigen Intervention eine informierte und gut begründete Entscheidung überhaupt in einem praktisch relevanten Ausmaß unterstützt und gefördert werden konnte. Folglich hat sich die Dialogstation „Gentest“ im Hinblick auf die Vermittlung eines kontroversen und gesellschaftlich relevanten Wissenschaftsthemas bewährt und erweist sich als geeignet zur

Unterstützung einer informierten und argumentativ gut begründeten Entscheidung im musealen Lernkontext.

Da Kuratoren und Ausstellungsgestalter in erster Linie an praktischen Empfehlungen für ihre Ausstellungsarbeit interessiert sind, sollen die in dieser Arbeit berichteten Ergebnisse auch unter einer praktischen Perspektive interpretiert werden, um Implikationen für die Optimierung der Lernumgebung identifizieren und ableiten zu können.

Sehr erfreulich war es zunächst, dass ein lernförderlicher Effekt des im Rahmen der Lernumgebung eingesetzten Präsentationsformats in Form eines *talking heads* nachgewiesen werden konnte, da die Erstellung dieser virtuellen Figuren sehr zeit- und kostenintensiv ist. Wie die Befunde zeigen konnten, stellte die Attraktivität der Medienfigur einen entscheidenden Einflussfaktor auf die Intensität der PSI-Prozesse dar. Um das Zustandekommen von parasozialen Interaktionen zu begünstigen, sollte deswegen bei der medialen Gestaltung von *talking heads* verstärkt darauf geachtet werden, möglichst attraktive Medienfiguren zu verwenden.

Die Ergebnisse zur Analyse der Argumentationsqualität offenbaren aber auch, dass durch die Auseinandersetzung mit den bereitgestellten Zusatzinformationen die Generierung und die Berücksichtigung von möglichen Gegenpositionen in der eigenen Entscheidung nur unzureichend unterstützt und gefördert werden konnte. Ferner gaben 6 Probanden weiteren Informationsbedarf an und es wurden zwei Argumente aus den Zusatzinformationen überhaupt nicht oder nur sehr wenig in die eigene Entscheidung mit einbezogen. Darüber hinaus konnten durch die zur Verfügung gestellten Zusatzinformationen auch bestehende Vorwissensunterschiede zwischen den Lernenden nicht ausreichend kompensiert werden. Für Lernende mit weniger Vorwissen war das verwendete Fallbeispiel zur PID-HLA wahrscheinlich zu komplex. Zusätzlich konnte im Rahmen einer formativen Evaluationsstudie im Feld, d. h. unter „regulären“ Lernbedingungen, anhand einer Logfile-Analyse festgestellt werden, dass die Zusatzinformationen nur sehr wenig von den Besuchern genutzt wurden.

Angesichts dieser Befunde liegt es auf der Hand, dass die Lernenden, insbesondere die vorwissensschwächeren, im Rahmen der multimedialen Lernumgebung einer besseren instruktionalen Unterstützung bedürfen (Fischer et al., 2009). Zur Optimierung der Lernumgebung erscheint es daher sinnvoll, andere Formen der instruktionalen Unterstützung in die Lernumgebung zu integrieren und zu erproben.

Ein konkreter Verbesserungsvorschlag bezieht sich auf die Darbietungsform der zusätzlichen und vertiefenden Informationen in Textform. Als effektivere Unterstützungsmaßnahme würde sich konkret in Anlehnung an den „Cognitive Apprenticeship“-Ansatz (vgl.

Collins et al., 1989) das *kognitive Modellieren* anbieten, indem zwei Experten (z. B. Humangenetiker, Arzt, Ethiker) modellhaft eine Beispiel- bzw. Musterlösung jeweils für die Pro- bzw. Contra-Position zur Dilemmageschichte demonstrieren. Jede Beispiellösung der Experten sollte dabei auch mögliche Gegenpositionen diskutieren und entkräften. Da das Lesen von Texten am Bildschirm generell sehr anstrengend ist, wäre es aus medialer Sicht überlegenswert, Ton- oder Video-Elemente anstatt On-Screen-Text zu verwenden. Indem zwei Audio- oder Video-Interviews von einem Experten für die Pro- und einem Experten für die Contra-Position auf dem Touchscreen abgespielt werden, könnte die kognitive Modellierung mit einem relativ geringen Aufwand implementiert werden. Im Hinblick auf die didaktische und mediale Gestaltung der Lernumgebung scheint diese Art der instruktionalen Unterstützung geeigneter zu sein, als die Bereitstellung zusätzlicher Informationen in Textform, die nicht nach Pro oder Contra, sondern nach Fakten, Bewertung und Gesetze geordnet sind.

Ein weiterer Verbesserungsvorschlag betrifft den Zeitpunkt des Abrufs der Zusatzinformationen. In Anlehnung an Knipfer (2009) und den Grundannahmen der Theorie der kognitiven Dissonanz (vgl. Festinger, 1957; Frey & Gaska, 1993) sollte zur Vermeidung eines *confirmation bias* die Auseinandersetzung mit den zusätzlichen Informationen erfolgen, bevor die eigene Entscheidung getroffen wird. Dementsprechend dürfte es grundsätzlich dysfunktional sein, die zusätzlichen Informationen erst nach der eigenen Positionierung bereitzustellen. Aus motivationaler Sicht sollten die Lernenden bei der Arbeit mit der Lernumgebung jedoch selbst bestimmen können, ob und zu welchem Zeitpunkt sie zusätzlicher Unterstützung bedürfen (vgl. Deci & Ryan, 1993). Um beiden Ansprüchen gerecht zu werden, würde es sich daher anbieten, dass die Zusatzinformationen sowohl vor als auch nach der Entscheidung frei bei Bedarf vom Besucher abgerufen werden können.

Ein weiterer möglicher Ansatzpunkt zur Optimierung der Lernumgebung besteht in der Feedbackgestaltung. Um dem Problem eines *confirmation bias* im Umgang mit den zusätzlichen Informationen und einer einseitigen Argumentation (*myside bias*) zu begegnen, wäre es auch überlegenswert, wie Knipfer (2009) elaborierteres Feedback zu geben, indem z. B. die Zusatzinformationen direkt an das Feedback gekoppelt werden. Auf diese Weise wäre es möglich, die Rückmeldungen auch adaptiv, d. h. antwortabhängig zu gestalten. Personen, deren eigene Entscheidung im Konflikt zum Abstimmungsergebnis der anderen Personen steht, könnten etwa automatisch die jeweilige präferenzinkonsistente (Pro- bzw. Contra-) Position des Experten eingespielt bekommen. Die Museumsbesucher könnten so zumindest implizit Auskunft darüber erhalten, warum die anderen Besucher vor ihnen eine andere Entscheidung trafen und auf welchen Argumenten möglicherweise ihre Entscheidung basierte.

Ob und inwieweit diese vorgeschlagenen Formen der instruktionalen Unterstützung tatsächlich effektiver im Hinblick auf die Unterstützung von kritischem Denken und dem Treffen einer reflektierten Entscheidung sind sowie bestehende Vorwissensunterschiede zwischen den Lernenden besser kompensieren können, sollte in weiteren Folgestudien noch geklärt werden.

Da die PID ein sehr komplexes gendiagnostisches Verfahren darstellt, zu dem auch nur wenig Vorwissen besteht, wäre es auch empfehlenswert, in die Lernumgebung zusätzliche Informationen zum Verfahren der PID zu integrieren. Hier würde sich konkret die Implementation eines Buttons „Was ist PID?“ anbieten, bei dessen Aufruf dem Modalitätsprinzip von Mayer (2005a) entsprechend mithilfe einer Grafik und erläuterndem gesprochenem Text das Verfahren und der technische Ablauf der PID beschrieben werden (vgl. auch Abbildung 1, S. 22). Allerdings muss auch hier eine kurze zeitliche Einheit realisiert werden, da sich Besucher generell mit einer gewissen Zeitökonomie durch eine Ausstellung bewegen.

Ein weiterer Ansatzpunkt zur Optimierung der Lernumgebung betrifft die implementierte Feedbackinformation. Gemäß dem Kontiguitätsprinzip von Mayer (2005a) sollten bei der Präsentation von Feedback die Frage, die Antwort und das Feedback auf derselben Bildschirmseite angeboten werden. In der derzeitigen Version wird lediglich das Umfrageergebnis („Danke für Ihre Abstimmung! So haben die anderen Besucher abgestimmt“) präsentiert. Folglich sollten zusätzlich auf derselben Bildschirmseite die Entscheidungsfrage („Würden Sie an ihrer Stelle für eine PID nach Amerika gehen?“) und die zuvor gegebene Antwort („Ja“ bzw. „Nein“) ergänzt werden.

Um die Reliabilität der Feedbackinformation zu erhöhen, empfiehlt es sich auch in Anlehnung an Hazlewood & Chaiken (1990), die Samplegröße mit einzubeziehen, also auf wie vielen Personen das Umfrageergebnis basiert. Neben den relativen Häufigkeiten (Prozentzahlen) sollten die natürlichen Häufigkeiten, also die Gesamtanzahl der Besucher, die für oder gegen eine PID-HLA in Amerika gestimmt haben, genannt werden (vgl. Gigerenzer, 2002).

Nichtsdestotrotz kann das in dieser Arbeit dargestellte Experiment als Bestätigung dafür gelten, dass durch die didaktische und mediale Gestaltung der musealen Lernumgebung Dialogstation „Gentest“ das Treffen einer informierten und gut begründeten Entscheidung im Bereich Medizinethik angemessen gefördert werden konnte. Allerdings konnte auch festgestellt werden, dass kritisches Denken und eine reflektierte Entscheidung, in der auch mögliche Gegenpositionen zum eigenen Standpunkt miteinbezogen werden, nur unzureichend durch die Bearbeitung der multimedialen Lernumgebung unterstützt werden konnten. In Folgestudien sollte daher überprüft werden, ob und inwieweit durch andere Formen

der instruktionalen Unterstützung kritisches Denken und eine reflektierte Entscheidung besser unterstützt werden kann. Darüber hinaus könnten weitere Untersuchungen, insbesondere im Feld zeigen, inwieweit die experimentell gewonnenen Befunde auch für andere Besuchergruppen, für die anderen Fallbeispiele der Lernumgebung und vor allem unter den „regulären“ Lernbedingungen im musealen Setting gelten. Mit Sicherheit sind hier in Zukunft noch spannende Ergebnisse zu erwarten.

12 Literatur

- Alexander, P. A., Kulikowich, J. M. & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, 31, 313-337.
- Allen, S. (2004). *Finding significance*. San Francisco, CA: Exploratorium.
- Allen, V. L. (1965). Situational factors in conformity. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 133-176). New York: Academic Press.
- Allport, F. H. (1924). *Social Psychology*. Cambridge, MA: Riverside Press.
- Anderson, D. (1999). *A common wealth: Museums in the learning age*. London: Department for Culture, Media and Sport.
- Arndt, D. & Obe, G. (2001). *Fortpflanzungsmedizin in Deutschland* (Schriftenreihe des Bundesministeriums für Gesundheit, Bd. 13). Baden-Baden: Nomos-Verlagsgesellschaft.
- Aro, A. R., Hakonen, A., Hietala, M., Lönnqvist, J., Niemelä, P., Peltonen, L. & Aula, P. (1997). Acceptance of genetic testing in a general population: Age, education and gender differences. *Patient Education and Counseling*, 31, 41-49.
- Aronson, E., Wilson, T. D. & Akert, R. M. (2008). *Sozialpsychologie* (6. aktualisierte Aufl.). München: Pearson Studium.
- Asch, S. E. (1956). Studies of independence and conformity: A minority of one against an unanimous majority. *Psychological Monographs*, 70 (9), 1-70.
- Astleitner, H. (1998). *Kritisches Denken. Basisqualifikation für Lehrer und Ausbilder*. Innsbruck: Studien Verlag.
- Atkinson, R. K., Mayer, R. E. & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, 30, 117-139.
- Ausubel, D. P. (1968). *Educational Psychology: A cognitive view*. New York: Holt, Rinehart & Winston.
- Axsom, D., Yates, S. M. & Chaiken, S. (1989). Audience response as a heuristic cue in persuasion. *Journal of Personality and Social Psychology*, 67, 5-19.
- Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, 40 (4), 199-209.
- Azevedo, R. & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, 13 (2), 111-127.
- Baddeley, A. D. (1999). *Essentials of Human Memory*. Hove: Psychology Press.
- Baetens, P., Van de Velde, H., Camus, M., Pennings, G., Van Steirteghem, A., Devroey, P. & Liebaers, I. (2005). HLA-matched embryos selected for siblings requiring haematopoietic stem cell transplantation: a psychological perspective. *Reproductive BioMedicine Online*, 10, 154-163.

- Bangert-Drowns, R. L., Kulik, C., Kulik, J. A. & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213-238.
- Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning*, 1 (3), 221-235.
- Baron, R. S., Vandello, J. A. & Brunzman, B. (1996). The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of Personality and Social Psychology*, 71, 915-927.
- Baruch, S., Kaufman, D. & Hudson, K. L. (2008). Genetic testing of embryos: practices and perspectives of US in vitro fertilization clinics. *Fertility and Sterility*, 89 (5), 1053-1058.
- Bayer, K. (1999, 2007). *Argument und Argumentation. Logische Grundanalysen der Argumentationsanalyse*. Opladen/Wiesbaden: Westdeutscher Verlag.
- Bayrhuber, H. (1992). Ethische Analyse der Genterapie von Keimbahnzellen im Unterricht. *Friedrich-Jahresheft X*, 128-131.
- Beier, H. M. (1999). Definition und Grenze der Totipotenz: Aspekte für eine Präimplantationsdiagnostik. *Ethik in der Medizin*, 11, 23-37.
- Bell, R. L. & Lederman, N. G. (2003). Understandings of the nature of science and decision making on science- and technology-based issues. *Science Education*, 87, 352-377.
- Beller, S. & Spada, H. (1996). Denken (thinking, reasoning). In G. Strube, B. Becker, C. Freksa, U. Hahn, O. Opwis & G. Palm (Hrsg.), *Wörterbuch der Kognitionswissenschaft* (S. 114-124). Stuttgart: Klett-Cotta.
- Bentham, J. (1789). *Eine Einführung in die Prinzipien der Moral und der Gesetzgebung*. London.
- Bereiter, C. (1995). A dispositional view of transfer. In A. McKeough & J. Lupart (Eds.), *Teaching for transfer: Fostering, generalization in learning* (pp. 21-34). Hillsdale, NJ: Erlbaum.
- Berth, H., Dinkel, A. & Balck, F. (2002). Gesundheit durch Gentests? Akzeptanz und Befürchtungen gegenüber genetischen Untersuchungen in einer deutschlandrepräsentativen Stichprobe. *Zeitschrift für Gesundheitspsychologie*, 10 (3), 97-107.
- Berth, H., Dinkel, A. & Balck, F. (2003). Chancen und Risiken genetischer Diagnostik. Ergebnisse einer Umfrage in der Allgemeinbevölkerung und bei Medizinstudierenden. *Zeitschrift für Medizinische Psychologie*, 4 (12), 177-185.
- Berth, H., Dinkel, A., Kreuz, F. R. & Balck, F. (2004). Der Genetische Wissensindex (GeWi) - Ein Instrument zur Erfassung des allgemeinen Wissens über Genetik. *Zeitschrift für Medizinische Psychologie*, 13, 21-28.
- Betsch, C. (2004). Präferenz für Intuition und Deliberation. Inventar zur Erfassung von affekt- und kognitionsbasiertem Entscheiden. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 25, 179-197.
- Betsch, T., Funke, J. & Plessner, H. (2011). *Denken - Urteilen, Entscheiden, Problemlösen*. Berlin: Springer.
- Betsch, T. & Haberstroh, S. (2005). *The Routines of Decision Making*. Mahwah, NJ: Erlbaum Associates.

- Bioethikkommission beim Bundeskanzleramt (2004). *Präimplantationsdiagnostik (PID)*. Verfügbar unter: <http://oesterreich.gv.at/DocView.axd?CobId=6415> [11.03.2010].
- Birnbacher, D. (1999). *Selektion am Lebensbeginn - ethische Aspekte*. Vortrag auf dem Deutschen Kongress für Philosophie. Konstanz, 7.10.1999. Verfügbar unter: <http://www.selbsthilfe-online.de/druckversion.php?id=67> [11.01.2011].
- Böcher, U. (2004). *Präimplantationsdiagnostik und Embryonenschutz. Zu den Problemen der strafrechtlichen Regelung eines neuen medizinischen Verfahrens*. Göttingen: Hubert & Co.
- Bögeholz, S., Höble, C., Langlet, J., Sander, E. & Schlüter, K. (2004). Bewerten - Urteilen - Entscheiden im biologischen Kontext: Modelle in der Biologiedidaktik. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 89-115.
- Bohner, G. (2003). Einstellungen. In K. Jonas, W. Stroebe & M. Hewstone (Hrsg.), *Sozialpsychologie. Eine Einführung* (4. Aufl., S. 265-315). Heidelberg: Springer.
- Bond, R. & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, 119, 111-137.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.
- Boyle, R. J. & Savulescu, J. (2001). Ethics of using preimplantation genetic diagnosis to select a stem cell donor for an existing person. *British Medical Journal*, 323, 1240-1243.
- Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, Ch. K. & Williams, S. M. (1990). Anchored Instruction: Why we need it and how technology can help. In D. Nix & R. J. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 115-141). Hillsdale: Erlbaum.
- Brähler, E. & Stöbel-Richter Y. (2004). Die Einstellungen der Deutschen zur Reproduktionsmedizin und PID. Verfügbar unter: www.lebenshilfe.de/wDeutsch/aus_fachlicher_sicht/downloads/reproduktionsmedizin.pdf [11.06.2011].
- Brehm, J. W. (1957). Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology*, 52, 384-389.
- Breitsameter, F., Hauser, B., Hauser, W. & Kampschulte, L. (2009). *Nano- und Biotechnologie im Zentrum Neue Technologien*. München: Deutsches Museum.
- Brown, J. S., Collins A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18 (1), 32-42.
- Brünken, R., Seufert, T. & Leutner, D. (2008). Lernen und Lehren mit neuen Medien. In A. Renkl (Hrsg.), *Lehrbuch Pädagogische Psychologie* (S. 299-338). Bern: Huber.
- Bruns, B. & Gajewski, P. (2000). *Multimediales Lernen im Netz. Leitfaden für Entscheider und Planer* (2. Aufl.). Berlin: Springer.
- Bundesärztekammer (BÄK) (2000). Diskussionsentwurf zu einer Richtlinie zur Präimplantationsdiagnostik. *Deutsches Ärzteblatt*, 97, A-525-A-528.
- Bundesgerichtshof (BGH) (2010). *Urteil vom 6. Juli 2010 - Az 5 StR 386/09*. Verfügbar unter: <http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=en&sid=922f945586b29833928ef5b3edc828f6&nr=52897&pos=0&anz=1> [22.11.2010].

- Bundesvereinigung Lebenshilfe (2010). *Position der Bundesvereinigung Lebenshilfe für Menschen mit geistiger Behinderung e. V. zur Präimplantationsdiagnostik*. Verfügbar unter: http://www.lebenshilfe.de/wDeutsch/aus_fachlicher_sicht/downloads/LH-PID-Positionspapier-20101216.pdf [12.01.2011].
- Burmester, G. R. & Pezzutto, A. (1998). *Taschenatlas der Immunologie*. Stuttgart, New York: Georg Thieme Verlag.
- Buselmaier, W. & Tariverdian, G. (2007). *Humangenetik* (4. Aufl.). Berlin: Springer.
- Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.
- Chadwick, A. (2004). Lebenslanges Lernen. Ein Überblick mit Bezug auf Museen und den Weiterbildungssektor. In H. John & J. Thinesse-Demel (Hrsg.), *Lernort Museum – neu verortet! Ressourcen für soziale Integration und individuelle Entwicklung. Ein europäisches Praxishandbuch* (S. 57-63). Bielefeld: transcript.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752-766.
- Chaiken, S. & Eagly, A. H. (1978). Communication modality as a determinant of message persuasiveness and message comprehensibility. *Journal of Personality and Social Psychology*, 34, 605-614.
- Chaiken, S. & Eagly, A. H. (1983). Communication modality as a determinant of persuasion: The role of communicator salience. *Journal of Personality and Social Psychology*, 45, 241-256.
- Chaiken, S., Liberman, A. & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In Uleman, J. S. & Bargh, J. A. (Eds.), *Unintended thought* (pp. 212-252). New York: Guilford Press.
- Chaiken, S. & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford.
- Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332.
- Chandler, P. & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233-246.
- Chang, S. N. & Chiu, M. H. (2008). Lakatos' Scientific Research Programmes as a Framework for Analysing Informal Argumentation about Socioscientific Issues. *International Journal of Science Education*, 30 (13), 1753-1773.
- Chen, S. & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social and cognitive psychology* (pp. 73-96). New York: Guilford Press.
- Cialdini, R. B. (2007). *Influence: The psychology of persuasion*. New York: Collins.
- Chittenden, D., Farmelo, G. & v. Lewenstein, B. (2004). *Creating connections. Museums and the public understanding of current research*. Walnut Creek, CA: Altamira Press.
- Clariana, R. B. (1993). A review of multiple-try feedback in traditional and computer-based instruction. *Journal of Computer-Based Instruction*, 20 (3), 67-74.

- Clariana, R. B. (2000). *Feedback in computer-assisted learning*. Netg White Papers. Verfügbar unter: <http://www.personal.psu.edu/faculty/r/b/rbc4/NETg.htm>[03.10.2011].
- Cognition and Technology Group at Vanderbilt (CTGV) (1997). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development*. London: Erlbaum.
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, 33, 107-112.
- Collani, G. von & Stürmer, S. (2007). Deutsche Skala zur Operationalisierung des Konstrukts Selbstüberwachung (Self-Monitoring) und seiner Facetten. In A. Glöckner-Rist (Hrsg.), *ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente*. ZIS Version 11.00. Bonn: GESIS.
- Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Erlbaum.
- Collins, M., Carnine, D. & Gersten, R. (1987). Elaborated corrective feedback and the acquisition of reasoning skills: A study of computer-assisted instruction. *Exceptional Children*, 54, 254-262.
- Cox, D. R. (1970). *Planning of experiments*. New York: Wiley.
- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and Instructional Methods. A Handbook for Research on Interactions*. New York: Irvington.
- Csikszentmihalyi, M. & Hermanson, K. (1999). Intrinsic motivation in museums: why does one want to learn? In E. Hooper-Greenhill (Ed.), *The Educational Role of the Museum* (second ed., pp. 146-160). New York: Routledge.
- D'Alessio, D. & Allen, M. (2002). Selective exposure and dissonance after decisions. *Psychological Reports*, 91, 527-532.
- Damásio, A. (1997). *Descartes' Irrtum - Fühlen, Denken und das menschliche Gehirn*. München: List.
- Damásio, A. (2011). *Selbst ist der Mensch: Körper, Geist und die Entstehung des menschlichen Bewusstseins*. München: Siedler.
- Damschen, G. & Schönecker, D. (2003). *Der moralische Status menschlicher Embryonen: Pro und Contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument*. Berlin: de Gruyter.
- Dawson, V. & Venville, G. J. (2009). High school students' informal reasoning and argumentation about biotechnology: An indicator of scientific literacy? *International Journal of Science Education*, 31 (11), 1421-1445.
- Deci, E. L., Koestner, R. & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, 71, 1-27.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39, 223-238.
- De Vries, N. K., De Dreu, C. K. W., Gordijn, E. & Schuurman, M. (1996). Majority and minority influence: A dual role interpretation. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 7, pp. 145-172). Chichester: Wiley.

- Dempsey, J. V., Driscoll, M. P. & Swindell, L. K. (1993). Text-based feedback. In J. V. Dempsey & G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 21-54). Englewood Cliffs, NJ: Educational Technology Publications.
- Deutsch, M. & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51, 629-636.
- Deutsche Fanconi-Anämie-Hilfe e. V. (2005). *Fanconi-Anämie: Ein Handbuch für Eltern, Patienten und ihre Ärzte*. Verfügbar unter: http://www.fanconi.de/_handbuch/50601_FA_Handbuch.pdf [12.03.2010].
- Deutscher Ethikrat (DER) (2010). *Präimplantationsdiagnostik. Stellungnahme*. Verfügbar unter: <http://www.ethikrat.org/dateien/pdf/stellungnahme-praeimplantationsdiagnostik.pdf> [10.03.2010].
- Deutsches IVF-Register (DIR) (2010). *Jahrbuch 2010*. Verfügbar unter: <http://www.deutsches-ivf-register.de/pdf-downloads/dirjahrbuch2010-d.pdf> [29.02.2012].
- Deutsches Museum (2009). *Jahresbericht 2009*. Verfügbar unter: http://www.deutsches-museum.de/fileadmin/Content/2009/01_Information/12_Wir/1_jahresbericht/jb_09.pdf [23.06.2010].
- Deutsches Referenzzentrum für Ethik in den Biowissenschaften (DRZE) (2010). *Präimplantationsdiagnostik*. Verfügbar unter: <http://www.drze.de/im-blickpunkt/pid> [12.09.2010].
- Devine, P. G., Tauer, J. M., Barron, K. E., Elliot, A. J. & Vance, K. M. (1999). Moving beyond attitude change in the study of dissonance-related processes. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology* (pp. 297-323). Washington, DC: American Psychological Association.
- Devolder, K. (2005). Preimplantation HLA typing: having children to save our loved ones. *Journal of Medical Ethics*, 3 (10), 582-586.
- De Wert, G. (2005). Preimplantation genetic diagnosis: the ethics of intermediate cases. *Human Reproduction*, 20 (12), 3261-3266.
- Dick, R. D. (1991). An empirical taxonomy of critical thinking. *Journal of Instructional Psychology*, 18, 79-92.
- Dietrich, J. (2003). Ethische Urteilsbildung - Elemente und Arbeitsfragen für den Unterricht. *Zeitschrift für Didaktik der Philosophie und Ethik*, 3, 269-278.
- Dietrich, J. (2004). Grundzüge ethischer Urteilsbildung. Ein Beitrag zur Bestimmung ethisch-philosophischer Basiskompetenzen und zur Methodenfrage der Ethik. In J. Rohbeck (Hrsg.), *Ethisch-philosophische Basiskompetenz* (S. 65-96). Thelem: Dresden.
- Dietrich, J. (2005). Ethisch-Philosophische Grundlagenkompetenzen: ein Modell für Studierende und Lehrende. In M. Mayring (Hrsg.), *Ethisch-Philosophisches Grundlagenstudium, Ein Studienbuch* (S. 15-32). LIT Verlag: Münster.
- Ditto, P. H. & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, 63, 568-584.

- Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovich, A. M. & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, 75, 53-69.
- Dochy, F. (1992). *Assessment of prior knowledge as a determinant for future learning. The use of prior knowledge state tests and knowledge profiles*. Utrecht: Uitgeverij Lemma B. V.
- Dochy, F., Segers, M., van den Bossche, P. & Gijbels, D. (2003). Effects of problembased learning: a meta-analysis. *Learning and Instruction*, 13 (5), 533-568.
- Dohmen, G. (2001). *Das informelle Lernen - Die internationale Erschließung einer bisher vernachlässigten Grundform menschlichen Lernens für das lebenslange Lernen*. Bonn: Bundesministerium für Bildung und Forschung.
- Duit, R. (1999). Conceptual change approaches in science education. In W. Schnotz, S. Vosniadou & M. Carretero (Eds.), *New perspectives on conceptual change* (pp. 263-282). Amsterdam: Pergamon.
- Dulitz, B. & Kattmann, U. (1990). *Bioethik. Fallstudien für den Unterricht*. Stuttgart: Metzler.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B. & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170-177.
- Durant, J. (2001). *Museums and the Public Understanding of Science*. London: Science Museum.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Hartcourt, Brace, Jovanovich.
- Edwards, K. & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, 71 (1), 5-24.
- Eggert, S. (2008). *Bewertungskompetenz für den Biologieunterricht - Vom Modell zur empirischen Überprüfung*. Verfügbar unter: <http://webdoc.sub.gwdg.de/diss/2008/eggert/eggert.pdf> [15.09.2009].
- Eisinger, F., Julian-Reynier, C., Sobol, H., Stoppa-Lyonnet, D., Lasset, C. & Nogues, C. (2000). Acceptability of prophylactic mastectomy in cancer-prone women. *Journal of the American Medical Association*, 283, 202-203.
- Elliot, A. J. & Devine, P. G. (1994). On the motivational nature of cognitive dissonance. Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, 67, 382-394.
- Embryonenschutzgesetz (ESchG) (1991). Verfügbar unter: <http://www.gesetze-im-internet.de/eschg/BJNR027460990.html> [09.04.2009].
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9-26). New York: W. H. Freeman.
- Enquete-Kommission (EK) (2002). *Schlussbericht der Enquete-Kommission „Recht und Ethik der modernen Medizin“*. Deutscher Bundestag, Drucksache 14/9020. Berlin.
- Enskat, R. (2003). Pro Identitätsargument. Auch menschliche Embryonen sind jederzeit Menschen. In G. Damschen & D. Schönecker (Hrsg.), *Der moralische Status*

- menschlicher Embryonen: Pro und Contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument* (S. 101-128). Berlin: de Gruyter.
- Epstein, S., Pacini, R., Denes-Raj, V. & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, 71, 390-405.
- Erb, H.-P. (1998). Sozialer Einfluss durch Konsens: Werbung mit Meinungsübereinstimmung. *Zeitschrift für Sozialpsychologie*, 29, 156-164.
- Erb, H.-P. & Bohner, G. (2002). Sozialer Einfluss durch Mehrheiten und Minderheiten. In D. Frey & M. Irle (Hrsg.), *Theorien der Sozialpsychologie, Bd. II: Gruppen-, Interaktions- und Lerntheorien* (2. Aufl., S. 47-61). Bern: Huber.
- Erb, H.-P., Bohner, G., Werth, L., Hewstone, M. & Reinhard, M.-A. (2006). Large minorities and small majorities: Interactive effects of inferred and explicit consensus on attitudes. *Basic and Applied Social Psychology*, 28, 221-231.
- Erduran, S., Simon, S. & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88, 915-933.
- Euler, D. (1992). *Didaktik des computerunterstützten Lernens. Praktische Gestaltung und theoretische Grundlagen*. Nürnberg: BW Bildung und Wissen Verlag und Software GmbH.
- Evans, J. St. B. T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, 128, 978-996.
- Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7, 454-459.
- Evans, J. St. B. T. & Thompson, V. A. (2004). Informal reasoning: Theory and method. *Canadian Journal of Experimental Psychology*, 58, 69-74.
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: The California Academic Press.
- Falk, J. & Dierking, L. (2000). *Learning from museums: Visitor experiences and the making of meaning*. Walnut Creek, CA: AltaMira Press.
- Falk, J. H. & Storksdieck, M. (2005). Using the contextual model of learning understand visitor learning from a science center exhibition. *Science Education*, 89 (5), 744-778.
- Faller, H. (1997). Genetische Testung bei familiären Brustkrebs - Psychosoziale Forschung und zukünftige Strategien. *Zeitschrift für Medizinische Psychologie*, 6, 109-116.
- Fengler, J. (1998). *Feedback geben. Strategien und Übungen*. Weinheim: Beltz.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-40.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row Peterson.
- Fischer, F., Mandl, H. & Todorova, A. (2009). Lehren und Lernen mit neuen Medien. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (2. Aufl., S. 753-771). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Fischer, F., Waibel, M. & Wecker, C. (2005). Nutzenorientierte Grundlagenforschung im Bildungsbereich: Argumente einer internationalen Diskussion. *Zeitschrift für Erziehungswissenschaft*, 8 (3), 427-442.
- Fischer, P. M. & Mandl, H. (1988). Improvement of the acquisition of knowledge by informing feedback. In H. Mandl & A. M. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 187-241). New York: Springer.
- Fischer, P. (2006). *Politische Ethik: Eine Einführung*. München: Wilhelm Fink.
- Foong, C.-C. & Daniel, E. G. S. (2010). Assessing students' arguments made in socio-scientific contexts: The considerations of structural complexity and the depth of content knowledge. *Procedia - Social and Behavioral Sciences*, 9, 1120-1127.
- Franklin S. & Roberts, C. (2006). *Born and Made. An Ethnography of Preimplantation Genetic Diagnosis*. Princeton: Princeton University Press.
- Frey, D. (1981). *Informationssuche und Informationsbewertung bei Entscheidungen*. Bern: Huber.
- Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 41-80). New York: Academic Press.
- Frey, D., Braun, S., Wesche, J. S., Kerschreiter, R. & Frey, A. (2010). Nichts ist praktischer als eine gute Theorie - Nichts ist theoriegewinnender als eine gut funktionierende Praxis: Zum Theorie-Praxis-Austausch in der Psychologie. In U. P. Kanning, L. von Rosenstiel & H. Schuler (Hrsg.), *Jenseits des Elfenbeinturms - Psychologie als nützliche Wissenschaft* (50-74). Göttingen: Vandenhoeck & Ruprecht.
- Frey, D., Dauenhaimer, D., Parge, O. & Haisch, J. (1993). Die Theorie sozialer Vergleichsprozesse. In D. Frey & M. Irle (Hrsg.), *Theorien der Sozialpsychologie, Band I: Kognitive Theorien* (S. 81-121). Bern: Huber.
- Frey, D. & Gaska, A. (1993). Die Theorie der kognitiven Dissonanz. In D. Frey & M. Irle (Hrsg.), *Theorien der Sozialpsychologie, Band I: Kognitive Theorien* (S. 275-324). Bern: Huber.
- Fuchs, M. & Lanzerath, D. (1998). Stichwort Eugenik. In W. Korff, L. Beck & P. Miktat (Hrsg.), *Lexikon der Bioethik, Bd. 1* (S. 701-704). Gütersloh: Gütersloher Verlagshaus.
- Gagné, R. M. (1968). *The conditions of learning*. New York: Holt, Rinehart & Winston.
- Galotti, K. (1989). Approaches to studying formal and everyday reasoning. *Psychological Bulletin*, 105, 331-351.
- Gerber-Hirt, S. & Noschka-Roos, A. (2012). Wissenschaftskommunikation und neue Technologien: Das Zentrum Neue Technologien als Forum zwischen Wissenschaft und Öffentlichkeit. In G. Herwig & S. Schleissing (Hrsg.), *Grüne Gentechnik: Zwischen Forschungsfreiheit und Anwendungsrisiko* (S. 329-347). Baden-Baden: Nomos.
- Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, 41 (6), 867-888.
- Geyer, C., Noschka-Roos, A. & Lewalter, D. (2005). *Abschlussbericht der Besucherbefragung in der Sonderausstellung „Leben mit Ersatzteilen“ des Deutschen Museums München (ZNT)*. Verfügbar unter: <http://www.deutsches->

- museum.de/fileadmin/Content/Haupthaus/Forschung/PDFs/struktur_medtec4.pdf [19.07.2010].
- Gibbons, F. X. & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, 76, 129-142.
- Gigerenzer, G. (2002). *Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken*. Berlin: Berlin Verlag.
- Gigerenzer, G. (2007). *Bauchentscheidungen: Die Intelligenz des Unbewussten und die Macht der Intuition*. München: Bertelsmann.
- Giles, D. C. (2002). Parasocial interaction: A review of the literature and a model for future research. *Media Psychology*, 4, 279-305.
- Giner-Sorolla, R. & Chaiken, S. (1997). Selective use of heuristic and systematic processing under defense motivation. *Personality and Social Psychology Bulletin*, 23 (1), 84-97.
- Gleich, U. (1997). *Parasoziale Interaktionen und Beziehungen von Fernsehzuschauern mit Personen auf dem Bildschirm: ein theoretischer und empirischer Beitrag zum Konzept des aktiven Rezipienten*. Landau: Verlag Empirische Pädagogik.
- Götz, C. (2001). *Förderung ethisch-moralischer Überzeugungen bei Jugendlichen*. Verfügbar unter: <http://geb.uni-giessen.de/geb/volltexte/2001/406/> [22.06.2008].
- Gräber, W. & Nentwig, P. (2002). Scientific Literacy - Naturwissenschaftliche Grundbildung in der Diskussion. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific Literacy, Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung* (S. 7-22). Opladen: Leske + Budrich.
- Gräsel, C. (1997). *Problemorientiertes Lernen*. Göttingen: Hogrefe.
- Graf, B. (2003). Ausstellungen als Instrument der Wissensvermittlung. Grundlagen und Bedingungen. *Museumskunde*, 68 (1), 73-81.
- Graf, B. & Noschka-Roos, A. (2009). Stichwort: Lernen im Museum oder: Eine Kamerafahrt mit der Besucherforschung. *Zeitschrift für Erziehungswissenschaft*, 12 (3), 7-27.
- Graumann, S. (2001). Gesellschaftliche Folgen der Präimplantationsdiagnostik. In D. Arndt & G. Obe (2001), *Fortpflanzungsmedizin in Deutschland* (Schriftenreihe des Bundesministeriums für Gesundheit, Bd. 132) (S. 215-220). Baden-Baden: Nomos-Verlagsgesellschaft.
- Green, D. W. (1994). Induction: Representation, strategy and argument. *International Studies in the Philosophy of Science*, 8, 45-50.
- Greeno, J. G. (1989). Situations, mental models and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 285-318). Hillsdale, NJ: Erlbaum.
- Greitemeyer, T. (2000). *Austausch von Informationen bei Entscheidungen in Kleingruppen*. Neuried: Ars Una.
- Greitemeyer, T., Schulz-Hardt, S. & Frey, D. (2003). Präferenzkonsistenz und Geteiltheit von Informationen als Einflussfaktoren auf Informationsbewertung und intendiertes Diskussionsverhalten bei Gruppenentscheidungen. *Zeitschrift für Sozialpsychologie*, 34, 9-23.

- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, S. 41-58). New York: Academic Press.
- Griesinger, G., Schultze-Mosgau, A., Finas, D., Felberbaum, R. E. & Diedrich, K. (2003). Präimplantationsdiagnostik: Methode und Anwendung aus reproduktionsmedizinischer Sicht. *Zeitschrift für medizinische Ethik*, 49, 325-42.
- Gruber, H. & Mandl, H. (1996). Expertise und Erfahrung. In H. Gruber & H. Ziegler (Hrsg.), *Expertiseforschung. Theoretische und methodische Grundlagen* (S. 18-34). Opladen: Westdeutscher Verlag.
- Guckes, B. (1997). *Das Argument der schiefen Ebene. Schwangerschaftsabbruch, die Tötung Neugeborener und Sterbehilfe in der medizinethischen Diskussion*. Stuttgart: Gustav Fischer.
- Haack, J. (2002). Interaktivität als Kennzeichen von Multimedia und Hypermedia. In L. Issing & P. Klimsa (Hrsg.) *Information und Lernen mit Multimedia und Internet* (3. Aufl., S. 127-136). Weinheim: Beltz PVU.
- Habermas, J. (1995). *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns*. Frankfurt: Suhrkamp.
- Hänle, M. (2008). *Formative Evaluation der Dialogstation „Entscheidungslabor Gentest“*. Unveröffentlichtes Manuskript.
- Hafer, C. L., Reynolds, K. & Obertynski, M. A. (1996). Message comprehensibility and persuasion: Effects of complex language in counterattitudinal appeals to lay people. *Social Cognition*, 14, 317-337.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgement. *Psychological Review*, 108, 814-834.
- Haker, H. (2002). *Ethik der genetischen Frühdiagnostik. Sozialethische Reflexionen zur Verantwortung am Beginn des menschlichen Lebens*. Paderborn: mentis.
- Haker, H. (2004). Pränataldiagnostik und Präimplantationsdiagnostik: ein Wertungswiderspruch? In W. Lenzen (Hrsg.), *Wie bestimmt man den „moralischen Status“ von Embryonen?* (S. 141-163). Paderborn: mentis.
- Halpren, D. F. (1989). *Thought and knowledge: An introduction to critical thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Hampel, J. & Renn, O. (1999). *Gentechnik in der Öffentlichkeit. Wahrnehmung und Bewertung einer umstrittenen Technologie*. Frankfurt/New York: Campus.
- Hancock, T., Thurman, R. & Hubbard, D. (1995). An expanded control model for the use of instructional feedback. *Contemporary Educational Psychology*, 20, 410-425.
- Handyside, A. H., Kontogianni, E. H., Hardy, K. & Winston, R. M. (1990). Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification. *Nature*, 344, 768-770.
- Harmon-Jones, E. (2000). Cognitive dissonance and experienced negative affect: Evidence that dissonance increases experienced negative affect even in the absence of aversive consequences. *Personality and Social Psychology Bulletin*, 26, 1490-1501.
- Harms, U. & Krombaß, A. (2008). Lernen im Museum - das Contextual Model of Learning. *Unterrichtswissenschaft*, 36 (2), 150-166.

- Harp, S. F. & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, 90, 414-434.
- Harper J. C., Delhanty, J. D. & Handyside, A. H. (2001). *Preimplantation Genetic Diagnosis*. Chichester: John Wiley & Son.
- Harper, J. C., Coonen, E., De Rycke, M., Harton, G., Moutou, C., Pehlivan, T., Traeger-Synodinos, J., Van Rij, M. C. & Goossens, V. (2010). ESHRE PGD Consortium data collection X: cycles from January to December 2007 with pregnancy follow-up to October 2008. *Human Reproduction*, 25 (11), 2685-2707.
- Hartmann, T. & Klimmt, C. (2005). Ursachen und Effekte Parasozialer Interaktionen im Rezeptionsprozess: Eine Fragebogenstudie auf der Basis des PSI-Zwei-Ebenen-Modells. *Zeitschrift für Medienpsychologie*, 17, 88-98.
- Hartmann, T. & Schramm, H. (2006). Logik der Forschung zu parasozialen Interaktionen und Beziehungen. In W. Wirth, A. Fahr & E. Lauf (Hrsg.), *Forschungslogik und -design in der empirischen Kommunikationswissenschaft. Bd. 2: Anwendungsfelder in der Kommunikationswissenschaft* (S. 264-291). Köln: Herbert von Halem Verlag.
- Hartmann, T., Schramm, H. & Klimmt, C. (2004a). Personenorientierte Medienrezeption: Ein Zwei-Ebenen-Modell parasozialer Interaktionen. *Publizistik*, 49 (1), 25-47.
- Hartmann, T., Schramm, H. & Klimmt, C. (2004b). *Vorbereitende Überlegungen zur theoretischen Modellierung parasozialer Interaktionen im Prozess der Medienrezeption*. Verfügbar unter: <http://www.ijk.hmt-hannover.de/psi/> [12.06.2008].
- Hashagen, U. (2001). Von Mäusen und Medien: "Neue Medien" im Museum. In M.-D. Weitze (Hrsg.), *Public Understanding of Science im deutschsprachigen Raum: Die Rolle der Museen* (S. 121-139). München: Deutsches Museum.
- Hattie, J. & Timperley, H. S. (2007). The power of feedback. *Review of Educational Research*, 77, 81-112.
- Hauser, B., Noschka-Roos, A. & Teichmann, J. (2007). Technische Museen und Science Centers: Innovative Forschung erfordert innovative Vermittlung. In E. Wagner & M. Dreykorn (Hrsg.), *Museum Schule Bildung. Aktuelle Diskurse, innovative Modelle, erprobte Methoden* (S. 131-135). München: kopaed.
- Hauser, W. (2005). Auf dem Weg zu einem "Public Understanding of Research": das Zentrum Neue Technologien des Deutschen Museums. In A. Noschka-Roos, W. Hauser & E. Schepers (Hrsg.), *Mit neuen Medien im Dialog mit den Besuchern? Berliner Schriften zur Museumskunde (Bd. 21)* (S. 9-20). Berlin: G + H Verlag.
- Hauser, W. (2006). Interaktion als Dialog: Mediale Räume für kontroverse Wissenschaftsthemen. In S. Schwan, H. Trischler & M. Prenzel (Hrsg.), *Die Rolle von Medien für die Resituierung von Exponaten: Mitteilungen und Berichte aus dem Institut für Museumsforschung, Bd. 38* (S. 89-99). Berlin: Institut für Museumsforschung.
- Hauser, W. (2009). Wie kommuniziert man aktuelle Forschung im Museum? In International Council of Museums (Hrsg.), *Wissenschaftskommunikation: Perspektiven der Ausbildung - Lernen im Museum* (S. 84-88). Frankfurt: Peter Lang.
- Hauser, W., Noschka-Roos, A., Reussner, E. & Zahn, C. (2009). Design-based research on digital media in a museum environment. *Visitor Studies*, 12 (2), 182-198.

- Hazlewood, J. D. & Chaiken, S. (1990). *Personal Relevance, Majority Influence, and the Law of Large Numbers*. Paper presented at the meeting of the American Psychological Association. Boston, MA.
- Heckhausen, H. (1974). *Leistung und Chancengleichheit*. Göttingen: Hogrefe.
- Heckhausen, H. (1989). *Motivation und Handeln* (2. Aufl.). Berlin: Springer.
- Hein, G. (1998). *Learning in the museum*. London, New York: Routledge.
- Heinecke, A. M. (2002). Evaluation of POI Systems by Logfile Recording. In H. Luczak, A. E. Çakir & G. Çakir (Eds.), *Proceedings of the 6th International Scientific Conference on Work With Display Units WWDU 2002 – World Wide Work* (pp. 452-454). Berlin: ERGONOMIC Institut für Arbeits- und Sozialforschung.
- Heinecke, A. M., Bumann, S. & Kerstan, T. (1995). Software-ergonomische Evaluation von Kiosksystemen im Museum. In H.-D. Böcker (Hrsg.), *Software-Ergonomie '95, Mensch-Computer-Interaktion - Anwendungsbereiche lernen voneinander* (S. 163-177). Berichte des German Chapter of the ACM, Bd. 45. Stuttgart: B. G. Teubner.
- Heinrichs, B. (2006). Ethische Aspekte. In P. Propping, S. Aretz, J. Schumacher, J. Taupitz, J. Guttmann & B. Heinrichs (Hrsg.), *Prädiktive genetische Testverfahren. Naturwissenschaftliche, rechtliche und ethische Aspekte* (S. 111-163). Freiburg: Alber.
- Heitkamp, I., Borchardt, H. & Witte, E. H. (2005). Zur simulierten Rechtfertigung wirtschaftlicher und medizinischer Entscheidungen in Ethikkommissionen: Eine empirische Analyse des Einflusses verschiedener Rollen. *Hamburger Forschungsberichte zur Sozialpsychologie (HAFOS)*, Nr. 55.
- Hengstschläger, M. (2006a). Die Genetik und das ungeborene Leben. In M. Fischer & K. S. Zänker (Hrsg.), *Medizin- und Bioethik* (S. 69-72). Frankfurt: Peter Lang.
- Hengstschläger, M. (2006b). Präimplantationsdiagnostik: der aktuelle Stand. *Speculum - Zeitschrift für Gynäkologie und Geburtshilfe*, 24 (1), 9-14.
- Hennen, L. & Sauter, A. (2004). *Präimplantationsdiagnostik: Praxis und rechtliche Regulierung im Ländervergleich*. Berlin: Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB-Arbeitsbericht, Bd. 94).
- Hense, H. (1990). *Das Museum als gesellschaftlicher Lernort*. Frankfurt: Brandes & Apsel.
- Hewstone, M. & Martin, R. (2007). Sozialer Einfluss. In Jonas, K., Stroebe, W. & Hewstone, M. R. (Hrsg.), *Sozialpsychologie* (5. Aufl., S. 360-408). Heidelberg: Springer.
- Hietala, M., Hakonen, A., Aro, A. R., Peltonen, L. & Aula, P. (1995). Attitudes towards genetic testing among the general population and relatives of patients with a severe genetic disease: A survey from Finland. *American Journal of Human Genetics*, 56, 1493-1500.
- Höble, C. (2001a). Ethische Dimensionen der Gentechnik im Unterricht - Teil 3: Ambivalenz des genetischen Tests. *Praxis der Naturwissenschaften - Biologie in der Schule*, 50 (7), 35-38.
- Höble, C. (2001b). *Moralische Urteilsfähigkeit. Eine Interventionsstudie zur moralischen Urteilsfähigkeit von Schülern zum Thema Gentechnik*. Studienverlag: Innsbruck.
- Höble, C. (2003a). Präimplantationsdiagnostik - medizinische und rechtliche Aspekte. *Praxis der Naturwissenschaften - Biologie in der Schule*, 52 (6), 5-9.

- Höble, C. (2003b). Der Embryo - Mensch von Anfang an? Schülervorstellungen zum Beginn menschlichen Lebens und zu dessen Schutzbedürftigkeit. In S. Albrecht, J. Dierken, H. Freese & C. Höble (Hrsg.), *Stammzellforschung - Debatte zwischen Ethik, Politik und Geschäft* (S. 43-75). Hamburg: University Press.
- Höble, C. & Lude, A. (2004). Bioethik im naturwissenschaftlichen Unterricht - ein Problemaufriss. In C. Höble, D. Höttecke & E. Kircher (Hrsg.), *Lehren und Lernen über die Natur der Naturwissenschaften* (S. 23-43). Baltmannsweiler: Schneider Verlag Hohengehren.
- Höffe, O. (1992). *Einführung in die utilitaristische Ethik*. Tübingen: Francke.
- Höffe, O. (1997). *Lexikon der Ethik*. München: Beck.
- Hogarth, R. M. (2005). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thoughts. In T. Betsch & S. Haberstroh (Eds.), *The routines of decision making* (pp. 67-82). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Holinski-Feder, E. (2006). DNA-Untersuchung - Diagnostische Anwendung beim Menschen. In J. Murken, T. Grimm & E. Holinski-Feder (Hrsg.), *Humangenetik* (S. 95-124). Stuttgart: Thieme.
- Hondefelder, L. (2002). Die Frage nach dem moralischen Status des Embryos. In O. Höffe, L. Hondefelder, J. Isensee & P. Kirchhof (Hrsg.), *Gentechnik und Menschenwürde* (S. 79-111). Köln: DuMont.
- Hondefelder, L. (2003). Pro Kontinuumsargument: die Begründung des moralischen Status des menschlichen Embryos aus der Kontinuität der Entwicklung des ungeborenen zum geborenen Menschen. In G. Damschen & D. Schönecker (Hrsg.), *Der moralische Status menschlicher Embryonen: Pro und Contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument* (S. 61-81). Berlin: de Gruyter.
- Hondefelder, L., Korff, W. & Kreß, H. (1998). Stichwort Ethik. In W. Korff, L. Beck & P. Mikat (Hrsg.), *Lexikon der Bioethik, Bd. 1* (S. 654-669). Gütersloh: Gütersloher Verlagshaus.
- Hornung, C. (1994). PC-basierte Multimedia-Systeme. In U. Glowalla, E. Engelmann & G. Rossbach (Hrsg.), *Multimedia 94: Grundlagen und Praxis* (S. 2-8). Berlin: Springer.
- Horster, D. (1999). *Jürgen Habermas*. Hamburg: Junius.
- Horton, D. & Wohl, R. R. (1956). Mass communication and para-social interaction: Observation on intimacy at a distance. *Psychiatry*, 19, 185-206.
- Human Genetics Quality Network (HGQN) (2010). *Krankheiten*. Verfügbar unter: <http://www.hgqn.org/> [16.09.2010].
- Hussy, W., Schreier, M. & Echterhoff, G. (2009). *Forschungsmethoden in Psychologie und Sozialwissenschaften*. Berlin: Springer.
- Huth, K. (2004). *Entwicklung und Evaluation von fehlerspezifischem informativem tutoriellem Feedback (ITF) für die schriftliche Subtraktion*. Verfügbar unter: http://deposit.ddb.de/cgi-bin/dokserv?idn=973482818&dok_var=d1&dok_ext=pdf&filename=973482818.pdf [08.10.2011].
- Iglesias, T. (1984). In vitro fertilisation: the major issues. *Journal of Medical Ethics*, 10 (1), 32-37.

- Ilgen, D. R., Fisher, C. D. & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64, 349-371.
- Imhoff, R. (2005). *Was macht Minderheitenmeinungen attraktiv?* Verfügbar unter: http://psydok.sulb.uni-saarland.de/volltexte/2005/492/pdf/DA_1_1_1.pdf [09.05.2009].
- Imhoff, R. & Erb, H.-P. (2009). What motivates nonconformity? Uniqueness Seeking blocks Majority Influence. *Personality and Social Psychology Bulletin*, 35, 309-320.
- Issing, L. J. (2002). Instruktions-Design für Multimedia. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (3. Aufl., S. 151-178). Weinheim: Beltz.
- Jachertz, N. (2000). Präimplantationsdiagnostik: Am Rande der schiefen Bahn. *Deutsches Ärzteblatt*, 97 (9), A-507.
- Jacobs, B. (2002). *Aufgaben stellen und Feedback geben*. Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2004/438/> [08.10.2011].
- Jacobson, M. J. & Spiro, R. J. (1992). Hypertext learning environments and cognitive flexibility: Characteristics promoting the transfer of complex knowledge. In L. Birnbaum (Ed.), *The International Conference on the Learning Sciences. Proceedings of the 1991 Conference* (pp. 240-248). Charlottesville: Association for the Advancement of Computing in Education.
- Jallinoja, P., Hakonen, A., Aro, A. R., Niemelä, P., Hietala, M., Lönnqvist, J., Peltonen, L. & Aula, P. (1998). Attitudes towards genetic testing: Analysis of contradictions. *Social Science and Medicine*, 46, 1367-1374.
- Jallinoja, P. & Aro, A. R. (2000). Does knowledge make a difference? The association between knowledge about genes and attitudes toward gene test. *Journal of Health Communication*, 5, 29-39.
- Janeway, C., Travers, P. Walport, M. & Shlomchik, M. (2002). Autoimmunität und Transplantation. In C. Janeway, P. Travers, M. Walport & M. Shlomchik (Hrsg.), *Immunologie* (S. 537-591). Heidelberg: Spektrum Akademischer Verlag.
- Jarvis, W. B. G. & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology*, 70, 72-194.
- Jiménez-Aleixandre, M., Rodríguez, A. & Duschl, R. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education*, 84 (6), 757-792.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: University Press.
- Johnson-Laird, P. N. & Byrne, R. M. L. (1991). *Deduction*. Hove: Lawrence Erlbaum.
- Jonas, E., Frey, D., Henninger, M., Pommer, M., Haeften, I. von, Schulz-Hardt, S. & Mandl, H. (2001). Erklärung des eigenen Handelns als Einflussfaktoren auf Informationssuche und Informationsbewertung in einer sequentiellen Lernsituation. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 33 (4), 242-252.
- Jonas, E., Graupmann, V., Fischer, P., Greitemeyer, T. & Frey, D. (2003). Schwarze Kassen, weiße Westen? Konfirmatorische Informationssuche und -bewertung im Kontext der Spendenaffaire der CDU. *Zeitschrift für Sozialpsychologie*, 34, 47-61.

- Jonas, E., Schulz-Hardt, S., Frey, D. & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on «selective exposure to information». *Journal of Personality and Social Psychology*, 80, 557-571.
- Jonas, K. J. & Mikula, G. (2006). *Deutsche Fassung der Skala zur Messung sozialer Vergleichsorientierung*. Manuscript in preparation.
- Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology: Research & Development*, 48 (4), 63-85.
- Jones, E. E. & Gerard, H. B. (1967). *Foundations of Social Psychology*. New York: Wiley.
- Jungermann, H., Pfister, H. R. & Fischer, K. (2005). *Psychologie der Entscheidung. Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag.
- Kagan, J. (1984). *The nature of the child*. New York: Basic Books.
- Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten*. Herausgegeben von Karl Vorländer (1994). Hamburg: Meiner (Philosophische Bibliothek 41).
- Kaufmann, M. (2003). Contra Kontinuumsargument: abgestufte moralische Berücksichtigung trotz stufenloser biologischer Entwicklung. In G. Damschen & D. Schönecker (Hrsg.), *Der moralische Status menschlicher Embryonen: Pro und Contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument* (S. 83-98). Berlin: de Gruyter.
- Keller, M. (2000). Gendiagnostik von hereditären Tumordispositionserkrankungen: Psychosoziale Aspekte. *Zeitschrift für psychosomatische Medizin*, 46, 80-97.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.
- King, P. M. & Kitchener, K. S. (2004). Reflective judgment: Theory and research on the development of epistemic assumptions through adulthood. *Educational Psychology*, 39, 5-18.
- Klein, A. (2003). *Besucherbindung im Kulturbetrieb. Ein Handbuch*. Wiesbaden: Westdeutscher Verlag.
- Klein, H. J. (1990). *Der gläserne Besucher: Publikumsstrukturen einer Museumslandschaft*. Berlin: Mann.
- Klein, H. J. (2000). *Go West. Die Besucher des Deutschen Museums und ihre Meinungen über das Neue Verkehrsmuseum*. Verfügbar unter: http://www.deutsches-museum.de/fileadmin/Content/endbericht-go_west.pdf [19.09.2009].
- Klein, H. J., Donecker, A., Hänle, M. & Herrmann, N. (2009). Besucherfeedback – Ein Planungskriterium? In H. Kunz-Ott, S. Kudorfer & T. Weber (Hrsg.), *Kulturelle Bildung im Museum. Aneignungsprozesse - Vermittlungsformen - Praxisbeispiele* (S. 169-178). Bielefeld: transcript.
- Klimsa, P. (2002). Multimediane Nutzung aus psychologischer und didaktischer Sicht. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (3. Aufl., S. 5-18). Weinheim: Beltz.
- Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254-284.

- Knipfer, K. (2009). *Pro or con nanotechnology? Support for critical thinking and reflective judgement at science museums. Pro oder Kontra Nanotechnologie? Unterstützung von kritischem Denken und reflektiertem Urteilen im Museum.* Verfügbar unter: <http://tobias-lib.ub.uni-tuebingen.de/volltexte/2009/3761/pdf/KnipferKristin.pdf> [26.09.2009].
- Knipfer, K., Zahn, C. & Hesse, F. W. (2008a). *Pro or con nanotechnology? A discussion terminal as scaffold for critical thinking and reflective judgement at science museums.* Manuscript submitted for publication.
- Knipfer, K., Zahn, C. & Hesse, F. W. (2008b). *Why do they think so? Disagreement fosters deliberative opinion formation in science museums.* Manuscript submitted for publication.
- Kochte-Clemens, B. (2003). *Fragen und Antworten zur Roten Gentechnik.* Stuttgart: Akademie für Technikfolgenabschätzung in Baden-Württemberg (TA-Akademie). Verfügbar unter: <http://elib.uni-stuttgart.de/opus/volltexte/2004/1807/pdf/KIrote-Gentechnik.pdf> [11.02.2009].
- Koehler, J. J. (1993). The influence of prior beliefs on scientific judgments of evidence quality. *Organizational Behavior and Human Decision Processes*, 56, 28-55.
- Kohlberg, L. (1964). Development of moral character and moral ideology. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of child development research, Vol. I* (pp. 381-431). New York: Russel Sage Foundation.
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive development and epistemology* (pp. 151-235). New York: Academic Press.
- Kollek, R. (2002). *Präimplantationsdiagnostik - Embryonenselektion, weibliche Autonomie und Recht.* Tübingen: Francke Verlag.
- Kolstø, S. D. (2001). Scientific literacy for citizenship: tools for dealing with the science dimension of controversial socio-scientific issues. *Science Education*, 85, 291-310.
- Kopp, B., Dvorak, S. & Mandl, H. (2003). *Evaluation des Einsatzes von Neuen Medien im Projekt "Geoinformation - Neue Medien für die Einführung eines neuen Querschnittsfachs" (Forschungsbericht Nr. 161).* München: Ludwig-Maximilians-Universität, Department Psychologie, Institut für Pädagogische Psychologie.
- Kopp, B. & Mandl, H. (2009). Gestaltung medialer Lernumgebungen. In M. Henninger & H. Mandl (Hrsg.), *Handbuch Medien- und Bildungsmanagement* (S. 55-72). Weinheim: Beltz.
- Kopp, B. & Mandl, H. (in Druck). Lerntheoretische Grundlagen von Rückmeldungen. In H. Ditton & A. Müller (Hrsg.), *Rückmeldungen und Feedback: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder.* Münster: Waxmann.
- Korff, W., Beck, L. & Miktat, P. (2000). *Lexikon der Bioethik, Bd. 1.* Gütersloh: Bertelsmann.
- Krapp, A. (1992). Das Interessenkonstrukt - Bestimmungsmerkmale der Interessenhandlung und des individuellen Interesses aus der Sicht einer Person-Gegenstands-Konzeption. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung* (S. 297-329). Münster: Aschendorff.

- Krause, U.-M. (2002, January). *Elaborated group feedback in virtual learning environments*. Paper presented at the Doctoral Consortium of the CSCL (Computer Support for Collaborative Learning) conference 2002 in Boulder, Colorado, USA.
- Krause, U.-M. (2007). *Feedback und kooperatives Lernen*. München: Waxmann.
- Krause, U.-M., Stark, R. & Mandl, H. (2003). *Förderung des computerbasierten Wissenserwerbs im Bereich empirischer Forschungsmethoden durch kooperatives Lernen und eine Feedbackmaßnahme (Forschungsbericht Nr. 160)*. München: Ludwig-Maximilians-Universität, Department Psychologie, Institut für Pädagogische Psychologie.
- Kreß, H. (2009). *Medizinische Ethik. Gesundheitsschutz - Selbstbestimmungsrechte - heutige Wertkonflikte* (2. Aufl.). Stuttgart: Kohlhammer.
- Kreuz, F. R. (1996). Attitudes of German persons at risk for Huntington's disease toward predictive and prenatal testing. *Genetic Counseling*, 7, 303-311.
- Krüger, D. & Vogt, H. (2007). *Theorien in der biologiedidaktischen Forschung*. Berlin: Springer.
- Kuhn, D. (1991). *The skills of argument*. New York: Cambridge University Press.
- Kuhn, D. (1993). Connecting scientific and informal reasoning. *Merrill-Palmer Quarterly*, 39, 74-93.
- Kuhn, D. (2001). How do people know. *Psychological Science*, 12 (1), 1-8.
- Kuhn, D., Shaw, V. & Felton, M. (1997). Effects of dyadic interaction on Argumentative Reasoning. *Cognition and Construction*, 15 (3), 287-315.
- Kulhavy, R. W. & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1, 279-308.
- Kulhavy, R. W., Yekovich, F. R. & Dyer, J. W. (1976). Feedback and response confidence. *Journal of Educational Psychology*, 68 (5), 522-528.
- Kulik, J. A. & Kulik, C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, 58, 79-97.
- Kulturausschuss der Kultusministerkonferenz (KMK) (1996). Handreichung des Kulturausschusses der Kultusministerkonferenz zu den Aufgaben der Museen. *Museumskunde*, 2 (61), 104-106.
- Kultusministerkonferenz (KMK) (2004). *Bildungsstandards im Fach Biologie für den Mittleren Bildungsabschluss*. Verfügbar unter: http://www.kmk.org/schul/Bildungsstandards/Biologie_MSA_16-12-04.pdf [16.04.2009].
- Lang, M. & Pätzold, G. (2002). *Multimedia in der Aus- und Weiterbildung: Grundlagen von Fallstudien zum netzbasierten Lernen*. Köln: Deutscher Wirtschaftsdienst.
- Langer, I., Schulz v. Thun, F. & Tausch, R. (1999). *Sich verständlich ausdrücken*. München: Ernst Reinhard Verlag.
- Lanzerath, D. (2010). Präimplantationsdiagnostik - Zentrale Fakten und Argumente. *Analysen & Argumente*, 85, 1-13.
- Lave, J. & Wenger, E. (1991). *Situated learning. Legitimate peripheral participation*. Cambridge: University Press.

- Leutner, D. (2002). Adaptivität und Adaptierbarkeit multimedialer Lehr- und Informationssysteme. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (3. Aufl., S. 115-126). Weinheim: Beltz.
- Lewalter, D. (1997). *Lernen mit Bildern und Animationen*. Münster: Waxmann.
- Lewalter, D. (2001). *Abschlußbericht Besucherbefragung in der Pharmazieabteilung des Deutschen Museums München*. Verfügbar unter: <http://www.deutsches-museum.de/fileadmin/Content/Haupthaus/Forschung/PDFs/pharma.pdf> [12.04.2009].
- Lewalter, D. (2009). Bedingungen und Effekte von Museumsbesuchen. In H. Kunz-Ott, S. Kudorfer & T. Weber (Hrsg.), *Kulturelle Bildung im Museum. Aneignungsprozesse - Vermittlungsformen - Praxisbeispiele* (S. 45-56). Bielefeld: transcript.
- Lewalter, D. & Geyer, C. (2003). *Abschlussbericht der Besucherbefragung in der Sonderausstellung „Klima“ des Deutschen Museums München*. Verfügbar unter: http://www.deutsches-museum.de/fileadmin/Content/Haupthaus/Forschung/PDFs/struktur_klima3_end2.pdf [12.04.2009].
- Lewalter, D. & Noschka-Roos, A. (2009). Museum und Erwachsenenbildung. In R. Tippelt & A. von Hippel (Hrsg.) *Handbuch Erwachsenenbildung/Weiterbildung* (S. 527-541). Wiesbaden: VS-Verlag.
- Linek, S. B. (2007). *Geschlechtsspezifisches Design von narrativen Animationen: „Speaker/Gender Effect“ und die Schema-Inkongruenz von Information*. (Bd. 6 der Reihe Wissensprozesse und digitale Medien). Berlin: Logos.
- Lipowsky, F. (2009). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 73-102). Berlin: Springer.
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116, 107-120.
- Lord, C. G., Ross, L. & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effect of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 54, 2098-2109.
- Ludwig, A. K., Diedrich, K., Ludwig, M. & Felberbaum, R. E. (2006). Fertilitätsstörungen und Sterilität. In M. Kaufmann, S. D. Costa & A. Scharl (Hrsg.), *Die Gynäkologie* (S. 163-193). Berlin: Springer.
- Lundgren, S. R. & Prislin, R. (1998). Motivated cognitive processing and attitude change. *Personality and Social Psychology Bulletin*, 24, 715-726.
- Mackie, D. M. (1987). Systematic and nonsystematic processing of majority and minority persuasive communication. *Journal of Personality and Social Psychology*, 53, 41-52.
- Maheswaran, D. & Chaiken, S. (1991). Promoting systematic processing in low motivation settings: The effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology*, 61, 13-25.
- Maio, G. (2001) Die Präimplantationsdiagnostik als Streitpunkt - Welche ethischen Argumente sind tauglich und welche nicht? *Deutsche Medizinische Wochenschrift*, 126, 889-895.
- Mandl, H., Kopp, B. & Dvorak, S. (2004). *Aktuelle theoretische Ansätze und empirische Befunde im Bereich der Lehr-Lern-Forschung - Schwerpunkt Erwachsenenbildung*.

Verfügbar unter: http://www.die-bonn.de/esprid/dokumente/doc-2004/mandl04_01.pdf [12.02.2007].

- Mandl, H., Gruber, H. & Renkl, A. (2002). Situiertes Lernen in multimedialen Lernumgebungen. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia und Internet* (3. Aufl., S. 138-148). Weinheim: Beltz.
- Mangold, M. & Woletz, J. (2007). Bildung, Wissen, Narrativität. Wissensvermittlung durch Digital Storytelling nicht nur für Museen. In M. Mangold, P. Weibel & J. Woletz (Hrsg.), *Vom Betrachter zum Gestalter. Neue Medien in Museen - Strategien, Beispiele und Perspektiven für die Bildung* (S. 33-48). Baden-Baden: Nomos Verlagsgesellschaft.
- Mayer, R. E. (1997). Multimedia Learning: Are we asking the right questions? *Educational Psychologist*, 32, 1-9.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer R. E. (2005a). *The Cambridge Handbook of Multimedia Learning*. Cambridge: University Press.
- Mayer, R. E. (2005b). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 31-48). Cambridge: University Press.
- Mayer, R. E. (2005c). Principles of multimedia learning based on social cues: Personalization, voice and image principles. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 201-212). Cambridge: University Press.
- Mayer, R. E., Fennell, S., Farmer, L. & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, 96, 389-395.
- Mayer, R. E. & Moreno, R. (2000). Engaging Students in Active Learning. *Journal of Educational Psychology*, 92 (4), 724-733.
- Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43-52.
- Mayer, R. E., Sobko, K. & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology*, 95, 419-425.
- Mayring, P. (2008). *Qualitative Inhaltsanalyse*. Weinheim: Beltz.
- McCroskey, J. C. & McCain, T. A. (1974). The measurement of interpersonal attraction. *Speech Monographs*, 41, 261-266.
- Means, M. L. & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability and knowledge levels. *Cognition and Instruction*, 14, 139-178.
- Meisert, A. & Kierdorf, H. (2001). Bioethische Themen im Biologieunterricht. Fachdidaktische Umsetzung am Beispiel der operativen Trennung Siamesischer Zwillinge. *MNU*, 54 (6), 359-363.
- Merkel, R. (2001). Rechte für Embryonen? Die Menschenwürde lässt sich nicht allein auf die biologische Zugehörigkeit zur Menschheit gründen. In C. Geyer (Hrsg.), *Biopolitik- Die Positionen* (S. 51-64). Frankfurt: Suhrkamp.
- Merkel, R. (2003). Contra Speziesargument: zum normativen Status des Embryos und zum Schutz der Ethik gegen ihre biologistische Degradierung In G. Damschen & D.

- Schönecker (Hrsg.), *Der moralische Status menschlicher Embryonen: Pro und Contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument* (S. 35-58). Berlin: de Gruyter.
- Middel, A. (2009). Rechtliche Aspekte der Präimplantationsdiagnostik. In Deutsches Referenzzentrum für Ethik in den Biowissenschaften (DRZE) (Hrsg.), *Präimplantationsdiagnostik* (Bd. 10 der Reihe Sachstandsberichte des DRZE) (S. 52-123). Freiburg: Alber.
- Mietzel, G. (2005). *Wege in die Psychologie*. Stuttgart: Klett-Cotta.
- Mill, J. S. (1871). *Der Utilitarismus*. London.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Mittelsten Scheid, N. & Höble, C. (2008). Wie Schüler unter Verwendung syllogistischer Elemente argumentieren - Eine empirische Studie zu Niveaus von Argumentation im naturwissenschaftlichen Unterricht. *ZfDN*, 14, 145-165.
- Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 507-523). Cambridge: University Press.
- Moreno, R. & Mayer, R. E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology*, 92, 724-733.
- Moreno, R. & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology*, 96, 165-173.
- Moreno, R., Mayer, R. E., Spires, H. & Lester, J. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177-213.
- Morris Hargreaves McIntyre (2004). *Learning Journeys: Using technology to connect the four stages of meaning making*. Manchester: Morris Hargreaves McIntyre.
- Mory, E. H. (2004). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology, Second Edition* (pp. 745-783). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Müller, H. J. (2005). *Gentests. Antworten zu Fragen der medizinischen Praxis*. Basel: Karger.
- Musch, J. (1999). Die Gestaltung von Feedback in computergestützten Lernumgebungen: Modelle und Befunde. *Zeitschrift für Pädagogische Psychologie*, 13 (3), 148-160.
- Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. Münster: Waxmann.
- Nationale Ethikkommission im Bereich der Humanmedizin (NEK-CNE) (2005). *Präimplantationsdiagnostik*. Verfügbar unter: [http://www.bag.admin.ch/nek-cne/04229/04232/index.html?lang=de&download=NHZLpZeg7t,lnp6I0NTU04212Z6ln1acy4Zn4Z2qZpnO2Yuq2Z6gpJCGenx5fWym162epYbg2c_JjKbNoKSn6A-- \[22.04.2009\]](http://www.bag.admin.ch/nek-cne/04229/04232/index.html?lang=de&download=NHZLpZeg7t,lnp6I0NTU04212Z6ln1acy4Zn4Z2qZpnO2Yuq2Z6gpJCGenx5fWym162epYbg2c_JjKbNoKSn6A-- [22.04.2009]).
- Nationale Ethikkommission im Bereich der Humanmedizin (NEK-CNE) (2007). *Präimplantationsdiagnostik II: Spezielle Fragen zur gesetzlichen Regelung und zur*

- HLA-Typisierung*. Stellungnahme Nr. 14. Verfügbar unter: http://studmed.unibe.ch/dozfiles/SGal260PIDII_d.pdf [22.04.2009].
- Nationaler Ethikrat (NER) (2003). *Stellungnahme zur genetischen Diagnostik vor und während der Schwangerschaft*. Verfügbar unter: http://www.ethikrat.org/dateien/pdf/Stellungnahme_Genetische-Diagnostik.pdf [07.03.2008].
- Nass, C., Moon, Y. & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology*, 29 (5), 1093-1110.
- Neubauer, M. (2009). *Medizinisch-naturwissenschaftliche, juristische und ethische Aspekte der Präimplantationsdiagnostik*. Hamburg: Igel.
- Neuer-Miebach, T. (1999). Welche Art von Prävention erkaufen wir uns mit der Zulässigkeit der Präimplantationsdiagnostik? *Ethik in der Medizin*, 1, 125-131.
- Nida-Rümelin, J. (2001). Wo die Menschenwürde beginnt. *Der Tagesspiegel vom 3.1.2001*.
- Niegemann, H. M. (2001). *Neue Lernmedien - Konzeption und Gestaltung multimedialer Lernumgebungen*. Bern: Hans Huber.
- Niegemann, H. M., Hessel, S., Hochscheid-Mauel, D., Aslanski, K., Deimann, M. & Kreuzberger, G. (2004) *Kompendium E-Learning*. Berlin: Springer.
- Noschka-Roos, A. (1994). *Besucherforschung und Didaktik. Ein museumspädagogisches Plädoyer*. Berliner Schriften zur Museumskunde, Bd. 11. Leske + Budrich, Opladen.
- Noschka-Roos, A. (2001). Bausteine eines besucherorientierten Informationskonzepts. In U. Schwarz & P. Teufel (Hrsg.), *Handbuch Museografie und Ausstellungsgestaltung* (S. 88-113). Ludwigsburg: avedition.
- Noschka-Roos, A. (2006). Rahmenbedingungen im Forschungsfeld Museum. In S. Schwan, H. Trischler & M. Prenzel (Hrsg.), *Die Rolle von Medien für die Resituierung von Exponaten: Mitteilungen und Berichte aus dem Institut für Museumsforschung*, Bd. 38 (S. 17-22). Berlin: Institut für Museumsforschung.
- Noschka-Roos, A. & Teichmann, J. (2006). Populäre Wissenschaft in Museen und Science Centers. In P. Faulstich (Hrsg.), *Öffentliche Wissenschaft. Neue Perspektiven der Vermittlung in der wissenschaftlichen Weiterbildung* (S. 87-104). Bielefeld: transcript.
- OECD (1999). *Measuring student knowledge and skills. A new framework for assessment*. Paris: OECD.
- Osborne, J., Erduran, S. & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41 (10), 994-1020.
- Packman, W. L. (1999). Psychosocial impact of pediatric BMT on siblings. *Bone Marrow Transplantation*, 24, 701-706.
- Paivio, A. (1986). *Mental Representations. A Dual-Coding Approach*. New York: Oxford University.
- Paris, S. G. (2002). *Perspectives on Object-Centered Learning in Museums*. Mahwah, NJ: Lawrence Erlbaum.
- Pennings, G. & de Wert, G. (2003). Evolving ethics in medically assisted reproduction. *Human Reproduction Update*, 9 (4), 397-404.

- Pennings, G., Schots, R. & Liebaers, I. (2002). Ethical considerations on preimplantation genetic diagnosis for HLA typing to match a future child as a donor of haematopoietic stem cells to a sibling. *Human Reproduction*, 17, 534-538.
- Perdue, B. C. & Summers, J. O. (1986). Checking the Success of Manipulations in Marketing Experiments. *Journal of Marketing Research*, 23 (4), 317-326.
- Perkins, D. N. (1989). Reasoning as it is and could be: An empirical perspective. In D. Topping, D. Crowell & V. Kabayashi (Eds.), *Thinking across cultures* (pp. 175-194). Hillsdale, NJ: Erlbaum.
- Perkins, D. N., Farady, M. & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 83-105). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perkins, D. N. & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, 18 (1), 16-25.
- Petty, R. E. & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46, 69-81.
- Petty, R. E. & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer.
- Petty, R. E., Cacioppo, J. T. & Goldman, R. (1981). Personal Involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41, 847-855.
- Petty, R. E., Cacioppo, J. T. & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, 10, 135-146.
- Petty, R. E., Wells, G. L. & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, 34, 874-884.
- Peus, C., Frey, D. & Stöger, H. (2006). Theorie der kognitiven Dissonanz. In H.-W. Bierhoff & D. Frey (Hrsg.), *Handbuch der Sozialpsychologie und Kommunikationspsychologie* (S. 373-379). Göttingen: Hogrefe.
- Pfeifer, V. (2003): *Didaktik des Ethikunterrichts. Wie lässt sich Moral lehren und lernen?* Stuttgart: Kohlhammer.
- Pfeifer, V. (2009). *Ethisch argumentieren: Eine Anleitung anhand von aktuellen Fallanalysen*. Paderborn: Schöningh.
- Pfister, H.-R., Böhm, G. & Jungermann, H. (2000). The cognitive representation of genetic engineering: Knowledge and evaluations. *New Genetics and Society*, 19, 295-316.
- Picoult, J. (2004). *Beim Leben meiner Schwester*. München: Piper.
- Plass, J. L., Chun, D. M., Mayer, R. E. & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology*, 90 (1), 25-36.
- Prenzel, M. (2009). Was man alles im Museum lernen kann: Lernvoraussetzungen, Prozesse und Ergebnisse. In International Council of Museums (Ed.), *Wissenschaftskommunikation: Perspektiven der Ausbildung - Lernen im Museum* (S. 137-142). Frankfurt: Peter Lang.

- Prenzel, M. & Thoma, G.-B. (2006). Bildungsforschung im Museum. In S. Schwan, H. Trischler & M. Prenzel (Hrsg.), *Die Rolle von Medien für die Resituierung von Exponaten: Mitteilungen und Berichte aus dem Institut für Museumsforschung*, Bd. 38 (S. 23-34). Berlin: Institut für Museumsforschung.
- Pridemore, D. R. & Klein, J. D. (1991). Control of feedback in computer-assisted instruction. *Educational Technology Research and Development*, 39 (4), 27-32.
- Priester, J. M. & Petty, R. E. (1995). Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. *Personality and Social Psychology Bulletin*, 21, 637-654.
- Ram, N. R. (2006). Britain's new preimplantation tissue typing policy: an ethical defence. *Journal of Medical Ethics*, 32, 278-282.
- Ratcliffe, M. & Grace, M. (2003). *Science education for citizenship*. Maidenhead: OUP.
- Rauprich, O. & Siegel, S. (2003). Der Natur den Weg weisen. Ethische Aspekte der Reproduktionsmedizin. In A. Ley & M. M. Ruisinger (Hrsg.), *Von Gebärdhaus und Retortenbaby. 175 Jahre Frauenklinik Erlangen* (S. 152-171). Nürnberg: Tümmels.
- Rechitsky, S., Kuliev, A., Tur-Kaspa, I., Morris, R. & Verlinsky, Y. (2004). Preimplantation genetic diagnosis with HLA matching. *Reproductive BioMedicine Online*, 9, 210-221.
- Redmon, R. B. (1986). How children can be respected as "ends" yet still be used as subjects in non-therapeutic research. *Journal of Medical Ethics*, 12, 77-82.
- Reeves, B. & Nass, C. (2006). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
- Rehmann-Sutter, C. (2007) Embryoselektion zur Gewebespende? Fälle von PID-HLA und ihre Analyse in individual- und sozialetischer Perspektive. *Ethica*, 15, 115-143.
- Reinmann, G. (2005). Gestaltung von E-Learning-Umgebungen unter emotionalen Gesichtspunkten. In D. Euler & S. Seufert (Hrsg.), *E-Learning in Hochschulen und Bildungszentren* (S. 351-371). München: Oldenbourg.
- Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie. Ein Lehrbuch*. (S. 613-658). Weinheim: Beltz.
- Reinmann-Rothmeier, G. & Mandl, H. (2000). Bedarfs- und implementationsorientierte Evaluation von Lernsoftware: Eine Feldstudie mit Meistern und Technikerschülern. In P. Schenkel, S.-O. Tergan & A. Lottmann (Eds.), *Qualitätsbeurteilung multimedialer Lern- und Informationssysteme. Evaluationsmethoden auf dem Prüfstand* (S. 243-261). Nürnberg: BW Bildung und Wissen.
- Reitschert, K. (2009). *Ethisches Bewerten im Biologieunterricht. Eine qualitative Untersuchung zur Strukturierung und Ausdifferenzierung von Bewertungskompetenz in bioethischen Sachverhalten bei Schülern der Sekundarstufe I*. Hamburg: Dr. Kovač.
- Reitschert, K. & Höhle, C. (2006). Die Struktur von Bewertungskompetenz. Ein Beitrag zur Dimensionierung eines Kompetenzmodells im Bereich der Bioethik. *Erkenntnisweg Biologiedidaktik*, 5, 99-114.
- Reitschert, K. & Höhle, C. (2007). Wie Schüler ethisch bewerten - Eine qualitative Untersuchung zur Strukturierung und Ausdifferenzierung von

- Bewertungskompetenz in bioethischen Sachverhalten bei Schülern der Sek.I-. *Zeitschrift für Didaktik der Naturwissenschaften (ZfDN)*, 13, 125-143.
- Renkl, A. (1996). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, 47, 62-78.
- Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, 16 (9), 13-20.
- Reussner, E. M., Schwan, S. & Zahn, C. (2007). New technologies for learning in museums? An interdisciplinary research project. In J. Trant & D. Bearman (Eds.), *Proceedings of the International Cultural Heritage Informatics Meeting (ICHIM07)*. Toronto, Canada: Archives & Museum Informatics.
- Rey, G. D. (2009). *E-Learning. Theorien, Gestaltungsempfehlungen und Forschung*. Bern: Hans Huber.
- Rheinberg, F. (2001). Bezugsnormen und schulische Leistungsbeurteilung. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 59-71). Weinheim: Beltz PVU.
- Rheinberg, F. (2006). Bezugsnormorientierung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 55-62). Weinheim: Beltz PVU.
- Rheinberg, F. & Krug, S. (1999). *Motivationsförderung im Schulalltag. Konzept, Realisation und Evaluation*. Göttingen: Hogrefe.
- Robertson, J. A. (2003) Ethical issues in new uses of preimplantation genetic diagnosis. *Human Reproduction*, 18, 465-471.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Ross, L., Greene, D. & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279-301.
- Runtenberg, C. (2001). *Didaktische Ansätze einer Ethik der Gentechnik: Produktionsorientierte Verfahren im Unterricht*. Freiburg: Alber.
- Sadler, T. D. (2004a). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, 41, 513-536.
- Sadler, T. D. (2004b). Moral and ethical dimensions of socioscientific decision-making as integral components of scientific literacy. *The Science Educator*, 13 (1), 39-48.
- Sadler, T. D., Barab, S. A. & Scott, B. (2007). What do students gain by engaging in socioscientific inquiry? *Research in Science Education*, 37 (4), 371-391.
- Sadler, T. D. & Donnelly, L. A. (2006). Socioscientific argumentation: The effects of content knowledge and morality. *International Journal of Science Education*, 28, 1463-1488.
- Sadler, T. D. & Fowler, S. (2006). A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education*, 90, 986-1004.
- Sadler, T. D. & Zeidler, D. L. (2003). The Morality of Socioscientific Issues: Construal and Resolution of Genetic Engineering Dilemmas. *Science Education*, 88, 4-27.
- Sadler, T. D. & Zeidler, D. L. (2004). The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education*, 89, 71-93.

- Sadler, T. D. & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision-making. *Journal of Research in Science Teaching*, 42, 112-138.
- Sales, G. C. (1993). Adapted and adaptive feedback in technology-based instruction. In J. V. Dempsey & G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 159-175). Englewood Cliffs, NJ: Educational Technology.
- Sandifer, C. (2003). Technological novelty and open-endedness: Two characteristics of interactive exhibits that contribute to the holding of visitor attention in a science museum. *Journal of Research in Science Teaching*, 40, 121-137.
- Sansone, C. (1986). A question of competence: The effects of competence and task feedback on intrinsic interest. *Journal of Personality and Social Psychology*, 51, 918-931.
- Schanze, S. & Saballus, U. (2009). CoReflect: webbasierte Unterstützung des forschenden Lernens. In D. Höttecke (Hrsg.), *Chemie- und Physikdidaktik für die Lehramtsausbildung* (S. 480-482). Münster: Gesellschaft für Didaktik der Chemie und Physik. GDCP Jahrestagung in Schwäbisch Gmünd 2008.
- Schaumburg, H. & Issing, L. J. (2004). Interaktives Lernen mit Multimedia. In R. Mangold & P. Vorderer (Hrsg.), *Lehrbuch der Medienpsychologie* (S. 717-742). Göttingen: Hogrefe.
- Scheersoi, A. (2006). Computer in Museumsausstellungen – Top oder Flop? In S. Schwan, H. Trischler & M. Prenzel (Hrsg.), *Die Rolle von Medien für die Resituierung von Exponaten: Mitteilungen und Berichte aus dem Institut für Museumsforschung, Bd. 38* (S. 47-58). Berlin: Institut für Museumsforschung.
- Schiefele, U. (1996). *Motivation und Lernen mit Texten*. Göttingen: Hogrefe.
- Schimmel, B. J. (1988). Providing meaningful feedback in courseware. In D. H. Jonassen (Ed.), *Instructional Designs for Microcomputer Courseware* (pp. 183-197). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schmidt, H. T. (2003). *Präimplantationsdiagnostik. Jenseits des Rubikons? Individual- und sozialethische Aspekte der PID/PGD*. Münster: LIT-Verlag.
- Schmitz, D. & Wiesing, U. (2008). *Ethische Aspekte der genetischen Diagnostik in der Arbeitsmedizin*. Köln: Deutscher Ärzte-Verlag.
- Schnotz, W. (2001). Wissenserwerb mit Multimedia. *Unterrichtswissenschaft*, 29, 292-318.
- Schockenhoff, E. (2003). Pro Speziesargument. Zum moralischen und ontologischen Status des Embryos. In G. Damschen & D. Schönecker (Hrsg.), *Der moralische Status menschlicher Embryonen: Pro und Contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument* (S. 11-33). Berlin: de Gruyter.
- Schockenhoff, E. (2005). Fortpflanzungsfreiheit und verantwortliche Elternschaft. Zur ethischen Problematik der Präimplantationsdiagnostik. In E. Schockenhoff, A. Buch, M. Volkenandt & V. Wetzstein (Hrsg.), *Medizinische Ethik im Wandel. Grundlagen - Konkretionen – Perspektiven* (S. 216-232). Freiburg im Breisgau: Schwabenverlag.
- Schöne-Seifert, B. (2003). Contra Potentialitätsargument: Probleme einer traditionellen Begründung für embryonalen Lebensschutz. In G. Damschen & D. Schönecker (Hrsg.), *Der moralische Status menschlicher Embryonen: Pro und Contra Spezies-*

- Kontinuums-, Identitäts- und Potentialitätsargument* (S. 169-185). Berlin: de Gruyter.
- Schräer, A. (2009). Ethische Aspekte der Präimplantationsdiagnostik. In Deutsches Referenzzentrum für Ethik in den Biowissenschaften (DRZE) (Hrsg.), *Präimplantationsdiagnostik* (Bd. 10 der Reihe Sachstandsberichte des DRZE) (S. 124-160). Freiburg: Alber.
- Schramm, H. (2008). Parasoziale Interaktion (PSI). In N. C. Krämer, S. Schwan & D. Unz & M. Suckfüll (Hrsg.), *Medienpsychologie – Schlüsselbegriffe und Konzepte* (S. 253-258). Stuttgart: Kohlhammer.
- Schramm, H. & Hartmann, T. (2008). Die Messung von parasozialen Interaktionen als mehrdimensionales Konstrukt. Entwicklung und Validierung von PSI-Prozess-Skalen auf Basis des Zwei-Ebenen-Modells parasozialer Interaktionen. In J. Matthes, W. Wirth, A. Fahr & G. Daschmann (Hrsg.), *Die Brücke zwischen Theorie und Empirie: Operationalisierung, Messung und Validierung in der Kommunikationswissenschaft* (S. 48-69). Köln: Halem.
- Schramme, T. (2002). *Bioethik*. Frankfurt/New York: Campus.
- Schröder, E., Schmutzer, G., Klaiberg, A. & Brähler, E. (2003). Ärztliche Sterbehilfe im Spannungsfeld zwischen Zustimmung zur Freigabe und persönlicher Inanspruchnahme - Ergebnisse einer repräsentativen Befragung in Deutschland. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 53, 334-343.
- Schütt, C. & Bröker, B. (2009). *Grundwissen Immunologie*. Heidelberg: Spektrum Akademischer Verlag.
- Schuh, H. (2000). Erbgut-Check für Embryonen. Die Zukunftstechnik PID könnte viel Leid lindern. *Die Zeit, Nr. 10, Bildung und Wissen*.
- Schulz, v. Thun, F. (1981). *Miteinander reden: Störungen und Klärungen*. Reinbek bei Hamburg: Rowohlt.
- Schwan, S. (2005). Die Bedeutung der neuen Medien für den Wissenserwerb in naturwissenschaftlich-technischen Museen. In A. Noschka-Roos, W. Hauser & E. Schepers (Hrsg.), *Mit Neuen Medien im Dialog mit den Besuchern? Grundlagen und Praxis am Beispiel des Zentrums Neue Technologien im Deutschen Museum. Berliner Schriften zur Museumskunde, Bd. 21* (S. 49-52). Berlin: G + H Verlag.
- Schwan, S. (2006). Lernen im Museum: Die Rolle der digitalen Medien für Wissenserwerb und Wissenskommunikation. In S. Schwan, H. Trischler & M. Prenzel (Hrsg.), *Die Rolle von Medien für die Resituierung von Exponaten: Mitteilungen und Berichte aus dem Institut für Museumsforschung, Bd. 38* (S. 1-8). Berlin: Institut für Museumsforschung.
- Schwan, S. (2009). Lernen und Wissenserwerb in Museen. In H. Kunz-Ott, S. Kudorfer & T. Weber (Hrsg.), *Kulturelle Bildung im Museum. Aneignungsprozesse - Vermittlungsformen - Praxisbeispiele* (S. 33-43). Bielefeld: transcript.
- Schwan, S., Trischler, H. & Prenzel, M. (2006). *Die Rolle von Medien für die Resituierung von Exponaten: Mitteilungen und Berichte aus dem Institut für Museumsforschung, Bd. 38*. Berlin: Institut für Museumsforschung.
- Schwan, S., Zahn, C., Wessel, D., Huff, M., Herrmann, N. & Reussner, E. (2008). Lernen in Museen und Ausstellungen - die Rolle digitaler Medien. *Unterrichtswissenschaft Zeitschrift für Lernforschung. Thema: Lernen im Museum*, 36 (2), 117-135.

- Schwarz, B., Neuman, Y., Gil, J. & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *Journal of the Learning Sciences*, 12 (2), 219-256.
- Schwarz, N., Frey, D. & Kumpf, M. (1980). Interactive effects of writing and reading a persuasive essay on attitude change and selective exposure. *Journal of Experimental Social Psychology*, 16, 1-17.
- Schweiger, W. (2007). *Theorien der Mediennutzung. Eine Einführung*. Wiesbaden: VS-Verlag für Sozialwissenschaften.
- Schwinger, E. (2004). Polkörperdiagnostik aus humangenetischer Sicht. *Medizinische Genetik*, 4, 391-392.
- Serrell, B. (1997). Paying attention: The duration and allocation of visitors' time in museum exhibitions. *Curator*, 40, 108-125.
- Serrell, B. (2002). Are They Watching? Visitors and Videos in Exhibitions. *Curator*, 45 (1), 50-64.
- Shaw, V. (1996). The cognitive processes in informal reasoning. *Thinking and Reasoning*, 2 (1), 51-80.
- Shaw, J. S. & Bassi, K. L. (2001). Lay attitudes toward genetic testing for susceptibility to inherited diseases. *Journal of Health Psychology*, 6, 405-423.
- Shenfield, F., Pennings, G., Cohen, J., Devroey, P. & Tarlatzis, B. (2005). Taskforce 9: the application of preimplantation genetic diagnosis for human leukocyte antigen typing of embryos. *Human Reproduction*, 20 (4), 845-847.
- Shettel, H. (1968). An evaluation of existing criteria for judging the quality of science exhibits. *Curator*, 11 (2), 137-135.
- Siegrist, M. & Bühlmann, R. (1999). Die Wahrnehmung verschiedener gentechnischer Anwendungen: Ergebnisse einer MDS-Analyse. *Zeitschrift für Sozialpsychologie*, 30, 32-39.
- Simon, A. & Schenker, J. G. (2005). Ethical consideration of intended preimplantation genetic diagnosis to enable future tissue transplantation. *Reproductive BioMedicine Online*, 10 (3), 320-324.
- Singer, P. (1990). *Praktische Ethik*. Stuttgart: Reclam.
- Singer, E., Corning, A. D. & Antonucci, T. (1999). Attitudes towards genetic testing and fetal diagnosis, 1990-1996. *Journal of Health and Social Behavior*, 40, 429-445.
- Snyder, M. (1979). Self-monitoring processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, pp. 85-128). New York: Academic.
- Snyder, C. R. & Fromkin, H. L. (1980). *Uniqueness, the human pursuit of difference*. New York: Plenum Press.
- Specht, I. & Lewalter, D. (2011). *Abschlussbericht der Besucherbefragung in der Ausstellung zur Nano- und Biotechnologie im Deutschen Museum München*. Verfügbar unter: http://www.deutsches-museum.de/fileadmin/Content/2009/06_Forschung/04_Projekte/Bericht_ZNT_Besucherstruktur_Februar_2011.pdf [30.06.2011].
- Spiewak, M. (2005). *Wie weit gehen wir für ein Kind? Im Labyrinth der Fortpflanzungsmedizin*. Frankfurt: Eichborn.

- Spiro, R. J., Feltovich, P. J., Coulson, R. L. & Anderson, D. K. (1989). Multiple analogies for complex concepts: Antidotes for analogy-induced misconception in advanced knowledge acquisition. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 498-531). Cambridge: Cambridge University Press.
- Spiro, R. J., Feltovich, P. J., Jacobson, M. J. & Coulson, R. L. (1991). Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology*, 31, 24-33.
- Spiro, R. J., Feltovich, P. J., Jacobson, M. J. & Coulson, R. L. (1992). Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. Duffy & D. Jonassen (Eds.), *Constructivism and the technology of instruction* (pp. 57-76). Hillsdale, NJ: Erlbaum.
- Spiro, R. J. & Jehng, J. C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. J. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 163-205). Hillsdale, NJ: Erlbaum.
- Spiro, R. J., Vispoel, W. P., Schmitz, J. G., Samarapungavan, A. & Boerger, A. E. (1987). Knowledge acquisition for application. In B. K. Britton & S. M. Glynn (Eds.), *Executive control processes in reading* (pp. 177-199). Hillsdale, NJ: Erlbaum.
- Stark, R. (2001). *Analyse und Förderung beispielbasierten Lernens: Anwendung eines integrativen Forschungsparadigmas*. München: Unveröff. Habilitationsschrift, Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.
- Stark, R., Gruber, H., Renkl, A. & Mandl, H. (1998). Instructional effects in complex learning: Do objective and subjective learning outcomes converge? *Learning & Instruction*, 8 (2), 117-129.
- Stark, R. & Mandl, H. (2001a). *Entwicklung, Implementation und Evaluation eines beispielbasierten Instruktionsansatzes zur Förderung von Handlungskompetenz im Bereich empirischer Forschungsmethoden (Forschungsbericht Nr. 141)*. München: Ludwig-Maximilians-Universität, Department Psychologie, Institut für Pädagogische Psychologie.
- Stark, R. & Mandl, H. (2001b). *Die Kluft zwischen Wissenschaft und Praxis – ein unlösbares Problem für die pädagogisch-psychologische Forschung (Forschungsbericht Nr. 118)*. München: Ludwig-Maximilians-Universität, Department Psychologie, Institut für Pädagogische Psychologie.
- Steck, T. (2001). *Praxis der Fortpflanzungsmedizin*. Stuttgart: Schattauer.
- Stein, N. L. & Albro, E. R. (2001). The origins and nature of arguments: Studies in conflict understanding, emotion, and negotiation. *Discourse Processes*, 32 (2), 113-133.
- Steinke, V. & Rahner, N. (2009). Medizinisch-naturwissenschaftliche Aspekte der Präimplantationsdiagnostik. In Deutsches Referenzzentrum für Ethik in den Biowissenschaften (DRZE) (Hrsg.), *Präimplantationsdiagnostik* (Bd. 10 der Reihe Sachstandsberichte des DRZE) (S. 13-51). Freiburg: Alber.
- Steinmetz, R. (2000). *Multimedia-Technology* (2. Aufl.). Berlin: Springer.
- Stephoe, P. C. & Edwards, R. G. (1978). Birth after the reimplantation of a human embryo. *The Lancet*, 312 (8085), 366.

- Stockdale, A. (1999). Public understanding of genetics and Alzheimer disease. *Genetic Testing*, 3, 139-145.
- Stoecker, R. (2003). Contra Identitätsargument: mein Embryo und ich. In G. Damschen & D. Schönecker (Hrsg.), *Der moralische Status menschlicher Embryonen: pro und contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument* (S. 129-145). Berlin: de Gruyter.
- Stollorz, V. (2000). Erbgut-Check für Embryonen. Die PID beschwört eine neue Eugenik herauf. *Die Zeit*, Nr. 10, *Bildung und Wissen*.
- Strack, F. & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8 (3), 220-247.
- Sweller, J. (1999). *Instructional Design in Technical Areas*. ACER: Melbourne.
- Sweller, J., Van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.
- Tesser, A., Campbell, J. & Mickler, S. (1983). The role of social pressure, attention to the stimulus, and self doubt in conformity. *European Journal of Social Psychology*, 13, 217-233.
- Thoben, D. F. & Erb, H. P. (2010). *Wie es euch gefällt: Sozialer Einfluss durch Mehrheiten und Minderheiten*. Verfügbar unter: <http://de.in-mind.org/content/wie-es-euch-gef%C3%A4hlt-sozialer-einfluss-durch-mehrheiten-und-minderheiten> [11.12.2010].
- Todt, E. & Götz, C. (1997). Hoffnungen und Befürchtungen gegenüber Gentechnik. *Zeitschrift für Didaktik der Naturwissenschaften (ZfDN)*, 2, 15-22.
- Todt, E. & Götz, C. (1998). Interesse von Jugendlichen an der Gentechnologie. *Zeitschrift für Didaktik der Naturwissenschaften (ZfDN)*, 1, 3-11.
- Töpfer, J. (2009). *Filmische Personalisierung von Ausstellungsinhalten. Einfluss narrativer Interviews auf den Wissenserwerb beim selbstgesteuerten Lernen im informellen Setting*. Hamburg: Dr. Kovač.
- Toplak, M. E. & Stanovich, K. E. (2003). Associations between myside bias on an informal reasoning task and amount of post-secondary education. *Applied Cognitive Psychology*, 17, 851-860.
- Toulmin, S. (1958). *The uses of argument*. New York: Cambridge University Press.
- Toulmin, S. (1996). *Der Gebrauch von Argumenten*. Weinheim: Beltz.
- Traut-Mattausch, E., Schulz-Hardt, S., Greitemeyer, T. & Frey, D. (2004). Expectancy confirmation in spite of disconfirming evidence: The case of price increases due to the introduction of the Euro. *European Journal of Social Psychology*, 34, 739-760.
- Tsafir, A., Shufaro, Y., Simon, A. & Laufer, N. (2005). Preimplantation genetic diagnosis. In S. Blazer & E. Z. Zimmer (Eds.), *The embryo: scientific discovery and medical ethics* (pp. 166-201). Basel: Karger.
- Tymstra, T. J. (1989). The Imperative Character of Medical Technology and the Meaning of "Anticipated Decision Regret". *International Journal of Technology Assessment in Health Care*, 5, 207-213.
- Van den Daele, W. (2005). „Vorgeburtliche Selektion: Ist die Pränataldiagnostik behindertenfeindlich“? In W. van den Daele (Hrsg.), *Biopolitik. Leviathan*

- Sonderheft Nr. 23/2005* (S. 97-122). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Von Loewenich, V. (2005). Ethische Probleme bei der Präimplantationsdiagnostik (PID) aus pädiatrischer Sicht. In D. Ganten & K. Ruckpaul (Hrsg.), *Molekularmedizinische Grundlagen von fetalen und neonatalen Erkrankungen* (S. 121-128). Heidelberg: Springer.
- Verlinsky, Y., Rechitsky, S., Sharapova, T., Morris, R., Taranissi, M. & Kuliev, A. (2004). Preimplantation HLA Testing. *JAMA*, 291 (17), 2079-2085.
- Verlinsky, Y., Rechitsky, S., Schoolcraft, W., Strom, C. & Kuliev, A. (2001). Preimplantation diagnosis for Fanconi anemia combined with HLA matching. *JAMA*, 285 (24), 3130-3133.
- Vorderer, P. (1996). *Fernsehen als „Beziehungskiste“*. Parasoziale Beziehungen und Interaktionen mit TV Personen. Opladen: Westdeutscher Verlag.
- Voss, J. F., Blais, J., Means, M. L., Greene, T. R. & Ahwesh, E. (1986). Informal Reasoning and Subject Matter Knowledge in the Solving of Economics Problems by Naive and Novice Individuals. *Cognition and Instruction*, 3, 269-302.
- Voss, J. F. & Means, M. L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction*, 1, 337-350.
- Walton, D. (1992). *Slippery Slope Arguments*. Oxford: Clarendon Press.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.
- Watson, D. & Clark, L. A. (1994). *The PANAS-X. Manual for the Positive and Negative Affect Schedule-Expanded Form*. Verfügbar unter: <http://www.psychology.uiowa.edu/faculty/Clark/PANAS-X.pdf> [24.09.2008].
- Weidenmann, B. (2002). Multicodierung und Multimodalität im Lernprozess. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia* (3. Aufl., S. 45-64). Weinheim: Beltz.
- Werth, L. & Mayer, J. (2008). *Sozialpsychologie*. Heidelberg: Spektrum Akademischer Verlag.
- West, R. F., Toplak, M. E. & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, 100, 930-941.
- Wieland, W. (2003). Pro Potentialitätsargument. Moralfähigkeit als Grundlage von Würde und Lebensschutz. In G. Damschen & D. Schönecker (Hrsg.), *Der moralische Status menschlicher Embryonen: pro und contra Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument* (S. 149-168). Berlin: de Gruyter.
- Whitehead, A. N. (1929). *The aims of education*. New York: MacMillan.
- Winkler, K. & Mandl, H. (2002). Knowledge-Master. Wissensmanagement-Weiterbildung mit WBT. In U. Dittler (Hrsg.), *E-Learning. Erfolgsfaktoren und Einsatzkonzepte mit interaktiven Medien* (S. 205-216). München: Oldenbourg.
- Wirth, W. & Schramm, H. (2005). Medienrezeption. In H. Bonfadelli, O. Jarren & G. Siegert (Hrsg.), *Einführung in die Publizistikwissenschaft* (S. 527-559). Bern: UTB.

- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.
- Wood, W. (1982). Retrieval of attitude-relevant information from memory: Effects on susceptibility to persuasion and on intrinsic motivation. *Journal of Personality and Social Psychology*, 42, 798-810.
- Wood, W., Kallgren, C. A. & Preisler, R. M. (1985). Access to attitude-relevant information in memory as a determinant of persuasion: The role of message attributes. *Journal of Experimental Social Psychology*, 21, 73-85.
- Wu, Y-T. & Tsai, C.-C. (2007). High School Students' Informal Reasoning on a Socio-scientific Issue: Qualitative and quantitative analyses. *International Journal of Science Education*, 29 (9), 1163-1187.
- Yaron, Y., Gamzu, R. & Malcov, M. (2004). Genetic analysis of the embryo. In D. K. Gardner, A. Weissmann, C. M. Howles & Z. Shoham (Eds.), *Textbook of Assisted Reproductive Techniques* (pp. 319-332). New York: Taylor & Francis.
- Zahn, C. (2006). Forschung zur Rolle neuer Medien im Museum - psychologische Perspektiven und Methoden am IWM. In S. Schwan, H. Trischler & M. Prenzel (Hrsg.), *Die Rolle von Medien für die Resituierung von Exponaten, Bd. 38* (S. 19-24). Berlin: Institut für Museumsforschung.
- Zahn, C. J. & Hopper, R. (1985). Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology*, 4, 113-122.
- Zerres, K. (1993). Prädiktive Diagnostik und genetisches Screening in der Bevölkerung. In K. Zerres & R. Rüdell (Hrsg.), *Selbsthilfegruppen und Humangenetiker im Dialog. Erwartungen und Befürchtungen* (S. 100-107). Stuttgart: Enke.
- Ziegler, U. (2004). *Präimplantationsdiagnostik in England und Deutschland. Ethische, rechtliche und praktische Probleme*. Frankfurt, New York: Campus.
- Zimbardo, P. G. & Gerrig, R. J. (2008). *Psychologie* (18. Aufl.). Pearson Studium: München.
- Zoglauer, T. (2004). Dammbrechargumente in der Bioethik. In H. Friesen & K. Berr (Hrsg.), *Angewandte Ethik im Spannungsfeld von Begründung und Anwendung* (S. 309-326). Frankfurt: Peter Lang.
- Zoglauer, T. (2005). *Einführung in die formale Logik für Philosophen*. Göttingen: Vandenhoeck & Ruprecht.
- Zohar, A. & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39, 35-62.
- Zumbach, J. (2007). *Instruktionspsychologische Grundlagen zum Lernen mit Neuen Medien*. Stuttgart: Kohlhammer.
- Zumbach, J., Haider, K. & Mandl, H. (2008). Fallbasiertes Lernen: Theoretischer Hintergrund und praktische Anwendung. In J. Zumbach & H. Mandl (Hrsg.), *Pädagogische Psychologie in Theorie und Praxis. Ein fallbasiertes Lehrbuch* (S. 1-11). Göttingen: Hogrefe.

13 Abbildungs- und Tabellenverzeichnis

Abbildungsverzeichnis

Abbildung 1: Ablauf einer PID an Blastomeren (Hengstschläger, 2006b, S. 10).....	22
Abbildung 2: Der Prozess des informellen Schlussfolgerns aus Sicht der Zwei-Prozess-Theorien (nach Wu & Tsai, 2007, p. 1166).....	74
Abbildung 3: Toulmin's Argumentschema (1958, 1996).....	82
Abbildung 4: Cognitive Theory of Multimedial Learning (Mayer, 2001; Übersetzung von Zumbach, 2007).....	97
Abbildung 5: Schematische Darstellung des Einflusses vorhandener bzw. fehlender sozialer Hinweisreize auf den Lernprozess in multimedialen Lernumgebungen im Rahmen der Social Agency Theory (Mayer, 2005c; Rey, 2009, S. 84).	101
Abbildung 6: Schema von Festinger's (1957) Prozessmodell der Dissonanz (Devine et al., 1999, p. 298).	130
Abbildung 7: Meinungsterminal in der Ausstellung „Leben mit Ersatzteilen“ des ZNT zum Thema Stammzellforschung, Eingabe-PC (Screenshot Touchscreen-Interface).....	141
Abbildung 8: Prototyp Dialogstation „Gentest“, Vorderansicht, Touchscreen-Monitor zur Bedienung (unten Mitte, Foto: Deutsches Museum).....	154
Abbildung 9: Prototyp Dialogstation „Gentest“, seitliche Ansicht von links, 3-D-Effekt (Teilausschnitt, Foto: Deutsches Museum und Charlotte Kaiser, Berlin)..	154
Abbildung 10: Einstiegsseite bzw. Startseite der Dialogstation mit sieben frei auswählbaren Fallbeispielen zum Thema Gentest (Screenshot Touchscreen).....	155
Abbildung 11: Aufforderung zur aktiven Positionierung über Ja/Nein-Abstimmung (Screenshot Touchscreen).....	156
Abbildung 12: Feedback in Form einer sozialen Vergleichsinformation (Konsensinformation) über die Gesamtentscheidung der anderen Besucher (Screenshot Touchscreen).....	156
Abbildung 13: Aufenthaltsdauer (<i>holding power</i>) der Besucher an dem Prototypen der Dialogstation „Gentest“ in Minuten, absolute Häufigkeiten, Verteilung der Werte (N=300).....	164
Abbildung 14: Finale Dialogstation „Gentest“ in der ZNT-Ausstellung zur Nano- und Biotechnologie im Deutschen Museum (Foto: Deutsches Museum, Anmerkung: Gesichter der beiden Ausstellungsbesucher wurden aus Datenschutzgründen unkenntlich gemacht).....	167
Abbildung 15: Für die vorliegende Studie verwendetes Fallbeispiel zur PID-HLA, Transkript des Films.	184

Abbildung 16: Frontalansicht <i>talking head</i> „Margot Kreidler“ (Foto: Deutsches Museum).	185
Abbildung 17: Seitliche Ansicht <i>talking head</i> „Margot Kreidler“ (Foto: Deutsches Museum).	185
Abbildung 18: <i>talking head</i> „Margot Kreidler“, unten Touchscreen-Monitor zur Bedienung der Lernumgebung (Foto: Deutsches Museum).	185
Abbildung 19: Startseite zum Aufruf des personalisierten Films (Screenshot Touchscreen).	186
Abbildung 20: Aufforderung zur aktiven Positionierung durch die Lernumgebung zur Erhebung der getroffenen Entscheidung per Logfile-Recording (Screenshot Touchscreen).	187
Abbildung 21: Aufforderung zur Bearbeitung von Frageblock 1 durch die Lernumgebung (Screenshot Touchscreen).	187
Abbildung 22: Beispiel Zusatzinformation Bewertung, Contra-Position der Nationalen Ethikkommission der Schweiz zur PID-HLA (Screenshot Touchscreen).	188
Abbildung 23: Überblick über die drei implementierten Messzeitpunkte und Frageblöcke (Zeitstrahl).	189
Abbildung 24: Beispiel Konsensinformation nach „Nein-Entscheidung“ für Bedingung „inkongruentes Feedback“ (Screenshot Touchscreen).	191
Abbildung 25: Überblick über den gesamten Versuchsablauf des Experiments.	193
Abbildung 26: Instruktion zur Bearbeitung der Lernumgebung im zweiten Untersuchungsteil.	194
Abbildung 27: Erlebtes psychologisches Unbehagen bzw. Dissonanzerleben bei beiden Experimentalgruppen ($n = 48$) (z-Werte).	220
Abbildung 28: Richtung der Entscheidungen (Pro PID = „Ja“ vs. Contra PID = „Nein“) im zeitlichen Verlauf (Messzeitpunkte t_1 bis t_3), absolute Häufigkeiten (Gesamtgruppe).	225
Abbildung 29: Zeitlicher Verlauf der Sicherheit der Entscheidung für alle drei Gruppen (Messzeitpunkte t_1 - t_3), Mittelwerte, Liniendiagramm.	228
Abbildung 30: Argumentationsqualität zum ersten Messzeitpunkt: Verteilung der Punktwerte, absolute Häufigkeiten (Gesamtgruppe).	236
Abbildung 31: Argumentationsqualität zum dritten Messzeitpunkt: Verteilung der Punktwerte, absolute Häufigkeiten (Gesamtgruppe).	237
Abbildung 32: Argumentationsqualität zum ersten und dritten Messzeitpunkt (vor und nach Bearbeitung der Zusatzinformationen) für die Gesamtgruppe (Mittelwerte und Ergebnisse der Signifikanzprüfung, Balkendiagramm).	239
Abbildung 33: Anzahl der generierten Gegenargumente für alle drei Lernbedingungen, (Mittelwerte, Säulendiagramm).	241

Tabellenverzeichnis

Tabelle 1: Auszüge aus dem Embryonenschutzgesetz (ESchG) von 1991.....	32
Tabelle 2: Positionen der Nationalen Ethikkommission der Schweiz bezüglich der ethischen Einschätzung der PID-HLA (NEK-CNE, 2007, S. 15-16).	60
Tabelle 3: Formale Grundstruktur eines Arguments im theoretischen Syllogismus (Bayer, 1999; Pfeifer, 2009).....	80
Tabelle 4: Normativer Syllogismus am Beispiel der PID (Dietrich, 2004, 2005; Höhle & Reitschert, 2006, S. 105).	80
Tabelle 5: Dimensionen zur Beschreibung medialer Angebote (Weidenmann, 2002, S. 47).	94
Tabelle 6: PSI-Dimensionen: drei PSI-Teilprozesse und ihre Unterdimensionen (Hartmann et al., 2004a).....	105
Tabelle 7: PSI-relevante Eigenschaften von Medienpersonae und Mediennutzern/ Rezipienten (Hartmann et al., 2004a; 2004b).	108
Tabelle 8: Übersicht über das Untersuchungsdesign und Stichprobengröße (N = 72).	182
Tabelle 9: Pretest zur Auswahl des Fallbeispiels, Chorea Huntington vs. PID-HLA, Entscheidungsrelevanz und Verständlichkeit der Zusatzinformationen für beide Fallbeispiele (Gesamtmittelwerte und Standardabweichungen in Klammer).	183
Tabelle 10: Kategorienschema zur Operationalisierung der Argumentationsqualität (Stufen, Beschreibungen und Ankerbeispiele).	207
Tabelle 11: Kategorienschema zur Erfassung der Argumentationsmuster (Kategorien mit Beschreibung und Ankerbeispielen).	211
Tabelle 12: Themenspezifisches Vorwissen zur Humangenetik für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.	213
Tabelle 13: Subjektives Vorwissen zum Thema Gentest, PID und Bioethik für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).	213
Tabelle 14: Persönlicher Bezug zum Thema Gentest, Interesse am Thema Gentest, Präferenz für Deliberation und Intuition für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).	214
Tabelle 15: Soziale Vergleichsorientierung und Orientierung an anderen für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.	215
Tabelle 16: Einstellungen (Positive Aspekte, Negative Aspekte und Befürchtungen) zu genetischen Untersuchungen für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).	215

Tabelle 17: Individuelle Lernzeit bzw. Bearbeitungszeit (in Minuten) der Lernumgebung für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung).....	217
Tabelle 18: Soziale Projektion, selbst erzeugter Konsens und Bedürfnis nach sozialem Vergleich für alle drei Lernbedingungen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.	218
Tabelle 19: Psychologisches Unbehagen nach dem Feedback für beide Experimentalgruppen: Positiver und negativer Affekt sowie Überraschung. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.	219
Tabelle 20: Wahrgenommener Konflikt für beide Experimentalgruppen. Mittelwerte (Standardabweichungen in Klammern) und Ergebnisse der Signifikanzprüfung.	221
Tabelle 21: Attraktivität der Medienfigur und parasoziale Interaktion mit der Medienfigur, insgesamt, d. h. über alle drei Gruppen hinweg. Mittelwerte (Standardabweichungen in Klammern), Minima, Maxima.	223
Tabelle 22: Korrelation nach Pearson zwischen der Attraktivität der Medienfigur und der parasozialen Interaktion mit der Medienfigur.	223
Tabelle 23: Richtung der Entscheidung (Ja vs. Nein) für alle drei Lernbedingungen zu allen drei Messzeitpunkten (zeitlicher Verlauf) (absolute Häufigkeiten).....	224
Tabelle 24: Entscheidungssicherheit im zeitlichen Verlauf (Messzeitpunkte t1 bis t3) für alle drei Lernbedingungen und gesamt. Mittelwerte (Standardabweichungen in Klammern).....	227
Tabelle 25: Entscheidungsänderung nach Feedback (t2) für alle drei Lernbedingungen und insgesamt. Absolute Häufigkeiten.	230
Tabelle 26: Entscheidungsänderung nach Bearbeitung der Zusatzinformation (t3) für alle drei Lernbedingungen und insgesamt. Absolute Häufigkeiten.	233
Tabelle 27: Argumentationsqualität zum ersten und dritten Messzeitpunkt (vor und nach Bearbeitung der Zusatzinformationen) (theoretisches Maximum: 5) für alle drei Lernbedingungen und insgesamt. Mittelwerte (Standardabweichungen in Klammern).....	238
Tabelle 28: Gesamtanzahl der generierten Argumente, der entscheidungsunterstützenden Argumente, der Gegenargumente und <i>myside bias</i> Index für alle drei Lernbedingungen (Mittelwerte, Standardabweichungen und Ergebnisse der Signifikanzprüfung).....	240
Tabelle 29: Verwendete Argumentationsmuster (rational und emotional-intuitiv) zur Lösung der Dilemmageschichte. Absolute Häufigkeiten, Gesamt und nach Geschlecht sowie Ergebnisse der Signifikanztestung des Geschlechtervergleichs.	243
Tabelle 30: Verwendete deontologische Argumente (Anzahl und Beispiel).	244
Tabelle 31: Verwendete utilitaristische Argumente (Anzahl und Beispiel).....	245

Tabelle 32: Korrelationen nach Pearson zwischen Lernvoraussetzungen und der Argumentationsqualität zum ersten und dritten Messzeitpunkt ($N =$ zwischen 71 und 72).....	247
Tabelle 33: Korrelation nach Pearson zwischen der parasozialen Interaktion (PSI) mit der Medienfigur, den affektiven und kognitiven PSI-Teilprozessen und der Argumentationsqualität zum ersten und dritten Messzeitpunkt ($N = 72$).	248
Tabelle 34: Subjektives Vorwissen zur PID und Sicherheit der Entscheidung für die beiden Gruppen Entscheidungsrevision ($n = 12$) und keine Entscheidungsrevision ($n = 60$) (Mittelwerte, Standardabweichungen in Klammer und Ergebnisse der Signifikanzprüfung).....	249

Abkürzungsverzeichnis

BÄK:	Bundesärztekammer
BGH:	Bundesgerichtshof
DIR:	Deutsches IVF-Register
DNA:	Deoxyribonucleic acid
EK:	Enquete-Kommission des 14. Deutschen Bundestages „Recht und Ethik der modernen Medizin“
ESchG:	Embryonenschutzgesetz
ESHRE:	European Society of Human Reproduction and Embryology
GG:	Grundgesetz
HLA:	Human Leukocyte Antigen
HGQN:	Human Genetics Quality Network
ICSI:	Intrazytoplasmatische Spermieninjektion
IVF:	In-vitro-Fertilisation: künstliche Befruchtung
NEK-CNE:	Nationale Ethikkommission im Bereich der Humanmedizin
NER:	Nationaler Ethikrat
PSI:	Parasoziale Interaktion (mit Medienfiguren)
PID:	Präimplantationsdiagnostik
PID-HLA:	Präimplantationsdiagnostik-Human Leukocyte Antigen: PID zur Erzeugung eines immunologisch geeigneten Spenders zum Zwecke der Heilung eines kranken Geschwisterkindes
PND:	Pränataldiagnostik
PUR:	Public understanding of research
PUS:	Public understanding of science
SKIP:	Spezies-, Kontinuums-, Identitäts- und Potentialitätsargument
SSI:	Socio-Scientific Issues
TMCKT:	Threshold Model of Content Knowledge Transfer
ZNT:	Zentrum Neue Technologien, Abteilung im Deutschen Museum

Anhang

A) Verwendete Lernumgebung: PID-HLA

Fallbeispiel/Film (Transkript)

Margot Kreidler, 36 Jahre, möchte ein „Retterkind“

Unser Sohn Max ist 5 und hat die Fanconi-Anämie. Uns war aufgefallen, dass er irgendwie immer öfter krank wurde und schwächer wirkte als früher, und da sind wir zum Arzt gegangen. Und nun wissen wir sicher, er hat diese angeborene Bluterkrankung, und seine Knochenmarkszellen bauen sich langsam immer weiter ab. Da kann man zwar Medikamente geben, aber richtig geheilt wird er dadurch nie. Aber wenn er einen ganz genau passenden Knochenmarksspender bekäme, könnte er ganz gesund werden! Aber bis sich da mal einer findet! Mein Mann oder ich passen nicht gut genug, sonst hätten wir das natürlich gleich gemacht. Jetzt haben wir uns gedacht, wo wir uns sowieso noch ein Kind wünschen, wäre es doch für alle am besten, wenn wir eine künstliche Befruchtung machen und am Embryo noch in der Schale testen lassen würden, welche Embryos am besten zu unserem Max passen würden. In Deutschland darf man so was ja nicht machen, das hat mir der Arzt auch gleich gesagt, aber in Amerika gab es das schon, da wurden Kinder mit Fanconi-Anämie mit dem Nabelschnurblut von ihrem ausgesuchten Geschwisterchen völlig geheilt! Sollte ich nicht auch nach Amerika gehen für so eine Präimplantationsdiagnostik? Würden Sie das an meiner Stelle tun?

Feedback

- 1) Kongruentes Feedback: Danke für Ihre Abstimmung! So haben die bisherigen Teilnehmer an dieser Studie abgestimmt: Mehrheit (86 %) entscheidet genauso, Minderheit (14 %) entscheidet sich anders.
- 2) Inkongruentes Feedback: Danke für Ihre Abstimmung! So haben die bisherigen Teilnehmer an dieser Studie abgestimmt: Mehrheit (86 %) entscheidet sich anders, Minderheit (14 %) entscheidet sich genauso.
- 3) Kein Feedback: Danke für Ihre Abstimmung! Die anderen Studienteilnehmer haben wie Sie auch abgestimmt, das Abstimmungsergebnis steht aber bislang leider noch nicht fest!

(siehe auch Kapitel 8.4.2; Abbildung 24, S. 191)

Zusatzinformationen

Fakten

„PID-Tourismus“

Es wird geschätzt, dass etwa 50-100 deutsche Paare pro Jahr im Ausland eine Präimplantationsdiagnostik (PID) durchführen lassen. Dabei werden vor allem Chromosomenstörungen oder bestimmte krankheitsrelevante Gene untersucht, aber auch Gewebetypisierungen zur Auswahl von „Retterkindern“ vorgenommen.

Bewertung

Besser erst gar keine Auswahlmöglichkeiten schaffen

Der Nationale Ethikrat in Deutschland lehnte in seiner Stellungnahme 2003 die Präimplantationsdiagnostik (PID) zur Erzeugung eines „Retterkindes“ grundsätzlich ab: wegen der Verwerfung von Embryonen und weil sie eine Auswahl und Auslese ermöglicht, die sich ungünstig auf das Menschenbild und das Selbstverständnis auswirken und zur Diskriminierung und Stigmatisierung behinderter Menschen führen kann.

Lebensrettung wiegt schwerer als psycho-soziale Risiken

Die Aussicht auf Lebensrettung eines kranken Kindes wog für die Hälfte der Nationalen Ethikkommission der Schweiz schwerer als gewisse psycho-soziale Risiken für das „Retterkind“. Auch die andere Hälfte, die eine PID in diesem Fall ablehnte, gelangte zu dem Schluss, dass Eltern, die eine PID zu diesem Zweck im Ausland durchführen lassen, aus nachvollziehbaren und ehrenvollen ethischen Motiven handeln und keine moralischen oder ethischen Vorwürfe verdienen.

Eltern und „Retterkinder“ können unter Druck geraten

Die PID ist in vielen Ländern zwar zum Ausschluss von schweren Erkrankungen zugelassen, aber nicht zur Erzeugung eines „Retterkindes“: Viele befürchten, dass sich Eltern bei einem solchen Angebot unter Druck gesetzt fühlen könnten, eine künstliche Befruchtung und die PID als „gute Eltern“ nutzen zu müssen. Außerdem wird das „Retterkind“ nicht um seiner selbst willen, sondern als Mittel zum Zweck erzeugt. Diese Instrumentalisierung kann bei erfolgreicher wie erfolgloser Behandlung schwer abschätzbare psycho-soziale Folgen haben (siehe Kapitel 8.3; Abbildung 22, S. 188)

Alternativen zur PID sind denkbar

Eine PID und die Erzeugung eines „Retterkindes“ würden sich erübrigen, wenn das Knochenmarkspendenetz weiter ausgebaut und ein System öffentlicher Nabelschnurblutbanken aufgebaut würde.

Gesetze

Das Deutsche Embryonenschutzgesetz verbietet die PID

Das Gesetz verbietet, eine Eizelle zu einem anderen Zweck künstlich zu befruchten, als eine Schwangerschaft der Frau herbeizuführen. Jede nicht der Erhaltung des Embryos dienende Handlung wird unter Strafe gestellt. Da die Präimplantationsdiagnostik (PID) die Verwerfung von Embryonen in Kauf nimmt, ist sie in Deutschland verboten.

In Frankreich ist PID für „Risiko-Paare“ und eingeschränkt zur Auswahl von „Retterkindern“ erlaubt

In Frankreich war die Präimplantationsdiagnostik (PID) zunächst nur für Paare erlaubt, die ein stark erhöhtes Risiko für ein Kind mit einer schweren genetischen Erkrankung haben. Die Neufassung des Bioethikgesetzes von 2003 erlaubt in Ausnahmefällen auch eine PID zur Auswahl eines „Retterkindes“.

In den USA gibt es keine einheitliche Gesetzeslage zur PID

In den USA ist die Präimplantationsdiagnostik (PID) nicht bundesrechtlich geregelt: In den meisten Bundesstaaten ist die PID in bestimmten Fällen zulässig oder nicht gesetzlich geregelt, so dass dort auch die Auswahl von Embryonen nach Gewebe-Eigenschaften oder dem Geschlecht möglich ist; in einigen Bundesstaaten ist die PID ausdrücklich verboten.

B) Instrumente

1. Untersuchungsabschnitt: Eingangsdiagnostik (t0)

Vortest zum Vorwissen über Humangenetik (Berth et al., 2004)

Bitte entscheiden Sie, ob die folgenden Aussagen über Humangenetik und genetisch bedingte Erkrankungen richtig oder falsch sind. Bitte kreuzen Sie in jeder Zeile an, was Ihrer Meinung nach zutrifft.

1. Einige genetische Erkrankungen kommen in bestimmten ethnischen Gruppen häufiger vor.	<input type="checkbox"/> richtig [*]	<input type="checkbox"/> falsch
2. Die meisten genetischen Krankheiten werden durch ein einzelnes Gen verursacht.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
3. Wenn ein genetischer Marker für eine Erkrankung bei einer Person identifiziert ist, kann die Erkrankung verhindert oder geheilt werden.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
4. Das Down-Syndrom ist eine genetisch bedingte Erkrankung, die vererbt werden kann.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
5. Viren können Ursachen für Veränderungen im menschlichen Erbgut sein.	<input type="checkbox"/> richtig [*]	<input type="checkbox"/> falsch
6. Die menschliche DNS kommt nur in den Chromosomen vor.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
7. Das X-Chromosom enthält ausschließlich Gene, die die Geschlechtsentwicklung beeinflussen.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
8. Jede Zelle des menschlichen Körpers enthält Chromosomen.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
9. Als „Allel“ wird die Gesamtheit aller Gene eines Menschen bezeichnet.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
10. Jede Zelle eines Organismus besitzt die gleiche Anzahl von Chromosomen.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
11. Beim Auftreten von genetisch bedingten Krankheiten wird immer eine Generation übersprungen.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]
12. Die Mukoviszidose (Cystische Fibrose) wird nur über die Mutter vererbt.	<input type="checkbox"/> richtig	<input type="checkbox"/> falsch [*]

Anmerkungen: * Lösung.

Subjektives Vorwissen zum Thema Gentest, PID und Bioethik

Was verbinden Sie mit dem Wort "Gentest" bzw. „gentische Untersuchung am Menschen"? Bitte nennen Sie uns einige Stichworte:

Wie viel wissen Sie über das Thema Gentest?

sehr wenig	①	②	③	④	⑤	sehr viel
------------	---	---	---	---	---	-----------

Wie viel wissen Sie über das gendiagnostische Verfahren der Präimplantationsdiagnostik (PID)?

sehr wenig	①	②	③	④	⑤	sehr viel
------------	---	---	---	---	---	-----------

Wie viel wissen Sie über das Thema Bioethik?

sehr wenig	①	②	③	④	⑤	sehr viel
------------	---	---	---	---	---	-----------

Persönlicher Bezug zum Thema Gentest

Wie schätzen Sie Ihren persönlichen Bezug zum Thema Gentest ein? (Zum Beispiel weil Sie privat oder sonst jemanden kennen, der von der Thematik betroffen ist?)

keinen	①	②	③	④	⑤	sehr stark
--------	---	---	---	---	---	------------

Skala zum Interesse am Thema Gentest (5 Items, $\alpha = .70$)

Bitte geben Sie in den folgenden Fragen Ihr Interesse am Thema Gentest an.

	Trifft nicht zu				Trifft zu
1. Ich interessiere mich für das Thema genetische Untersuchungen.	①	②	③	④	⑤
2. Ich finde die Auseinandersetzung mit ethischen Problemen, die mit der Anwendung von genetischen Untersuchungen verbunden sind, spannend.	①	②	③	④	⑤

3. Es ist für mich nicht wichtig, mich mit der ethischen Beurteilung genetischer Untersuchungen auseinanderzusetzen.	①	②	③	④	⑤
4. Ich möchte mehr darüber erfahren, inwieweit man sicher ausschließen kann, dass gentechnische Methoden zur „Menschenzüchtung“ (Herstellung ein-eiiger Zwillinge, Züchtung von Menschen mit bestimmten Eigenschaften, Auslese erkrankter Embryonen usw.) eingesetzt werden.	①	②	③	④	⑤
5. Ich möchte mehr darüber erfahren, welche Vorteile und welche Nachteile genetische Untersuchungen haben.	①	②	③	④	⑤

Skala zur Präferenz für Deliberation (5 Items, $\alpha = .82$)

	Trifft nicht zu				Trifft zu
1. Wenn ich eine Entscheidung rechtfertigen muss, denke ich vorher besonders gründlich nach.	①	②	③	④	⑤
2. Ich denke erst nach, bevor ich handle.	①	②	③	④	⑤
3. Bevor ich Entscheidungen treffe, denke ich meistens erst mal über meine Ziele nach, die ich erreichen will.	①	②	③	④	⑤
4. Bevor ich Entscheidungen treffe, denke ich meistens erst mal gründlich nach.	①	②	③	④	⑤
5. Ich nehme bei einem Problem erst mal die harten Fakten und Details auseinander, bevor ich mich entscheide.	①	②	③	④	⑤

Skala zur Präferenz für Intuition (5 Items, $\alpha = .66$)

	Trifft nicht zu				Trifft zu
1. Bei meinen Entscheidungen spielen Gefühle eine große Rolle.	①	②	③	④	⑤
2. Bei den meisten Entscheidungen ist es sinnvoll, sich ganz auf sein Gefühl zu verlassen.	①	②	③	④	⑤

3. Ich ziehe Schlussfolgerungen lieber aufgrund meiner Gefühle, Menschenkenntnis und Lebenserfahrung.	①	②	③	④	⑤
4. Ich bin ein sehr intuitiver Mensch.	①	②	③	④	⑤
5. Ich mag emotionale Situationen, Diskussionen und Filme.	①	②	③	④	⑤

Skala zur sozialen Vergleichsorientierung (11 Items, $\alpha = .84$)

	Trifft nicht zu					Trifft zu				
1. Ich vergleiche die Leistung der Menschen, die mir nahe stehen (Freund oder Freundin, Familienmitglieder, usw.) oft mit der Leistung anderer.	①	②	③	④	⑤					
2. Ich vergleiche immer meine Leistung mit der Leistung anderer.	①	②	③	④	⑤					
3. Ich bin nicht der Typ Mensch, der sich oft mit Anderen vergleicht.(-)	①	②	③	④	⑤					
4. Ich vergleiche mich immer mit anderen Menschen, wie ich gesellschaftlich ankomme (soziale Fertigkeiten, Beliebtheit).	①	②	③	④	⑤					
5. Wenn ich wissen will, wie gut ich etwas gemacht habe, dann vergleiche ich das, was ich gemacht habe mit dem, was andere gemacht haben.	①	②	③	④	⑤					
6. Oft will ich wissen, was ich im Vergleich zu Anderen im Leben schon erreicht habe.	①	②	③	④	⑤					
7. Ich unterhalte mich oft und gerne mit Menschen über unterschiedliche Meinungen und Erlebnisse.	①	②	③	④	⑤					
8. Ich frage oft, was Andere denken, die ähnliche Probleme haben wie ich.	①	②	③	④	⑤					
9. Ich will immer wissen, was Andere in ähnlichen Situationen tun würden.	①	②	③	④	⑤					
10. Ich vergleiche meine Lebenslage nie mit der von Anderen (-).	①	②	③	④	⑤					
11. Wenn ich über etwas mehr erfahren will, dann frage ich, was Andere darüber denken	①	②	③	④	⑤					

Skala zur Orientierung an anderen (5 Items, $\alpha = .60$)

	Trifft nicht zu				Trifft zu
1. Wenn ich unsicher bin, wie ich mich in einer sozialen Situation verhalten soll, suche ich im Verhalten anderer nach Hinweisen.	①	②	③	④	⑤
2. Ich würde meine Meinungen (oder meine Art, Dinge zu tun) nicht ändern, nur um jemand anderen zufrieden stellen oder um ihm zu gefallen. (-)	①	②	③	④	⑤
3. Damit ich zu Recht komme und gemocht werde, richte ich mich hauptsächlich danach, was andere von mir erwarten.	①	②	③	④	⑤
4. Es ist mir ziemlich gleichgültig, wie andere mein Verhalten in einer Situation beurteilen. (-)	①	②	③	④	⑤
5. Ich richte mein Verhalten sehr wenig danach, was andere an meiner Stelle tun würden. (-)	①	②	③	④	⑤

Skala Einstellungen zu genetischen Untersuchungen (12 Items, $\alpha = .72$)

Mit Hilfe moderner genetischer Untersuchungen kann das Vorliegen eines Risikos für eine erblich bedingte Erkrankung bei Erwachsenen, aber auch ungeborenen Kindern/Embryonen untersucht werden. Solche Tests sind aber in der Öffentlichkeit nicht unumstritten.

Mit den folgenden Fragen möchten wir erfahren, wie Sie über das Durchführen genetischer Untersuchungen denken. Ihre Antwort ist korrekt, wenn es Ihre tatsächliche und persönliche Meinung ausdrückt.

Genetische Untersuchungen ...	Trifft nicht zu				Trifft zu
1. sollten allen zugänglich sein, die etwas über ihre Krankheiten erfahren möchten.	①	②	③	④	⑤
2. sind akzeptabel, weil sie helfen können, die Kosten im Gesundheitswesen zu reduzieren.	①	②	③	④	⑤
3. sind akzeptabel, weil alle das Recht haben, über ihre Gene Bescheid zu wissen und damit das eigene Leben und die Gesundheit beeinflussen können.	①	②	③	④	⑤

4. sind akzeptabel, weil es durch neue Technologien möglich geworden ist, die zugrunde liegenden Ursachen genetischer Erkrankungen zu entdecken.	①	②	③	④	⑤
5. sollten nicht in jedem Falle durchgeführt werden.	①	②	③	④	⑤
6. sind nicht akzeptabel, weil es wichtigere Probleme im öffentlichen Gesundheitswesen gibt, die Vorrang haben.	①	②	③	④	⑤
7. sind nicht akzeptabel, weil die natürliche Ordnung respektiert werden sollte.	①	②	③	④	⑤
8. sind nicht akzeptabel, weil die Ergebnisse zur Diskriminierung von Personen mit Krankheitsgenen führen könnten.	①	②	③	④	⑤
9. sind nicht akzeptabel, weil dadurch Abtreibungen häufiger werden könnten.	①	②	③	④	⑤
10. Ich befürchte, dass die Ergebnisse von genetischen Untersuchungen für wissenschaftliche Zwecke ohne die Zustimmung der Betroffenen benutzt werden können.	①	②	③	④	⑤
11. Ich befürchte, dass Ergebnisse von genetischen Untersuchungen in die Hände von Dritten gelangen könnten.	①	②	③	④	⑤
12. Ich befürchte, dass genetische Untersuchungen zu Eugenik („Erbhygiene“) führen könnten.	①	②	③	④	⑤

Soziodemographische Daten

Geschlecht: 1 männlich 2 weiblich

Alter: _____ Jahre

Studienfach:

Semesterzahl: _____

Religionszugehörigkeit:

Würden Sie von sich sagen, dass Sie eher religiös oder eher nicht religiös sind?

nicht religiös	①	②	③	④	⑤	sehr religiös
----------------	---	---	---	---	---	---------------

Schätzen Sie bitte, wie die anderen Studienteilnehmer sich im Vergleich zu Ihnen entschieden haben (Bitte nur eine Antwort ankreuzen)

- Die anderen Studienteilnehmer haben sich genauso wie ich entschieden
 Die anderen Studienteilnehmer haben sich anders wie ich entschieden.

Wie viel Prozent der anderen Studienteilnehmer haben sich Ihrer Meinung nach ähnlich wie Sie entschieden? (Bitte nur eine Antwort ankreuzen)

- 0 % haben sich ähnlich wie ich entschieden.
 25 % haben sich ähnlich wie ich entschieden.
 50 % haben sich ähnlich wie ich entschieden.
 75 % haben sich ähnlich wie ich entschieden.
 100 % haben sich ähnlich wie ich entschieden.

Wie interessant fänden Sie es zu erfahren, wie die anderen Studienteilnehmer sich in diesem Fall entschieden haben?

überhaupt nicht ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ sehr

In den folgenden Fragen geht es um die Person Margot Kreidler, der Sie gerade zugehört haben. Bitte kreuzen Sie auf der Skala von 1 bis 5 jeweils an, wie Margot Kreidler Ihrer Meinung nach einzustufen ist.

	Trifft nicht zu				Trifft zu
1. Margot Kreidler war mir sympathisch.	①	②	③	④	⑤
2. Ich war voller Mitgefühl mit Margot Kreidler.	①	②	③	④	⑤
3. Ich habe mir zu den Dingen, die Margot Kreidler gesagt hat, keine Meinung gebildet.	①	②	③	④	⑤
4. Ich habe sehr emotional auf Margot Kreidler reagiert.	①	②	③	④	⑤
5. Ich habe mir ein Urteil über Margot Kreidler gebildet.	①	②	③	④	⑤

	Trifft nicht zu					Trifft zu				
6. Ich ließ mich von Margot Kreidlers Gefühlen anstecken.	①	②	③	④	⑤					
7. Was Margot Kreidler gesagt hat, hat mich weder angeregt noch aufgeregt.	①	②	③	④	⑤					
8. Die Stimmung von Margot Kreidler hat sich kaum auf mich übertragen.	①	②	③	④	⑤					
9. Ich habe mich damit beschäftigt, was mich mit Margot Kreidler verbindet oder was mich von ihr unterscheidet.	①	②	③	④	⑤					
10. Was Margot Kreidler gesagt hat, löste keinerlei Emotionen bei mir aus.	①	②	③	④	⑤					

Anmerkungen: 10 Items; $\alpha = .76$.

Welchen Eindruck hat die Person Margot Kreidler auf Sie gemacht?

nicht nett	①	②	③	④	⑤	nett
kalt	①	②	③	④	⑤	warm
unfreundlich	①	②	③	④	⑤	freundlich
unangenehm	①	②	③	④	⑤	angenehm
nicht liebenswert	①	②	③	④	⑤	liebenswert

Anmerkungen: 5 Items; $\alpha = .83$.

Wenn Sie alle Fragen beantwortet haben, drücken Sie bitte den orangen Weiter-Button auf dem Touchscreen der Dialogstation!

Prozessdiagnostik, Messzeitpunkt t2 (Frageblock 2a für EG1 und EG 2)

Frageblock 2a

Als ich das Abstimmungsergebnis der anderen Studienteilnehmer erfahren habe...

	Trifft nicht zu				Trifft zu
1. ...war ich beunruhigt.	①	②	③	④	⑤
2. ...war ich verstört.	①	②	③	④	⑤
3. ...war ich glücklich.	①	②	③	④	⑤
4. ...war ich erstaunt.	①	②	③	④	⑤
5. ...war ich verwundert.	①	②	③	④	⑤
6. war ich wie vor den Kopf geschlagen.	①	②	③	④	⑤
7. ...war ich erleichtert.	①	②	③	④	⑤
8. ...war ich beruhigt.	①	②	③	④	⑤
9. ...war ich erfreut.	①	②	③	④	⑤
10. ..war ich überrascht.	①	②	③	④	⑤

Wie sehr haben sich Ihre eigene Entscheidung und das Abstimmungsergebnis der anderen Studienteilnehmer unterschieden?

überhaupt nicht	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	sehr
-----------------	---	---	---	---	---	---	---	---	---	---	------

Wenn Sie alle Fragen beantwortet haben, drücken Sie bitte den orangenen Weiter-Button auf dem Touchscreen der Dialogstation

Prozessdiagnostik, Messzeitpunkt t2 (Frageblock 2 für KG)**Frageblock 2****Wie sicher sind Sie sich in Ihrer Entscheidung?**

überhaupt nicht ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ sehr

Wenn Sie alle Fragen beantwortet haben, drücken Sie bitte den orangenen Weiter-Button auf dem Touchscreen der Dialogstation!

