

Der Begriff des Tiefencues in der  
Wahrnehmungsforschung:  
Eine historisch-systematische und empirische  
Untersuchung

DISSERTATION

zur Erlangung des

Doktorgrades der Philosophie (Dr. phil.)

vorgelegt

der Philosophischen Fakultät I

*Sozialwissenschaften und historische Kulturwissenschaften*

der Martin-Luther-Universität

Halle-Wittenberg,

von

Herr Robert Pagel

geboren am 04. 10. 1980 in Halle (Saale)

**Gutachter:**

Prof. em. Dr. D. Heyer

Prof. Dr. J. Lukas

Prof. Dr. O. Peters

Tag der Verteidigung: 31. März 2016



*das Schauen aber und das Schauende ist nicht mehr Vernunft, sondern größer als Vernunft, vor der Vernunft und über der Vernunft, ebenso wie das Geschaute*

Plotin, Enneade VI 9



# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>2. Die historische Entwicklung des Konzepts der Tiefencues</b>	<b>11</b>
2.1. Wahrnehmung als Deduktion: Alhazen . . . . .	12
2.2. Natürliche Geometrie und Nervenbewegungen: René Descartes . . . . .	20
2.3. Eine neue Theorie des Sehens: George Berkeley . . . . .	31
2.4. Empfindung und Wahrnehmung: Thomas Reid . . . . .	45
2.5. Physiologisch orientierte Ansätze im 19. Jahrhundert . . . . .	55
2.5.1. Müllers Gesetz der spezifischen Sinnesenergien . . . . .	57
2.5.2. Lotzes Theorie der Lokalzeichen . . . . .	64
2.6. Die Theorie der unbewussten Schlüsse: Hermann von Helmholtz . . . . .	74
2.6.1. Biographische und allgemeine Vorbemerkungen . . . . .	75
2.6.2. „Über das Sehen“ (1855) . . . . .	83
2.6.3. Das „Handbuch der physiologischen Optik“ . . . . .	94
2.6.4. „Die Tatsachen in der Wahrnehmung“ (1878) . . . . .	119
2.6.5. Zusammenfassung . . . . .	126
2.7. Der Komputationale Ansatz: David Marr . . . . .	130
<b>3. Der Begriff des Tiefencues in der heutigen Wahrnehmungsforschung</b>	<b>139</b>
3.1. Aktuelle Charakterisierungen des Konzepts der Tiefencues . . . . .	140
3.2. Weiterführende Erörterungen . . . . .	152
3.2.1. Das Informationskonzept im Kontext der Tiefencues . . . . .	152
3.2.2. Das Konzept der ‘internen Repräsentation von Tiefe’ . . . . .	159
3.3. Tiefencues und das Problem homunkularer Redeweisen . . . . .	168
<b>4. Empirische Untersuchung</b>	<b>181</b>
4.1. Exkurs: Bildwahrnehmung . . . . .	181
4.1.1. Grundlagen der zentralprojektiven Darstellung . . . . .	185
4.1.2. Perspektivenrobustheit . . . . .	193
4.1.3. Die Dualität der Bildwahrnehmung . . . . .	198

4.2. Einleitende Bemerkungen zur empirischen Untersuchung . . . . .	204
4.3. Ziele und inhaltliche Hypothesen . . . . .	215
4.4. Methode . . . . .	218
4.5. Ergebnisse . . . . .	230
4.6. Diskussion . . . . .	242
<b>5. Fazit</b>	<b>251</b>
<b>Appendix</b>	<b>261</b>
<b>A. Daten</b>	<b>261</b>
<b>Literaturverzeichnis</b>	<b>277</b>

# 1 | Einleitung

„Vision is a strange and wonderful business.“

---

*Gibson (1979/1986, S. xiii)*

Die vorliegende Arbeit ist das Ergebnis einer Faszination für das, was wir *Sehen* nennen, verbunden mit einer gewissen, sich beim Studium aktueller wahrnehmungspsychologischer Ansätze einstellenden Ernüchterung. Theorien visueller Wahrnehmung sollen begreiflich machen, um eine oft zitierte Wendung von Koffka (1935/2001) aufzugreifen, *warum* das von uns Gesehene so aussieht, wie es sich uns im Sehen darstellt, oder kurz: warum wir sehen, was wir sehen. Trotz der immensen theoretischen und empirischen Anstrengungen in diesem Forschungsgebiet scheinen wir dennoch von einer verständlichen Antwort auf diese Frage immer noch weit entfernt zu sein, die Fülle und stetig zunehmende Fülle an empirischen Befunden scheint sich geradezu disproportional zum theoretischen Verständnis des Sehens zu verhalten.

*Dass* wir sehen ist uns allen unmittelbar zugänglich, wenn wir die Augen öffnen. Wir *benutzen* diese Fähigkeit im Alltag mit Selbstverständlichkeit und Erfolg, in der Regel ohne tiefgreifende Überlegungen über das Zustandekommen unserer visuellen Wahrnehmungen anzustellen. Wir haben alle einen unmittelbaren Zugang zum Sehen. Wir beschreiben das Gesehene in der Regel aus einer realistischen Weltsicht heraus und verhalten uns auch dem Gesehenen gegenüber so. Wir nehmen sowohl uns als auch das von uns Gesehene als Teil einer unabhängig von unserer Wahrnehmung bestehenden Welt wahr, das Gesehene erscheint uns unmittelbar als existent. Dabei handelt es sich nicht um eine ‘bewusste Entscheidung’ des Wahrnehmenden, das Wahrgenommene als unabhängig von der Wahrnehmung existierend *anzusehen*, sondern es ist uns in dieser Erscheinungsweise schlicht *gegeben*.<sup>1</sup> Im Alltag fällt es uns auch nicht sonderlich schwer, uns mit anderen über das von uns Gesehene auszutauschen. Wir sprechen davon, dass wir *Dinge* sehen, die sich außerhalb von uns befinden, wir können das Gesehene in differenzierender und zergliedernder Weise beschreiben, indem wir uns darüber äußern, welche Dinge wir sehen, welche Form und Farbe diesen zukommt, ob

---

<sup>1</sup>Husserl (1913/2009) nennt dies im Rahmen seiner Phänomenologie auch die *natürliche Einstellung*.

und in welcher Weise sich diese bewegen und in welchen räumlichen Relationen diese zueinander stehen. Kurz: Im Alltag scheint uns das Sehen keinerlei theoretische oder praktische Schwierigkeiten zu bereiten, wenngleich festzuhalten ist, dass wir im Alltag nicht über das Sehen selbst, sondern stets nur über Gesehenes sprechen.

Beginnt man allerdings über dieses für uns so unmittelbare, eindringliche und für unser alltägliches Leben zentrale Phänomen nachzudenken, wird man relativ schnell auf Schwierigkeiten stoßen, die nicht zuletzt darauf zurückzuführen sind, dass wir im Alltag den Begriff des Sehens auf unzählige verschiedene Weisen verwenden und so zunächst gar nicht klar scheint, *worüber* man eigentlich nachdenken möchte. Eine der größten mit einer Behandlung dieses Phänomens verbundenen Schwierigkeiten scheint es zu sein, überhaupt angemessen erscheinende theoretische Vorstellungen zu entwickeln, um sinnvolle Fragen formulieren und so auch den Gegenstandsbereich selbst vorläufig umreißen zu können. Man fühlt sich unmittelbar an die berühmte Passage der „Bekenntnisse“ des Augustinus erinnert, der im Rahmen seiner Behandlung der Frage, was Zeit sei, äußert: „Wenn mich niemand darnach fragt, weiß ich es, wenn ich es aber einem, der mich fragt, erklären sollte, weiß ich es nicht“ (Augustinus, 1989, S. 278).

Der allgemeine konzeptuelle Rahmen, in dem wir über das Sehen nachdenken, bedingt die Art der Fragen, die wir zu beantworten suchen; so liegt etwa der Frage nach den ‘neuronalen Mechanismen’ der Wahrnehmung ein ganz bestimmter konzeptueller Rahmen zugrunde. Im Kontext der aktuellen wahrnehmungspsychologischen Forschung lässt sich jedoch leider zu selten eine ernsthafte Beschäftigung mit den zugrundeliegenden Konzepten auffinden. Üblicherweise wird innerhalb eines bestehenden konzeptuellen Rahmens versucht, anhand empirischer Befunde und deren Interpretation theoretische Ansätze zur Erklärung der bestehenden Datenlage zu entwickeln oder Hinweise darauf zu finden, ob sich ein bestimmter bestehender Ansatz (oder ein bestimmtes Modell) vor dem Hintergrund der vorliegenden Daten bewährt hat oder in Zweifel gezogen werden sollte. So nachvollziehbar es ist, im ‘wissenschaftlichen Alltagsgeschäft’ empirische Befunde zu spezifischen Teilgebieten zusammenzutragen, vermisst man doch gleichzeitig eine Auseinandersetzung mit den zur Erklärung herangezogenen Grundbegriffen, welche die empirischen Ergebnisse erst in einen gewissen ‘Sinnzusammenhang’ bringen sollen.<sup>2</sup> Ohne eine theoretische Auseinandersetzung mit den Grundannahmen der immer wieder in wahrnehmungstheoretischen Untersuchungen zur Erklärung herangezogenen zentralen Begriffe oder Konzepte besteht die Gefahr, dass ein solcher Sinnzusammenhang verloren geht und in der Tat ‘nur’ Daten zusammengetragen werden. Insbesondere bei wahrnehmungspsychologischen Fachartikeln steht man als Leser vor dem Problem, dass bei der Interpretation empirischer Befunde und deren theoretischer

---

<sup>2</sup>Baker und Hacker (1982, S. 227) formulieren dies in wesentlich schärferem Ton: „The endemic sin of the experimental psychologist ... is to neglect the conceptual investigations which are preconditions for fruitful, intelligible experiments.“

Erklärung Konzepte herangezogen werden, deren Bedeutung bei genauerer Betrachtung nicht klar wird. Der Gegenstand der vorliegenden Arbeit soll ein solcher, in der visuellen Wahrnehmungspsychologie weit verbreiteter, aber weithin wenig explizierter Grundbegriff sein. Die Motivation zu dieser Arbeit ist nicht zuletzt darin begründet, dass dem Autor im Rahmen seiner eigenen wahrnehmungspsychologischen Forschung ein solcher Sinnzusammenhang abhandeln zu kommen drohte, da in entsprechenden wahrnehmungstheoretischen Ansätzen immer wieder auf einen bestimmten Terminus zurückgegriffen wurde, dessen Bedeutung dem Autor jedoch nicht verständlich wurde.

Beschäftigt man sich näher mit wahrnehmungspsychologischer Fachliteratur, so wird man relativ schnell und auffallend häufig auf den Terminus „*cue*“ stoßen, der im deutschsprachigen Raum, falls er nicht schlicht übernommen wird, meist mit *Hinweisreiz* übersetzt wird.<sup>3</sup> In über der Hälfte aller der ca. 350 im Jahre 2014 im *Journal of Vision* erschienen Artikel taucht dieser Terminus auf, in den Zeitschriften *Vision Research* und *Perception* wird er in etwa der Hälfte bzw. einem Drittel aller 2014 erschienen Artikel im *Abstract* verwendet. Dabei erstrecken sich die Themenbereiche, in deren Zusammenhang von *cues* gesprochen wird, u.a. auf die Aufmerksamkeitsforschung (etwa Zhao & Heinke, 2014), die Farbforschung (etwa Chen & Cave, 2014), die Forschung zur Wahrnehmung von Figur und Grund (etwa Ramenahalli, Mihalas & Niebur, 2014), die Forschung zur Wahrnehmung von Oberflächeneigenschaften (Schlüter & Faul, 2014) und viele weitere – so behandelt etwa eine Studie (Fisher, Hahn, DeBruine & Jones, 2014) auch den Einfluss von *cues* für Adipositas auf die wahrgenommene Attraktivität und den Gesundheitszustand von Gesichtern. Die Feststellung scheint daher nicht unangemessen, dass es sich hierbei um ein in der Wahrnehmungspsychologie weithin verwendetes und daher auch zentrales Konzept handelt. Diese Beobachtung steht jedoch in einem deutlichen Kontrast zu der Tatsache, dass nur auffallend wenige bis keine Publikationen im wahrnehmungspsychologischen Kontext explizit dieses *Konzept* zum Gegenstand haben. In der Regel wird es in der Literatur schlicht verwendet, meist mit einem allgemeinen Hinweis, dass es sich bei *cues* um bestimmte ‘Informationen‘ handle, die vom visuellen System ‘genutzt’ werden könnten.

Einen zentralen Stellenwert nimmt das *cue*-Konzept in wahrnehmungspsychologischen Theorien der Tiefen- oder Raumwahrnehmung ein. Beschäftigt man sich mit der Entwicklungsgeschichte wahrnehmungstheoretischer Ansätze, so wird deutlich, dass der Ursprung des heute sogenannten *cue*-Ansatzes in diesem Themenbereich zu verorten ist. Gegenstand solcher Theorien ist die Frage, warum sich das von uns Gesehene in der

---

<sup>3</sup>Die etymologische Herkunft des Wortes „*cue*“ ist in der Sprache der Theaterpraxis des 16. Jahrhunderts verwurzelt (Skeat, 1893). Es bezeichnete dort eine festgelegte Stelle in einer Rede eines Schauspielers, die als Stichwort für den Einsatz eines anderen Schauspielers diente. Es wird vermutet, dass *cue* für das lateinische Wort „*quando*“ (wann) steht, eine Vermutung, die unter anderem dadurch als bekräftigt angesehen wird, dass ein *cue* im Theaterkontext des 16. Jahrhunderts auch durch den Buchstaben „*Q*“ abgekürzt wurde.

Regel in räumlicher Ausdehnung und als unterschiedlich weit von uns oder als unterschiedlich weit voneinander entfernt präsentiert, oder kurz: warum wir, wie wir heute sagen, ‘räumlich’ wahrnehmen. Theoretische Vorstellungen, die diesen Aspekt der visuellen Wahrnehmung verständlich machen wollen, lassen sich bis in die Ionische Periode der Antike (etwa bei Demokrit) zurückverfolgen (Howard, 2012). Insbesondere nach der von Johannes Kepler im 17. Jahrhundert vorgebrachten Theorie, dass das ins Auge einfallende Licht durch die Linse des Auges als invertiertes, verzerrtes und auf dem Kopf stehendes ‘Bild’ (*pictura*) der äußeren Gegenstände auf die Netzhaut projiziert werde, nahm das theoretische Interesse am Sehen, im Speziellen an der Räumlichkeit des Sehens, enorm zu, da Keplers Vorstellung auf eine fundamentale Schwierigkeit hinzuweisen schien.<sup>4</sup> Eine übliche Formulierung dieses vermeinten Problems lautet: Wie ist es möglich, dass wir räumlich sehen, wenn doch alles, was unseren Augen als Organen des visuellen Sinnes zur Verfügung steht, ein ‘zweidimensionales Bild’ ist? Diese oder dem Sinn nach ähnliche Formulierungen findet man auch heutzutage in der wahrnehmungspsychologischen Fachliteratur, etwa in beinahe sämtlichen Standardlehrbüchern der Wahrnehmungspsychologie:

The accomplishment involved in seeing objects in depth is quite amazing considering that the basic information available to the nervous system is just a flat image on the retina. (Coren, Ward & Ennis, 1999, S. 274)

All the information we have regarding the visual world comes from stimulation of our two retinae. The retina ... is equivalent to a flat, two-dimensional layer of tissue. How does stimulation of a two-dimensional retinal surface get translated by the brain into perceptions of depth? (Levine, 2000, S. 297)

Zieht man neben üblichen Lehrbüchern der Wahrnehmungspsychologie weitere Fachliteratur heran, so lässt sich feststellen, dass dies innerhalb der wahrnehmungspsychologischen Forschung (von wenigen Ausnahmen abgesehen) bis heute als *das* zentrale Problem der Wahrnehmung von Tiefe oder Räumlichkeit angesehen wird.<sup>5</sup>

Räumliches Sehen sei aufgrund dieses Umstandes nicht ‘unmittelbar’, sondern nur mit Hilfe sogenannter *Tiefencues* oder *cues* für Tiefe möglich, denen somit eine fundamentale Rolle bei der Tiefenwahrnehmung zugeschrieben wird. Tiefencues, so die übliche Charakterisierung in der aktuellen Literatur, seien ‘Informationen über Tiefe’, die vom visuellen System ‘genutzt’ werden könnten, um einen Tiefeneindruck zu

<sup>4</sup>Für eine ausführliche Darstellung der Keplerschen Wahrnehmungstheorie und deren historische Vorläufer siehe Lindberg (1976).

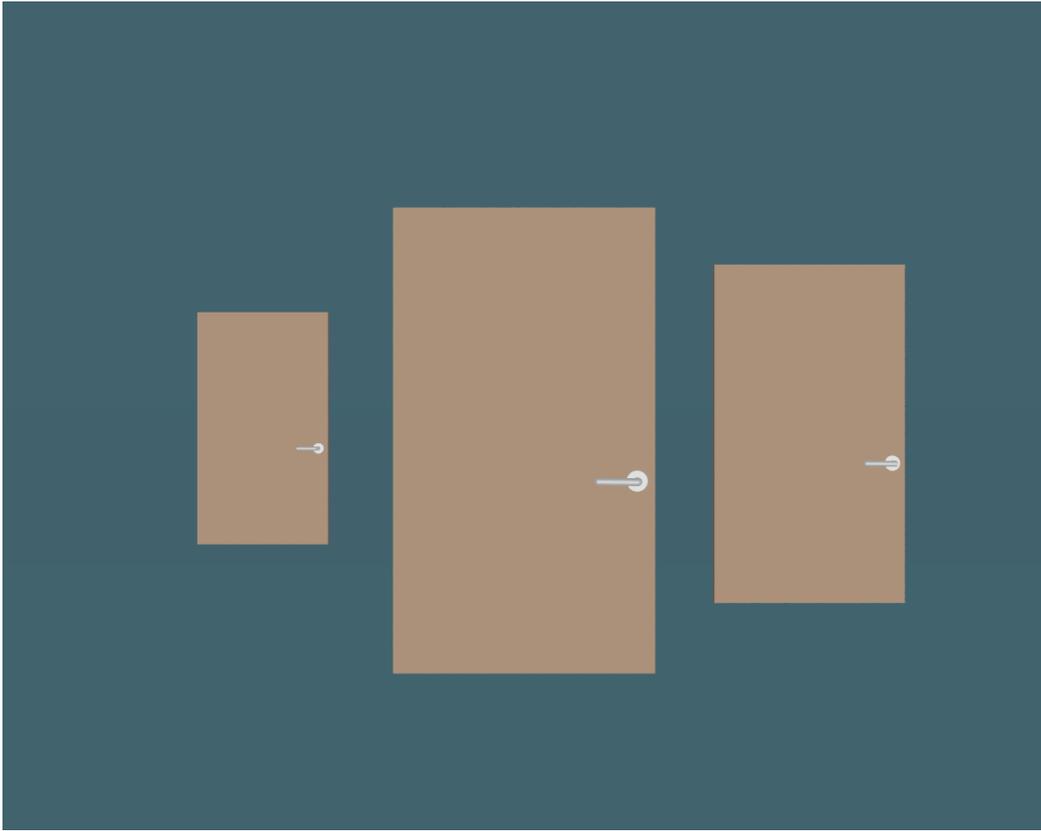
<sup>5</sup>Siehe etwa Bruce, Green und Georgeson (2003); Burge, Peterson und Palmer (2005); Foley und Matlin (2010); Goldstein (1999, 2010); Hershenson (1999); Hillis, Watt, Landy und Banks (2004); Howard (2002, 2012); Johnston, Cumming und Parker (1993); Marshall, Burbeck, Ariely, Rolland und Martin (1996); Mather (2009); Mon-Williams, Tresilian und Roberts (2000); O’Shea, Blackburn und Ono (1994); Palmer (1999); Seydell, Knill und Trommershäuser (2010); Troscianko, Montagnon, Le Clerk, Malbert und Chanteau (1991); Wolfe, Kluender und Levi (2009) oder Yantis (2014).

‘erzeugen’. Zur Veranschaulichung dieses Konzeptes wird nicht selten auf Bilder zurückgegriffen, man betrachte etwa die beiden in Abbildung 1.1 auf der nächsten Seite dargestellten Bilder und achte auf die mit diesen einhergehenden Wahrnehmungseindrücke.

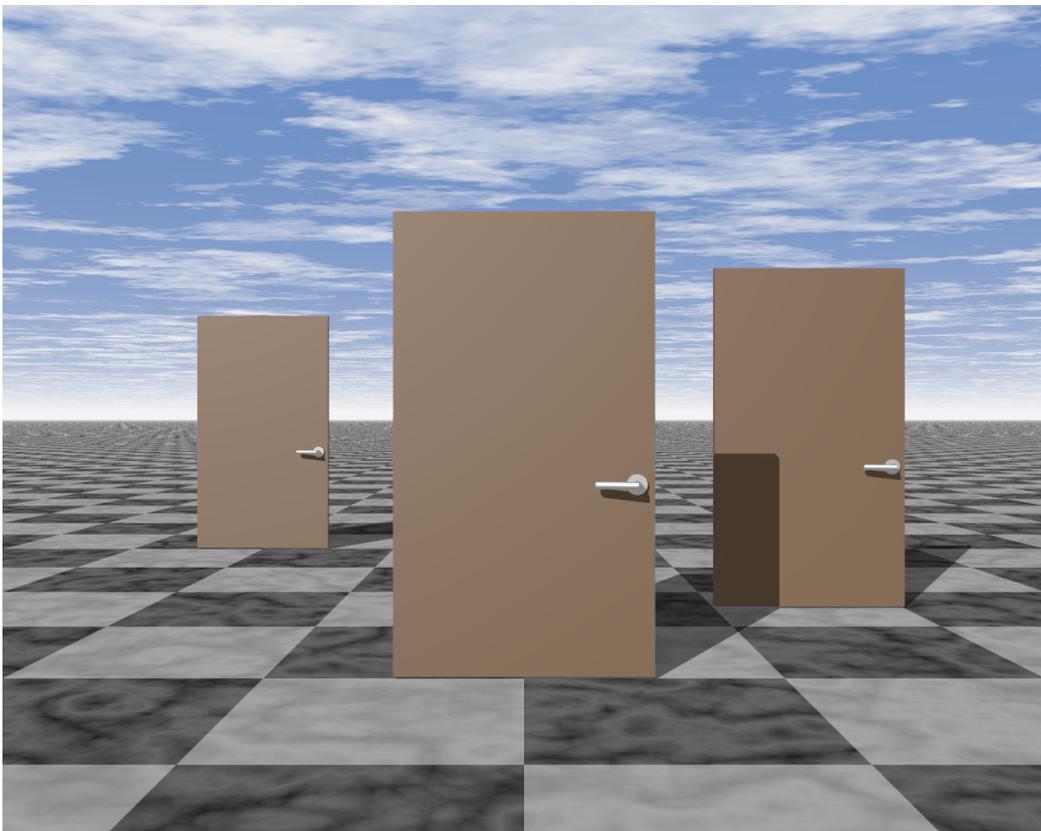
Die Betrachtung von Abbildung 1.1 (a) ist in der Regel mit keinem eindringlichen Eindruck von Tiefe oder Räumlichkeit verbunden. Die dargestellte Szene und deren Elemente wirken ‘flächhaft’, gleichsam so, als ob man zweidimensionale geometrische Formen ausgeschnitten und auf einer Ebene angebracht hätte. Wäre man aufgefordert, die räumlichen Relationen zwischen den dargestellten Elementen zu beschreiben, etwa welches dieser am ‘weitesten entfernt’ scheint, so fallen diesbezügliche Urteile schwer. Es könnte sich um drei etwa gleich große Gegenstände in unterschiedlicher Entfernung, oder um unterschiedlich große Gegenstände in identischer Entfernung handeln. Die Betrachtung von Abbildung 1.1 (b) ist hingegen mit einem qualitativ unterschiedlichen Wahrnehmungseindruck verbunden. Die dargestellte Szene wirkt ‘räumlich’, es scheint gleichsam so, als würde sie sich in die Tiefe erstrecken. Gleichzeitig fällt es hier wesentlich einfacher, Urteile über die räumlichen Relationen zwischen den dargestellten Elementen abzugeben. So wird man hier eher geneigt sein zu urteilen, dass es sich um etwa gleich große Elemente in unterschiedlicher Entfernung handelt und dass bei genauer Betrachtung (etwa durch Auszählen der ‘Bodenkacheln’) der Abstand zwischen der Tür in der Mitte des Bildes und der sich rechts davon befindenden geringer ist als der Abstand zwischen der Tür rechts im Bild und der Tür links im Bild.

Der unterschiedliche mit diesen Bildern verbundene Tiefeneindruck wird im Rahmen des *cue*-Ansatzes nun dadurch erklärt, dass Abbildung 1.1 (a) *keine* Tiefencues aufweist, Abbildung 1.1 (b) hingegen mehrere solcher Tiefencues. So sei etwa in Abbildung 1.1 (b) der von der Tür in der Mitte auf die Tür rechts im Bild geworfene Schatten ein solcher Tiefencue, den das visuelle System nun gleichsam als ‘Hinweis’ oder ‘Zeichen’ dafür ‘interpretiere’, dass sich die mittlere Tür *vor* der Tür rechts im Bild befinde. Ebenso seien bestimmte Elemente der linearperspektivischen Darstellung, wie die relative Höhe der dargestellten Objekte im Bezug auf den Horizont, als Tiefencues anzusehen, die als ‘Informationen über Tiefe’ genutzt würden und dazu führten, dass uns die in Abbildung 1.1 (b) dargestellte Szene als räumlich ausgedehnt erscheint und Urteile über die räumlichen Relationen der dargestellten Elemente überhaupt möglich oder ‘genauer’ werden.

In der Regel erschöpft sich die Darstellung des *cue*-Ansatzes im Bereich der Tiefenwahrnehmung in einer nach bestimmten Merkmalen geordneten Auflistung dieser wohl jedem an visueller Wahrnehmung Interessierten, allseits bekannten Tiefencues wie etwa *Bewegungsparallaxe*, *Verdeckung*, *Schattenwurf*, *binokulare Disparität*, *Akkommodation des Auges* oder *Konvergenz*. Die aufgelisteten Elemente scheinen jedoch selbst



(a)



(b)

Abbildung 1.1. Beispielhafte Veranschaulichung der sogenannten Tiefencues. Bild (a) enthalte keine Tiefencues, Bild (b) hingegen mehrere Tiefencues.

bei flüchtigem Blick verschiedenen Beschreibungsebenen zu entstammen. Manche der Tiefencues scheinen sich auf *physiologische Vorgänge* zu beziehen (etwa die sich verändernde Linsenkrümmung im Falle der Akkommodation), andere scheinen eher dem Wahrgenommenen selbst entnommen zu sein (etwa die Verdeckung), bei vielen Tiefencues bleibt man eher ratlos, welcher Beschreibungsebene diese zugehörig sein sollen. Aus welchen Gründen die aufgelisteten Elemente unter der gemeinsamen Kategorie der Tiefencues zusammengefasst werden, bleibt, bis auf die Bemerkung, dass sie allesamt ‘Hinweise für Tiefe’ oder ‘Informationen für Tiefe’ seien, unklar. Eine klare, über eine bloße Aufzählung und Darstellung der einzelnen Tiefencues hinausgehende Beschreibung des Konzeptes sowie dessen Stellenwert innerhalb der theoretischen Erklärung der Tiefenwahrnehmung lässt sich in der Regel nicht finden. Die übliche Charakterisierung der Tiefencues ist allerdings ohne weitere Klärung der in dieser verwendeten Begriffe nicht verständlich. In welchem Sinne wird hier von Informationen oder Hinweisen gesprochen? Was kann es heißen, dass das visuelle System diese Informationen ‘nutzt’? Und welche Beziehung soll zwischen dem ‘Nutzen’ dieser Informationen durch das visuelle System und *meiner Wahrnehmung* bestehen? In der Regel finden diese für das Verständnis dieses Konzeptes doch zentralen Punkte keine weitere Beachtung in den Darstellungen. Die vorliegende Arbeit stellt den Versuch dar, diesem Konzept und dessen Funktion in wahrnehmungstheoretischen Ansätzen ausführlicher nachzugehen und kritisch zu beleuchten.

## Ziele und Gliederung der vorliegenden Arbeit

Diese Aufgabe soll in mehreren, aufeinander aufbauenden Schritten verfolgt werden. Zunächst soll die Genese und der Gebrauch dieses Konzeptes anhand ausgewählter historischer theoretischer Ansätze aufgezeigt werden, welche die Wahrnehmung von Tiefe explizit zum Gegenstand haben und sich dieses Konzeptes zu bedienen scheinen (Kapitel 2). Die meisten in aktuellen Darstellungen aufgelisteten Tiefencues haben eine lange Geschichte, sie wurden bereits in wahrnehmungstheoretischen Ansätzen (etwa bei Descartes) verwendet, bevor es eine akademische Disziplin der Psychologie gegeben hat. Die Hoffnung ist, dass sich durch einen Blick auf diese historischen Verwendungsweisen und deren Veränderungen im Laufe fortschreitender Theoriebildung ein näheres Verständnis dieses Konzeptes erlangen lässt. Diese historischen Analysen weisen einen eher *suchenden* Charakter auf.

Anschließend soll vor dem Hintergrund der in diesen historischen Betrachtungen herausgearbeiteten Erkenntnisse die Verwendung und Funktion des *cue*-Konzeptes in der aktuellen wahrnehmungspsychologischen Literatur aus eher systematischer Perspektive analysiert werden (Kapitel 3). Es soll insbesondere untersucht werden, ob sich der modernen Formulierung des *cue*-Konzeptes durch weitergehende Analyse der in dieser

verwendeten Konzepte (etwa das Informationskonzept) eine kohärente Bedeutung geben lässt. In diesem Zusammenhang sollen auch die impliziten Voraussetzungen, die diesem Konzept zugrunde liegen, herausgearbeitet, sowie der Frage nachgegangen werden, ob dieses Konzept zu einem Verständnis der Tiefenwahrnehmung beitragen kann.

Im letzten Teil dieser Arbeit (Kapitel 4) soll anhand einer eigens durchgeführten empirischen Untersuchung zum Bereich der Bildwahrnehmung *exemplifiziert* werden, auf welche Weise das Konzept der Tiefencues üblicherweise herangezogen wird, um im Bereich der Tiefenwahrnehmung Hypothesen zu formulieren und experimentelle Resultate zu interpretieren. Anhand dieses konkreten Beispiels sollen in Bezug auf die vorangegangenen Erörterungen im Speziellen die Fragen behandelt werden, welchen Erklärungswert das *cue*-Konzept in der Interpretation empirischer Ergebnisse aufweisen kann und warum diesem Konzept eine so zentrale Rolle innerhalb der wahrnehmungspsychologischen Theoriebildung zugeschrieben wird.

Der grundsätzliche Ansatz dieser Arbeit ist ein für die aktuelle Wahrnehmungspsychologie, die sich selbst wesentlich als eine empirische Wissenschaft versteht, durchaus ungewöhnlicher; es wird nicht versucht, anhand empirischer Ergebnisse Hinweise auf die Brauchbarkeit einer oder mehrerer bestehender theoretischer Vorstellungen oder Modelle zu erlangen. Gegenstand dieser Arbeit ist im Wesentlichen eine theoretische oder konzeptuelle Analyse, sie betrifft den allgemeinen konzeptuellen Rahmen, in dem wahrnehmungstheoretische Vorstellungen, die sich des *cue*-Ansatzes bedienen, formuliert und in dem zugehörige empirische Untersuchungen durchgeführt und interpretiert werden. Wahrnehmungspsychologie ist ein wahres Minenfeld potentieller begrifflicher Verwirrungen (Hacker, 1991) und bereits Merleau-Ponty (1934/2003, S. 13) wies darauf hin, dass diese „mit philosophischen Vorannahmen belastet [sei], die sich schon bei den augenscheinlich unschuldigsten Begriffen einschleichen“. Es scheint daher durchaus geboten, im Rahmen einer theoretischen Auseinandersetzung mit dem Konzept der Tiefencues, das eine *zentrale* Rolle in der aktuell üblichen Erklärung der Tiefenwahrnehmung einnimmt, der generellen Frage nachzugehen, ob dieses ein kohärentes und sinnvolles Konzept ist.

Diese Aufgabe ist mit einigen Schwierigkeiten verbunden, wie die historischen und systematischen Analysen zeigen werden. So werden wir im Rahmen des Versuchs der Explikation und Analyse des *cue*-Konzeptes immer wieder auf notorische theoretische Probleme und problematische Konzepte stoßen, die heutzutage meist unter dem Stichwort einer *Philosophie des Geistes* behandelt werden. Gleichzeitig wird sich häufig das Problem stellen, bestimmten, in wahrnehmungstheoretischen Vorstellungen verwendeten Wendungen oder Charakterisierungen eine klare und sinnvolle Bedeutung zu verleihen, weil sie, im Fall historischer wahrnehmungstheoretischer Ansätze, in den spezifischen philosophischen und sprachlichen Kontext der Zeit eingebettet sind, oder

weil sie aufgrund der mit dem Thema Wahrnehmung verbundenen theoretischen Probleme keine klare Bedeutung haben oder möglicherweise aus prinzipiellen Gründen eine solche auch gar nicht haben können. Die vorliegende Arbeit hat in keiner Weise den Anspruch, diese Probleme ‘lösen’ zu wollen, sondern sie verfolgt das Ziel, dem Leser zu zeigen, dass das *cue*-Konzept aufgrund der mit diesem verbundenen theoretischen Probleme für ein tieferes Verständnis des Phänomens visueller Wahrnehmung letztlich nicht geeignet erscheint.

Diese Arbeit berührt sowohl genuin philosophische, als auch wissenschaftshistorische Probleme, ist jedoch nicht in erster Linie an Vertreter dieser Disziplinen gerichtet. Dem Autor – weder akademischer Philosoph, noch Wissenschaftshistoriker, sondern Wahrnehmungspsychologe – möge daher nachgesehen werden, wenn der Umgang mit diesen Problemen nicht den gängigen Standards der akademischen Philosophie bzw. der Wissenschaftsgeschichte entsprechen sollte.



## 2 | Die historische Entwicklung des Konzepts der Tiefencues

„Auch ist wenig Aussicht, dass zum Ziele der Erkenntniss kommen wird, wer nicht mit dem Anfang anfängt.“

---

*Helmholtz (1896/2002a, S. 343)*

Gegenstand dieses Kapitels sollen ausgewählte wahrnehmungstheoretische Ansätze sein, in deren Rahmen dem Konzept der Tiefencues ähnliche Vorstellungen zur Erklärung der Tiefenwahrnehmung herangezogen werden. Diese Darstellung soll keine erschöpfende sein, es wurden selektiv Theorien ausgewählt, die neuartige Impulse in der allgemeinen Theoriebildung setzen konnten und auf den Tiefencues ähnliche Vorstellungen zurückgreifen, um so die historische Entwicklung dieses Konzeptes nachzuzeichnen. Da der Begriff des *cues* erst zu Beginn des letzten Jahrhunderts zu einem festen Bestandteil wahrnehmungspsychologischer Theorien wurde (Harper & Boring, 1948), wird in den hier darzustellenden historischen Theorien dieser Begriff selbst nicht auftauchen, es wird sich jedoch zeigen, dass die in der aktuellen wahrnehmungspsychologischen Literatur als Tiefencues bezeichneten Elemente zentrale Bestandteile dieser historischen Theorien sind.

Die Darstellung eines bestimmten theoretischen Ansatzes soll im Allgemeinen ebenfalls nicht den Anspruch erheben, eine umfassende und erschöpfende zu sein, sondern es soll sich vornehmlich auf die für unsere Diskussion relevanten Punkte beschränkt werden; in der Hoffnung, durch diese Beschränkung kein verzerrtes Bild dieser Theorien zu erstellen. Da diese Theorien enorm vom Sprachgebrauch und in der Regel auch von bestimmten philosophischen Diskussionen der jeweiligen Zeit geprägt sind, ist eine solche Darstellung selbstverständlich nicht ohne Schwierigkeiten. Wir wollen einerseits nicht einer gelegentlich anzutreffenden Praxis folgen, die in diesen Theorien auftauchenden Begriffe ihres historischen Kontextes zu berauben, indem sie in moderne theoretische Konzepte ‘übersetzt’ werden, andererseits soll sich die Darstellung auch nicht in einer bloßen Aneinanderreihung unkommentierter Zitate erschöpfen. Es soll jeweils versucht werden, eine bestimmte Lesart dieser Theorien vorzuschlagen. Um dem Leser die Mög-

lichkeit zu geben, die Stimmigkeit dieser Lesart zu überprüfen, wird in einem hoffentlich nicht als unangemessen zu bezeichnendem Maße diese durch Zitate unterstützt werden.

## 2.1. Wahrnehmung als Deduktion: Alhazen

Der persische Mathematiker, Optiker und Astronom *Abū Alī al-Ḥasan ibn al-Haiṭham* (965 bis etwa 1039), latinisiert Alhazen, kann laut Lindberg (1967) als einer der bedeutendsten und wirkungsmächtigsten Wahrnehmungstheoretiker im Zeitraum von der Antike bis ins siebzehnte Jahrhundert angesehen werden.<sup>1</sup> Das für unsere Darstellung relevante und vermutlich 1021 vollendete Werk *Kitāb-al-Manāzīr* (übersetzt etwa „Schatz der Optik“ oder „Buch des Sehens“) enthält ausführliche Erörterungen zu Themen der Optik, der Anatomie und Physiologie des Auges, die in wegweisende theoretische Vorstellungen visueller Wahrnehmung münden. Eine lateinische Übersetzung dieses Werkes („*De aspectibus*“) erschien erstmals 1572.

Alhazen gibt sich in diesem Werk als ein entschiedener Gegner der u.a. von Plato, Galen und Ptolemäus vertretenen sogenannten *Extramissions-* oder *Sendetheorie* visueller Wahrnehmung zu erkennen. Grundannahme dieses Ansatzes ist, dass sich Sehen vollzieht, indem von den Augen sogenannte *Sehstrahlen*, gleichsam wie Fühler ‘ausgesendet’ werden, welche in direkten Kontakt mit den Objekten der Umwelt treten und diese gleichsam ‘abtasten’. Im Gegensatz dazu gehen sogenannte *Intromissions-* oder *Empfangstheorien* des Sehens (zu denen sich Alhazens Theorie zählen lässt) davon aus, dass die Augen nicht etwas aussenden, sondern etwas von den Objekten ausgesendet und von den Augen empfangen wird.<sup>2</sup> In frühen Formulierungen solcher Intromissionstheorien wurde in der Regel davon ausgegangen, dass die Objekte kleine ‘Bilderchen’ (εἰδωλα) von sich selbst aussenden, diese ins Auge gelangen und sich dem Sehsinn zur Wahrnehmung ‘darbieten’. Alhazen lehnt auch diesen Ansatz ab und führt einen für seine Zeit neuartigen Ansatz ein, nämlich dass Sehen nur über eine Einwirkung von *Licht* auf die Augen geschehe: „it is a property of light to affect sight, whereas it is in the nature of sight to be affected by light“ (Alhazen, 2001, S. 355).<sup>3</sup> Damit wir Objekte sehen können, müssten diese entweder selbst Licht aussenden oder Licht reflektieren und dieses Licht auf unsere Augen treffen. Insbesondere geht Alhazen davon aus, dass von *jedem* Punkt der Oberfläche eines Objektes Licht geradlinig in sämtliche Richtungen ausgestrahlt wird. Trifft dieses von jedem Punkt eines Objektes ausgesen-

<sup>1</sup>Für eine Bewertung der Rolle von Alhazens Wirkung auf die Entwicklung von Wahrnehmungstheorien in der westlichen Welt siehe Lindberg (1967).

<sup>2</sup>Für eine äußerst lohnenswerte Darstellung der historischen Entwicklung der Sende- und Empfangstheorien visueller Wahrnehmung siehe Lindberg (1976).

<sup>3</sup>Aufgrund einer mangelnden deutschen Übersetzung wird im Folgenden die von der *American Philosophical Society* herausgegebene englische Übersetzung des *Kitāb-al-Manāzīr* als Quelle verwendet.

detete Licht auf die Augen, spricht Alhazen auch davon, dass dieses „forms of light and color“ empfangen, was wir im Folgenden mit *Formen aus Helligkeit und Farbe* übersetzen wollen. Was anschließend mit diesen *Formen* passiert, wird auf sehr komplexe und an vielen Stellen mehrdeutige Weise beschrieben (siehe Lindberg, 1976, Kap. 4), ist jedoch für unsere Diskussion nicht von zentraler Bedeutung, weshalb hier auf eine ausführliche Diskussion verzichtet werden soll.<sup>4</sup> Grob vereinfacht handelt es sich um einen mehrstufigen Prozess, der dazu führe, dass diese über verschiedene anatomische Strukturen des Auges „quasi-optischen Regeln“ folgend (Lindberg, 1976, S. 87) zu einer physiologischen Struktur, die auch als *final sensor (ultimum sentiens)* bezeichnet wird, weitergeleitet werden oder sich zu diesem hin ‘ausbreiten’, wo *qua* einer dieser innewohnenden ‘wahrnehmenden Kraft’ das Objekt wahrgenommen werde.<sup>5</sup>

Von Relevanz für die weitere Diskussion ist die immer wieder im *Kitāb-al-Manāzīr* auftauchende Vorstellung Alhazens, dass sich das, was wir sehen, nämlich Objekte in bestimmter Größe, Farbe, Form und räumlicher Anordnung, nicht allein durch den Effekt des Lichts auf unsere Augen erklären lasse. Wenn Licht in unsere Augen trifft, so werde es bereits im Auge gewissermaßen ‘wahrgenommen’, nämlich als Punkte bestimmter Helligkeit, Farbe und Anordnung (die *forms of light and color*), um jedoch zu einer Wahrnehmung im eigentlichen Sinne – einer Wahrnehmung von *Objekten* mit verschiedenen Eigenschaften – zu gelangen, müssten diese bzw. deren Effekte vom *final sensor* gewissermaßen ‘interpretiert’ werden. Diese Vorstellung wird sich (in unterschiedlichen Varianten) als ein zentrales Element der nachfolgenden Theorien herausstellen – das, was von außen auf die Sinne einwirkt, genüge *per se* nicht, um im üblichen Sinne zu sehen, sondern es müsse zusätzlich ‘interpretiert’, ‘elaboriert’ oder ‘ausgewertet’ werden.

---

<sup>4</sup>Für eine kompakte Darstellung siehe etwa Eastwood (1986) oder Smith (2005).

<sup>5</sup>An (mindestens) einer Stelle beschreibt Alhazen das, was zum *final sensor* gelangt nicht als Form, sondern als *Effekt* dieser *Formen* auf bestimmte Nerven:

We might claim that the forms arriving at the eye do not reach the common nerve, but that the sensible effect [of those forms] will instead extend from the eye to the common nerve, just as the sense of pain and the sense of touch [extend through the nerves], and that at this time the final sensor perceives that sensible effect. (Alhazen, 2001, S. 378)

An einer späteren Stelle spricht Alhazen davon, dass sowohl die *Formen* als auch deren Effekt, der auch als *sensation (sensus)* bezeichnet wird, gemeinsam zum *final sensor* gelangen:

Therefore, the forms reach the vitreous humor [der Glaskörper des Auges, von mir] arranged as they actually are on the surface of the visible subject, and this body receives them and senses them. ... The resulting sensation, as well as the resulting form, will then extend through this body until the sensations and form reach the final sensor. But the passage of the sensation and the passage of the form through the body of the vitreous and through the sensitive body that fills the hollow of the optic nerve to the final sensor will be like the passage of the sensation of touch or the sensation of pain to the final sensor. (Alhazen, 2001, S. 422)

Im dritten Kapitel des zweiten Buches des *Kitāb-al-Manāzir*, welches der Frage gewidmet ist, auf welche Weise es zur Wahrnehmung spezifischer Eigenschaften eines Objektes, etwa Form, Farbe, Größe, räumliche Ausdehnung und Entfernung komme, lassen sich weitergehende Erläuterungen Alhazens finden. Die Wahrnehmung der meisten dieser Eigenschaften, insbesondere die der räumlichen Ausdehnung und Entfernung, sei nicht durch die Effekte des Lichts auf den Sehsinn allein zu erklären, sondern beruhe auf einem *Urteil*: „Not everything that is perceived by sight is perceived through brute sensation; instead, many visible characteristics will be perceived through judgement“ (Alhazen, 2001, S. 431). Das Zustandekommen dieser Urteile geschehe üblicherweise in einer sehr kurzen Zeit, so dass der Wahrnehmende diesen Urteilsprozess selbst nicht wahrnehme:

Several of the visible characteristics that are perceived through judgement and differentiation are perceived in an extraordinarily short time. ... It is not apparent that their perception involves judgement and differentiation because of the speed of the inferential process, through which these characteristics are perceived. But there is no perception at that time that their perception involves judgement. (Alhazen, 2001, S. 433)

Alhazen führt weiter aus, dass sich der ‘perzeptuelle Urteilsprozess’ mit einem üblichen syllogistischen Schluss vergleichen ließe. Die enorme Geschwindigkeit des ‘perzeptuellen Urteils’ lasse sich dadurch erklären, dass zwar Urteile gefällt bzw. Schlussfolgerungen gezogen werden, dies aber aufgrund von Erfahrung und einer ‘Offensichtlichkeit der Gegebenheiten’ auf eine gleichsam *automatisierte* Weise geschehe und so dem Wahrnehmenden nicht bewusst werde:

Since the perception of these characteristics does involve judgement, however, it is only because of the obviousness of their interrelationships and the faculty of discrimination’s familiarity with such characteristics [that the process of judgement goes unnoticed by the perceiver]. Accordingly, as soon as this form reaches [the eye], sight perceives all the characteristics it possesses, and so they will be differentiated by it at the moment of perception. (Alhazen, 2001, S. 433)

Dies sei keine allein dem Sehen eigene Weise des Urteilens, sondern sie lasse sich auch im Rahmen des üblichen logischen Schlussfolgerens finden, nämlich dann, wenn die Prämissen selbst evident sind.<sup>6</sup> Liegt ein solcher Fall vor, müsse die schlussfolgernde Instanz

---

<sup>6</sup>„And the same applies to logical argument and all forms of reasoning when the premises are evident and general; the faculty of discrimination does not require much time to reach the conclusions entailed by them but, instead, will understand the conclusions immediately after grasping the premises“ (Alhazen, 2001, S. 433).

die Prämissen nicht logisch ordnen und das Argument Schritt für Schritt durchgehen, sondern könne, auf Basis wiederholter Erfahrungen mit diesen Prämissen, unmittelbar zur korrekten Schlussfolgerung ‘springen’, die dann selbst zu einer evidenten Gegebenheit werde, zu der man in Zukunft nicht mehr auf schlussfolgernde Weise gelangen müsse.<sup>7</sup>

Alhazen verdeutlicht diesen Punkt anhand der Aussage „Das Ganze ist mehr als ein Teil desselben“. Diese gelte uns allen offensichtlich und ohne bewusst aufwendigen Schlussfolgerungsprozess als wahr. Dennoch beruhe die offensichtliche Wahrheit dieser Aussage auf einer syllogistischen Schlussfolgerung, die erst einmal durchgeführt werden müsse. Ist sie hingegen einmal durchgeführt und die evidente Wahrheit der Aussage festgestellt worden, bekomme diese Aussage den Status der Selbstevidenz, so dass ihr Wahrheitsgehalt nicht jedes mal wieder deduziert werden müsse und ein Urteil darüber, gleichsam über ein ‘Wiedererkennen’, ohne Zeitaufwand möglich sei.<sup>8</sup> Dieser Prozess selbst werde im Fall des Sehens, wie bereits erwähnt, vom Wahrnehmenden selbst nicht wahrgenommen, es bleibe ihm verborgen, *wie* diese Urteile zustande kamen, sie erscheinen einfach als unmittelbar gegeben. Es sei allerdings auch möglich, sich als Wahrnehmender diesem *Wie* der Wahrnehmung selbst zuzuwenden, also das Wahrnehmen selbst zum Gegenstand einer Wahrnehmung zu machen, dies wäre dann ebenfalls ein deduktiver Prozess, der nun aber wesentlich aufwendiger sei und demnach auch mehr Zeit benötige.<sup>9</sup>

---

<sup>7</sup>„The reason is that the faculty of discrimination does not proceed by juxtaposing and ordering premises in the way that an argument based on terms does, for its conclusions will not be based on words or on the arrangement of premises. ... For the arrangement of words in an argument is only one way in which the faculty of discrimination reaches a conclusion, but to reach a perceptual conclusion the faculty of discrimination does not need [this particular] mode of reasoning or [this particular] arrangement of [premises leading to] a perceptual conclusion. ... And the same holds for all deductions that are made through judgements when their premises are evident and their conclusions true; for when the soul realizes that the conclusion is true and reaches that conclusion frequently afterward, the conclusion will be transformed into an evident premise. Thus, when the soul sees the premise, it will immediately reach the conclusion without having to go through the steps of argumentation“ (Alhazen, 2001, S. 433f.).

<sup>8</sup>„And since the truth of the conclusion of this syllogism is absolutely certain in the soul and exists in memory, when the proposition occurs to it, the understanding accepts it without having to go through the steps of argumentation, so it realizes it by means of recognition alone. ... Therefore, syllogisms whose premises are universal and obvious are grasped in an imperceptible amount of time. Then, if the syllogism is frequently reiterated, the intellect will grasp it in such a way that the truth of its conclusion will be assimilated or certified in the soul, at which time the conclusion will become an evident premise. In this way the faculty of discrimination will grasp numerous deductions that are reached by means of judgement in an imperceptible amount of time without having to go through the steps of argumentation“ (Alhazen, 2001, S. 435).

<sup>9</sup>„Furthermore, this second deductive process through which the faculty of discrimination perceives how it perceives what it perceives is not a process that occurs terribly quickly; instead it requires deliberation. ... Therefore, the perception of how the perception occurs and that it is of such-and-such a kind is reached only through a deductive procedure and a differentiation that is not swift“ (Alhazen, 2001, S. 436).

## Tiefenwahrnehmung

Alhazen widmet sich im Anschluss an diese eher allgemein gehaltenen Ausführungen dem speziellen Problem der Wahrnehmung von Entfernung bzw. Tiefe. Es wird mehrfach betont, dass die Wahrnehmung der Distanz eines Objektes nicht allein auf den von den Objekten ausgehenden Lichtstrahlen und deren Effekt auf die Augen beruhe, sondern zusätzliche Urteilsprozesse der eben geschilderten Art notwendig seien („the distance of a visible object from the eye will not be perceived by sight through brute sensation“, Alhazen, 2001, S. 448). In seinen Ausführungen zum Problem der Tiefenwahrnehmung trifft Alhazen zunächst eine subtile Unterscheidung zwischen der Wahrnehmung von Entfernung oder Distanz *per se* und der Wahrnehmung des *Ausmaßes* der Entfernung eines Objektes.<sup>10</sup> Wahrgenommene Distanz *per se* bezieht sich auf den Umstand, dass wir Objekte überhaupt als außerhalb unseres eigenen Körpers gelegen wahrnehmen, die wahrgenommene Größe der Entfernung hingegen darauf, dass wir Gegenstände als *unterschiedlich* weit entfernt wahrnehmen. Bemerkenswert ist diese Unterscheidung Alhazens, da in der aktuellen Wahrnehmungspsychologie eine solche Differenzierung nicht vorzufinden ist. Dass wir Objekte überhaupt als sich außerhalb von uns befindend wahrnehmen, wird in der Regel nicht problematisiert oder als erklärungsbedürftig angesehen. Kurz gefasst erklärt Alhazen die Wahrnehmung von Distanz *per se* ebenfalls über einen Schlussmechanismus, der durch ständige Wiederholung automatisiert werde und so außergewöhnlich schnell ablaufe. Der durchzuführende Schluss habe dabei in etwa folgende Form (vgl. Alhazen, 2001, S. 450f.): Wenn ich meine Augenlider öffne, nehme ich Objekte wahr, wenn ich sie wieder schließe, verschwinden die von mir gesehenen Objekte. Bewege ich zudem meine geöffneten Augen, bewegen sich die von mir gesehenen Objekte nicht in dieselbe Richtung, in die ich meine Augen bewege. Daraus kann ich schließen, dass das von mir Gesehene sich nicht in meinen Augen befinden kann, sondern durch etwas außerhalb von mir hervorgerufen wird.

Wie lässt sich laut Alhazen nun verständlich machen, dass wir üblicherweise Objekte als unterschiedlich weit von uns entfernt wahrnehmen? Für Sendetheorien visueller Wahrnehmung stellte dies kein zentrales und eigenständiges theoretisches Problem dar, die in dieser Vorstellung angenommenen ‘Sehstrahlen’ müssten einen bestimmten Weg von den Augen ausgehend zurücklegen, bis sie das Objekt erreichen, in direkten Kontakt mit diesem treten und gleichsam ‘abtasten’ können. Die so konzipierte Situation des Sehens unterscheidet sich daher nicht wesentlich von der Situation, in der etwa ein Arm ausgestreckt wird, um ein Objekt zu ergreifen und hinsichtlich seiner räumlichen Ausdehnung zu erfühlen. Für Alhazen wird dies jedoch zu einem zentralen Problem. In

---

<sup>10</sup> „The magnitude of distance is different from the fact of distance itself, insofar as it is distance, because distance [per se] means an absence of contact between two bodies, and an absence of contact means that there is some space between these bodies“ (Alhazen, 2001, S. 448).

seiner Konzeption treffen die von den Objekten ausgehenden Lichtstrahlen auf das Auge und es lässt sich nicht sehen, wie sich die Wirkungen des von zwei unterschiedlich weit entfernten Objekten reflektierten Lichts unterscheiden sollten. Ein unmittelbarer Zugang zur Entfernung der Objekte durch das auf die Augen einwirkende Licht sei laut Alhazen nicht möglich, diese müsste ‘erschlossen’ werden. Hier bringt Alhazen nun etwas ins Spiel, was auffallende Ähnlichkeit zum modernen Konzept der Tiefencues aufweist. Alhazen differenziert in einem ersten Schritt zwei verschiedene Fälle, die sich hinsichtlich der ‘Präzision’ der wahrgenommenen Distanz eines Objektes unterscheiden sollen, eine Differenzierung, die sich ebenfalls im Rahmen moderner Formulierungen der Tiefencues finden lässt. Es gebe Situationen, in denen die Wahrnehmung der Entfernung eines Objekts ‘wenig präzise’ sei und Situationen, in denen sie (zumindest innerhalb bestimmter Grenzen) ‘akkurat’ sei:

The perception of the magnitude of a distance from the eye varies, for some distances are perceived by the sense of sight, and their magnitudes are accurately determined, but others are perceived without having their magnitudes accurately determined. (Alhazen, 2001, S. 451)

Zu einer ‘präzisen’ Wahrnehmung der Distanz eines Objektes (bzw. eines Urteils über diese) komme es allein in Situationen, in denen sich zwischen diesem Objekt und dem Wahrnehmenden „a continuous, ordered range of bodies“ (Alhazen, 2001, S. 451) auf dem Boden befinde. Was kann damit gemeint sein? Der Übersetzer äußert im Kommentar zu diesem Teil des *Kitāb-al-Manāzīr*, dass sich aus dem Zusammenhang eigentlich nur eine Lesart ergebe: Eine ‘akkurate’ Wahrnehmung der Distanz eines Objektes sei nur in einer Situation möglich, in der sich zwischen Wahrnehmendem und dem in Frage stehenden Objekt weitere Objekte auf dem Boden befinden, deren Größe und Abstände untereinander dem Wahrnehmenden bekannt oder zumindest von diesem bestimmbar sein müssen. Diese Informationen würden dann genutzt, um über einen Schlussmechanismus der oben dargestellten Art die Distanz des Objektes zu bestimmen:

Now sight perceives the magnitude of the distance of a visible object only through deduction. And sight deduces any measure *only by comparing that measure to another measure already known to sight or to some measure perceived at the same time*; but without an ordered range of bodies spanning the distance of a visible object, sight has no means of measuring the distance of the visible object or of subjecting it to comparison in order to perceive its measure correctly. Therefore, the magnitude of the distance of a visible object is not perceived by the sense of sight unless its distance is spanned by a range of continuous, ordered bodies, and sight perceives those bodies as well as their measures. (Alhazen, 2001, S. 454)

Alhazens Ausführungen lassen sich zu (mindestens) zwei der in der aktuellen Literatur aufgelisteten Tiefencues in Beziehung bringen, nämlich der sogenannten *familiar size as a cue to depth* und der sogenannten *relative size as a cue to depth* (etwa Goldstein, 1999). Mit diesen Tiefencues ist im Wesentlichen die Vorstellung verbunden, dass die Wahrnehmung der Distanz von Objekten aufgrund auf Erfahrungen basierenden Wissens um die ‘tatsächliche Größe’ dieser Objekte in Verbindung mit Erfahrungen, wie groß diese in verschiedenen Entfernungen ‘erscheinen’, möglich sei. Von zwei Objekten gleicher ‘tatsächlicher Größe’ würden wir jenes als weniger weit entfernt wahrnehmen, das uns kleiner erscheint. Betrachte ich etwa mehrere, unterschiedlich groß erscheinende ausgewachsene Bäume gleicher Art, so sei dies laut diesem Ansatz gleichsam ein ‘Hinweis’ darauf, dass sich die kleiner erscheinenden Bäume *weiter* von mir entfernt befinden, als die größer erscheinenden, da ich aufgrund von Erfahrungen wisse, dass Bäume dieser Art in etwa gleich groß seien.

Wichtig in Alhazens Konzeption ist dabei, dass sich *mehrere* Objekte (im Idealfall mit gleicher ‘tatsächlicher Größe’) in meinem Sehfeld befinden. Diese fungierten in Verbindung mit meinem Wissen über die ‘tatsächliche Größe’ dieser Objekte gleichsam als ‘Referenzrahmen’. Ist ein solcher Referenzrahmen gegeben, komme es laut Alhazen zu einer ‘akkuraten’ Entfernungswahrnehmung, indem aufgrund des vorhandenen Wissens und entsprechender Erfahrungen zunächst auf die Entfernungen des sich zwischen dem Wahrnehmenden und dem fraglichen Objekt befindenden „continuous, ordered range of bodies“ geschlussfolgert werde und anschließend auf Basis dieser Schlussfolgerungen die Entfernung des fraglichen Objektes (das mir möglicherweise unbekannt sein kann) deduziert werde. Auch diese Schlussfolgerungen würden nicht bewusst wahrgenommen werden. Liegt ein solcher Referenzrahmen nicht vor, sei keine ‘akkurate’ Distanzwahrnehmung durch Deduktion möglich, Alhazen verdeutlicht dies anhand des Beispiels der wahrgenommenen Entfernung von Wolken über einer Steppe. Diese würden viel weiter entfernt erscheinen, als sie tatsächlich sind, da sich in einer Steppe keine weiteren, den Raum zwischen Wahrnehmendem und der Wolke gleichsam strukturierenden Objekte befinden, die so ‘Hinweise’ auf die Distanz der Wolken geben könnten.

Alhazen legt in seinen Ausführungen sehr viel Wert darauf, dass es bei der ‘akkuraten’ Distanzwahrnehmung nicht zu einer Schätzung der Größen und Entfernungen der zwischen Objekt und Wahrnehmenden sich befindenden Körper komme, sondern zu einer Art *Kalkulation* oder *Berechnung*, was etwa in Alhazens Beschreibung von Fällen nicht ‘akkurater’ Distanzwahrnehmung deutlich wird:

When sight perceives visible objects the magnitudes of whose distances are not accurately determined by sight, the faculty of discrimination immediately apprehends the measures of their distances according to estimation rather than true reckoning. (Alhazen, 2001, S. 455)

## Zusammenfassung

Ist man mit Alhazens wahrnehmungstheoretischen Vorstellungen nicht vertraut, wird man vermutlich erstaunt sein, wie viele Ähnlichkeiten sich zwischen diesem und moderneren Ansätzen finden lassen.<sup>11</sup> Zentral für die weitere Entwicklung wahrnehmungstheoretischer Ansätze ist dabei Alhazens Feststellung, dass visuelle Wahrnehmung (im Gegensatz etwa zur haptischen Wahrnehmung) nicht ‘unmittelbar’ durch Kontakt mit den Objekten geschehe, sondern dass das, was durch unsere Augen empfangen wird, in einem von diesem als passiv zu charakterisierenden Empfangen deutlich abzugrenzenden Prozess zusätzlich ‘interpretiert’ oder ‘erschlossen’ werden müsse, um von Wahrnehmung im eigentlichen Sinne reden zu können.

Der Status des vom Auge Empfangenen ist bei Alhazen nicht mit aller Deutlichkeit auszumachen. Zunächst soll es sich um Licht handeln, trifft dieses Licht auf das Auge, spricht Alhazen von *Formen aus Helligkeit und Farbe*, die sowohl als Formen durch das Auge weitergeleitet werden, als auch einen Effekt auf bestimmte Nerven haben, der dann ebenfalls weitergeleitet wird. Alhazen fasst den angenommenen Prozess der Interpretation des Empfangenen als eine dem Wahrnehmenden nicht bewusste Schlussfolgerung auf, wobei auf Erfahrung beruhendes Wissen und Gewöhnung eine zentrale Rolle spielen. Die wahrgenommene Entfernung der gesehenen Objekte sei ebenfalls auf einen inferentiellen Prozess zurückzuführen, die Entfernung eines Objektes könne über eine Schlussfolgerung ‘akkurat ermittelt’ werden, wenn sich zwischen diesem und dem Wahrnehmenden weitere Objekte befinden, deren ‘tatsächliche Größe’ und ‘gesehene Größe’ bei unterschiedlichen Entfernungen bekannt sei. Insgesamt sind die Ausführungen Alhazens zur Tiefenwahrnehmung zu großen Teilen jedoch sehr vage gehalten.

In der vorgeschlagenen Lesart wirft diese Konzeption allerdings Fragen auf. Wenn es für die (‘akurate’) Distanzwahrnehmung notwendig sei, dass sich zwischen dem Objekt und dem Wahrnehmenden andere, auf bestimmte Weise angeordnete Objekte befinden sollen, aus deren gesehener Größe und Distanz dann auf die Entfernung des in Frage stehenden Objektes geschlossen werden soll, wie kommt es im ersten Schritt überhaupt zu einer Wahrnehmung der Größen und Entfernungen der sich dazwischen befindenden Objekte? Alhazens Konzeption scheint nur verständlich zu sein, wenn diese Objekte als Objekte bereits *wahrgenommen* wurden. Die nicht interpretierten *Formen aus Helligkeit und Farbe* scheinen selbst keine wahrgenommenen Objekte zu sein, erst der inferentielle Prozess soll sie zu solchen machen. In dieser Vorstellung müssen wahr-

---

<sup>11</sup>So schreibt etwa Bauer (1911, S. 54) knapp 900 Jahre später bezüglich Alhazens Theorie der Tiefenwahrnehmung: „Hier hat der mittelalterliche Gelehrte eine Reihe der wichtigsten psychologischen Probleme berührt, und seine Erklärungen antizipieren in überraschender Weise Gedanken, die erst in der jüngsten Entwicklung der Psychologie wieder aufgenommen wurden“. Bauer spielt hier u.a. auf die sogenannte *Theorie der unbewussten Schlüsse* von Hermann von Helmholtz an, die, wie wir in Abschnitt 2.6 sehen werden, deutliche Parallelen zur Theorie des Alhazen aufweist.

genommene Objekte gleichsam als ‘Prämissen’ in den inferentiellen Prozess eingehen, die Schlussfolgerung soll aber die Wahrnehmung von Objekten und deren Entfernung erst begreiflich machen. Im *Kitāb-al-Manāzīr* scheint dieser Punkt häufig dadurch umgangen zu werden, indem davon gesprochen wird, dass der *final sensor* wahrnehme. Diese Beschreibung könnte auch so aufgefasst werden, dass zunächst der *final sensor* das Empfangene ‘interpretiere’ oder wahrnehme, diese Interpretation dann als ‘Prämisse’ in den inferentiellen Prozess eingehe, der dann dazu führe, dass *wir* Objekte in bestimmter Entfernung wahrnehmen. Aber dies würde offensichtlich nicht zu einer Erklärung unserer Wahrnehmung genügen, denn nun müsste verständlich gemacht werden, auf welche Weise der *final sensor* wahrnimmt und wie das Wahrnehmen des *final sensors* zu *meiner* Wahrnehmung führt.

Ein weitere Schwierigkeit besteht in der generellen Konzeption von visueller Wahrnehmung als Ergebnis einer Schlussfolgerung. Ohne weitere Erläuterungen ist nicht unmittelbar verständlich, wie diese Beschreibung gemeint sein soll. Eine übliche logische Schlussfolgerung besteht aus *Propositionen*, wie Alhazens Beispiel („Das Ganze ist mehr als ein Teil desselben“) ebenfalls demonstriert. Wenn ich visuell wahrnehme, sehe ich allerdings keine Propositionen, sondern *Objekte*. Es kann sich hier demnach nur um eine metaphorische Beschreibung handeln. Dieser Punkt wird in der Darstellung der Wahrnehmungstheorie von Helmholtz, der Wahrnehmung wie Alhazen als Ergebnis eines Schlusses auffasst, ausführlicher behandelt werden. Wir werden im weiteren Verlauf dieses historischen Teils sehen, dass die angeführten Schwierigkeiten nicht allein auf eine mögliche ‘Antiquiertheit’ der Theorie Alhazens oder den möglicherweise als bilderreich zu charakterisierenden Stil, in dem diese von Alhazen vorgebracht wird, zurückzuführen sind, sondern in ähnlicher Form immer wieder auftauchen und sich zuspitzen werden.

## 2.2. Natürliche Geometrie und Nervenbewegungen: René Descartes

Neue Impulse für eine wahrnehmungstheoretische Behandlung der Tiefenwahrnehmung konnten nach Alhazen durch die intensive Auseinandersetzung mit Fragen der perspektivischen Darstellung in der Kunst der Renaissance sowie dem Aufkommen der ersten systematischen anatomischen Studien des Auges in der späten Renaissance bzw. der anstehenden Aufklärungsbewegung gesetzt werden. Ein exemplarischer Fall einer Synthese geometrischer und physiologischer Überlegungen zur Erklärung visueller Wahrneh-

mung ist der wahrnehmungstheoretische Ansatz von René Descartes (1596 bis 1650).<sup>12</sup> Auch bei Descartes lassen sich wahrnehmungstheoretische Ausführungen finden, in denen visuelle Wahrnehmung als ein mehrstufiger Prozess beschrieben wird, der mit dem Empfangen von Licht durch die Augen beginnt.

Die wohl aussagekräftigsten Ausführungen zu einem gestuften Wahrnehmungsprozess lassen sich in den Erwidernungen Descartes' auf die 1642 verfasste sechste Gruppe an Einwänden gegen seine „*Meditationes de prima philosophia*“ finden (Atherton, 2002). Im 9. Einwand, der dem Verhältnis von Sinnen und Geist gewidmet ist und im Speziellen die von Descartes vertretene Position betrifft, dass den Sinnen generell zu misstrauen und nur durch den Intellekt oder Geist eine 'epistemologische Gewissheit' möglich sei, fragen seine Kritiker, wie dies zu verstehen ist, da doch der Intellekt mögliche 'Fehler' eines Sinnessystems (z.B. des visuellen) allein durch Informationen aus einem anderen Sinnessystem (etwa dem haptischen) 'korrigieren' könne. Als Antwort auf diesen Einwand schreibt Descartes (1641/1972):

Um recht zu verstehen, wie es mit der Gewißheit der Sinnesauffassung steht, muß man sozusagen drei Stufen derselben unterscheiden: Zur ersten gehört nur das, wodurch unmittelbar das körperliche Organ von den äußeren Objekten affiziert wird, und das kann nichts anderes sein als die Bewegung der kleinen Teile dieses Organs und die Veränderung der Gestalt und Lage, die von dieser Bewegung herrührt. Die zweite Stufe umfaßt alles das, was unmittelbar im Geiste dadurch sich ergibt, daß er mit dem so affizierten Organe vereinigt ist; hierher gehören die Perzeptionen des Schmerzes, des Kitzels, des Durstes, des Hungers, der Farben, des Schalles, des Geschmackes, des Geruches, der Wärme, der Kälte und dergl., die ... aus der Vereinigung und sozusagen Vermischung des Geistes mit dem Körper entstehen. Die dritte schließlich umfaßt alle die Urteile, die wir bei Gelegenheit von Bewegungen des körperlichen Organs über die Dinge außer uns zu fällen von Jugend auf gewohnt sind. (S. 378)

Descartes verdeutlicht dies anschließend anhand des Beispiels der Wahrnehmung eines Stockes. Zunächst würden von diesem Stock Lichtstrahlen reflektiert, welche bei Eintritt in die Augen den optischen Nerv auf mechanische Weise 'in Bewegung' setzten. Die Regelmäßigkeit der Reflektion des Lichts in das Auge, der Brechung des Lichts durch die Linse usw. seien durch die optischen Gesetze gegeben und beschreibbar, die Regelmäßigkeit der 'Bewegungen des Nervenapparates' Sache der Anatomie oder Phy-

---

<sup>12</sup>Ob es angemessen erscheint, die theoretischen Vorstellungen zur visuellen Wahrnehmung des Descartes als *eine* Theorie zu bezeichnen, ist umstritten, da sich in Descartes' umfangreichen Werk durchaus heterogene Ansätze finden lassen (siehe Wee, 2014).

siologie. Diese erste Stufe des Wahrnehmungsprozesses lasse sich bei allen mit Augen ausgestatteten Lebewesen finden.

Descartes hatte, vermutlich angeregt durch Johannes Keplers Vorstellung des ‘Netzhautbildes’ (Lindberg, 1976, Kap. 9), in selbst durchgeführten anatomischen Untersuchungen von Kuhaugen festgestellt, dass sich auf dem Augenhintergrund ein kleines, verzerrtes, invertiertes und auf dem Kopf stehendes ‘Bild’ der äußeren Gegenstände *betrachten* lässt. Die Vorstellung, dass Wahrnehmung durch ein Betrachten dieses ‘Netzhautbildes’ durch eine Art ‘inneren Beobachter’ zu erklären sei, weist er jedoch an anderer Stelle, in seiner Schrift „*La Dioptrique*“ („Die Optik“), entschieden zurück:

Now although this picture, in being transmitted into our head, always retains some resemblance to the objects from which it proceeds, nevertheless, as I have already shown, we must not hold that it is by means of this resemblance that the picture causes us to perceive the objects, as if there were yet other eyes in our brain with which we could apprehend it. (Descartes, 1637/2001, S. 101)<sup>13</sup>

Das ‘Netzhautbild’ *affiziere* das Auge auf mechanische Weise, was dazu führe, dass dieses in ein bestimmtes ‘Bewegungsmuster’ des optischen Nervs umgewandelt wird, das schließlich an die Zirbeldrüse weitergeleitet werde und dann an dieser Stelle direkt auf den Geist ‘einwirke’.

Auf der nächsten, der zweiten Stufe des von Descartes angenommenen Wahrnehmungsprozesses komme es durch diese Einwirkung zu einer *Perzeption* „der durch den Stock reflektierten Farbe oder des reflektierten Lichts“ (Descartes, 1641/1972, S. 379). Diese sei unmittelbar und rühre daher, dass Körper und Geist auf ganz bestimmte, durch die Natur vorgegebene Weise miteinander verbunden seien, so dass die verschiedenen Bewegungen des Nervenapparates den Geist affizieren und verschiedene Perzeptionen wie etwa Farbe – wohlgermerkt nicht *die Farbigkeit* eines sich außerhalb des Wahrnehmenden befindenden Stockes – zur Folge haben. Descartes betont mit Nachdruck, dass, wenn man eine scharfe Trennlinie zwischen den (körperlichen) Sinnen und dem Geist oder Verstand ziehen wollte, diese ersten beiden Stufen den Sinnen zuzuordnen habe („und nichts anderes wäre auf den Sinn zu beziehen, wenn wir ihn genau vom Verstande unterscheiden wollen“, Descartes, 1641/1972, S. 379). Auf diesen beiden ersten Stufen sei es auch nicht sinnvoll zu sagen, dass man sich als Wahrnehmender irre. Bei einer bloßen Perzeption von Farbe (oder auch Schmerz, Lust, Hunger usw.) könne man sich nicht irren, da es sich um eine *unmittelbare* Erfahrung ohne Beteiligung des Geistes handle.

---

<sup>13</sup>Da es nach Wissen des Autors keine vollständige deutsche Übersetzung der *Dioptrique* gibt, werden entsprechende Stellen aus der englischen Übersetzung zitiert.

Grundsätzlich Verschiedenes geschehe dann allerdings auf der dritten Stufe des Wahrnehmungsprozesses: Die verschiedenen Perzeptionen der zweiten Stufe würden laut Descartes nun vom Geist *gedeutet*, es würden Urteile gefällt bzw. Schlüsse aus den Perzeptionen gezogen. Die bloße Perzeption von Farbe auf der zweiten Stufe führe nun etwa dazu, dass diese als ein *farbiges, nicht meiner Innen-, sondern der Außenwelt angehörendes Objekt* gedeutet werde. Ebenso werde anhand des reflektierten Lichtes und den dadurch unmittelbar ausgelösten Perzeptionen die Größe, Form und Entfernung dieses Objektes *berechnet*.<sup>14</sup>

Laut Descartes lasse sich jedes visuelle Wahrnehmungsobjekt auf sechs Qualitäten reduzieren: Helligkeit, Farbe, Position im Raum, Entfernung, Form und Größe (Descartes, 1637/2001), wobei nur die Wahrnehmung der ersten beiden Qualitäten dem Sehsinn selbst (der zweiten Stufe) zuzuschreiben sei und die restlichen allein dem Geist.<sup>15</sup> Die Urteilsprozesse auf der dritten Stufe seien, wie bei Alhazen, im Wesentlichen nicht von den (logischen) deduktiven Leistungen zu unterscheiden. Der einzige Unterschied bestehe darin (und auch hier klingt wieder Alhazen an), dass in einem Großteil der Fälle die ‘perzeptuellen Urteile’ mit viel größerer Geschwindigkeit getroffen würden als die herkömmlichen logischen Urteile, da wir die meisten dieser Urteile schon mehrfach getroffen hätten und so auf unsere Erfahrung bauen könnten oder uns lediglich an die früheren Urteile erinnern und somit nicht den möglicherweise komplexen Urteilsprozess in Gänze vollziehen müssten. Dies sei auch der Grund, warum im Alltag die Wahrnehmung externer Objekte mit bestimmter Form und Position im Raum den Sinnen und nicht dem Geist zugeschrieben werde.<sup>16</sup>

<sup>14</sup>In Descartes (1641/1972) heißt es:

Denn daß ich auf Grund dieser Wahrnehmung der Farbe, durch die ich affiziert werde, urteile, daß der außer mir befindliche Stab farbig sei, imgleichen, daß ich auf Grund der Ausdehnung dieser Farbe, ihrer Begrenzung und Lagebeziehung zu den Teilen des Gehirns einen Schluss ziehe auf die Größe, Gestalt und Entfernung eben dieses Stabes, das hängt ... doch offenbar allein vom Verstand ab. (S. 379)

Die Übersetzung scheint an dieser Stelle etwas unglücklich oder zumindest einschränkend, da es in der Originalpassage lautet: „itemque quòd ex istius coloris extensione, terminatione, ac fitis relatione ad partes cerebri, de eiusdem baculi magnitudine, figurâ & distantia ratiociner“ (Descartes, 1641/1996, S. 437) und „*ratiocinari*“ üblicherweise eher oder zumindest auch im Sinne von „berechnen“ verwendet wird. In der von Atherton (2002, S. 6) zitierten Übersetzung ins Englische heißt es auch: „I make a rational calculation about the size, shape and distance of the stick“.

<sup>15</sup>Die Wahrnehmung von Farbe nimmt dabei allerdings eine Sonderrolle ein, da Descartes diesen Begriff in zwei verschiedenen Bedeutungen gebraucht, einmal als unmittelbare Wahrnehmung von Farbe (die der zweiten Stufe zuzuordnen wäre) und der wahrgenommenen Farbigkeit eines vom Wahrnehmenden als extern angesehenen Objektes (die der dritten Stufe angehört).

<sup>16</sup>„Der Unterschied ist hierbei nur der, daß wir solche Urteile, die wir jetzt zum ersten Male infolge irgendeiner neuen Wahrnehmung fällen, dem Verstande zuschreiben; solche dagegen, die wir von Jugend an gerade auf die Weise wie jetzt, in Betreff dessen, was unsere Sinne affiziert, gefällt, sowie auch die Schlüsse, die wir gezogen haben, den Sinnen zuschreiben, weil wir nämlich infolge *der Gewohnheit* hierüber so *rasche Schlüsse* ziehen und urteilen, – oder vielmehr uns der früher bereits von uns über ähnliche Dinge gefällten Urteile *erinnern*, – daß wir diese Fähigkeiten von der einfachen Sinnesauffassung nicht unterscheiden“ (Descartes, 1641/1972, S. 379).

Allein auf dieser dritten Stufe, so Descartes, könne man überhaupt sinnvoll davon sprechen, dass man sich irre, wobei darunter lediglich verstanden werden soll, dass man etwa als wenig reflektierendes Kind oder als in seinen Schlussfolgerungen allzu unbedacht vorgehender Erwachsener zu Urteilen komme (etwa, dass ein in Wasser getauchter Stab krumm *ist*), denen weniger zu vertrauen sei, als die Urteile, die man als vernünftiges Wesen treffen könne (dass der Stab uns so *erscheint*, dass man denken könnte, er wäre gekrümmt). Dabei betont Descartes noch einmal ausdrücklich und gegen seine Kritiker gewandt, dass in diesem Fall das vom Intellekt getroffene Urteil eines krummen Stabes nicht durch den Tastsinn, sondern allein durch den Geist (wenngleich über Informationen aus dem Tastsinn), ‘korrigiert’ werden könne (vgl. Descartes, 1641/1972, S. 380). Descartes gibt in dieser relativ kurzen Passage leider nur wenig Anhaltspunkte, welche Beziehungen zwischen diesen drei Stufen bestehen, insbesondere, auf welche Weise oder über welche Prozesse der Wahrnehmende auf der dritten Stufe zu den Schlüssen über Größe, Entfernung usw. komme.

### Tiefenwahrnehmung

In Descartes’ vor den Er widerungen auf die Einwände gegen die *Meditationes* verfassten Hauptschrift zum Sehen, der *Dioptrique*, auf die er explizit auch in den Antworten auf die Einwände gegen die *Meditationes* verweist, lassen sich diesbezüglich auch keine Hinweise finden. In dieser Schrift stellt sich die Cartesische Auffassung des Sehens anders dar, enthält jedoch weitergehende Ausführungen zur Tiefenwahrnehmung, weshalb sie hier herangezogen werden soll. Die oben dargestellte strenge Trennung zwischen verschiedenen Wahrnehmungsstufen, die den Sinnessystemen bzw. dem Geist zuzuordnen seien, lässt sich hier nicht deutlich ausmachen, Descartes besteht aber auch hier darauf, dass es der Geist sei, der wahrnehme: „First of all, it is the mind which sees, not the eye; and it can see immediately only through the intervention of the brain“ (Descartes, 1637/2001, S. 108). Der ‘Mechanik’ der durch das ins Auge einfallende Licht ausgelösten Nervenbewegungen kommt jedoch eine wesentlich bedeutendere Rolle zu, als möglichen Urteilsprozessen des Geistes. An vielen Stellen scheint es gar, als würde Descartes den gesamten Wahrnehmungsprozess als einen rein mechanischen auffassen, wobei der kausale Zusammenhang zwischen Nervenbewegungen und Wahrnehmung schlicht naturgegeben sei (Wolf-Devine, 2000).

Die Wahrnehmung von Licht und Farbe sei auf durch das einfallende Licht ausgelöste (Nerven-)Bewegungen in jenen Teilen des Hirns, mit denen die Sehnerven verbunden sind, zurückzuführen. Die ‘Stärke’ (*force*) dieser Bewegungen verursache die Wahrnehmung unterschiedlicher Helligkeit des Lichts und der unterschiedliche ‘Charakter’ dieser Bewegungen verursache die Wahrnehmung unterschiedlicher Farben. Diese Beziehun-

gen seien naturgegeben, die Wahrnehmung von Licht und Farbe daher nichts anderes als eine Konsequenz des Aufbaus des visuellen Sinnessystems.<sup>17</sup>

Etwas Vergleichbares geschehe nun aber auch bei der Wahrnehmung der Distanz, eine der vier wahrgenommenen Qualitäten die in den Einwänden gegen die *Meditationes* als allein durch ein Urteil des Geistes zustande kommend angesehen werden. Das von den Objekten in die Augen reflektierte Lichtmuster (das ‘Netzhautbild’) sei wesentlich zweidimensional, so dass die durch dieses ausgelösten Nervenbewegungen allein nicht hinreichend wären, um die Distanz eines Objektes wahrzunehmen. Um die Wahrnehmung von Distanz zu ermöglichen, bedürfe es verschiedener ‘Hilfsmechanismen’. Descartes schlägt verschiedene solcher Hilfsmechanismen vor, die jedoch zusammen eine recht heterogene Gruppe bilden („the ones that Descartes provides form a rather heterogenous and ill-assorted group“, Wolf-Devine, 2000, S. 511). Manche scheinen mit Descartes’ eher mechanistischer Vorstellung der Wahrnehmung in Übereinstimmung zu bringen zu sein, andere scheinen hingegen eine aktive Tätigkeit des Geistes zu erfordern. Sämtliche der von Descartes vorgeschlagenen Hilfsmechanismen sind heute als Tiefencues bekannt.

In erster Linie beruhe die Tiefenwahrnehmung auf ‘Veränderungen des Augenkörpers’ bei Fixation unterschiedlich weit entfernter Objekte, also dem, was in der aktuellen Literatur (in leicht anderer Formulierung) auch als *Akkommodation* bezeichnet wird. Die sich bei Fixation eines Objektes einstellenden Veränderungen des Augenkörpers führten zu Veränderungen in bestimmten Hirnbereichen, die aufgrund naturgegebener Beziehungen zur Folge haben, dass die Entfernung eines Objektes schlicht wahrgenommen werde, ohne Beteiligung von Urteilsprozessen.<sup>18</sup>

An zweiter Stelle sei Distanzwahrnehmung durch die bei binokularer Fixation unterschiedlich weit entfernter Objekte unterschiedliche Stellung der beiden Augen zueinander möglich, also durch das, was heutzutage als *Konvergenz* bezeichnet wird. Descartes verdeutlicht dies anhand einer Analogie. Man stelle sich einen blinden Mann vor, der zwei Stöcke in seinen Händen  $A$  und  $B$  hält, die sich im ihm unbekanntem Punkt  $X$  kreuzen. Dem Mann sei die Länge der Strecke  $\overline{AB}$  (also der Abstand seiner beiden Hände) sowie die beiden zwischen  $\overline{AB}$  und  $\overline{AX}$  sowie  $\overline{AB}$  und  $\overline{BX}$  eingeschlossenen Winkel bekannt. Wie über *natürliche Geometrie* („as if by natural geometry“, Descartes,

<sup>17</sup>„The nature of our mind is such that the force of movements in the areas of the brain where the small fibers of the optic nerves originate cause it to perceive light; and the character of these movements cause it to have the perception of color“ (Descartes, 1637/2001, S. 101).

<sup>18</sup>„The seeing of distance depends no more than does the seeing of location upon any images emitted from objects; but in the first place upon the shape of the body of the eye. For as we have said, for us to see things close to our eyes this shape must be slightly different from the shape which enables us to see things farther away; and as we adjust the shape the shape of the eye according to the distance of objects, we change a certain part of our brain in a manner that is ordained by nature to make our soul perceive this distance. Ordinarily this happens without our reflecting upon it“ (Descartes, 1637/2001, S. 106).

1637/2001, S. 106) sei es dem blinden Mann nun ohne Weiteres möglich, mit Hilfe dieser Informationen die Entfernung des Punktes  $X$  zu *wissen*. Etwas Analoges geschehe nun auch bei der Tiefenwahrnehmung. Die Stöcke des blinden Mannes sollen nun den Sehlinien der beiden Augen des Wahrnehmenden entsprechen, die sich im Punkt  $X$ , dem fixierten Objekt unbekannter Entfernung, kreuzen. Die Länge der Strecke  $\overline{AB}$  (die Distanz zwischen den beiden Ursprüngen der Sehlinien) sowie die beiden zwischen  $\overline{AB}$  und den Sehlinien eingeschlossenen Winkel seien uns als Wahrnehmenden bekannt. Über *natürliche Geometrie* sei es uns nun wie dem blinden Mann möglich, mit Hilfe dieser Informationen die Distanz des Objektes zu wissen.<sup>19</sup>

Geometrisch kann man sich das, was Descartes hier vor Augen hat, vereinfacht folgendermaßen veranschaulichen (vgl. Abbildung 2.1 auf der nächsten Seite). Der Wahrnehmende fixiere das ihm in unbekannter Entfernung  $d$  vor ihm liegende Objekt  $X$  mit seinen beiden Augen  $A$  und  $B$ , d.h. die Sehlinien der beiden Augen kreuzen sich im Punkt  $X$ . Die Länge der Strecke  $\overline{AB}$ , aus Darstellungsgründen in Abbildung 2.1 mit  $i$  (für interokularer Abstand) bezeichnet, sei dem Wahrnehmenden bekannt, ebenso wie die Winkel  $\alpha$  und  $\beta$ . In dieser *geometrischen Abstraktion* der Sehsituation muss zunächst gelten, dass  $\gamma = 180 - (\alpha + \beta)$ . Weiterhin gilt aufgrund der Eigenschaften der Tangensfunktion:

$$\tan(\gamma/2) = \frac{i}{2d}. \quad (2.1)$$

Durch einfaches Umstellen erhält man nun die Länge  $d$  folgendermaßen:

$$d = \frac{i}{2 \tan(\gamma/2)}. \quad (2.2)$$

Mit anderen Worten: In dieser *geometrischen Abstraktion* der Sehsituation lässt sich die Länge  $d$  unter den Voraussetzungen, die Descartes angibt, ohne Weiteres *berechnen*. Descartes Auffassung scheint es zu sein, dass eben dies bei der Distanzwahrnehmung mit Hilfe der Konvergenz geschehe, über *natürliche Geometrie*.

Eine weitere, dritte Möglichkeit der Distanzwahrnehmung, die wohl von der zu Descartes' Zeiten weit verbreiteten Vorstellung des Auges als einer *camera obscura* herührt, bestehe darin, dass bei Fixation eines Objektes  $X$  andere Objekte, die (grob gesprochen) vor oder hinter diesem Objekt liegen, weniger scharf und unterschiedlich

---

<sup>19</sup>„And similarly, when our two eyes A and B are turned towards point X, the length of the line AB and the size of the two angles XAB and XBA enable us to know where the point X is“ (Descartes, 1637/2001, S. 106).

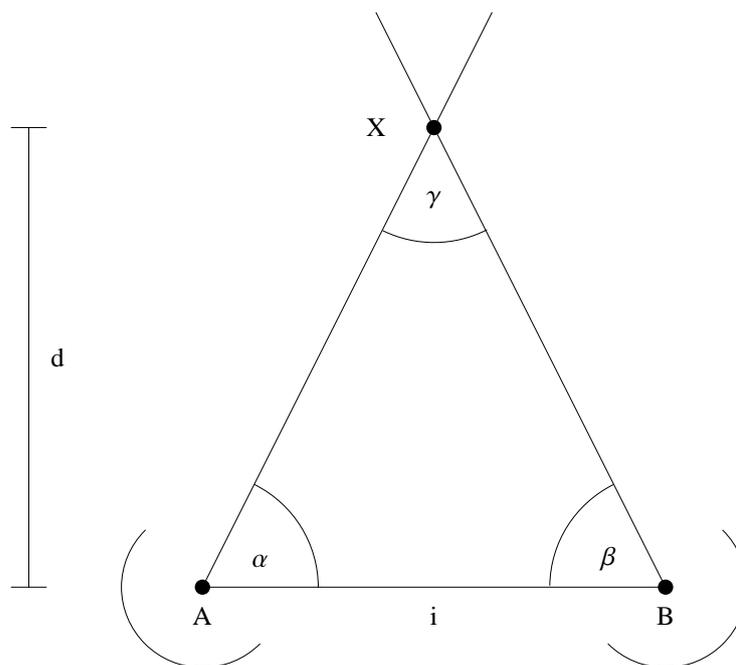


Abbildung 2.1. Geometrische Veranschaulichung der Cartesischen Auffassung der Konvergenz als Mittel zur Tiefenwahrnehmung. Erläuterungen im Text.

lichtstark auf die Netzhaut abgebildet werden, als dieses Objekt  $X$ .<sup>20</sup> In der aktuellen Literatur wird dieser Tiefencue auch als *image blur* bezeichnet (Howard, 2012). Interessanterweise spricht Descartes davon, dass wir aufgrund dieses Tiefencues *Urteile* über die Distanz dieser Objekte fällen würden („we judge it to be farther away“, Descartes, 1637/2001, S. 107). Die Wirkungsweise dieses Tiefencues wird nicht wie bei der Akkommodation als rein mechanistisch beschrieben, sondern soll offenbar eine aktive Tätigkeit des Geistes beinhalten.

Daran anschließend schlägt Descartes eine letzte Methode vor, die es erlauben soll, Distanz wahrzunehmen. Diese weist eine deutliche Ähnlichkeit zu der von Alhazen behaupteten auf, nämlich dass ein Wissen um die ‘tatsächliche’ und die ‘erscheinende Größe’ von Gegenständen in verschiedenen Entfernungen es erlaube, die Distanz anhand eines Urteils zu ermitteln:

Finally, when moreover we already imagine the size of an object, or its position, or the distinctness of its shape and of its colors, or merely the strength of the light which comes from it, this enables us *not actually to see, but to imagine its distance*. Thus, looking from afar at some body which we are used to seeing close at hand, we *judge* its distance much better than we would if its size were not so well known to us. ... Looking at two ships on

<sup>20</sup>„We have yet another way of perceiving distance, which is through the distinctness or indistinctness of the shape seen, together with the strength or weakness of the light. ... From this we see that they are farther from us, or nearer to us“ (Descartes, 1637/2001, S. 106).

the sea, one of which is smaller than the other but proportionately closer, so that they *appear* equal in size, we will be able to judge which is farther away by the difference in their shapes [sämtliche Hervorhebungen von mir].

(Descartes, 1637/2001, S. 106)

Descartes spricht hier davon, dass wir anhand der angeführten Informationen in der Lage seien, die Distanz eines Objektes uns *vorzustellen*, wenngleich wir die Distanz selbst nicht sehen können. Er scheint demnach eine Unterscheidung zwischen einem *Sehen der Entfernung* und einem Wahrnehmen von Entfernung über *Imagination* treffen zu wollen, worin jedoch der Unterschied zwischen diesen beiden Formen der Tiefenwahrnehmung bestehen soll, wird nicht weiter erläutert. Soll der Unterschied möglicherweise (wie bei Alhazen) den ‘Grad der Genauigkeit’ betreffen, etwa in der Form, dass die Wahrnehmung über Imagination wesentlich ‘ungenauer’ sei? Descartes’ weitere Ausführungen deuten nicht darauf hin, denn er hält *jegliche* Form der Tiefenwahrnehmung für ‘ungenau’: „It is also to be noted that all the means that we have for knowing distance are very uncertain“ (Descartes, 1637/2001, S. 110).

### Zusammenfassung

Zusammenfassend ergeben die beiden hier dargestellten Texte ein wenig kohärentes Bild.<sup>21</sup> In den Antworten auf die Einwände gegen die *Meditationes* formuliert Descartes eine Drei-Stufen-Theorie der (Tiefen-)Wahrnehmung, in der er streng zwischen Funktionen des visuellen Sinnes und denen des Geistes trennt. Die Funktionen des Geistes, die auf der dritten Stufe zu einer Wahrnehmung (im Sinne des Sehens von Objekten) führen sollen, sollen sich im Wesentlichen nicht von üblichen Deduktionsprozessen unterscheiden, die uns jedoch – wie bei Alhazen – nicht bewusst wären, da wir diese unzählige Male durchgeführt hätten. Descartes scheint in diesem Text die ausgezeichnete Rolle des Geistes bei der Tiefenwahrnehmung hervorheben zu wollen. Die wahrgenommene Distanz eines Objektes, so könnte man es kurz fassen, sei das Ergebnis eben eines solchen *Urteilsprozesses*, einer geistigen Leistung. Liest man jedoch Descartes’ *Dioptrique*, so bietet sich einem ein anderes Bild dar, wenngleich in diesem Text ausführlicher auf die Wahrnehmung von Tiefe eingegangen wird. Eine explizite Stufentheorie der Wahrnehmung lässt sich hier nicht ausmachen, Descartes behauptet zwar, dass es der Geist sei, der sehe, an vielen Stellen entsteht allerdings der Eindruck, dass der Geist lediglich auf naturgegebene (mechanische) Weise auf die an die Zirbeldrüse gelangenden Nervenbewegungen gleichsam reagiere und nicht Urteile fällt oder Schlussfolgerungen vornimmt.

<sup>21</sup>Eine Feststellung, die etwa auch Hatfield (1990) und Wolf-Devine (2000) in ihren Behandlungen der Cartesischen Wahrnehmungstheorie treffen.

Die Wahrnehmung von Tiefe und Entfernung sei nicht allein durch das auf die Augen projizierte ‘Netzhautbild’ bzw. die dadurch ausgelösten Nervenbewegungen möglich, sondern es bedürfe zusätzlicher Mittel. Descartes erweitert nun diese bereits bei Alhazen aufzufindende Vorstellung, indem er, auf anatomische und optische Erkenntnisse seiner Zeit aufbauend, Hilfsmittel wie die Akkommodation des Auges und die Konvergenz postuliert, die einen deutlich physiologischen Charakter aufweisen. Insgesamt bilden die von Descartes aufgezählten Mittel der Tiefenwahrnehmung jedoch eine eher heterogene Gruppe, die keine übergreifende theoretische Vorstellung zu verbinden scheint.

Bezüglich der Akkommodation wird die Wirkungsweise deutlich mechanistisch beschrieben, was auch der allgemeinen Ausrichtung der Cartesischen Naturphilosophie entspricht (siehe dazu etwa den Sammelband von Gaukroger, Schuster & Sutton, 2000). Im Fall der Konvergenz ist die Lage weitaus unklarer. Descartes behauptet, dass wir die Distanz eines Objektes über die Stellung der Augen zueinander *wüssten* „as if by natural geometry“ (Descartes, 1637/2001, S. 106). Ob ein solches Wissen der Distanz einen grundsätzlich anderen Status haben soll als eine Wahrnehmung der Distanz, zu der es etwa über die Akkommodation komme, lässt sich nicht ersehen. Descartes spricht im Kontext der Konvergenz nicht davon, dass wir tatsächlich Geometrie *benutzen* würden, die Ausführungen deuten aber darauf hin, dass nur eine tatsächliche Anwendung geometrischer Gesetzmäßigkeiten zum gewünschten Ergebnis führt. Die Annahme, dass wir selbst auf bewusste Weise solche geometrischen Gesetze anwenden, um Distanzen zu ermitteln, erscheint nicht plausibel, da der mathematische Laie sicher nicht ohne weiteres diese Gesetze angeben könnte, aber dennoch problemlos Objekte in unterschiedlicher Distanz wahrnimmt – darüber hinaus kann sicher nicht einmal der mathematische Experte die für Descartes’ Vorstellung benötigten Winkel bei Fixation eines Objektes auch nur annähernd angeben.

Aber wer oder was soll dann diese Berechnungen vornehmen und woher stammt das Wissen um die für eine solche Berechnung notwendigen Größen? Mentale Tätigkeiten dieser Art kommen laut Descartes allein dem Geist zu. Aber selbst wenn man dem Geist diese Berechnungen zuschreiben und davon ausgehen würde, dass diese dem Wahrnehmenden aufgrund oftmaliger Durchführung nicht bewusst seien, wie lässt sich dann eine Verbindung zwischen der Berechnung der Entfernung und dem Sehen eines sich in bestimmter Entfernung befindenden Objektes herstellen? Weiterhin scheint es aus eher phänomenologischer Perspektive unplausibel, die Veränderungen der Stellung der Augen zueinander bei Fixation unterschiedlich weit entfernter Objekte als Erklärung für die wahrgenommene Distanz der Objekte heranzuziehen. Wenn ich in einer Alltagssituation meine Augen öffne und ein bestimmtes Objekt fixiere, sehe ich dennoch die Gegenstände in der Peripherie meines Sehfelds erstens *als* Gegenstände und zweitens

als Gegenstände in bestimmter Entfernung von mir, wenngleich dieser Eindruck nicht so deutlich sein mag, wie bei dem fixierten Objekt. Wende ich nun meinen Blick auf ein sich vorher in der Peripherie befindendes Objekt und fixiere dieses, so zeigt sich in meiner Wahrnehmung nicht plötzlich ein gleichsam in eine bestimmte Distanz ‘springendes’ Objekt, wie es die Auffassung der Konvergenz nahelegt, sondern das Objekt erscheint lediglich deutlicher, in klareren Umrissen, aber nicht als in seiner Distanz sich verändernd.

Bezüglich der ‘erscheinenden Größe’ eines Objektes als Mittel zur Tiefenwahrnehmung stellen sich analoge Probleme wie bei den recht ähnlichen Ausführungen des Alhazen ein, die sich bei Descartes aufgrund seiner Konzeption des ins Auge projizierten ‘Netzhautbildes’ jedoch zuspitzen. Dieses bzw. die durch dieses ausgelösten Nervenbewegungen haben in Descartes’ Konzeption einen privilegierten Status für das Sehen, es wird jedoch nicht wirklich klar, worin dieser besteht. Descartes lehnt die Auffassung, dass der Geist das ‘Netzhautbild’ sehe, in aller Deutlichkeit ab, da es impliziere, dass der Geist selbst Augen habe. Gleichzeitig deuten viele seiner Ausführungen jedoch daraufhin, dass der Geist auf irgendeine Weise dennoch *Zugang* zu diesem haben muss. Besonders deutlich wird dies in seinen Ausführungen zur ‘erscheinenden Größe’ eines Objekts, die in direktem Zusammenhang mit der wahrgenommenen Distanz und der Größe des auf die Augen projizierten ‘Bildes’ diese Objektes stehen soll:

Their size is estimated according to the knowledge, or the opinion, that we have of their distance, *compared with the size of the images that they imprint on the back of the eye; and not absolutely by the size of these images* [Hervorhebung von mir]. (Descartes, 1637/2001, S. 107)

Wie soll ein solcher Vergleich von statten gehen, wenn der Geist keinen Zugang zum ‘Netzhautbild’ als Bild der äußeren Objekte hätte? Er ist nur dann möglich, wenn das ‘Netzhautbild’ *gesehen* wird. Der Wahrnehmende hat nun aber keinen Zugang zu dem auf die Augen projizierten Lichtmuster, er kann es nicht sehen. Descartes Konzeption erscheint nur dann sinnvoll, wenn man eine personenähnliche Instanz, einen Homunkulus, annimmt, der das ‘Netzhautbild’ sehen und diesbezügliche Vergleiche anstellen kann. Würde man diese Annahme machen, bestünde aber immer noch das Problem, dass erstens das Sehen des Homunkulus verständlich gemacht werden müsste und zweites die Beziehung zwischen dem Sehen des Wahrnehmenden und dem Sehen dieses Homunkulus unklar bleibt. Es lässt sich daher feststellen, dass auch bei Descartes’ wahrnehmungstheoretischem Ansatz ein Akt der Wahrnehmung schon vorausgesetzt werden muss, um Wahrnehmung verständlich machen zu können.

## 2.3. Eine neue Theorie des Sehens: George Berkeley

Der irische Theologe und Bischof George Berkeley (1685-1753), der in vielen Einführungen in die Geschichte der Philosophie als Hauptvertreter der Strömung des sogenannten *Immaterialismus* gehandelt wird, ist wohl am ehesten für die ihm zugeschriebene und oft auch gleichsam als Motto des Immaterialismus ausgegebene Wendung „*esse is percipi*“ bekannt und berühmt geworden, welche die Philosophie seiner Zeit in Aufruhr versetzte. Berkeley veröffentlichte Zeit seines Lebens drei stilistisch unterschiedlich ausgerichtete theoretische Abhandlungen über das Sehen des Menschen, den „*Essay Towards a New Theory of Vision*“, die „*Theory of Vision, or Visual Language Vindicated and Explained*“ und den vierten Dialog seiner Schrift „*Alciphron*“. Das erste Werk behandelt das Problem visueller Wahrnehmung in einer analytischen oder induktiven Weise, das zweite in einer synthetischen oder deduktiven Weise und das dritte schließlich in der Form eines Dialoges (vgl. Turbayne, 1963). Wir wollen uns in diesem Abschnitt im Wesentlichen auf das erste veröffentlichte Werke Berkeleys zu diesem Thema konzentrieren, da ein zentraler Teil dieses Werkes der Distanzwahrnehmung gewidmet ist und Berkeley explizit auf verschiedene Aspekte der im letzten Abschnitt dargestellten Theorie der Distanzwahrnehmung des Descartes Bezug nimmt.

In dem 1709 verfassten „*Essay Towards a New Theory of Vision*“ (im Folgenden kurz als *Essay* bezeichnet), der enorm wichtige Beiträge zur theoretischen Behandlung visueller Wahrnehmung, insbesondere der Distanzwahrnehmung lieferte, aber leider oft nur unter metaphysischem Blickwinkel rezipiert wurde (Atherton, 1990), gibt der zu dieser Zeit vierundzwanzigjährige Berkeley zu Beginn das Ziel seines Versuches an:

Mein Ziel ist es zu zeigen, auf welche Weise wir durch den Gesichtssinn die Entfernung, Größe und Lage von Objekten wahrnehmen, auch den Unterschied zu betrachten, der zwischen den Vorstellungen des Gesichtssinnes und denen des Tastsinnes besteht, und zu sehen, ob es irgendeine Vorstellung gibt, die beiden Sinnen gemeinsam ist. (Berkeley, 1987, S. 13)

Aufgrund der Ausrichtung dieser Arbeit wollen wir uns im Folgenden auch hier im Wesentlichen mit Berkeleys Theorie der Distanzwahrnehmung beschäftigen und den zweiten großen Teil des *Essay*, der sich mit der Frage beschäftigt, „ob es irgendeine Vorstellung gibt, die beiden Sinnen gemeinsam ist“, vernachlässigen.

Große Teile des *Essay* lassen sich als eine Auseinandersetzung Berkeleys mit theoretischen Vorstellungen lesen, die einen *optisch-geometrischen Ansatz* zur Erklärung der Distanzwahrnehmung verfolgen, also die wahrgenommene Entfernung von Objekten durch geometrische Regularitäten der abstrakten Sehsituation zu erklären suchen (Grush, 2007). Wie im vorangehenden Abschnitt dargestellt, behauptete etwa Descartes, dass bei binokularer Fixation eines Objektes über *natürliche Geometrie* die

Entfernung des fixierten Objektes ermittelt werden könne, was voraussetzt, dass in irgendeiner Weise ein Zugang zur Länge jener Strecke, die sich aus Verbindung beider Augen ergibt, und der Größe der zwischen den Sehlinien und dieser Strecke eingeschlossenen Winkel besteht.<sup>22</sup>

Die Motivation Berkeleys, der *geometrischen Theorie* der Tiefenwahrnehmung eine eigene, neue gegenüberzustellen, ist darin begründet, dass Berkeley im geometrischen Ansatz und insbesondere dessen Schlussfolgerungen bezüglich des ontologischen Status wahrgenommener Objekte im Raum inakzeptable Probleme zu finden meinte. Es ist in diesem Zusammenhang wichtig, dass sich Berkeley nicht gegen jegliche optisch-geometrischen Überlegungen wandte, sondern allein gegen Theorien, die aus optischen und geometrischen Gesetzmäßigkeiten ableiten wollen, *wie* wir wahrnehmen:

Zu erklären, wie der Geist oder die Seele des Menschen einfach nur sieht, ist eine Sache und gehört zur Philosophie. Teilchen zu betrachten, wie sie sich auf bestimmten Linien bewegen, Lichtstrahlen, wie sie gebrochen oder reflektiert werden, sich kreuzen oder Winkel bilden, ist eine ganz andere Sache und gehört zur Geometrie. Den Gesichtssinn durch den Mechanismus des Auges zu erklären, ist eine dritte Sache, die zur Anatomie und zur Experimentalwissenschaft gehört. Diese beiden letzten Forschungsbereiche sind in der Praxis nützlich, um in Übereinstimmung mit den Naturgesetzen, die in diesem Weltsystem gelten, bei Fehlern des Gesichtssinns zu helfen und seine Störungen zu beheben. Die erstgenannte Theorie ist aber die, die uns die wahre Natur des Sehens, aufgefaßt als eine Fähigkeit der Seele verstehen lässt. (Berkeley, 1987, S. 122)

Auf welche Theorie Berkeley sich genau bezieht, wenn er von einer geometrischen Theorie der Wahrnehmung spricht, ist unter Wissenschafts- und Philosophiehistorikern nicht unumstritten, es gibt jedoch ernstzunehmende Hinweise, dass es sich um eine bestimmte, zu Berkeleys Zeit offenbar weit verbreitete Variante der in der *Dioptrique* des Descartes ausgeführten Theorie der Tiefenwahrnehmung (bzw. bestimmte Aspekte dieser) handelt.<sup>23</sup>

---

<sup>22</sup>Die Ursprünge solcher geometrischer Erwägungen lassen sich bis in Platos Akademie in Athen und dort insbesondere auf die Geometer Eudoxos von Knidos (ca. 408-355 v.Chr.) und Theaitetos (ca. 417-369 v.Chr.) zurückführen (Howard, 2012). Als erste systematische Behandlung eines geometrischen Ansatzes zur Erklärung des Sehens gelten die beiden Euklid von Alexandria (ca. 323-285 v.Chr.) zugeschriebenen optischen Schriften *ὀπτικά* und *κατοπτρικά*, welche das Fundament für die weitere Entwicklung einer geometrischen Optik über Alhazen bis hin zu Kepler darstellen (siehe dazu Lindberg, 1976 sowie Burton, 1945)

<sup>23</sup>So schreibt Berkeley im Anhang des *Essay* als Reaktion auf Kritiker, die ihm vorwarfen, dass es die Theorie, gegen die er sich wende, überhaupt nicht gebe: „Und, um zu zeigen, daß ich hier nicht mit meinem eigenen Schatten fechte, will ich hier eine Stelle des berühmten Descartes wiedergeben“ (Berkeley, 1987, S. 90), worauf ein Zitat aus dem sechsten Diskurs aus Descartes' *Dioptrique* folgt. Es muss allerdings auch festgestellt werden, dass Berkeley hier dem Leser *selektiv* bestimmte Aspekte

Laut Atherton (1990) ist es in der Literatur über den *Essay* eine weit verbreitete Ansicht, dass Berkeley dort überhaupt keine positiv ausformulierte Theorie der Tiefen- bzw. Distanzwahrnehmung ausarbeite, da er davon ausgehe, dass wir die Entfernung eines Objektes überhaupt nicht wahrnehmen könnten, wobei als Beleg meist die folgende, berühmt gewordene Passage aus Abschnitt 2 des *Essay* herangezogen wird:

Alle stimmen, wie ich meine, darin überein, daß Entfernung nicht an sich und unmittelbar gesehen werden kann, denn Entfernung ist eine mit dem Ende auf das Auge gerichtete Linie. Ihre Projektion auf dem Augenhintergrund ergibt nur einen Punkt, der unverändert derselbe bleibt, ob die Entfernung größer oder kleiner ist. (Berkeley, 1987, S. 13)

Berkeley verweist hier auf die geometrische Tatsache, dass alle Punkte, die auf einer Sehlinie liegen, auf denselben Punkt der Retina abgebildet oder projiziert werden.<sup>24</sup> An einer anderen Stelle des *Essay* stellt Berkeley darüber hinaus richtigerweise fest, dass ein beliebiges Lichtmuster auf der Retina von unendlich vielen möglichen Objekten unterschiedlicher Größe, Entfernung und Form stammen kann. Geometrisch ist dies unmittelbar klar, da es sich um eine Projektion von Punkten eines dreidimensionalen Raumes auf eine Ebene handelt.<sup>25</sup> Man beachte, dass in diesem Argument implizit starke Annahmen über die zugehörige Sehsituation sowie die Funktionsweise visueller Wahrnehmung überhaupt enthalten sind. So wird vorausgesetzt, dass sich der Wahrnehmende nicht bewegt, abstrakte Punkte betrachtet, mit nur einem Auge sieht und auch dieses nicht bewegt wird. Weiterhin wird vorausgesetzt, dass ‘etwas unmittelbar sehen’ nichts anderes bedeutet, als das auf die Retina projizierte Lichtmuster zu sehen. Merkwürdigerweise scheint sich Berkeley hier fundamentaler Annahmen eben jenes geometrischen Ansatzes zu bedienen, den er zu kritisieren sucht. Es ist letztlich nicht zu entscheiden, ob Berkeley hier selbst ein Argument *gegen* eine geometrische Theorie des Sehens vorbringen will oder auf ein geometrisches Problem referiert, welches den Vertretern einer geometrischen Theorie bewusst war („Alle stimmen ... darin überein“) und sie, wie im Fall des Descartes, dazu führte, zusätzliche Mittel zur Wahrnehmung von Entfernung annehmen zu müssen.

Bemerkenswert und für die weitere Diskussion von Wichtigkeit ist allerdings der Status, den Berkeley dem auf die Retina projizierten Lichtmuster zuweist. Berkeley spricht in der Regel nicht von einem ‘Netzhautbild’, in dem sich mögliche Formen

---

dieser Theorie des Descartes, insbesondere das Konzept der Konvergenz als Mittel zur Distanzwahrnehmung, als einheitliche und eigenständige Theorie verkaufen möchte.

<sup>24</sup>Dieser Punkt wurde schon vor Berkeley, etwa von Nicolas Malebranche in seiner 1674 erschienen Schrift „*De la recherche de la vérité*“ (Malebranche, 1674/1968) festgestellt, allerdings im Zusammenhang mit der Wahrnehmung von Bewegung (vgl. Atherton, 1990).

<sup>25</sup>Dieses Problem der nicht eindeutig bestimmbar Quelle eines auf die Retina projizierten Lichtmusters ist in der aktuellen Literatur auch als *inverse projection problem* oder *inverse optics problem* bekannt (siehe etwa Epstein, 1995; Pizlo, 2001).

oder gar Objekte identifizieren ließen, sondern charakterisiert dieses eher als eine mehr oder minder chaotische Ansammlung von Farben verschiedener Helligkeiten und dies allein sei es, was ‘unmittelbar’ oder eigentlich mit dem Sehsinn wahrgenommen werden könne:

All das, was mit dem Sehvermögen eigentlich wahrgenommen wird, läuft auf nichts anderes hinaus als auf Farben mit ihren Variationen und verschiedenen Verhältnissen von Licht und Schatten. Die dauernde Unbeständigkeit und Flüchtigkeit dieser unmittelbaren Objekte des Gesichtssinnes macht es unmöglich, sie nach Art der geometrischen Figuren zu behandeln. (Berkeley, 1987, S. 87)

Es sei aber nun auch nicht von der Hand zu weisen, wie Berkeley in Abschnitt 11 des *Essay* feststellt, *dass* wir Objekte als verschieden weit von uns entfernt sehen:

Nun ist es aus Abschnitt 2 klar, daß Entfernung ihrer eigenen Natur nach unwahrnehmbar ist, und doch wird sie durch den Gesichtssinn wahrgenommen. Es bleibt daher nur übrig, daß sie mittels einer anderen Vorstellung, die selbst unmittelbar im Akt des Sehens wahrgenommen wird, sichtbar gemacht wird. (Berkeley, 1987, S. 15)

Welche Erklärung bietet Berkeley für die wahrgenommene Distanz von Objekten an? Da Distanz „ihrer eigenen Natur nach unwahrnehmbar ist“, wird Berkeley ebenfalls auf eine Art Hilfsmittel der Distanzwahrnehmung zurückgreifen müssen. Es wird sich herausstellen, dass Berkeley auf Ideen des Descartes, insbesondere die Konvergenz als Mittel der Tiefenwahrnehmung, zurückgreifen wird, die jedoch auf eine deutlich andere Weise aufgefasst werden.

Die nächsten Abschnitte des *Essay* sind der weitergehenden Kritik der geometrischen Theorie der Distanzwahrnehmung gewidmet. Nachdem Berkeley zu Beginn eine ‘unmittelbare’ Wahrnehmung der Distanz (als Strecke auf der Sehlinie zwischen *einem* Auge und dem Objekt) durch den Sehsinn ausgeschlossen hat, widmet er sich nun der von Descartes in der *Dioptrique* vorgebrachten Konvergenz als Mittel zur Distanzwahrnehmung, also die Auffassung, dass die Entfernung eines Objektes über die Größe des zwischen den Sehlinien *beider* Augen eingeschlossenen Winkels berechnet werden kann, da sich diese Größe in Abhängigkeit der Entfernung des mit beiden Augen fixierten Objektes verändert. Darüber hinaus diskutiert Berkeley den ebenfalls bei Descartes auffindbaren Punkt, dass wir Entfernung nicht nur über eingeschlossene Winkel berechnen, sondern auch Entfernungsschätzungen mit Hilfe bestimmter anderer ‘Hinweise’ vornehmen könnten, etwa der aufgrund bestimmter Erfahrungen bekannten Größe der Objekte.<sup>26</sup>

---

<sup>26</sup>„Ich stelle fest, man gibt auch zu, daß die Schätzung, die wir über die Entfernung eines erheblich entfernteren Objektes machen, eher ein auf Erfahrung als auf die Sinne gegründeter Urteilsakt ist.“

In diesem Zusammenhang trifft Berkeley eine für den weiteren Verlauf der Argumentation des *Essay* wichtige Unterscheidung, indem er die zu seiner Zeit diskutierten Mittel der Distanzwahrnehmung in zwei Kategorien unterteilt. Bei den ‘rein geometrischen Hinweisen’ bestehe, etwa im Fall der Konvergenz, eine *notwendige* Verknüpfung zwischen der Distanz des fixierten Objektes und dem durch die Sehlinien eingeschlossenen Winkel, bei allen anderen bestehe hingegen keine notwendige Verknüpfung, sie würden sich allein durch Erfahrung ergeben.<sup>27</sup> In einem allgemeineren Rahmen entsprechen diese beiden Kategorien zudem zwei verschiedenen Annahmen über die allgemeine ‘Funktionsweise’ der Tiefenwahrnehmung. Im ersten Fall ist dies die Annahme, dass Distanzen erfahrungsunabhängig, gleichsam ‘direkt’, unter Verwendung *geometrischer Gesetze* wahrgenommen werden können, im zweiten Fall ist es die Annahme, dass Tiefenwahrnehmung ‘erlernt’ werden müsse.

Berkeleys Ziel ist es, in einem ersten Schritt diese erfahrungsunabhängige, rein geometrische Theorie der Entfernungswahrnehmung zu widerlegen, indem er in den Abschnitten 9 bis 15 des *Essay* folgendes Argument aufbaut: Entfernung, die aufgrund der oben genannten Gründe selbst nicht unmittelbar wahrnehmbar sei, müsse mit Hilfe der Wahrnehmung von etwas davon Verschiedenem, einer „anderen Vorstellung, die selbst unmittelbar im Akt des Sehens wahrgenommen wird“ (Berkeley, 1987, S. 15), wahrgenommen werden.<sup>28</sup> Nichts, was selbst nicht unmittelbar wahrgenommen wird, könne hierfür in Frage kommen. Nun verhalte es sich aber so, dass wir die geometrischen Regularitäten der Sehsituation, die Linien und eingeschlossenen Winkel, überhaupt nicht wahrnehmen *können*, also könne die Wahrnehmung der Distanz auch nicht über diese geschehen.

Berkeley sieht dies unter anderem auch darin begründet, dass andernfalls Annahmen über ein umfangreiches geometrisches Wissen des Wahrnehmenden gemacht werden

---

Wenn ich zum Beispiel eine große Anzahl von dazwischenliegenden Objekten wahrnehme ... , von denen ich aus Erfahrung weiß, daß sie einen erheblichen Raum einnehmen, dann urteile ich oder ziehe den Schluß, daß sich das Objekt, das ich hinter ihnen sehe, in einer großen Entfernung befindet“ (Berkeley, 1987, S. 13).

<sup>27</sup>„Zwischen dieser und der vorausgegangenen Art der Entfernungsschätzung gibt es folgenden bemerkenswerten Unterschied: Während es keine offensichtliche, notwendige Verknüpfung zwischen kleiner Entfernung und einer großen und starken Erscheinung gab, oder zwischen großer Entfernung und einer kleinen und schwachen Erscheinung, scheint ein notwendiger Zusammenhang zwischen einem stumpfen Winkel und geringer Entfernung wie auch zwischen einem spitzen Winkel und größerer Entfernung zu bestehen. Es hängt nicht im geringsten von der Erfahrung ab, sondern kann von jedem, bevor er davon eine Erfahrung hat, erkannt werden“ (Berkeley, 1987, S. 14).

<sup>28</sup>Ein Beispiel, das Berkeley (1987) in diesem Zusammenhang immer wieder anführt ist das folgende:

So sind zum Beispiel die Gemütsregungen, die sich im Geiste eines anderen befinden, an sich unsichtbar. Ich kann sie trotzdem durch den Gesichtssinn wahrnehmen, wenn auch nicht unmittelbar, so doch mittels der Farben, die sie im Gesicht hervorrufen. Wir sehen oft Scham oder Furcht im Aussehen eines Menschen, indem wir die Veränderungen seines Gesichtes wahrnehmen, wenn er errötet oder blaß wird. (S. 15)

müssten, die ihm untragbar scheinen. Der alltäglich Wahrnehmende sei üblicherweise niemand, der sich (wie die Theoretiker der Optik) in optischen und geometrischen Gesetzen auskenne, und dennoch habe er keinerlei Probleme, ohne Anstrengung verschiedene Objekte als unterschiedlich weit von ihm entfernt wahrzunehmen. Darüberhinaus könne jeder an sich selbst feststellen, dass er beim Sehen wohl kaum geometrische Berechnungen anstellt, sondern es geschehe ‘einfach so’, ohne unser Zutun.<sup>29</sup>

Doch jene Linien und Winkel, durch die einige Leute vorgeben, die Entfernungswahrnehmung zu erklären, werden selbst überhaupt nicht wahrgenommen, und sie werden von denen, die in der Optik nicht bewandert sind, niemals wirklich in Erwägung gezogen. Ich berufe mich auf jedermanns Erfahrung bezüglich der Frage, ob er beim Sehen eines Objektes dessen Entfernung durch die Größe eines Winkels berechnet, der durch die beiden optischen Achsen gebildet wird, oder ob er jemals an die größere oder geringere Divergenz der Strahlen denkt, die von irgendeinem Punkt zu seiner Pupille gelangen. (Berkeley, 1987, S. 15)

Das Hauptgewicht dieses Argumentes liegt aber offenbar in dem Punkt, dass wir die abstrakte Geometrie der Sehsituation selbst nicht wahrnehmen könnten und diese daher zur Erklärung der Distanzwahrnehmung prinzipiell nicht in Frage komme. Die Linien und Winkel, über die wir laut dem geometrischen Ansatz die Entfernung von Objekten wahrnehmen würden, hätten „keine wirkliche Existenz in der Natur“ (Berkeley, 1987, S. 16), sondern seien lediglich von Mathematikern in die Optik eingeführte *Hypothesen*, um die Distanzwahrnehmung *unter geometrischen Gesichtspunkten* behandeln zu können. Wenn ich etwa zwei relativ nahe, aber unterschiedlich weit entfernte Objekte mit beiden Augen fixiere, so könne ich zwar feststellen, dass sich die Stellung meiner beiden Augen zueinander in Abhängigkeit vom fixierten Objekt verändere und ich könne, entsprechendes geometrisches Wissen vorausgesetzt, eine abstrakte Zeichnung der geometrischen Situation anfertigen, in der Sehlinien und eingeschlossene Winkel vorkommen, ich könne aber keinerlei eingeschlossene Winkel und Linien über den Sehsinn *wahrnehmen*, da sie diesem *als solche* nicht verfügbar seien. Und selbst wenn man zugeben würde, dass die geometrischen Linien und Winkel tatsächlich in der Natur auffindbar wären *und* es auch möglich wäre, diese wahrzunehmen, so gäbe es laut Berkeley dennoch verschiedene Wahrnehmungsphänomene, die sich auf diese Weise nicht erklären ließen.<sup>30</sup>

<sup>29</sup>Man beachte, dass ein wesentlicher Teil dieses Argumentes darauf basiert, dass die von Berkeley kritisierte Theorie die Annahme machen würde, dass der Wahrnehmende sich des ‘Berechnungsprozesses’ in irgendeiner Weise bewusst wäre. Wie bei den Ausführungen zu Descartes’ Theorie dargestellt wurde, wird dies jedoch nicht explizit angenommen, sondern bleibt eher im Dunkeln.

<sup>30</sup>Berkeley spielt hier unter anderem auf die sogenannte *Barrow-Illusion* an, die sich in Kürze folgendermaßen beschreiben lässt (vgl. *Essay* Abschnitt 29-40, siehe auch Atherton, 1990, S. 86ff): Bei

Es bleibt also die Frage, wie Berkeley die Wahrnehmung von Distanz zu erklären sucht. Was ist jene „andere Vorstellung“, die unmittelbar wahrgenommen werden muss und Distanzwahrnehmung ermöglichen soll? Sein Vorschlag läuft darauf hinaus, dass er den von Descartes angenommenen Mitteln der Distanzwahrnehmung einen anderen konzeptuellen Status gibt. Essentiell ist dabei Berkeleys Behauptung, dass Distanzwahrnehmung nicht über den Sehsinn geschehe, sondern *allein* über den Tastsinn.

In erster Linie sei Distanzwahrnehmung über die Wahrnehmung der *Bewegungen der Augenmuskeln* möglich. Dass sich die Stellung der Augen zueinander bei Fixation verschieden weit entfernter Objekte verändere, könne – im Gegensatz zu den Sehlinien und eingeschlossenen Winkeln – unmittelbar wahrgenommen werden. Unterschiedlich wahrgenommene Stellungen der Augen würden zur Wahrnehmung unterschiedlich weit entfernter Objekte führen. Es ist an diesem Punkt wichtig, sich klar zu machen, worauf sich die Kritik Berkeleys an der geometrischen Theorie genau bezieht, nämlich die Behauptung, dass wir über eine direkte Wahrnehmung der *abstrakten Geometrie der Sehsituation* in der Lage seien, verschiedene Distanzen wahrzunehmen. Berkeley nimmt nun zwar auch an, dass die unterschiedliche Stellung der Augen zueinander bei Fixation unterschiedlich weit entfernter Objekte ein Mittel zur Distanzwahrnehmung ist, interpretiert diese physiologische Tatsache aber nicht abstrakt geometrisch. Wir nehmen eben nicht, so Berkeley, die eingeschlossenen Winkel zwischen den Sehlinien beider Augen über den Sehsinn wahr und ‘berechnen’ dann die Distanz des fixierten Objektes, sondern was wir allein wahrnehmen seien verschiedene Bewegungen der Augenmuskeln bei Fixation verschiedener Objekte.<sup>31</sup> Die Wahrnehmung unterschiedlicher Augenstellungen bei unterschiedlicher Fixation sei eine unmittelbare, es müsse weder etwas berechnet oder erschlossen werden. Der Vorteil dieser Herangehensweise liegt auf der Hand: Wenn man den Wahrnehmenden, das Auge, den Sehsinn oder den Cartesischen Geist nicht zum Lösen geometrischer Probleme heranziehen muss, entbindet man sich der naheliegenden Frage, wie die entsprechende Instanz zu den benötigten geometrischen Gesetzen kommt und weshalb uns diese geometrische Analyse beim Wahrnehmungsprozess nicht zugänglich ist.

Diese Auffassung wirft allerdings die Frage auf, wie man sich die Verbindung zwischen der Wahrnehmung unterschiedlicher Muskelaktivitäten und dem *Sehen* verschieden weit entfernter Objekte vorzustellen habe. Berkeley betont mehrfach, dass es sich hierbei nicht um eine notwendige Verbindung zwischen Wahrnehmung der Augenbe-

---

Betrachtung eines Objektes durch ein Vergrößerungsglas nimmt die wahrgenommene Entfernung des Objektes ab, wenn die sich die Distanz zwischen Augen und Vergrößerungsglas erhöht, ein für eine rein geometrische Theorie unerklärbares Phänomen.

<sup>31</sup>„Diese Neigung oder Drehung der Augen ist von einer Empfindung begleitet, die mir das zu sein scheint, was in diesem Falle *die Vorstellung größerer oder geringerer Entfernung mit sich bringt*.“ (Berkeley, 1987, S. 16)

wegungen und Wahrnehmung näher oder weiter entfernt liegender Objekte handle, sondern um eine rein *gewohnheitsmäßige*:

Weil der Geist durch beständige Erfahrung gefunden hat, daß die verschiedenen Empfindungen die den verschiedenen Neigungen der Augen entsprechen, jeweils von einem anderen Entfernungsgrad im Objekt begleitet sind, ist da eine habituelle und gewohnheitsmäßige Verknüpfung zwischen diesen beiden Vorstellungsarten entstanden, so daß der Geist, sobald er die Empfindung wahrnimmt, die daraus entsteht, daß er den Augen eine unterschiedliche Drehung erteilt, um die Pupillen näher zusammen oder weiter auseinander zu bringen, dabei sogleich die entsprechende Entfernungsvorstellung wahrnimmt, die mit dieser Empfindung verknüpft zu sein pflegte. (Berkeley, 1987, S. 16)

Durch ein wiederholtes *gemeinsames Auftreten* der Wahrnehmung bestimmter Augenbewegungen und verschieden weit vom Wahrnehmenden entfernter Objekte sei eine Verknüpfung derart entstanden ist, dass die Wahrnehmung dieser Augenbewegungen die Vorstellung unterschiedlich weit entfernter Objekte zur Folge habe, oder diese „suggeriere“, wie Berkeley sich auch ausdrückt. Dies gelte auch für alle anderen Mittel der Tiefenwahrnehmung. Keines dieser Mittel weise notwendigerweise eine Verbindung zur Wahrnehmung unterschiedlich weit entfernter Objekte auf, sondern die Verbindung sei allein dadurch entstanden, dass diese wiederholt gemeinsam mit der Wahrnehmung unterschiedlich weit entfernter Objekte aufgetreten wären.<sup>32</sup>

Diese Erklärung Berkeleys wirkt auf den ersten Blick verwirrend, da hier offenbar die Wahrnehmung von Distanz damit erklärt werden soll, dass die Wahrnehmung verschiedener Bewegungen der Augenmuskeln wiederholt mit der *Wahrnehmung unterschiedlicher Distanzen* aufgetreten wäre und sich so eine nicht notwendige, gewohnheitsmäßige Verbindung zwischen beiden gebildet habe. An diesem Punkt zeigt sich nun das Revolutionäre an Berkeleys Ansatz. Die Verbindung zwischen den verschiedenen wahrgenommenen Bewegungen der Augenmuskeln und der wahrgenommenen Entfernung komme allein mittels der Wahrnehmung verschiedener Distanzen *über den Tastsinn* zustande. Durch das wiederholte gemeinsame Auftreten der Wahrnehmung verschiedener Stellungen der Augen zueinander und den über den Tastsinn wahrgenommenen unterschiedlichen Distanzen der Objekte komme es dazu, dass sich eine Verbindung

---

<sup>32</sup>„Es ist in den meisten Fällen wahr, daß noch verschiedene andere Umstände zur Bildung unserer Vorstellung von Entfernung beitragen, nämlich die spezielle Anzahl, Größe, Art usw. der gesehenen Dinge. Über sie, wie auch alle anderen Anlässe, welche Entfernung suggerieren, will ich nur bemerken, daß keiner von ihnen aufgrund seiner eigenen Natur irgendeine Beziehung zu ihr oder eine Verknüpfung mit ihr besitzt. Es ist auch nicht möglich, daß sie jemals die verschiedenen Grade der Entfernung bezeichnen, wenn nicht durch Erfahrung gefunden wurde, daß sie mit ihnen verknüpft sind“ (Berkeley, 1987, S. 20).

zwischen diesen Wahrnehmungen herausbilde, die sich dem Wahrnehmenden so darstelle, als ob Entfernung gesehen werde. Berkeley geht sogar noch einen Schritt weiter, indem er behauptet, dass „weder Entfernung noch Dinge, die sich in einer Entfernung befinden, ... wirklich durch den Gesichtssinn wahrgenommen [werden], auch nicht ihre Vorstellungen“ (Berkeley, 1987, S. 32). Was kann damit gemeint sein?

Wenn wir davon sprechen, dass wir ein Objekt mit einer bestimmten Form in einer bestimmten Entfernung *sehen*, so sei das, worauf wir uns da beziehen, laut Berkeley nicht etwas, was wir im eigentlichen Sinne sehen, d.h. allein über unseren Sehsinn wahrnehmen, sondern es handle sich um eine durch Wahrnehmungen aus dem Tastsinn *suggestierte* Vorstellung. Der Wahrnehmende, so Berkeley, falle einem auf ungenauer Verwendung unserer Sprache im Alltag sowie ununterbrochener Gewohnheit basierenden Vorurteil zum Opfer, wenn er glaube, dass das, was er *sieht*, ein Objekt sei, welches sich in einer Entfernung von ihm befinde. Die Vorstellung von Objekten, die sich in einer gewissen Entfernung von uns befinden, sei, ebenso wie die Vorstellung eines ‘Außerhalb’ überhaupt, streng genommen nichts, was sich durch den Sehsinn wahrnehmen lasse.

Zur Verdeutlichung führt Berkeley ein Beispiel an, das er auf äußerst subtile Weise analysiert. Wenn jemand etwa davon spreche, dass er gerade den Mond *sehe* als sich so und so weit von ihm entfernt befindend, dann könne er damit nicht das von ihm *gesehene* Objekt meinen. Bewege sich nämlich derjenige nun immer weiter auf das von ihm Gesehene zu, so verändere sich das, was er sehe, auf eklatante Weise, so dass es jenes von ihm Gesehene, von dem er vorher behauptet hatte, dass es so und so weit entfernt sei, gar nicht mehr gebe, oder es, wie Berkeley sich ausdrückt, „verschwunden“ sei. Der Ausdruck, dass man den Mond so und so weit entfernt *sehe*, sei irreführend. Was dieser Ausdruck eigentlich bedeute sei, dass man eine so und so weite Strecke zurücklegen müsse, um mit dem Gesehenen in *haptischen Kontakt* zu kommen. Entfernung sei allein durch den Tastsinn unmittelbar wahrnehmbar, indem ich etwa meine Hände ausstrecke und das Objekt *fühle* oder die Strecke zwischen mir und dem Objekt zurücklege und dann das Objekt fühlen kann. Allein der Tastsinn, so die verblüffende Hypothese Berkeleys, sei in der Lage, Entfernungen tatsächlich wahrzunehmen. Wenn jemand gleichsam unbedacht sagt, er *sehe* ein bestimmtes Objekt in einer gewissen Entfernung, dann könne damit eigentlich nur gemeint sein,

daß das, was er sieht, seinem Verstand nur suggeriert, nach dem Durchlaufen einer gewissen Entfernung, die durch die mit dem Tastsinn wahrnehmbare Bewegung seines Körpers zu messen ist, werde er zur Wahrnehmung der und der tastbaren Vorstellungen gelangen, die gewöhnlich mit den und den sichtbaren Vorstellungen verknüpft waren. (Berkeley, 1987, S. 32)

Die Verknüpfung sei dabei eine rein arbiträre, auf Gewohnheit basierende. Berkeley verwendet in diesem Zusammenhang eine Analogie zur *Sprache*, die im Kontext wahrnehmungstheoretischer Ansätze noch häufiger auftauchen wird. Äußere etwa jemand das Wort „Kutsche“, dann sei das, was wir über unser Gehör unmittelbar wahrnehmen könnten, nichts anderes als ein Geräusch oder ein Klang. Als Sprecher der deutschen Sprache habe sich jedoch durch wiederholtes gemeinsames Auftreten eine Verknüpfung zwischen diesem durch unser Gehör unmittelbar wahrnehmbaren Geräusch und dem *Objekt*, was das Wort *bezeichnen* soll, gebildet, so dass das Hören des Geräuschs unserem Verstand das *Objekt* „Kutsche“ suggeriere. Es sei uns geradezu unmöglich, dass wir in diesem Fall dieses Geräusch hören, *ohne* dass gleichzeitig und ohne unser Zutun das Bezeichnete, die Bedeutung des Wortes gleichsam ‘mitgedacht’ werde:

Sobald wir die Worte, die in einer uns vertrauten Sprache ausgesprochen werden, in unseren Ohren hören, bieten sich unserem Geist sofort von selbst die ihnen entsprechenden Vorstellungen dar. Der Klang und die Bedeutung kommen in ein und demselben Augenblick in unseren Verstand. Sie sind so eng vereint, daß es nicht in unserer Macht steht, das eine fernzuhalten, es sei denn, wir schließen auch das andere aus. Wir handeln sogar in jeder Hinsicht so, als hörten wir wirklich die Gedanken selbst. (Berkeley, 1987, S. 34)

Dass es sich hier aber um zwei sorgfältig auseinanderzuhaltende Konzepte handle, sei evident, wenn wir etwa Lautäußerungen in einer uns fremden Sprache hörten. Was wir dann nämlich hören würden sei in der Tat das, was durch unser Gehör unmittelbar wahrnehmbar sei, nämlich für uns bedeutungslose Geräusche. Da in diesem Fall keine gewohnheitsmäßige Verknüpfung zwischen diesen Geräuschen und dem, was sie bezeichnen sollen, bestehe, stelle sich beim Hören auch keine zugehörige Bedeutung oder eine Vorstellung des Bezeichneten ein.

Die Verbindung zwischen der Zeichen- oder Lautkombination „Kutsche“ und dem Bezeichneten sei vollkommen willkürlich, es bestehe keine notwendige Verknüpfung oder auch nur eine Ähnlichkeit zwischen der Lautäußerung und dem, was diese bezeichnen solle.<sup>33</sup> Und ebenso wie das Wort „Kutsche“ eine solche bezeichne und man beim Hören oder Lesen dieses Wortes gleichsam automatisch von der Bezeichnung auf das Bezeichnete übergehe, das Wort oder die entsprechende Lautäußerung aber keinesfalls mit dem gemeinten Objekt identisch sei oder auch nur Ähnlichkeit mit ihm haben müsse, so gehe man im Fall der Distanzwahrnehmung auch automatisch von dem, was unmittelbar über den Sehsinn wahrgenommen werden könne zu den durch den Tastsinn

<sup>33</sup>Das Berkeleysche Verständnis von Sprache als einer im Wesentlichen rein willkürlichen Verknüpfung zwischen Zeichen und Bezeichnetem findet in der heutigen Sprachwissenschaft nur noch geringen Anklang, siehe dazu etwa Greenberg (1963).

vermittelten und mit den entsprechenden Wahrnehmungen des Sehsinnes wiederholt gemeinsam aufgetretenen Entfernungsvorstellungen über. Es sei jedoch für „jeden, der die wahre Natur des Sehens verstehen will“, äußerst wichtig, diese beiden Sachen konzeptuell auseinanderzuhalten, da wir aufgrund unseres ungenauen Sprachgebrauches sowie Mangel an Reflexion „dazu neigen, zu den ersteren zu rechnen, was nur zu den letzteren gehört“ (Berkeley, 1987, S. 35). Dies werde laut Berkeley aber unter anderem dadurch erschwert, dass wir üblicherweise das Bezeichnende und das Bezeichnete mit denselben Namen benennen.

Die durch Seh- und Tastsinn vermittelten Vorstellungen seien also wie Zeichenkombination und Bezeichnetes vollkommen verschieden voneinander und allein durch wiederholtes gemeinsames Auftreten miteinander verbunden. Aber dadurch, dass wir ihr gleichzeitiges Auftreten gewohnt seien, sei uns nicht möglich, beide zu trennen und wir würden sie wie *ein- und dasselbe* behandeln, was erkläre, weshalb die sprachliche Verwirrung entstanden sei, dass wir Entfernung *sehen* würden. Berkeley verdeutlicht dies zusätzlich, indem er anführt, dass sich die Wahrnehmung von Entfernung ‘über den Sehsinn’ in nichts von der Wahrnehmung von Entfernung ‘über den Hörsinn’ unterscheide, man beide in der Alltagssprache aber unterschiedlich behandle. Streng genommen könne man weder vom Gesichtssinn noch vom Hörsinn sagen, dass sie Entfernung wahrnehmen, da in beiden Fällen lediglich eine gewohnheitsmäßige Verknüpfung mit entsprechenden Vorstellungen des Tastsinns vorliege, der eigentlichen Instanz der Entfernungswahrnehmung. Wenn ich etwa, so das Beispiel Berkeleys, eine Kutsche die Straße entlang fahren höre, so könne ich die verschiedenen Entfernungen der Kutsche über die verschiedenen Geräusche ‘wahrnehmen’, da diese mit bestimmten Vorstellungen des Tastsinns verbunden sind (etwa dass ich jetzt so und so weit von meinem Standpunkt aus laufen müsste, um die Kutsche zu erreichen). Dennoch würde ich in dem Fall nicht im gleichen Sinne wie beim Sehen sagen: „Ich höre Entfernung“ und niemand würde behaupten, die durch Hörsinn und Tastsinn vermittelten Vorstellungen seien *dieselben*. Im Gegensatz zum Sehen sei man beim Hören viel weniger geneigt, anzuzweifeln, dass „Körper und äußere Dinge nicht die eigentlichen Objekte des Hörens sind, sondern dies bloß Laute sind, durch deren Vermittlung unserem Denken die Vorstellung dieses oder jenes Körpers bzw. dieser oder jener Entfernung suggeriert wird“ (Berkeley, 1987, S. 33). Die besondere sprachliche Verwirrung im Falle des Sehsinnes sei darin begründet, dass die über den Sehsinn vermittelten Vorstellungen so leicht mit denen des Tastsinns zu verwechseln seien, was noch dadurch unterstützt werde, dass wir ‘gefühlte Dinge’ und ‘gesehene Dinge’ mit denselben Begriffen bezeichnen. Es stehe

aber ohne Zweifel fest: „man sieht und fühlt nicht ein und dasselbe Ding mehr als man ein und dasselbe Ding hört und fühlt“ (Berkeley, 1987, S. 33).<sup>34</sup>

### Zusammenfassung

Zusammenfassend formuliert Berkeley im *Essay* einerseits eine wirkungsvolle Kritik an einem rein geometrischen Ansatz zur Erklärung der Distanzwahrnehmung und entwickelt andererseits selbst eine Theorie der Distanzwahrnehmung, die die zu dieser Zeit diskutierten Mittel der Distanzwahrnehmung auf radikale Weise neu interpretiert. Die Kritik am geometrischen Ansatz zielt im Wesentlichen darauf ab, dass erstens dieser fälschlicherweise davon ausgehe, dass wir die abstrakten geometrischen Regularitäten der Sehsituation selbst wahrnehmen würden und dass zweitens, selbst wenn wir diese wahrnehmen würden, ein rein geometrischer Ansatz erklären müsse, welche Instanz für die Lösung der geometrischen Probleme zuständig sein soll, woher diese die optischen Gesetzmäßigkeiten kenne und wieso der Lösungsprozess dem Wahrnehmenden nicht zugänglich sei. Nur was durch unsere Sinne wahrnehmbar sei könne wahrgenommen werden, die von den Augen ausgehenden Linien und die eingeschlossenen Winkel seien jedoch nicht wahrnehmbar. Die Ausführungen der Vertreter geometrischer Theorien seien zwar aus geometrischer Sicht vollkommen richtig und auch von theoretischem und praktischem Wert (etwa bei der Herstellung von Vergrößerungsgläsern) hätten aber nichts mit der Situation des Wahrnehmenden zu tun, sondern seien eine *geometrische Abstraktion*. Wir könnten keine von unseren Augen ausgehenden Linien, die sich in einem bestimmten Punkt kreuzen, wahrnehmen und anhand dieser Wahrnehmung die Distanz berechnen, alles, was wir allein über den Sehsinn wahrnehmen könnten, sei eine mehr oder minder chaotische Ansammlung von Licht und Farbe.

Auf dieser Kritik aufbauend formuliert Berkeley eine eigene, innovative Theorie, welche darauf abzielt, zu zeigen, dass wir Entfernung streng genommen überhaupt nicht sehen könnten, sondern unsere Vorstellung unterschiedlich weit entfernter Objekte

---

<sup>34</sup>Diese Passage, wohl eine der am häufigsten aus dem *Essay* zitierten, lässt sich verschiedenartig auslegen. Die bezüglich der Konsequenzen harmloseste Interpretation (die u.a. von Atherton, 1990 vorgeschlagen wird) wäre, dass Berkeley an dieser Stelle nur deutlich machen will, dass die ‘Informationen’ (oder die Vorstellungen), die wir über den Tastsinn bekommen eben verschieden von denen sind, die wir über den Sehsinn bekommen. Unsere verschiedenen sensorischen Modalitäten liefern quasi qualitativ unterschiedliche ‘Informationen’, die uns zusammengenommen helfen, uns in der Welt zurechtzufinden und eine adäquate Grundlage bilden, um uns Wissen über die räumlichen Eigenschaften der physikalischen Welt anzueignen. Eine hinsichtlich ontologischer und metaphysischer Konsequenzen wesentlich gravierendere Interpretation wäre hingegen, dass Berkeley an dieser Stelle meint, dass wir in der Tat zwei *verschiedene Objekte* über Tast- und Sehsinn wahrnehmen würden, die aufgrund von Gewohnheit und sprachlichen Unbedachtheiten für ein und dasselbe Objekt gehalten werden. Ohne Frage hängt die Plausibilität der verschiedenen Interpretationen von dem ab, was Berkeley genau mit „Ding“, „gefühltes Ding“ und „gesehenes Ding“ meint. Eine Entscheidung über diese Frage würde eine ausführlichere Sichtung des Gesamtwerkes von Berkeley erfordern, auf die aufgrund der Ausrichtung dieser Arbeit jedoch verzichtet werden soll.

auf einer gewohnheitsmäßigen Verknüpfung bestimmter Wahrnehmungen des Sehens mit bestimmten Wahrnehmungen des Tastsinns beruhe, der allein in der Lage sei, Entfernungen wahrzunehmen. Die Mittel zur Tiefenwahrnehmung, die von Alhazen und Descartes vorgeschlagen wurden, werden somit bei Berkeley radikal umgedeutet. Sie sind nicht länger Bestandteil eines Prozess des Schlussfolgerns oder Berechnens, sondern haben eher den Status eines arbiträren *Zeichens*, so wie etwa ein Wort einen Gegenstand bezeichnet und die Beziehung zwischen Wort und Gegenstand arbiträr ist.

Man könnte diesen Ansatz, in Anlehnung an Turbayne (1963), auch als *linguistischen Ansatz* des Sehens bezeichnen. Sehen müsse, ebenso wie jedwede Sprache, *gelernt* werden.<sup>35</sup> Eine Sprache könne nur dann verstanden werden, wenn man die willkürlichen Verbindungen zwischen Zeichen und Bezeichnetem gelernt habe.<sup>36</sup> Wenn diese Verknüpfung durch wiederholtes gemeinsames Auftreten etabliert wurde, sei es zudem zunehmend schwierig, im alltäglichen Gebrauch Zeichen und Bezeichnetes zu trennen, da beim Hören eines erlernten Wortes unmittelbar dessen Bedeutung suggeriert werde, was oft dazu führe, dass man beides für ein und dasselbe halte. Ähnlich verhalte es sich nun auch beim Wahrnehmen von Entfernung: Der Wahrnehmende verknüpfe durch wiederholtes gemeinsames Auftreten die über den Tastsinn unmittelbar wahrgenommene Entfernung (die auch nur vom Tastsinn unmittelbar wahrgenommen werden könne) mit der Wahrnehmung bestimmter Bewegungen der Augenmuskeln. Dies führe dann dazu, dass bei Auftreten unterschiedlicher Bewegungen der Augenmuskeln die damit verknüpften unterschiedlichen Entfernungen dem Verstand suggeriert würden. Auch hier komme es dann zunehmend zu einer Verwechslung von Zeichen und Bezeichnetem, was unter anderem dazu führe, dass man (fälschlicherweise) davon spreche, dass man Entfernung *sehe*.

Dieser Ansatz ist in der Hinsicht elegant, als dass er die in den Abschnitten zu Alhazen und Descartes dargestellten Probleme zunächst vermeidet. Bei Alhazen und Descartes kommt es gleichsam zu einer gewissen ‘Intellektualisierung der Sinne’, im Rahmen der Berkeleyschen Theorie muss hingegen keine Urteile fällende oder Berechnungen vornehmende Instanz sowie die dafür notwendigen Voraussetzungen erklärt werden. Berkeley scheint zwar davon auszugehen, dass wir (im Prinzip) das durch unseren Sehsinn unmittelbar wahrnehmen, was auf die Retinae projiziert wird, er ist aber nicht auf geometrische Eigenschaften eines ‘Netzhautbildes’ angewiesen, um Distanz-

---

<sup>35</sup>Wir werden später sehen, dass insbesondere Helmholtz diesen Aspekt der Berkeleyschen Theorie in seinem theoretischen Ansatz wieder aufgreifen wird

<sup>36</sup>Berkeley verwendet zwar im *Essay* an nur wenigen Stellen den Begriff des Zeichens, aber es lässt sich an vielen Stellen der analogische Charakter zwischen Sehen und Sprache herauslesen. Augenfällig wird er in Anbetracht des Titels seines zweiten, dem Sehen gewidmeten Werkes, der „*Theory of Vision or Visual Language*“, in dem dann auch wesentlich häufiger von *signs* die Rede ist. Siehe zu diesem Punkt auch Copenhaver (2014) sowie McGowan (1982).

wahrnehmung erklärbar zu machen, denn Distanzwahrnehmung geschehe nur über den Tastsinn.

Diese Konzeption birgt allerdings auch neue Schwierigkeiten. So ist beispielsweise unklar, wie genau die Wahrnehmung von Entfernung über den Tastsinn geschehen soll, die dann mit Wahrnehmungen des Sehsinnes verknüpft werden sollen. Wenn ich mich auf ein weiter entferntes Objekt zubewege und es erfühle, ‘berechne’ oder ‘messe’ ich dann die zurückgelegte Distanz? Hierzu äußert sich Berkeley nicht, er scheint vorauszusetzen, dass eine Entfernungswahrnehmung über den Tastsinn nicht weiter erklärungsbedürftig sei. Das Problem der Tiefenwahrnehmung wird von Berkeley auf den Tastsinn verschoben. Zudem bleibt unklar, wie die Berkeleysche Auffassung einer *haptischen* Entfernungswahrnehmung, die durch gewohnheitsmäßige Verknüpfung das *Sehen* eines entfernten Objektes *suggestiere*, letztendlich aufzufassen ist. Das Sehen unterschiedlich weit entfernter Objekte wird in Berkeleys Ansatz nicht wirklich verständlich, da dieser für die Theorie essentielle Vorgang der ‘Suggestion’ nicht weiter ausgeführt wird.<sup>37</sup>

Damit es überhaupt zu einer gewohnheitsmäßigen Verknüpfung zwischen den Wahrnehmungen des Sehsinns und den Entfernungswahrnehmungen des Tastsinnes kommen kann, die uns erlaubt, in vielfältigen Situationen Objekte als verschieden weit entfernt wahrzunehmen, müsste darüber hinaus *jede mögliche* diesbezügliche Wahrnehmung des Sehsinns mit den entsprechenden Entfernungswahrnehmungen des Tastsinns mehrmals aufgetreten sein, was absurd erscheint. Es sei denn, Berkeley versteht unter einer gewohnheitsmäßiger Verknüpfung gerade nicht, dass die Elemente, um verbunden zu werden, mehrmals gleichzeitig aufgetreten sein müssen, worauf der *Essay* allerdings nicht hindeutet. Damit verbunden ist auch die Frage, wie es kommen kann, dass wir Objekte, deren Entfernung wir unmöglich über den Tastsinn wahrgenommen haben können (etwa den Mond) trotzdem als von uns so und so weit entfernt wahrnehmen. Das kann nur plausibel gemacht werden, wenn man einer bestimmten Instanz wiederum bestimmte ‘Schätzprozesse’ zuschreibt, was einen jedoch wieder zu Problemen der oben dargestellten Art führt.

Bezüglich der Mittel zur Entfernungswahrnehmung stellen sich nun ebenfalls neue Fragen. Berkeley ist der Auffassung, dass Konvergenz und Akkommodation – hier verstanden als Wahrnehmung entsprechender Muskelbewegungen – Distanzwahrnehmung ermöglichen, da sie wiederholt gemeinsam mit bestimmten Entfernungswahrnehmungen des Tastsinnes aufgetreten seien. In Fällen ‘relativ naher’ Objekte mag dies plausibel sein, fixiert man jedoch nacheinander mehrere, unterschiedlich weit entfernte Objekte, die allesamt hinreichend weit vom Betrachter entfernt sind, lassen sich die zugehörigen

---

<sup>37</sup>So stellt etwa auch Hatfield (1990, S. 41) diesbezüglich fest: „Despite the crucial role played by suggestion in Berkeley’s theory, he provided precious little analysis of it“.

Muskelbewegungen nicht mehr als unterschiedlich wahrnehmen, so dass es hier zum Sehen von in etwa gleich weit entfernten Objekten kommen sollte, was jedoch nicht mit unserem alltäglichen Erleben in Übereinstimmung zu bringen ist.

Eine weitere Schwierigkeit besteht darin, dass sich Berkeley zwar ausführlich darüber äußert, wie Konvergenz und Akkommodation mit Entfernungswahrnehmungen des Tastsinnes verbunden werden, sich bezüglich der anderen von Descartes vorgeschlagenen Mittel der Entfernungswahrnehmung aber eher bedeckt hält, obgleich Berkeley sie als solche anzusehen scheint (vgl. Fussnote 32 auf Seite 38). Die ‘erscheinende Größe’ eines Objektes etwa, die Berkeley auch als ein Mittel zur Entfernungswahrnehmung anspricht, scheint nicht in seine Konzeption zu passen, da das, was wir laut Berkeley unmittelbar über den Sehsinn wahrnehmen würden, lediglich eine chaotische Anordnung von Licht und Farbe sei und so nicht verständlich wird, wie sich in diesem Rahmen sinnvoll von der ‘erscheinenden Größe von Objekten’ sprechen lässt. Ein Punkt, der ein genaues Verständnis dieses Aspektes der Berkeleyschen Theorie erschwert, ist, dass Berkeley gelegentlich auch von Tiefenwahrnehmung als Ergebnis eines *Urteils* spricht, wobei in der Regel nicht deutlich wird, ob dies seine eigene Auffassung oder eine Beschreibung der Auffassung der Vertreter der geometrischen Theorie ist (siehe etwa Fußnote 26 auf Seite 35). Falls es seine eigene Auffassung sein sollte, so ist diese nur schwer mit der generellen Ausrichtung von Berkeleys Theorie der Distanzwahrnehmung in Übereinstimmung zu bringen. Wenn unterschiedlich weit entfernte Objekte gesehen werden, da sich aufgrund wiederholten gemeinsamen Auftretens eine Verknüpfung zwischen bestimmten Bewegungen der Augenmuskeln und bestimmten Wahrnehmungen des Tastsinnes ausgebildet hätte, so dass bei Auftreten einer bestimmten Augenmuskelbewegung unmittelbar die zugehörige über den Tastsinn vermittelte Distanzwahrnehmung suggeriert wird, worüber müsste dann noch ein Urteil gefällt werden? Laut Hatfield (1990) lassen sich in anderen Werken Berkeleys zwar Hinweise finden, dass dieser ein Suggestieren von einem Urteil durchaus trennen möchte, konkrete Analysen, worin dieser Unterschied bestehen soll, seien jedoch nicht aufzufinden.

## 2.4. Empfindung und Wahrnehmung: Thomas Reid

Thomas Reid (1710-1796), Zeitgenosse Berkeleys, war eine zentrale Figur der schottischen Aufklärung und Begründer der schottischen Schule der sogenannten *common-sense*-Philosophie. Wahrnehmungstheoretische Überlegungen lassen sich sowohl in Reids 1764 verfassten „*Inquiry into the human mind on the principles of common sense*“, als auch in den etwa 20 Jahre später ausgearbeiteten „*Essays on the intellectual powers of man*“ finden. Reids Abhandlungen zu wahrnehmungstheoretischen Fragen sind zu großen Teilen in eine philosophische Diskussion seiner Zeit einzuordnen, die, unter an-

derem von Berkeleys provokantem „*esse is percipi*“ ausgelöst, den epistemologischen und ontologischen Status des Wahrgenommenen zum Gegenstand hatte. Diese Diskussion und Reids Rolle in dieser soll jedoch nicht Gegenstand dieser Darstellung sein (siehe dazu etwa Greco, 1995), sondern es soll sich in diesem Abschnitt insbesondere auf eine von Reid getroffene konzeptuelle Unterscheidung im Kontext visueller Wahrnehmung konzentriert werden, welche einen nachhaltigen Einfluss auf die nachfolgende Theoriebildung ausüben sollte.

Die fundamentale Unterscheidung, die Reid als Erster explizit trifft (Atherton, 2002; Boring, 1942) und die bis heute in der Wahrnehmungspsychologie Verwendung findet, ist die zwischen *sensation* und *perception*, für die sich im Deutschen die Begriffe „Empfindung“ und „Wahrnehmung“ eingebürgert haben.<sup>38</sup> Wahrnehmung und Empfindung seien laut Reid Zustände oder Akte des Geistes und nicht des Körpers, es bestehe jedoch ein prinzipieller Unterschied zwischen beiden, der im Rahmen einer sorgfältigen Analyse der Wahrnehmung, insbesondere bei epistemologischen und ontologischen Überlegungen, beachtet werden sollte.<sup>39</sup>

Reid verdeutlicht diese Unterscheidung in einer überaus modern anmutenden Passage folgendermaßen (vgl. Reid, 1813, S. 376ff.): Ein Beispiel für den Ausdruck einer Empfindung sei etwa die Aussage „Ich fühle einen Schmerz“, wohingegen die Aussage „Ich sehe einen Baum“ eine Wahrnehmung bezeichne. Grammatikalisch sind beide Ausdrücke identisch, sie bestehen beide aus einem Subjekt, einem aktiven Verb und einem Objekt. Allerdings sei im ersten Fall die Unterscheidung zwischen dem Akt und dem Objekt des Aktes keine reale, im zweiten schon. Es bestehe laut Reid kein Unterschied zwischen dem Gefühl des Schmerzes und dem Schmerz selbst, obwohl die grammatikalische Form der Aussage dies nahelege („*feeling a pain* signifies no more than *being pained*“, Reid, 1813, S. 377). Die Aussage „Ich fühle einen Schmerz“ beziehe sich nicht auf ein von meinem Empfinden unabhängiges Objekt, sondern allein auf mein Empfin-

---

<sup>38</sup>Hatfield (1990) merkt an, dass sich in Malebranches „*De la recherche de la vérité*“ aus dem Jahre 1674 zwar bereits eine (inhaltlich) ähnliche Unterscheidung, nämlich die zwischen sogenannten *perceptions sensibles* und *perceptions pures* finden lässt, dass Reid jedoch diese Unterscheidung ausführlicher und konzeptuell wesentlich klarer ausarbeitet. Auf die von Hatfield (1990) in diesem Zusammenhang durchgeführten philosophiehistorisch faszinierenden Analysen zur konzeptuellen Ähnlichkeit zwischen der Unterscheidung zwischen Wahrnehmung und Empfindung und den durch hauptsächlich Locke prominenten sogenannten *primary* und *secondary qualities* soll hier nicht weiter eingegangen werden, da sie für die Zielstellung dieser Arbeit nicht von Relevanz sind. Der an einer Übersicht der historischen Entwicklung der Unterscheidung zwischen Wahrnehmung und Empfindung von der Antike bis in die Neuzeit interessierte Leser sei auf das erste Kapitel von Boring (1942) sowie den Aufsatz von Atherton (2002) verwiesen, für einen kompakten Überblick der Wahrnehmungstheorie von Malebranche siehe Atherton (1990).

<sup>39</sup>„Sensation, and the perception of external objects by the senses, though very different in their nature, have commonly been considered as one and the same thing. ... But, without attending carefully to this distinction, it is impossible to have any just conception of the operation of our senses“ (Reid, 1813, S. 376).

den; Akt und Objekt des Aktes seien identisch.<sup>40</sup> Dieses Merkmal sei charakteristisch für sämtliche Empfindungen, allerdings sei es schwer, eine Vielzahl an weiteren Beispielen zu benennen, da die wenigsten Empfindungen einen eigenen Namen hätten. Es sei zudem nicht möglich, dass man sich in seinen Empfindungen täuschen könne.<sup>41</sup> Für eine Wahrnehmung sei es hingegen bezeichnend, dass der Akt vom Objekt des Aktes unterschieden sei. Wahrnehmung sei ein intentionaler Akt, Empfindung hingegen ein nicht intentionaler Akt. Wenn ich davon spreche, dass ich einen Baum wahrnehme, so beziehe ich mich laut Reid einerseits auf ein *von mir unabhängiges* Objekt, das wahrgenommen werde, und andererseits auf einen Akt des Geistes, durch den es wahrgenommen werde. Diese beiden Sachen seien nicht dasselbe, es bestehe nicht einmal eine Ähnlichkeit zwischen diesen, sie seien wesentlich verschieden. Die Wahrnehmung eines jeden Objektes bestehe aus zwei Elementen, nämlich erstens aus einem Konzept des Objektes, das ich als Wahrnehmender habe, und zweitens aus der unmittelbaren und festen *Überzeugung (belief)* an die gegenwärtige Existenz des wahrgenommenen Objektes. Diese Überzeugung beruhe nicht auf Verstandestätigkeiten, ich schlussfolge nicht etwa auf die Existenz eines von mir unabhängigen Objektes, sondern sie trete unmittelbar mit der Wahrnehmung auf und sei in der Natur des Menschen selbst begründet.<sup>42</sup> Wir könnten aufgrund unserer Natur gleichsam nicht anders, als die von uns wahrgenommenen Dinge als unabhängig von uns existierende Objekte einer Aussenwelt wahrzunehmen, was sich sowohl in unserer Sprache, als auch in unserem alltäglichen Verhalten den von uns wahrgenommenen Objekten gegenüber äußere.<sup>43</sup>

---

<sup>40</sup>„Sensation is ... an act of mind, which may be distinguished from all others by this, that it hath no object from the act itself“ (Reid, 1814, S. 27).

<sup>41</sup>„It is impossible that there can be any fallacy in sensation: for we are conscious of all our sensations, and they can neither be any other in their nature, not greater or less in their degree that we feel them. It is impossible that a man should be in pain, when he does not feel pain; and when he feels pain, it is impossible that his pain should not be real, and in its degree what it is felt to be; and the same thing may be forgot when it is past, but when it is present, it can be nothing but what we feel. If there be any fallacy in our senses, it must be in the perceptions of external objects“ (Reid, 1814, S. 311).

<sup>42</sup>„In perception we not only have a notion more or less distinct of the object perceived, but also an irresistible conviction and belief of its existence. ... This conviction is not only irresistible, but it is immediate; that is, it is not by a train of reasoning and argumentation that we come to be convinced of the existence of what we perceive; we ask no argument for the existence of the object, but that we perceive it“ (Reid, 1814, S. 113).

<sup>43</sup>Auf den skeptischen Einwand, warum man ohne vernünftige Gründe an die Existenz der wahrgenommenen Objekte glauben müsse, müsse man, so Reid, antworten, dass diese Überzeugung eben nicht auf bestimmten (Verstandes-)Tätigkeiten des Wahrnehmenden beruht, also nicht von ihm selbst stammt, sondern natur-, oder wie Reid später präzisiert, *gottgegeben* ist. Darüberhinaus sieht Reid keinen Grund, mittels Verstandestätigkeit an dieser Überzeugung Zweifel zu hegen. Verstand wie Fähigkeit zur Wahrnehmung seien beide naturgegeben und warum solle man an dem einen zweifeln und das andere als Maßstab nehmen („Why ... should I believe the faculty of reason more than that of perception; they came both out of the same shop“, Reid, 1813, S. 378)? Zudem liege es aufgrund dieser Naturgegebenheit überhaupt nicht in der Macht des Wahrnehmenden, die Überzeugung an die Existenz des wahrgenommenen Objektes zu beseitigen, egal wie stark man sich auch anstrengen möge. In einer daran anschließenden, durchaus humoristischen Passage erläutert Reid, welche Konsequenzen

Welche Verbindung soll nun zwischen Empfindungen und Wahrnehmungen bestehen? Auch Reid fasst Wahrnehmung allgemein als einen Prozess auf, der sich in mehrere verschiedene Stufen unterteilen lasse: Zuerst müsse, insofern kein direkter Kontakt zwischen Objekt und Sinnesorgan bestehe, ein Medium zwischen diesen bestehen; im Falle visueller Wahrnehmung das von den Objekten reflektierte Licht. Daraufhin gebe es einen Effekt oder eine Einwirkung (*impression*) auf das Sinnesorgan, es werde affiziert, entweder direkt durch das Objekt oder indirekt über das vermittelnde Medium. Diese Einwirkung habe wiederum einen bestimmten Effekt auf jene Nerven, die das affizierte Sinnesorgan mit dem Hirn verbinden und schließlich wird über jene Nerven auf das Gehirn selbst eingewirkt. Hierbei handle es sich um rein physiologische Prozesse, „which affect the body only“ (Reid, 1813, S. 385). Der Einwirkung auf das Sinnesorgan, die Nerven und das Gehirn folge dann in der Regel eine Empfindung, welche schließlich in eine Wahrnehmung eines externen Objektes übergehe. Empfindung und Wahrnehmung seien keine physiologischen Prozesse, sondern Akte des Geistes:

We perceive no object, unless some impression is made upon the organ of sense. ... In order to our perceiving objects, the impressions made upon the organs of sense must be communicated to the nerves, and by them to the brain. ... Here the material part ends ... the rest is all intellectual. (Reid, 1814, S. 80f.)

Nicht unsere Augen, Nerven, das Gehirn oder alle drei gemeinsam würden etwas wahrnehmen oder empfinden, sondern *wir*, als mit Geist ausgestatte Wesen seien es, die empfinden und wahrnehmen.<sup>44</sup> Wieso es aufgrund der über die Augen und die Nerven weitergeleiteten Einwirkung auf das Gehirn zu einer Empfindung komme, verschleiße sich unserer Erkenntnis. Alles, was wir diesbezüglich sagen könnten, wäre, dass Empfindungen und Wahrnehmungen nicht ohne Einwirkung auf unsere Sinnesorgane zustande kämen und dass bestimmten Einwirkungen auf unsere Sinnesorgane bestimmte Empfindungen und Wahrnehmungen folgen.<sup>45</sup> Physiologische und anatomische Untersuchungen könnten daher nichts zu einer Aufklärung unserer Empfindungen und Wahr-

---

ein ständiger Zweifel an der Existenz wahrgenommener Objekte hätte: Man würde eine Vielzahl physischer Unfälle erleiden (mit dem Kopf gegen Wände stoßen, in einen Kanal fallen etc.) und letztendlich im Irrenhaus landen. Jeder Skeptiker, der behauptet, durchgängig die Existenz wahrgenommener Objekte zu bezweifeln, sei entweder ein Heuchler oder jemand, der sich selbst täusche, denn er verhalte sich doch letztendlich so, wie jemand, der an die Existenz wahrgenommener Objekte glaubt.

<sup>44</sup>„Another necessary caution in this matter its, that we ought not confound the organs of perception with the being that perceives. Perception must be the act of some being that perceives. ... The eye is a machine most admirably contrived for refracting the rays of light, and forming a distinct pictures of objects upon the retina; but it sees neither the object nor the picture“ (Reid, 1814, S. 78).

<sup>45</sup>„But how are the sensations of the mind produced by impressions of the body? Of this we are absolutely ignorant, having no means of knowing how the body sets upon the mind, or the mind upon the body. ... There is a deep and dark gulf between them, which our understanding cannot pass. Experience teaches us, that certain impressions upon the body are constantly followed by certain sensations of the mind“ (Reid, 1813, S. 387).

nehmungen beitragen, eine physiologische Beschreibung der Einwirkung eines spitzen Gegenstandes auf unsere Haut, der dadurch ausgelösten Nervenerregungen und deren Weiterleitung an das Hirn könne uns nichts über unser Empfinden von Schmerz sagen.

Zur näheren Charakterisierung der Beziehung zwischen Empfindung und Wahrnehmung greift Reid (wie auch schon Berkeley) auf das Konzept des *Zeichens* zurück. Unsere Empfindungen, so Reid, weisen weder eine Ähnlichkeit mit den von uns wahrgenommenen Objekten auf, noch lasse sich eine notwendige Verbindung zwischen Empfindung und Wahrnehmung auffinden; die Empfindungen seien bloße Zeichen für die wahrgenommenen externen Objekte.<sup>46</sup> Um diesen Punkt genau zu verstehen, sollte man sich nochmals in Erinnerung rufen, was Reid genau unter Empfindung und Wahrnehmung verstehen möchte. Wenn ich etwa eine Rose rieche, so ein weiteres Beispiel Reids, so müsse man in dem durch diese Wendung ausgedrückten Vorgang Empfindung und Wahrnehmung strikt trennen. Der Geruch selbst, ohne Bezug auf ein externes Objekt, sei eine Empfindung, diese sei nicht mehr, als das, ‘wonach es sich anfühlt’, es gebe hier keine Trennung zwischen der Empfindung und dem, was empfunden wird, „its very essence consists in being felt ... there is no difference between the sensation and the feeling of it“ (Reid, 1814, S. 243). Unsere Wahrnehmung des Geruchs der Rose sei aber etwas grundsätzlich Verschiedenes, denn Wahrnehmung beziehe sich stets auf ein externes Objekt. Die Wahrnehmung des Geruchs der Rose beziehe sich daher auf eine bestimmte Qualität des von mir als unabhängig von meiner Wahrnehmung aufgefassten Objekts Rose.

Im Wahrnehmungsprozess gehen wir, so Reid, unmittelbar von der Empfindung zur Wahrnehmung über, so wie wir, man erinnere sich an den identischen Vergleich bei Berkeley, beim Lesen eines Textes unmittelbar vom Zeichen zum Bezeichneten übergehen, wobei Empfindung und Wahrnehmung, ebenso wie Zeichen und Bezeichnetes, keinerlei Ähnlichkeit miteinander haben müssten. In unserem Alltag würden wir Empfindung und Wahrnehmung nicht trennen, was sich auch in unserem Sprachgebrauch manifestiere, für eine gründliche theoretische Analyse unserer Wahrnehmung sei dies jedoch essentiell:

Sensation, taken by itself, implies neither the conception nor belief of any external object. It supposes a sentient being, and a certain manner in which that being is affected; but it supposes no more. Perception implies an immediate conviction and belief of something external; something different both from the mind that perceives, and from the act of perception. Things so different in their nature ought to be distinguished; but by our constitu-

---

<sup>46</sup>„Our sensations have no resemblance to external objects; nor can we discover, by reason, any necessary connection between the existence of the former, and that of the latter. ... There is no necessity of a resemblance between the sign and the thing signified: and indeed no sensation can resemble any external object“ (Reid, 1813, S. 388f.).

tion they are always united. Every different perception is conjoined with a sensation that is proper to it. The one is the sign, the other the thing signified. (Reid, 1814, S. 249)

Reid unterscheidet zudem allgemein zwischen Zeichen, die *naturgegeben* auf das zu Bezeichnende verweisen (sogenannte *natural signs*) und Zeichen, bei denen die Verbindung zwischen Zeichen und Bezeichnetem erlernt werden müsste (sogenannte *artificial* oder *acquired signs*). Bei den Wörtern einer Sprache etwa handle es sich um *artificial signs*, bestimmte durch den Tastsinn gegebene Empfindungen seien hingegen *natural signs*, da sie naturgegeben beispielsweise auf die Festigkeit eines vom Wahrnehmenden unabhängigen Objektes verweisen würden.

Auf welche Weise wendet Reid nun sein Konzept von Empfindungen als Zeichen für die Wahrnehmung externer Objekte im speziellen Fall der visuellen Wahrnehmung an? Unglücklicherweise werden hier die Beschreibungen deutlich ungenauer, Reid gibt weder im *Inquiry* noch in den *Essays* an, wie ‘visuelle Empfindungen’ zu charakterisieren seien. Stattdessen spricht er von der *sichtbaren Erscheinung* von Objekten („visible appearance of objects“ oder „the appearance that objects make to the eye“, Reid, 1813, S. 263), die ein Zeichen für die externen Objekte sei. Die *sichtbare Erscheinung* eines Gegenstandes werde vom Wahrnehmenden nicht registriert, sondern es werde auch hier von diesem Zeichen unmittelbar zu dem Bezeichneten, dem externen Objekt, übergegangen:

The visible appearances of objects are intended by nature only as signs or indications; and that the mind passes instantly to the things signified, without making the least reflection upon the sign, or even perceiving that there is any such thing. (Reid, 1813, S. 362)

Der Status der *sichtbaren Erscheinung* eines Objektes ist einer der am stärksten diskutierten Punkte in der Literatur über Reids wahrnehmungstheoretischen Ansatz, es lassen sich aber sowohl im *Inquiry* als auch in den *Essays* mehrere Textstellen finden, die stark darauf hindeuten, dass Reid sich auf verschiedene geometrische Eigenschaften des auf die Retinae projizierten Lichtmusters bezieht (für eine ausführliche Diskussion siehe Nichols, 2007, Kap. 4). Reid charakterisiert die *sichtbare Erscheinung* als ein physikalisches Objekt, als „a real and external object which is presented to the eye“ (Reid, 1813, S. 282). Werde ein Gegenstand etwa aus unterschiedlichen Entfernungen betrachtet, so schrumpfe bzw. vergrößere sich laut Reid die *sichtbare Erscheinung*; werde derselbe Gegenstand aus unterschiedlichen Winkeln betrachtet, so verändere sich die *sichtbare Erscheinung* ebenfalls, was Reid anhand rein geometrischer Analysen darlegt. Die *sichtbare Erscheinung* eines Objektes sei durch (optische) Gesetze vorgegeben.<sup>47</sup> In

<sup>47</sup> „The appearance of things to the eye, always corresponds to the fixed laws of nature“ (Reid, 1813, S. 405).

Reids allgemeinem Rahmen ist diese Charakterisierung nicht unproblematisch, da im Fall der visuellen Wahrnehmung nach der vorgeschlagenen Lesart keine Empfindung ein Zeichen für die externen Objekte sein soll, sondern die physikalischen Gesetzen folgende *materielle* Einwirkung von Licht auf das Auge.<sup>48</sup>

Wie der Zusammenhang zwischen der *sichtbaren Erscheinung* und der Wahrnehmung externer Objekte aufgefasst werden soll, lässt sich Reids Ausführungen nicht deutlich entnehmen. Die *sichtbaren Erscheinungen* seien Zeichen für die externen Objekte, die vom Wahrnehmenden selbst nicht wahrgenommen werden. Sie werden aber sowohl als naturgegebene *natural signs* („the visible appearances of objects are intended by nature only as signs“, Reid, 1813, S. 362), als auch als durch Erfahrungen erlernte *artificial signs* beschrieben („but the thing signified by that appearance, is that which is presented to the mind by custom“, Reid, 1813, S. 374). Im Kontext der Wahrnehmung von Tiefe bzw. Entfernung greift Reid jedoch ausschließlich auf das Konzept der *artificial signs* zurück.

### Tiefenwahrnehmung

Reid geht ebenso wie alle anderen bisher dargestellten Autoren davon aus, dass die Entfernung eines Objektes kein ‘unmittelbarer Gegenstand’ der Wahrnehmung sei. Die wahrgenommene Entfernung eines Objektes könne nur über zusätzliche Mittel verständlich werden, Reid verweist dabei insbesondere auf verschiedene Eigenschaften der *sichtbaren Erscheinung* eines Objektes:

Distance from the eye is no immediate object of sight. There are certain things in the visible appearance, which are signs of distance from the eye, and from which ... we learn by experience to judge of that distance. (Reid, 1813, S. 266)

Reid spricht in diesem Zusammenhang ebenfalls von Zeichen, wobei die Bedeutung dieser Zeichen durch Erfahrung gelernt werden müsse („of the signs by which we learn to perceive distance from the eye“, Reid, 1813, S. 390), in Reids Terminologie handelt es sich demnach um *artificial signs*. Insgesamt zählt Reid fünf solcher Zeichen für die Distanz eines Objektes auf, die mit Ausnahme des zuletzt angeführten in der bisherigen Darstellung bereits aufgetreten sind:

1. Akkommodation und Konvergenz, die Reid ebenso wie Berkeley nicht geometrisch auffasst, sondern als bestimmte Muskelkontraktionen der Augen,
2. weitere Objekte, die sich zwischen dem Wahrnehmenden und dem Objekt befinden,

---

<sup>48</sup>„There seems to be no sensation that is appropriated to visible figure ... it seems to be suggested immediately by ... a material impression upon the retina“ (Reid, 1813, S. 288).

3. die ‘erscheinende Größe’ eines Objektes in Verbindung mit einem Wissen um die ‘tatsächliche Größe’ des Gegenstandes und
4. das Verwaschen und Ineinanderfließen der Farben sowie eine abnehmende Schärfe der Konturen weiter entfernt liegender Objekte.

Der zuletzt aufgeführte Punkt wird in der aktuellen Literatur auch als *atmospheric view* oder *aerial perspective* bezeichnet. In der Regel geschehe die Wahrnehmung der Distanz eines Objektes nicht über eines dieser Zeichen allein, sondern über mehrere, wobei in bestimmten Situationen es auch vorkommen kann, dass verschiedene Zeichen auf *unterschiedliche* Distanzen verweisen. Beim Betrachten eines linearperspektivischen Bildes etwa würden Konvergenz und Akkommodation darauf verweisen, dass sämtliche im Bild dargestellten Objekte dieselbe Distanz zum Wahrnehmenden aufweisen, wohingegen der Maler explizit die verbleibenden aufgelisteten Zeichen nutzen würde, um den Eindruck unterschiedlich weit entfernter Objekte beim Betrachter zu erzeugen. Reid spricht hier bereits etwas an, das in der aktuellen Wahrnehmungspsychologie als *cue-Integration* bekannt ist, nämlich die Frage, wie mögliche *unterschiedliche* ‘Hinweise’ aus verschiedenen Tiefencues integriert werden, um eine *einheitliche* Wahrnehmung ‘hervorzurufen’.

Wie bereits erwähnt, sieht Reid die Verbindung zwischen den Zeichen und der Distanz eines Objektes als eine durch Erfahrung oder Gewohnheit entstandene an, die Bedeutung eines Zeichens für Distanz müsse, ähnlich wie die Bedeutung der Wörter einer fremden Sprache, über Erfahrung gelernt werden. Durch wiederholtes gleichzeitiges Auftreten des Zeichens und des Bezeichneten entstehe eine Verknüpfung dergestalt, dass bei Auftreten des Zeichens nicht dieses selbst, sondern unmittelbar dessen Bedeutung wahrgenommen werde. Interessant ist nun, dass Reid die Akkommodation und die Konvergenz als Empfindungen auffasst, die restlichen drei hingegen nicht. Wir seien uns im Fall der Akkommodation und Konvergenz der zugehörigen Bewegungen unserer Augenmuskeln nicht bewusst, der Anstrengung (*effort*), um diese zu erzeugen hingegen schon. Diese Anstrengung gehe mit einer bestimmten Empfindung einher, welche mit der Distanz des Objektes, das diese Bewegungen bzw. die damit verbundenen Muskelanstrengungen bei Fixation desselben ‘hervorgerufen’ habe, assoziiert werde.<sup>49</sup> Diese Art des Zeichens sei nur in gewissen Grenzen ‘reliabel’, im Falle ‘sehr weit’, aber dennoch unterschiedlich weit entfernter Objekte (etwa Sonne und Mond) seien wir uns der Unterschiede in den jeweiligen Muskelanstrengungen aufgrund ihrer äußerst geringen

---

<sup>49</sup>„But although we are not conscious of the motions we perform in order to make the eyes right for the distance of the object, we are conscious of the effort employed in producing those motions; and probably have some sensation which accompanies them, to which we give as little attention as to other sensations. And thus an effort consciously exerted, or a sensation consequent upon that effort, comes to be conjoined with the distance of the object which gave occasion to it; and by this conjunction becomes a sign of that distance“ (Reid, 1813, S. 392).

Ausgeprägtheit nicht mehr bewusst, hätten dadurch keine unterschiedlichen Empfindungen und würden daher diese Objekte als gleich weit entfernt wahrnehmen. Die drei verbleibenden Zeichen für Distanz seien laut Reid hingegen Eigenschaften der *sichtbaren Erscheinung* der Objekte. So wisse ich etwa aus Erfahrung, dass ein- und dasselbe Objekt in unterschiedlicher Distanz unterschiedlich groß ‘erscheine’, so dass die ‘erscheinende Größe’ eines Objektes ein Zeichen für die Distanz des Objektes sei. Ebenso könne man durch Erfahrung lernen, dass weit entfernte Objekte als weniger deutlich umrissen bzw. gleichsam ‘verschwommen’ erscheinen. Die *sichtbare Erscheinung* werde so über gewohnheitsmäßige Verknüpfung zu einem Zeichen eines externen Objektes in bestimmter Distanz.

### Zusammenfassung

Wie eingangs angedeutet, wurde Reids theoretischer Ansatz hauptsächlich aufgrund der von ihm getroffenen und für die weiteren theoretischen Entwicklungen wichtigen Unterscheidung zwischen Empfindung und Wahrnehmung in diese Darstellung aufgenommen. Empfindungen wie Wahrnehmungen seien Akte des Geistes und strikt von physiologischen Funktionen der Sinnesorgane zu trennen. Empfindungen und Wahrnehmungen seien zwar nur möglich, wenn die Sinnesorgane affiziert werden, eine physiologische Analyse dieser Sinnesorgane könne aber nicht zu einer Aufklärung jener Akte des Geistes beitragen. Es bestehe keinerlei Ähnlichkeit zwischen (materieller) Einwirkung auf die Sinnesorgane und der dieser folgenden Empfindung, beide seien gleichsam durch eine unüberbrückbare Kluft getrennt. Kennzeichnend für eine Empfindung sei, dass diese im Gegensatz zu einer Wahrnehmung keinen vom Akt des Empfindens unterschiedenen Gegenstand aufweisen. Wird hingegen etwas wahrgenommen, so lasse sich hierbei der Akt der Wahrnehmung von dem wahrgenommenen Gegenstand unterscheiden, der stets als ein vom Wahrnehmen *unabhängiger* wahrgenommen werde. Empfindungen seien zudem als natürliche Zeichen für Wahrnehmungen anzusehen, die Verbindung zwischen Zeichen und Bezeichnetem entziehe sich jedoch unserer Erkenntnis, sie sei natur- bzw. gottgegeben.

Im Fall visueller Wahrnehmung wird dieser konzeptuelle Rahmen von Reid jedoch aufgebrochen. Hier seien es keine Empfindungen, welche auf externe Objekte verweisen, sondern die sogenannte *sichtbare Erscheinung* der Objekte, deren Status und Beziehung zur Wahrnehmung externer Objekte in Reids Ansatz jedoch weitgehend unklar bleibt. Man fragt sich, warum Reid für die visuelle Wahrnehmung eine solche Sonderlösung vorschlägt, behauptet er doch, dass „every different perception is conjoined with a sensation that is proper to it“ (Reid, 1814, S. 249). Eine mögliche Erklärung wäre, dass Reid das Konzept visueller Empfindungen selbst für unplausibel hält. Das immer wieder von Reid angeführte paradigmatische Beispiel einer Empfindung ist der

Schmerz. Gelingt es ihm noch anhand des Beispiels des Geruchs der Rose die Unterscheidung zwischen einem Empfinden dieses Geruchs und der Wahrnehmung einer bestimmten Qualität eines externen Objektes zumindest auf den ersten Blick nicht unplausibel erscheinen zu lassen, sieht er möglicherweise im Fall visueller Wahrnehmung eine Schwierigkeit, eine analoge Unterscheidung treffen zu können („There seems to be no sensation that is appropriated to visible figure ... it seems to be suggested immediately by ... a material impression upon the retina“, Reid, 1813, S. 288). Wir werden in den anschließend darzustellenden Theorien, in denen explizit von ‘visuellen Empfindungen’ gesprochen wird, ausführlicher auf diese Frage eingehen. Plausibler erscheint indes, dass Reid auf die optischen Erkenntnisse seiner Zeit nicht verzichten wollte und diese in seine Theorie zu integrieren suchte, indem er die durch die optischen Gesetze beschreibbaren geometrischen Relationen des auf die Retinae projizierten Lichts selbst als Zeichen für die externen Objekte auffasst.

Durch diese Konzeption stellt sich auch für Reid die Tiefenwahrnehmung als ein eigenständiges Problem dar. Aufgrund ähnlicher geometrischer Argumente, die Berkeley vorbrachte, lasse sich laut Reid die Wahrnehmung der Distanz eines Objektes nur über besondere Zeichen für die Distanz, deren Funktion als Zeichen erlernt werden müsse, verständlich machen. Die von Reid vorgeschlagenen Mittel der Tiefenwahrnehmung stellen im Wesentlichen eine Zusammenstellung der von Alhazen, Descartes und Berkeley bereits vorgeschlagenen dar, wobei die Akkommodation und die Konvergenz von Reid eher im Sinne Berkeleys, d.h. als Bewegungen der Augenmuskeln, mit denen eine bestimmte Empfindung einhergeht, aufgefasst wird. Berkeley wollte die Auffassung dieser Augenbewegungen oder Muskelanstrengungen als Zeichen für Distanz dadurch verständlich machen, dass diese wiederholt gemeinsam mit Wahrnehmungen aus dem Tastsinn, über den allein Distanzwahrnehmung möglich sei (die aber unerklärt bleibt), aufgetreten wären, bei Reid hingegen lassen sich keine expliziten Hinweise auf die Beteiligung des Tastsinns finden. Es bleibt bezüglich der Konvergenz und der Akkommodation letztlich unklar, auf welche Weise es in Reids Ansatz zur wahrgenommenen Distanz kommt, so dass diese Wahrnehmung mit den bei Akkommodation und Konvergenz einhergehenden Muskelanstrengungen verknüpft werden könnte.

Bezüglich der verbleibenden, von Reid vorgeschlagenen Zeichen für Distanz ist man etwas ratlos, da diese sich auf Eigenschaften der *sichtbaren Erscheinung* beziehen sollen, die aber laut Reid selbst nicht wahrgenommen werde. Die ‘erscheinende’ oder ‘sichtbare Größe’ eines Objektes etwa verändere sich, so Reid, in Abhängigkeit der Distanz des Objektes. Ein Objekt in geringer Distanz erscheine größer als dasselbe Objekt in größerer Distanz, daher sei die erscheinende Größe als ein Zeichen für die Distanz des Objektes anzusehen. Aber wie kann der Wahrnehmende diese Erfahrung überhaupt machen? Reid scheint in seinen Beschreibungen auf eine andere Auffassung der *sichtbaren*

*Erscheinung* zurückgreifen zu müssen, damit diese verständlich werden. Wenn Reid mit der *sichtbaren Erscheinung*, wie oben argumentiert wurde, das auf die Retinae projizierte Lichtmuster als ein physikalisches Objekt meint, so sind diese Beschreibungen irreführend, da sich in diesem keine ‘erscheinende Größe’ eines Gegenstandes auffinden lässt, insofern es nicht *gesehen* wird. Wir stoßen hier auf ein Problem, das sich bereits bei den theoretischen Ansätzen von Alhazen und Descartes zeigte. Reids Ansatz scheint nur dann verständlich zu sein, wenn das auf die Retinae projizierte Licht selbst als ‘Bild der äußeren Objekte’ gesehen wird. Wenn der Wahrnehmende die *sichtbare Erscheinung* aber selbst, wie Reid zu meinen scheint, nicht sieht, so müsste auch hier eine personenähnliche Instanz, ein Homunkulus, angenommen werden.

## 2.5. Physiologisch orientierte Ansätze im 19. Jahrhundert

Zu Beginn des 19. Jahrhunderts kam es im Zuge zunehmender Erkenntnisse, die primär auf experimentellem Wege durch Verwendung chemischer, physikalischer und anatomischer Methoden gewonnen wurden, zur Herausbildung einer eigenständigen und wirkungsreichen wissenschaftlichen Disziplin, der Physiologie. Damit verbunden war auch eine in der Psychologie der Zeit deutlich spürbare Tendenz, das sich stetig erweiternde physiologische Wissen in psychologische Theorien zu integrieren oder auch psychologische Theorien gänzlich zu verwerfen, wenn sie mit neuen Erkenntnissen aus der Physiologie nicht kompatibel zu sein schienen. Psychologische Problemstellungen, so könnte man es auch formulieren, seien stets im Kontext der Physiologie zu behandeln. Diese Haltung lässt sich auch exemplarisch an dem zu dieser Zeit weit verbreiteten, von dem Physiologen Johannes Peter Müller geprägten Motto „*Nemo psychologus nisi physiologus*“ ablesen. Generell lässt sich feststellen, dass ein Großteil der psychologischen Theoriebildung dieser Zeit, insbesondere jedoch im Bereich der so genannten *Sinnespsychologie* (also dem, was heute als Wahrnehmungspsychologie bezeichnet wird), stark physiologisch bzw. stark an physiologischen Erkenntnissen ausgerichtet war. Die Motivation dafür scheint gewesen zu sein, der bis dahin oft als subjektiv und spekulativ verbrämten ‘Seelenlehre’ durch Unterfütterung mit physiologischen Befunden eine ‘objektive Basis’ zu geben und so die Psychologie in die von der Physik angeführten Reihe der *exakten Naturwissenschaften* einzureihen.<sup>50</sup>

<sup>50</sup>Eine ähnliche Tendenz lässt sich im Aufkommen der experimentellen Psychologie zu dieser Zeit feststellen. Gurwitsch (1966) umschreibt das Selbstverständnis der Psychologie zu dieser Zeit mit dem Begriff einer *Physik der Seele* („physics of the soul“) und beschreibt die damalige Intention zu dieser Auffassung folgendermaßen:

If physics is the more or less realized ideal of what a true science should be – the most perfect form and most adequate realization of scientific tendencies and intentions – then

Eine für die Sinnespsychologie dieser Zeit entscheidende und einflussreiche Entdeckung war das dann später sogenannte *Gesetz von Bell-Magendie*, welches zu Beginn des 19. Jahrhunderts unabhängig von Charles Bell in England und, knapp ein Jahrzehnt später, von François Magendie in Frankreich formuliert wurde. Dieses besagt grob, dass bei Rückenmarksnerven die anteriore Nervenwurzel allein motorisch efferente Nervenfasern und die posteriore Nervenwurzel allein afferente, d.h. sensorische Nervenfasern enthält und somit Efferenzen und Afferenzen räumlich getrennt vorliegen.<sup>51</sup> Für viele Physiologen dieser Zeit wurde diese über Tierexperimente gewonnene Entdeckung als empirischer Beleg dafür angesehen, dass Empfindungen und Nervenreizungen auf kausale Weise miteinander verbunden seien. Dadurch vollzieht sich in der Psychologie eine deutliche ideologische Wende hin zum *Physiologismus*, dessen Grundthese man in der Wendung zusammenfassen könnte, dass sich sämtliche psychischen Vorgänge auf physiologische Prozesse zurückführen lassen und nur durch diese hinreichend verstehbar werden (siehe dazu ausführlicher Figlio, 1975, sowie Heidelberger, 1997).

Darüber hinaus wurden diese Ergebnisse auch als ein Hinweis dafür angesehen, dass aufgrund der räumlichen Isoliertheit der Nervenfasern ‘sensorische Signale’ aus verschiedenen Teilen des Körpers auch räumlich isoliert übertragen werden, was nicht selten mit der Idee einherging, dass verschiedene Bereiche des Hirnes isoliert für verschiedene Sinnesempfindungen ‘zuständig’ seien. Dadurch kommt es nun auch zu einer deutlichen Bedeutungsverschiebung des Empfindungsbegriffs. Während Reid sowohl unter Empfindungen als auch unter Wahrnehmungen Akte des Geistes und *nicht* des Körpers versteht, gehen die physiologisch orientierten Theoretiker des 19. Jahrhunderts dazu über, Empfindungen als Zustände des Körpers, insbesondere als Zustände des Nervenapparates anzusehen. Manche sprechen dann zwar noch davon, dass diese Zustände des Nervenapparates über das Gehirn ‘zu Bewusstsein kommen’, und so eine gewisse Nähe zur Reidschen Konzeption gewahrt bleibt, es lässt sich allerdings auch häufig die Auffassung finden, dass Empfindungen eben genau diese Zustände des Nervenapparates *seien*. Insgesamt werden die Begriffe „Empfindung“ und „Wahrnehmung“ zu dieser Zeit auf vielfältigste Weise verwendet. Uphues (1888), der in Halle im späten 19. Jahrhundert Philosophie lehrte, beschreibt die zu dieser Zeit vorherrschenden Schwankungen in der Verwendungsweise des Empfindungsbegriffs folgendermaßen:

Wer die Schriften der meisten unserer Physiologen und Psychologen, auch der hervorragenden, liest, der wird sich oft eines Gefühls der Unsicherheit

---

psychology will have to keep itself as close as possible to physics if it is to be raised to the rank of science in the strict sense of the word. (Gurwitsch, 1966, S. 57)

<sup>51</sup>Mittlerweile hat man wenige Ausnahmen von dieser Regel feststellen können, etwa die für die Schmerzwahrnehmung wichtigen unmyelinisierten *Gruppe C*-Nervenfasern (vgl. Birbaumer & Schmidt, 2010).

nicht erwehren können, wenn er dem Worte Empfindung begegnet. Es ist merkwürdig, in welcher vielfältigen Bedeutung dieses Wort von im übrigen sehr sorgfältigen Schriftstellern häufig in demselben Abschnitt, ja auf derselben Seite gebraucht wird. Bald bezeichnet es etwas Physisches, Zustände der Nerven, die, wie es scheint, als solche ohne weitere psychische Vermittelung zum Bewußtsein kommen sollen; bald bezeichnet es etwas Psychisches, aber merkwürdig ebensowohl eine Art Erkenntnisvorgang, wie den Gegenstand dieses Erkenntnisvorganges; bald wird unter Empfindung eine Art Gefühlszustand verstanden, in dem sich Bewußtsein und Inhalt in keiner Weise unterscheiden läßt. Daneben spricht man dann trotzdem von unbewußten Empfindungen. (Uphues, 1888, S. 5f)

William James, der das 19. Jahrhundert dominierende Psychologe im nordamerikanischen Raum, kommt zu einem ähnlichen, wenngleich weniger elaboriertem Fazit: „The words sensation and perception do not carry very definitely discriminated meanings in popular speech, and in psychology also their meanings run into each other“ (James, 1890/2007, S. 1).

Eine gründliche Untersuchung der konzeptuellen Unterscheidung zwischen Wahrnehmung und Empfindung und deren Entwicklung von Reid bis in die heutige Zeit wäre ein wichtiges und sicher lohnenswertes Unterfangen, müsste aber aufgrund des Umfangs Gegenstand einer eigenständigen Arbeit sein. Für eine hervorragende Darstellung und Analyse der unterschiedlichen Verwendungsweisen des Empfindungs- und Wahrnehmungsbegriffes in der deutschen Psychologie des ausgehenden neunzehnten und beginnenden zwanzigsten Jahrhunderts siehe neben Uphues (1888) auch Hofmann (1913), eine ebenfalls lohnenswerte, wenngleich knappe Darstellung der zunehmenden ‚Physiologisierung‘ des Empfindungsbegriffs gibt Boring (1942). Wir wollen uns nun zwei zu dieser Zeit formulierten Ansätzen zuwenden, die eine deutliche Wirkung auf die weitere Entwicklung wahrnehmungstheoretischer Vorstellungen ausüben sollten.

### 2.5.1. Müllers Gesetz der spezifischen Sinnesenergien

Eine für die weitere Entwicklung der Sinnespsychologie wegweisende Idee war das von Johannes Müller (1801-1858) erstmals 1826 formulierte sogenannte *Gesetz der spezifischen Sinnesenergien*, das von Hermann von Helmholtz, auf dessen eigene Wahrnehmungstheorie dieses einen großen Einfluss ausüben sollte, als der „bedeutsamste Fortschritt, den die Physiologie der Sinnesorgane in neuerer Zeit gemacht hat“ (Helmholtz, 1896/2002d, S. 98) bezeichnet wurde. Müller, der als der bedeutendste deutsche Physiologe des zweiten Viertels des 19. Jahrhunderts gelten kann (Hatfield, 1990) und Hermann von Helmholtz, Emil du Bois-Reymond, Ernst Haeckel und Wilhelm Wundt

zu seinen Schülern zählen konnte, verfasste mit seinem umfangreichen, in zwei Bänden erschienenen „*Handbuch der Physiologie des Menschen*“ das Standardwerk der Physiologie seiner Zeit. In diesem widmet sich Müller in mehreren Abschnitten auch der Natur der Empfindungen und der Frage, ob, und wenn ja, welche Art Zusammenhang zwischen den auf die Sinnesorgane einwirkenden Objekten der Außenwelt und den Empfindungen und Wahrnehmungen bestehe.

Unter Empfindung versteht Müller allgemein die „allen Sinnesnerven gleiche Leitung auf das Sensorium“ (Müller, 1840, S. 249), wobei der Terminus „Sensorium“ von Müller im Sinne des Gehirns als dem „Organe des Bewusstseyns“ (Müller, 1837, S. 695) verwendet wird. Allerdings schwankt – das obige Zitat von Uphues deutet es schon an – auch Müller häufig in der Verwendungsweise des Empfindungsbegriffes, so heißt es an mehreren Stellen (etwa Müller, 1837, S.705), dass die *Nerven selbst* Empfindungen haben, an anderen Stellen (etwa Müller, 1837, S. 698 oder Müller, 1840, S. 249) werden Empfindungen beschrieben als Zustände der Nerven, die uns „zu Bewusstsein kommen“. In Müllers Schriften überwiegt jedoch insgesamt die Charakterisierung, dass die Nerven selbst empfinden.

Bis zum Durchbruch der Physiologie als eigenständiger wissenschaftlicher Disziplin sei es laut Müller (1837, S. 779) Konsens gewesen, „die Sinnesnerven als blosse Leiter für die Qualitäten der äußeren Dinge“ anzusehen, die passiv die Eigenschaften der Objekte der Aussenwelt an das Bewusstsein „weiterleiten“. Erst die neueren physiologischen Entdeckungen hätten gezeigt, dass dies eine irriige Ansicht sei; die Empfindungen seien als Zustände der von äußeren Reizen affizierten Nerven aufzufassen:

Die Empfindung ist also nicht die Leitung einer Qualität oder eines Zustandes der äußeren Körper zum Bewusstseyn, sondern die Leitung einer Qualität, eines Zustandes unserer Nerven zum Bewusstseyn, veranlasst durch eine äußere Ursache. Wir empfinden nicht das Messer, das uns Schmerz verursacht, sondern den Zustand unserer Nerven schmerzhaft. (Müller, 1837, S. 780)

Anhand des angeführten Zitates lässt sich die angesprochene physiologische Interpretation des Reidschen Empfindungsbegriffes gut erkennen. Die Qualität der Empfindungen bezieht sich zwar auch bei Müller nicht auf ein vom Empfindenden unabhängiges Objekt, aber sie wird als physiologischer Zustand unserer Nerven aufgefasst, der an das Gehirn weitergeleitet wird; Empfinden sei kein psychologischer, sondern ein physiologischer Vorgang. Damit zusammenhängend führe laut Müller derselbe Reiz nicht stets zur selben Empfindung (oder Wahrnehmung), sondern die Empfindungsqualität hänge wesentlich davon ab, *welche* Nerven, oder allgemeiner, welche Sinnesmodalität, durch diesen erregt werde. Ein Lichtreiz habe etwa eine ‘Lichtempfindung’ zur Folge, wenn

er auf die Retina treffe, derselbe Reiz führe hingegen (hinreichende Energie vorausgesetzt) zu einer Wärmeempfindung, wenn er auf eine Hautstelle treffe.<sup>52</sup> Umgekehrt führten unterschiedliche Reize, die denselben Nerv affizieren, zu gleicher Empfindungsqualität. Werde etwa der optische Nerv durch mechanische Einwirkung, elektrische Reizung oder Licht affiziert komme es stets zu einer 'Lichtempfindung'. Die Qualität einer Empfindung (etwa Wärme, Farbe, Schmerz usw.) sei demnach nicht von der Art des Reizes, sondern von der Art der durch diesen affizierten Nerven bestimmt, wobei jeder Typus sensorischer Nerven durch Reizung in der Lage sei, eine ganz bestimmte Empfindungsqualität zu erzeugen, die Müller die „Sinnesenergie“ der Nerven nennt. Die Sinnesenergien seien *spezifisch*, da die Reizung eines bestimmten Typs sensorischer Nerven stets nur eine Empfindungsqualität zur Folge habe.<sup>53</sup> Diese Punkte zusammenfassend formuliert Müller das „Gesetz der spezifischen Sinnesenergien“ schließlich wie folgt:

Die Sinnesempfindung ist nicht die Leitung einer Qualität oder eines Zustandes der äusseren Körper zum Bewusstsein, sondern die Leitung einer Qualität, eines Zustandes eines Sinnesnerven zum Bewusstsein, veranlasst durch eine äussere Ursache, und diese Qualitäten sind in den verschiedenen Sinnesnerven verschieden, die Sinnesenergien. (Müller, 1840, S. 254)

Die Empfindungen als Zustände der Nerven würden anschließend in einem weiteren Schritt durch „die Vorstellung“ in wesentlicher Weise verändert, wobei nicht klar wird, worauf sich dieser Begriff beziehen soll. Müller formuliert in diesem Zusammenhang auch Ansätze einer Theorie der visuellen Wahrnehmung, die allerdings insgesamt als äußerst vage zu charakterisieren ist. Nach umfangreichen Darstellungen der Anatomie des Auges und der mit dem Lichteinfall ins Auge verbundenen optischen Gesetze stellt Müller fest, dass wir bei Erregung verschiedener, räumlich getrennter Stellen der Netzhaut unterschiedlich lokalisierte Lichtpunkte empfinden würden, da die räumliche Beziehung zwischen den erregten Netzhautstellen über die räumliche Anordnung der optischen Nerven, mit denen diese Netzhautstellen verbunden seien, erhalten bleibe. Die räumliche Anordnung der erregten Netzhautstellen werde über die räumlich getrennten optischen Nerven an räumlich getrennte Areale des Gehirns weitergeleitet. Bei Erregung der Netzhaut durch von einem Körper reflektierten Licht würden wir daher zunächst lediglich „flächenhafte Bilder“ (Müller, 1840, S.268) *empfinden*, was wir

---

<sup>52</sup>„Dieselbe äußere Ursache erregt in den verschiedenen Sinnen verschiedene Empfindungen, nach der Natur jedes Sinnes, nämlich das Empfindbare des bestimmten Sinnesnerven“ (Müller, 1840, S. 251).

<sup>53</sup>„Ein Sinnesnerv scheint nur einer bestimmten Art der Empfindung und nicht derjenigen der übrigen Sinnesorgane fähig zu seyn, und es kann daher auch keine Vertretung eines Sinnesnerven durch einen andern davon verschiedenen stattfinden“ (Müller, 1840, S. 258).

jedoch schließlich *sehen* würden, seien körperliche Gegenstände außerhalb von uns.<sup>54</sup> Dies sei auf eine die Empfindungen ‘modifizierende Tätigkeit’ der Vorstellung zurückzuführen, die selbst durch Erfahrungen veränderlich sei:

Es liegt nicht in der Natur der Nerven selbst, den Inhalt ihrer Empfindungen ausser sich gegenwärtig zu setzen, die unsere Empfindungen begleitende, durch Erfahrung bewährte Vorstellung ist die Ursache dieser Versetzung. ... Die Vorstellung wirkt aber bald so modificierend und herrschend in die Acte des Gesichtssinnes ein, dass der Gesichtssinn uns nach aussen zu wirken scheint, dass an die Stelle der flächenhaften Bilder in der Vorstellung die körperlichen Gegenstände selbst treten. (Müller, 1840, S. 268)

Müller scheint an dieser Stelle seiner Theorie vor dem Problem zu stehen, dass ich Empfindungen als Zustände *meiner* Nerven eben auch als *meinem Körper zugehörige* Empfindungen empfinden sollte (wie etwa Schmerz); wenn ich sehe tritt mir jedoch, wenn man sich so ausdrücken möchte, eine Welt und sich gegenwärtig darin befindende Objekte *außerhalb* von mir entgegen. Darüber hinaus sollte das, was aufgrund der Erregungszustände der verschiedenen Netzhautstellen laut Müller empfunden werde, „flächenhafte Bilder“ sein, was allerdings gesehen wird, sind Körper. Müller bedient sich nun einer zugegeben recht einfachen wie vagen Lösung dieses Problems, indem er sowohl das Versetzen von ‘Innen’ nach ‘Außen’ als auch die hinzukommende Räumlichkeit einer nicht weiter explizierten Instanz zuschreibt, die er *Vorstellung* nennt. Wie genau diese ‘modifizierende Tätigkeit’ geschehen soll oder warum überhaupt eine solche Trennlinie zwischen ‘Innen’ und ‘Außen’ in unserer Wahrnehmung besteht, dazu äußert sich Müller nicht weiter, abgesehen davon, dass er der Ansicht zu sein scheint, dass die Trennung zwischen Innen- und Außenwelt über ein im Mutterleib bereits vorhandenes elementares Gefühl der räumlichen Ausdehnung des eigenen Körpers in gewisser Weise *angeboren* sei (siehe dazu auch Lenoir, 1993).<sup>55</sup>

Ein Kind sei allerdings nicht in der Lage, Objekte als *unterschiedlich* weit wahrzunehmen, für es „liegt Alles in gleicher Ferne“ (Müller, 1840, S. 362). Diese Fähigkeit müsse erst erlernt werden, Müller spricht diesbezüglich auch von der „Beurtheilung der Nähe und Ferne“, die keine Sache der Empfindung, sondern des Verstandes sei (Müller,

<sup>54</sup>Anhand dieser Passage lässt sich erneut gut ablesen, auf welche unterschiedlichen Arten der Empfindungsbegriff von Müller gebraucht wird. In diesem Beispiel etwa ist vollkommen unklar, was die Wendung ‘flächenhafte Bilder empfinden’ bedeuten soll, wenn doch Empfindungen *Zustände der Nerven* seien. Darüber hinaus widerspricht in diesem Fall die Alltagsbeobachtung der Müllerschen Konzeption. Ich habe zwar eine Vorstellung davon, was es heißt, Schmerz zu empfinden, aber ich empfinde nicht im selben Sinne „flächenhafte Bilder“, die dann modifiziert werden, sondern ich sehe stets räumliche Gegenstände.

<sup>55</sup>„Das Kind wird schon mit dunklen Vorstellungen von Aussendungen ausser seinem lebenden Körper geboren, mit Vorstellungen von ihrer Realität als Ursache von Empfindungen“ (Müller, 1840, S. 271).

1840, S. 362). Was man über den Sehsinn empfinde, seien lediglich „flächenhafte Bilder“, es bedürfe daher zusätzlicher Mittel der Distanzwahrnehmung. Müller zählt als wichtigstes Mittel zur Beurteilung der Entfernung die Größe des ‘retinalen Bildes’ des Objekts auf. Ein Objekt werde als umso weiter entfernt beurteilt, je kleiner das ‘retinale Bild’ im Vergleich zu jenem ‘retinalen Bild’ sei, das entstehe, wenn das Objekt sich unmittelbar vor einem befinde. Weiterhin werde ein Objekt *A* als weiter entfernt von einem anderen Objekt *B* beurteilt, wenn *A* von *B* „zum Theil bedeckt“ (Müller, 1840, S. 362) werde. Müller bringt hier den heute auch als *Verdeckung* bezeichneten Tiefencue ins Spiel, der in den bisher dargestellten Ansätzen noch keine Erwähnung fand. Weiterhin sei auch die Konvergenz der Augen ein Mittel, das Urteile über die Distanz eines Objektes erlaube, welches jedoch allgemein „überschätzt“ werde (Müller, 1840, S. 362). Müller fasst die Konvergenz eher im Sinne Descartes’, also in geometrischen Termini, und weniger als Empfindungen der zugehörigen Augenmuskeln, wie Reid und Berkeley, auf.

Aufgrund der bisherigen Ausführungen könnte man vermuten, dass wir laut Müller über unsere Empfindungen nichts über eine mögliche Außenwelt, sondern allein etwas über die Zustände unserer Nerven erfahren würden und unsere Wahrnehmungen in diesem Sinne ‘subjektiv’ seien. Da die sensorischen Nervenfasern allerdings nun selbst Bestandteil der physikalischen Welt seien und daher räumliche Ausdehnung, eine bestimmte Anordnung etc. aufweisen und bei Erregung auf physikalische und/oder chemische Weise verändert würden, würden unsere Empfindungen oder Wahrnehmungen, da sie Wahrnehmungen der (durch die äußere Reizung) veränderten Zustände der Nerven seien, auch auf (physikalische) Zustände der Außenwelt verweisen:

Die Sinnesnerven empfinden zwar zunächst nur ihre eigenen Zustände, oder das Sensorium empfindet die Zustände der Sinnesnerven; aber dadurch dass die Sinnesnerven als Körper die Eigenschaften anderer Körper theilen, dass sie im Raume ausgedehnt sind, dass ihnen eine Erzitterung mitgetheilt werden kann und dass sie chemisch, durch die Wärme, und die Electricität verändert werden können, zeigen sie bei ihrer Veränderung durch äussere Ursachen, dem Sensorium ausser ihrem Zustande auch Eigenschaften und Veränderungen der Aussenwelt an, in jedem Sinne verschieden nach dessen Qualitäten oder Sinnesenergien (Müller, 1840, S. 262).

Es wird nicht wirklich klar, wie dieses Argument Müllers zu verstehen ist, es ist eine der wenigen, durch Nummerierung deutlich hervorgehobenen Passagen des die Sinne behandelnden Teils des *Handbuchs*, die weithin unkommentiert bleibt. An verschiedenen Stellen (etwa Müller, 1840, S.268) spricht Müller zwar davon, dass wir mit „einiger Sicherheit“ auf die Eigenschaften der die Nerven affizierenden Körper „schliessen“ würden, führt diesen Punkt jedoch nicht weiter aus. Das obige Zitat scheint mit dem Gesetz

der spezifischen Sinnesenergien auch nur schwer in Übereinstimmung zu bringen sein. Ein Wärmeempfindung etwa könne laut diesem eine Folge der Einwirkung vollkommen unterschiedlicher Reize sein (Licht, eine chemische Substanz, Reibung), es lässt sich daher nicht ohne Weiteres ersehen, auf welche Weise sich anhand einer bestimmten Empfindung mit 'einiger Sicherheit' Rückschlüsse auf Eigenschaften der diese verursachenden Aussenwelt ziehen lassen sollen.

### Zusammenfassung

Insgesamt sind die von Müller dargelegten Ausführungen zur visuellen Wahrnehmung (abgesehen von den anatomischen Beschreibungen des Auges) mehr oder weniger schemenhaft und bleiben erklärungsbedürftig. Der Übergang von einer äußeren Einwirkung auf die Sinnesnerven zu einer Empfindung 'flächenhafter Bilder', die als *Zustand* räumlich getrennter Nerven aufgefasst wird, die mit räumlich getrennten Stellen der Netzhaut verbunden seien, bleibt unklar, ebenso wie die „modifizierende Tätigkeit der Vorstellung“, welche zu einer Wahrnehmung nicht dieser Empfindungen, sondern von Objekten *außerhalb* meiner selbst führe. Müller zählt ebenfalls mehrere Mittel zur Distanzwahrnehmung auf, u.a. die Konvergenz und die 'retinale Größe' eines Objektes. Interessant ist, dass im Fall der 'retinalen Größe' nun offensichtlich ist, dass sich Müller damit auf das retinale Lichtmuster bezieht, was in den vorangehenden Ansätzen, in denen von 'erscheinender Größe' gesprochen wurde, nicht so offensichtlich war. Auch in Müllers Konzeption ist diese Auffassung jedoch problematisch. Das retinale Lichtmuster kann vom Wahrnehmenden selbst nicht wahrgenommen werden, es ist selbst kein 'flächenhaftes Bild'. Durch die Behauptung, dass die räumliche Beziehung zwischen den erregten Netzhautstellen über die räumliche Anordnung der optischen Nerven, mit denen diese Netzhautstellen verbunden seien, erhalten bleibe, scheint Müller diesen Punkt in gewisser Weise umgehen zu wollen, indem er das Ergebnis der Lichteinwirkung als Empfindung eines 'flächenhaften Bildes' bezeichnet, dessen Status aber unklar ist. Es gibt hier mehrere Möglichkeiten der Interpretation, die jedoch allesamt unbefriedigend erscheinen. Das auf die Retina projizierte Lichtmuster könnte bereits von den Nerven selbst gleichsam *als* 'flächenhaftes Bild' verschiedener Objekte *wahrgenommen* und als ein solches an das Gehirn 'weitergeleitet' werden, was jedoch nur wenig Erklärungswert hätte. Das retinale Lichtmuster könnte aber auch lediglich als bloßes 'Erregungsmuster' durch die Nerven weitergeleitet werden, dann müsste allerdings verständlich gemacht werden, an welcher Stelle und durch welche Instanz es zu einer 'Interpretation' dieses Erregungsmusters *als* 'flächenhaftes Bild' komme, so dass die 'retinale Größe' eines *Objektes* als Mittel der Distanzwahrnehmung genutzt werden könne. Der Übergang vom retinalen Lichtmuster zur Empfindung 'flächenhafter Bilder' bleibt bei Müller im Dunkeln, ist aber für seine Konzeption der Mittel zur Distanzwahr-

nehmung essentiell. Ähnliche Bemerkungen lassen sich bezüglich der Verdeckung als ein von Müller vorgeschlagenes Mittel der Distanzwahrnehmung machen. Damit dieses überhaupt eine solche Funktion erfüllen kann, müssten die ‘flächenhaften Bilder’ den Status von Bildern sich verdeckender Objekte haben; das retinale Lichtmuster selbst ist kein Bild sich verdeckenden Objekte. Möglicherweise ist es auch Müllers Ansicht, dass eine solche ‘Interpretation’ die Aufgabe der Vorstellung sei, aber dann würde das Problem lediglich verschoben werden, da nun erstens zu erklären wäre, wie die Vorstellung dies anstellt und zweitens, welche Verbindung zwischen der Wahrnehmung oder Interpretation der Vorstellung und dem, was *vom Wahrnehmenden* gesehen wird, besteht. Auch in Müllers Ansatz muss demnach bereits ein Wahrnehmungsakt bzw. eine schon ‘wahrnehmende Instanz’ vorausgesetzt werden, damit die vorgeschlagenen Mittel der Distanzwahrnehmung verständlich werden. Um im retinalen Lichtmuster die ‘retinale Größe’ eines Objektes oder sich verdeckende Objekte identifizieren zu können, muss es als ein Bild *gesehen* werden.

Da Müller in seinen Ausführungen explizit von ‘visuellen Empfindungen’ spricht und dies eine zentrale Annahme sowohl Müllers als auch der im Folgenden darzustellenden Theorien ist, soll an dieser Stelle der im Abschnitt über Reid angesprochene Punkt wieder aufgegriffen werden, ob diesem Terminus überhaupt eine sinnvolle Bedeutung verliehen werden kann. Reid führt ebenso wie Müller als paradigmatischen Fall einer Empfindung den Schmerz an. Wie Reid richtig feststellt, sind Schmerzen etwas, das man *hat*. Bezüglich Schmerzen hat es keinen Sinn zu sagen, dass man diese wahrnimmt oder nicht wahrnimmt, dass man weiss oder nicht weiss, dass man sie hat oder dass Schmerzen einem bewusst oder nicht bewusst sind. Üblicherweise lässt sich das Haben von Schmerzen auch lokalisieren, wenn ich etwa Zahnschmerzen habe, so ist der Ort des Schmerzes ein anderer als im Fall einer Fussverletzung. Es ist meines Erachtens nicht einzusehen, wie sich eine solche Beschreibung auf das Feld des Sehens übertragen lassen sollte. Die ‘visuellen Empfindungen’, die Müller als ‘flächenhafte Bilder’ beschreibt, kann ich nicht im selben Sinn *haben*, wie ich Schmerzen habe. Werde ich aufgefordert, meine Schmerzen zu beschreiben, kann ich sie beispielsweise als stechend, brennend, pochend oder stärker bei Bewegung und schwächer bei Ruhe charakterisieren. Würde ich aufgefordert werden, meine ‘visuellen Empfindungen’ zu beschreiben, wüsste ich überhaupt nicht, was ich beschreiben sollte, außer das, was ich *sehe*, nämlich *Objekte*.<sup>56</sup>

<sup>56</sup>Merleau-Ponty (1966, S. 22) argumentiert aus phänomenologischer Perspektive ähnlich. Der Begriff der Empfindungen, insbesondere auf den Sehsinn bezogen, finde keine Unterstützung durch unsere Erfahrung, die „angebliche Evidenz des Empfindens gründet sich nicht auf ein Zeugnis unseres Bewusstseins, sondern auf ein Vorurteil“. Was wir perzeptuell erleben, seien keine ‘internen Empfindungen’, sondern externe *Gegenstände*. Wir könnten nicht nicht-intentional sehen. Wenn man davon spreche, dass man etwa ‘Rot empfinde’, so habe das eine ganz andere Bedeutung als die Aussage, dass man Schmerzen empfinde. Ersteres sei keine Empfindung, sondern etwas *Empfundenes*, nämlich Eigenschaften eines Gegenstandes. Ich könne nicht die Qualität „Rot“ empfinden unabhängig von *etwas*, das mir rot erscheint. Die Begriffe, mit denen üblicherweise ‘visuelle Empfindungen’ beschrieben

Ich kann meine visuellen Wahrnehmungen als Wahrnehmungen *von etwas* beschreiben, aber nicht von diesen unterschiedene ‘visuelle Empfindungen’, die, wie Schmerz, keinen intentionalen Charakter aufweisen sollen. Zu sagen, dass ‘visuelle Empfindungen’ in unserem Erleben nicht auffindbar wären, da sie uns üblicherweise nicht bewusst seien, ist ebenfalls nicht sinnvoll, da Empfindungen doch gerade dadurch gekennzeichnet sein sollen, dass man sie *hat*. Von Schmerzen zu reden, die uns nicht bewusst seien, hat keinen Sinn. Der Terminus der ‘visuellen Empfindungen’ ist äußerst erklärungsbedürftig, wird aber von Müller nicht erklärt, sondern es wird mit Verweis auf Schmerz oder Wärme suggeriert, dass dem Begriff der Empfindung im Fall des Sehens die gleiche Bedeutung zukommen würde, was aber nicht plausibel gemacht werden kann. Bei ‘visuellen Empfindungen’ scheint es sich eher um einen durch Wahrnehmungstheoretiker eingeführten ‘Kunstbegriff’ zu handeln, dem nicht die gleiche Bedeutung zukommen kann wie Schmerz als einer Empfindung, mögliche andere Bedeutungen aber auch nicht eingeführt werden. Zudem ist bei Müller die Kennzeichnung der Empfindungen als bestimmter Zustände der Nerven kritisch zu betrachten. Dass bei spezifischer Reizung bestimmter Nerven ein Schmerzempfinden auftritt, bedeutet zunächst nicht mehr, als dass eine Reizung der Nerven eine kausale Bedingung dafür ist, dass jemand Schmerzen hat, nicht dass wir den ‘Zustand’ der Nerven empfinden würden, insofern man dieser Wendung überhaupt eine sinnvolle Bedeutung geben kann.

### 2.5.2. Lotzes Theorie der Lokalzeichen

Rudolph Hermann Lotze (1817-1881), der in Leipzig sowohl Physiologie als auch Philosophie studiert hatte und heutzutage fast in Vergessenheit geraten ist, gehörte in der zweiten Hälfte des 19. Jahrhunderts zu den international bekanntesten und wichtigsten Vertretern deutscher Philosophie, insbesondere der Metaphysik.<sup>57</sup> Anhänger unterschiedlicher philosophischer Schulen, etwa Brentano, Husserl, Frege oder Cassirer, knüpften an verschiedene seiner Positionen an (Plümacher, 2004). Das umfangreiche Werk Lotzes ist durchzogen von dem Anspruch, die Philosophie und die Naturwissenschaften (insbesondere die Physiologie) in einer Weise zusammenzuführen, so dass

---

werden (rot, hell etc.), seien der Sprache über *von uns wahrgenommenen Dingen* entnommen. Die Qualitäten der von uns wahrgenommenen Dinge würden so einfach dem Bewusstsein zugeschrieben:

Man begeht den von der Psychologie sogenannten „experience error“, indem man, was wir von den Dingen wissen, unserem unmittelbaren Bewusstsein zuschreibt. Aus Wahrgenommenem macht man Wahrnehmung. Und da das Wahrgenommene selbst natürlich nicht zugänglich ist ohne Wahrnehmung, begreift man schließlich das eine so wenig wie das andere. (Merleau-Ponty, 1966, S. 23)

<sup>57</sup>Eine ausführliche Darstellung des Lebensweges Lotzes und der Grundzüge seines Denkens im Kontext seiner Zeit bietet Pester (1997), eine kompaktere Darstellung lässt sich Hall (1912, Kap. 2) finden.

einerseits die Philosophie die stetig wachsenden Erkenntnisse aus den Naturwissenschaften in die Theoriebildung einbeziehen zu habe, andererseits die Naturwissenschaften aber nicht das letzte Wort bezüglich genuin philosophischer Fragen haben könne. Allein die Titel vieler Werke Lotzes, wie etwa die 1852 erschienene „*Medizinische Psychologie oder Physiologie der Seele*“ lassen diesen Anspruch einer Zusammenführung philosophischer und naturwissenschaftlicher Konzepte erahnen. Lotzes wichtigster und wirkungsmächtigster Beitrag zum Thema der visuellen Wahrnehmung speziell der Tiefenwahrnehmung war das Konzept der sogenannten *Lokalzeichen*, welches für die Entwicklung der Zeichentheorie von Hermann von Helmholtz eine wichtige Rolle spielte und das im Folgenden im Vordergrund der Darstellung stehen soll.<sup>58</sup>

Lotzes Ausführungen zu den Lokalzeichen, das soll vorangestellt werden, beziehen sich – unter der Voraussetzung, dass sich erstens Objekte außerhalb von uns in einer räumlichen Relation zueinander befinden und zweitens wir als Menschen von Natur aus räumlich wahrnehmen – allein auf die Frage, *wie* wir die räumlichen Relationen der äußeren Objekte in der Wahrnehmung eines einheitlichen Raumes sozusagen ‘wiederfinden’. Mit anderen Worten soll die Theorie der Lokalzeichen keine Erklärung dafür liefern, warum wir *überhaupt* räumlich wahrzunehmen in der Lage sind, sie soll vielmehr zur Klärung der Frage beitragen, *nach welchen Prinzipien* die von uns wahrgenommenen räumlichen Relationen ‘erzeugt’ werden, so dass wir Objekte als unterschiedlich im Raum lokalisiert wahrnehmen.

Lotze bedient sich ebenfalls des Empfindungskonzeptes, weist aber die Müllersche Auffassung von Empfindungen als Zuständen der Nerven entschieden zurück. Es sei ein fundamentaler Irrtum, „den Nervenprozess für einen psychischen Vorgang, für eine schon zur Empfindung gewordene Erregung zu halten, die nur noch der Fortleitung bis zum Gehirn bedürfe“ (Lotze, 1852, S. 176). Die physische Erregung der Nerven werde zwar an das Gehirn geleitet, müsse aber dort in „Wechselwirkung mit der Seele“ treten, um Empfindungen als *psychische* Gegebenheiten zur Folge zu haben, die Lotze als „das bewusste Empfinden einer einfachen Sinnesqualität, eines Tons, einer Farbe“ (Lotze, 1852, S. 178) charakterisiert. Die Seele (Lotze spricht auch von dem Bewusstsein) könne man sich nur als etwas nicht räumlich Ausgedehntes, Immaterielles vorstellen, sie sei nicht mit einem physischen Organ zu identifizieren; das Bewusstsein sei kein Objekt im üblichen Sinne, sondern wird von Lotze eher als eine Art Tätigkeit oder Aktivität aufgefasst. Der Übergang von physiologischen Prozessen zu psychischen Gegebenheiten könne „nur in der Seele selbst, nie in irgend einem körperlichen Organ stattfinden“ (Lotze, 1852, S. 178), wobei Lotze nicht näher auf diesen Übergang eingeht, sondern

---

<sup>58</sup>Das Konzept der Lokalzeichen war nicht auf die visuelle Wahrnehmung begrenzt, sondern wurde von Lotze und anderen Theoretikern auch zur Erklärung der Wahrnehmung anderer Sinnesmodalitäten, insbesondere des Tastsinns herangezogen (siehe etwa Wundt, 1862).

sie eher als eine unbestreitbare, wenngleich nicht weiter zu erklärende Tatsache zu halten scheint.

Visuelle Wahrnehmung, das Sehen externer Objekte, beruhe laut Lotze auf einer zusätzlichen 'vorstellenden Tätigkeit' des Bewusstseins. Es bedürfe zunächst des Mediums Licht, welches mit den Objekten der Außenwelt interagiere und, gewissen geometrischen Gesetzen folgend, auf unsere Augen einwirke. Im durch Reflektion von den Objekten auf die Netzhaut projizierten Lichtmuster seien zwar geometrische Relationen enthalten, die denen der äußeren Objekte zueinander in gewisser Weise entsprächen und durch geometrische bzw. optische Gesetzmäßigkeiten beschreibbar sind, das retinale Lichtmuster selbst sei aber keinesfalls hinreichend, um die Wahrnehmung von in räumlichen Relationen zueinander stehenden Objekten zu erklären:

Im Auge hat die Natur sorgsame Anstalt getroffen, daß die Lichtstrahlen, die von einem leuchtenden Punkte kommen, sich auf der Netzhaut wieder in *einem* Punkte sammeln, und daß die verschiedenen Bildpunkte, welche hier entstehen, dieselbe gegenseitige Lage einnehmen, wie die Objectpunkte außer uns, denen sie entsprechen. Ohne Zweifel ist dieses so sorgfältig vorbereitete sogenannte ‚Bild des Gegenstandes‘ eine unentbehrliche Bedingung dafür, daß wir den Gegenstand in seiner wahren Gestalt und Lage vorstellen können. Allein es ist der Grund aller Irrthümer in dieser Sache, zu glauben, daß das bloße Dasein dieses Bildes ohne weiteres schon unsere Vorstellung von der Lage seiner Theile erkläre (Lotze, 1881, S. 27f.).

Das retinale Lichtmuster errege die verschiedenen Nervenfasern zwar in einer gewissermassen 'räumlichen Weise', nämlich derart, dass Lichtreizung auf verschiedenen Punkten der Retina zur Erregung verschieden räumlich lokalisierter Nervenfasern führe. Dass man aber der etwa von Müller vertretenen Position Glauben schenke, dass die durch räumlich verschieden lokalisierte Lichtpunkte auf der Retina räumlich unterschiedliche Erregung der Nervenfasern erklären könne, warum wir verschieden lokalisierte Objekte wahrnehmen, sei lediglich eine feinere Form der überwunden geglaubten antiken Seheorie, dass sich kleine 'Bildchen' (εἰδωλα) von den Objekten in der Welt lösen und diese Bilder einfach vom Bewusstsein, einem Spiegel gleich, gleichsam 'aufgenommen' würden:

Dieses Vorurteil nun, dass die Form, in welcher eine Anzahl gleichzeitiger *Erregungen* im Nervensystem räumlich neben einander verlaufen, unmittelbar der Grund für eine ähnliche räumliche Disposition der *Empfindungen* enthalte, ist der Irrtum, der in den allermannigfachsten Gestalten wiederkehrend, die Erklärung der sinnlichen Weltauffassung überall verdirbt. (Lotze, 1852, S. 328)

Die im retinalen Lichtmuster enthaltenen geometrischen Relationen, die zu räumlich getrennter Erregung der Nervenfasern führten, könnten nicht einfach in das Bewusstsein 'übergehen', da dieses kein Objekt im herkömmlichen Sinne sei und daher auch keine räumliche Ausdehnung habe.<sup>59</sup> Selbst wenn räumlich verschiedene Erregung der Netzhaut räumlich verschieden lokalisierte Nervenirregung zur Folge hätte, so müssten diese räumlichen Relationen doch notwendigerweise verlorengehen und in ein *Gleichzeitiges*, ohne räumliche Bestimmung, übergehen, wenn Empfindungen auf einer Tätigkeit des Bewusstseins beruhten, was Lotze für nicht widerlegbar hält:

Unfehlbar muss auch für sie [die räumlich verschieden lokalisierten Erregungen der Nervenfasern] in ihrer Ueberleitung zur Seele der Punkt eintreten, wo ihre ganze geometrische Gestalt spurlos zerstört wird, und wo sie in der Seele durch eine *Summe intensiver Erregungen* ersetzt werden, die gleich einer Vielheit von Tönen, keine Andeutung räumlicher Ausdehnung oder Lage enthält. (Lotze, 1852, S. 328)

An dieser Stelle wird Lotze in seinen Beschreibungen etwas unschärfer, er greift jedoch des Öfteren auf das Bild eines Akkordes zurück, welcher aus mehreren verschiedenen, gleichzeitig vorhandenen, aber in keinerlei räumlichen Relationen zueinander stehenden Tönen bestehe.<sup>60</sup>

Um die Wahrnehmung verschieden lokalisierter und verschieden weit entfernter Objekte zu ermöglichen, müsse das Bewusstsein daher aus dieser „Summe intensiver Erregungen“ ohne jegliche räumliche Bestimmungen die räumlichen Relationen der Objekte aktiv *rekonstruieren*, oder *wiedererzeugen*. Lotzes weitere Beschreibungen legen nahe, dass es sich bei den von uns wahrgenommenen räumlichen Relationen um eine 'veridikale Repräsentation' der äußeren Gegebenheiten handeln soll. Die räumlichen Relationen der Objekte der Außenwelt seien in dem retinalen Lichtmuster veridikal repräsentiert, könnten jedoch nicht unmittelbar als solche in das Bewusstsein übergehen. Daher müsse es zu einer anderen, 'nicht räumlichen Repräsentation' dieser kommen, aus denen das Bewusstsein die räumlichen Verhältnisse der Außenwelt rekonstruieren könne.

Sollen wir daher eine Anschauung der wirklichen Lage äusserer Objecte gewinnen, so kann es nicht auf dem Wege der *Auffassung*, sondern auf dem der *Wiedererzeugung der Räumlichkeit* sein. Ueberall wird das Extensive in ein Intensives verwandelt, und aus diesem erst muss die Seele eine neue innerliche Raumwelt reconstruieren, in welcher die Bilder der äusseren Objecte

<sup>59</sup>„Die Seele ist weder ein widerstandsloses, noch ein ausgedehntes Mittel, in welchem die räumlichen Bilder der Objekte Platz nehmen können“ (Lotze, 1852, S. 327).

<sup>60</sup>„Viele Eindrücke sind in der Seele zugleich, aber nicht räumlich neben einander, sondern bloß so zusammen, wie die gleichzeitigen Töne eines Accords, d.h. qualitativ verschieden, aber nicht neben, über oder unter einander“ (Lotze, 1881, S. 29).

ihre entsprechenden Stelle finden. So werden auch die gesehenen Eindrücke in der Seele sich wieder zu einer Raumwelt ausbreiten, nicht indem sie eine verborgene Räumlichkeit in das Bewusstsein eingeschwärzt, sondern weil sie vermocht haben, zwischen den intensiven Erregungen der Seele, die sie erzeugten, Relationen zu unterhalten, aus denen in der reconstruierenden Thätigkeit der Anschauung das Bild der veranlassenden Objekte wiedererstehen muss. (Lotze, 1852, S. 328f.)

Dies soll nun die Theorie der Lokalzeichen begreifbar machen, welche durch das folgende, an die Ausführungen Lotzes angelehnte Beispiel erläutert werden soll. Ein beliebiges Licht  $l$  führe idealiter an sämtlichen Stellen der Retina, die durch dieses erregt werden, zu einer von den physikalischen Eigenschaften dieses Lichts bestimmten Nervenerrregung  $\lambda_l$ , welche die Wahrnehmung eines bestimmten Lichtpunktes mit einer bestimmten Helligkeit, Farbe usw. zur Folge habe.<sup>61</sup> Man stelle sich nun vor, dass dieses Licht  $l$  auf zwei räumlich getrennte Stellen  $a$  und  $b$  der Retina projiziert werde. Anhand der bisherigen Ausführungen sollte deutlich geworden sein, dass für Lotze weder die bloße Tatsache, dass  $a$  und  $b$  sich auf der Retina in einer bestimmten geometrischen Relation zueinander befinden, noch die Annahme, dass durch  $l$  an den Stellen  $a$  und  $b$  räumlich verschiedene lokalisierte Nervenfasern erregt werden, eine Erklärung dafür sein könne, dass in dieser Situation zwei unterschiedlich lokalisierte Lichtpunkte wahrgenommen werden, da das Bewusstsein, welches sich für unsere Wahrnehmungen verantwortlich zeichne, selbst keine räumlichen Bestimmtheiten aufweise. Die räumlichen Relationen auf der Retina müssten demnach auf eine Weise in nicht räumliche Intensitätsverhältnisse gleichsam ‘übersetzt’ werden, aus denen das Bewusstsein die räumlichen Verhältnisse der erregten Netzhautstellen rekonstruieren könne. Dies geschehe mit Hilfe der sogenannten Lokalzeichen. Das auf  $a$  und  $b$  projizierte Licht  $l$  bringe laut Lotze jeweils einen *weiteren* Nervenprozess – oder wie er sich später (Lotze, 1884) auch ausdrückt, einen gewissen „Nebeneindruck“ oder eine gewisse „eigenthümliche Färbung“ –  $\alpha$  bzw.  $\beta$  mit sich, nämlich das jeweilige Lokalzeichen, welches die räumliche Bestimmung des wahrgenommenen Lichts festlege, selbst aber nicht räumlicher Natur sei.

Das Lokalzeichen sei dabei vollkommen unabhängig vom Licht  $l$ , sondern allein abhängig von der erregten Netzhautstelle. Eine Nervenerrregung  $\lambda_l\alpha$  entstehe demnach, wenn  $l$  auf Stelle  $a$  der Netzhaut projiziert wird, eine Nervenerrregung  $\lambda_l\beta$ , wenn  $l$  auf  $b$  projiziert wird usw. Aus diesen ‘zusammengesetzten Nervenerrregungen’ könne das Bewusstsein schließlich die Lokalisation der diese Erregungen verursachenden Lichtrei-

---

<sup>61</sup>Aus Ergebnissen der Farbwahnehmungsforschung weiß man, dass dem nicht so ist, sondern die wahrgenommene Farbe unter anderem davon abhängig ist, auf welche Stelle der Retina das Licht projiziert wird. Für die Theorie der Lokalzeichen ist dieser Punkt allerdings unerheblich, da sie ja erklären soll, auf welche Weise wir Objekte als unterschiedlich lokalisiert wahrnehmen, unabhängig davon ob diese nun dieselbe wahrgenommene Helligkeit und Farbe haben oder nicht.

ze rekonstruieren. Um dies zu ermöglichen, müssten die den jeweiligen Netzhautstellen zugehörigen Lokalzeichen  $\alpha, \beta, \gamma, \dots$  jedoch nicht als *bloß verschieden* angenommen werden, sondern sie müssten, etwa Zahlenpaaren vergleichbar, eine gewisse Ordnungsrelation aufweisen. Ansonsten könne man zwar mit Hilfe der Lokalzeichen erklären, wie es möglich sei, dass die durch Projektion auf  $a$  und  $b$  ausgelösten Nervenregungen nicht die Wahrnehmung eines, sondern zweier *verschiedener* Lichtpunkte zur Folge habe, also erklären, wie es möglich sei, dass wir Objekte als bloß verschieden lokalisiert wahrnehmen. Es bleibe aber unverständlich, wie verschiedene räumliche Relationen zwischen den Objekten – etwa „Objekt  $B$  befindet sich etwa halb so weit links von Objekt  $A$  wie Objekt  $C$ “ – wahrgenommen werden könnten.<sup>62</sup>

Ein solches *geordnetes System* von Lokalzeichen entstehe laut Lotze durch einen physiologischen, reflexartigen Mechanismus, nämlich eine bestimmte Systematik der Augenbewegungen. Fällt ein Lichtpunkt auf einen peripheren Punkt der Netzhaut, so führe dies laut Lotze zu einer unwillkürlichen, reflexartigen Bewegung des Auges (oder einer Tendenz zu dieser Bewegung) in Richtung dieses Lichtpunktes, so dass dieser Punkt fixiert werde. Mit jedem (gereizten) Punkt auf der Netzhaut sei daher eine, nur diesem Punkt allein zukommende, spezifische Augenbewegung verbunden, die wiederum mit einem spezifischen Erregungsmuster der motorischen Fasern einhergehe.<sup>63</sup> Dieses spezifische motorische Erregungsmuster habe nun eine spezifische Empfindung zur Folge, das sogenannte *Stellungsgefühl*, das für jeden Punkt der Retina spezifisch sei und das bei Erregung der Netzhaut auftretende sensorische Erregungsmuster ‘begleite’ oder ‘einfärbe’ und so als Lokalzeichen wirke. Die Verbindung zwischen dem spezifischen motorischen Erregungsmuster und dem sensorischem Erregungsmuster sei, darauf weist Lotze ausdrücklich hin, *keine* durch Erfahrung oder Assoziation zustande gekommene, sondern eine *a priori* vorhandene, welche uns nicht bewusst werde und unwillkürlich ablaufe. Es sei zudem auch ohne tatsächliche Durchführung der Fixationsbewegungen (also bei ruhendem Auge) möglich, mit Hilfe der Lokalzeichen Objekte als unterschiedlich lokalisiert wahrnehmen zu können, da bei Reizung einer bestimmten Stelle der Retina selbst die bloße Bewegungstendenz das entsprechende Stellungsgefühl auslöse und so als Lokalzeichen fungiere.

---

<sup>62</sup> „So lange wir nämlich die Localzeichen nur als *verschieden* ansehen ... sie dagegen nicht als *Glieder einer geordneten Reihe* betrachten, so dass wir sie durch vergleichbare Zahlen bezeichnen könnten: so lange wird zwar ein Auseinandertreten der von ihnen begleiteten Empfindungen statthaben, aber ohne dass daraus irgend eine deutliche Raumschauung erwüchse; denn man würde durch solche ihrer gegenseitigen Distanz nach unbestimmten Localzeichen zwar zur Trennung genöthigt, zur Anweisung bestimmter relativer Lagen der empfundenen Elemente dagegen nicht befähigt sein“ (Lotze, 1852, S. 332).

<sup>63</sup> „Bezeichnen wir daher mit  $s$  die Summe aller dieser Bewegungen, so ist sie für jeden Punkt der Netzhaut eine unvertauschbare und spezifische Combination, und eben deshalb glauben wir an ihr das Localzeichen zu besitzen, welches die Erregung jedes dieser Punkte von der jedes anderen unterscheidet“ (Lotze, 1852, S. 357).

Lotze schlägt mit seiner Theorie der Lokalzeichen etwas vor, was in heutigen kognitionspsychologischen Ansätzen, in denen das Hirn wesentlich als ein ‘Informationsverarbeitungssystem’ aufgefasst wird, unter dem Stichwort der ‘Kodierung’ behandelt wird. Das motorische Erregungsmuster bei Fixation auf einen bestimmten Punkt (oder der Tendenz zu dieser Bewegung) werde, so könnte man Lotzes Konzept in den heutigen Duktus übersetzen durch etwas ‘kodiert’, das bestimmte Ordnungsrelationen aufweise, aus denen dann die räumlichen Verhältnisse durch das Bewusstsein ‘rekonstruiert’ werden könnten.

Wie kommt es nun zu der von Lotze behaupteten aktiven ‘Rekonstruktion’ der räumlichen Relationen aus den (trotz Lokalzeichen) immer noch bloßen Nervenerregungen durch das Bewusstsein? Die Theorie der Lokalzeichen selbst könne darauf keine Antwort geben, wie Lotze selbst zugibt:

Gesetzt nun auch, der Seele werde mit jeder qualitativen Empfindung zugleich jenes Lokalzeichen zugeführt, das nun bestehen möge, worin es wolle: ist es nicht dennoch eine Erschleichung, wenn wir behaupten, dass sie hierdurch allein befähigt und genöthigt werde, ihre Empfindungen nicht nur *überhaupt auseinanderzuhalten*, sondern sie auch *räumlich* voneinander zu trennen? In der That müssen wir diese Frage zum Teil bejahen. (Lotze, 1852, S. 334)

Die Lokalzeichen könnten zwar erklären, auf welche Weise das Bewusstsein in der Lage sei, Reizungen unterschiedlicher Punkte der Retina anhand der unterschiedlichen Lokalzeichen auseinanderzuhalten, d.h. als geordnet verschiedene, aber nicht *räumlich* verschiedene, anzusehen. Dass das Bewusstsein *überhaupt* in der Lage ist, anhand der mit den Lokalzeichen versehenen Nervenerregungen räumliche Relationen zu rekonstruieren und diese Rekonstruktion schließlich dazu führe, dass wir räumliche Verhältnisse wahrnehmen, müsse man einfach an- oder besser hinnehmen, nicht zuletzt aufgrund der Tatsache, *dass* wir eben räumlich wahrnehmen.<sup>64</sup> Wenn man aber voraussetze, dass diese Nervenerregungen durch das Bewusstsein ständig in räumliche Vorstellungen gleichsam ‘umgewandelt’ würden, so ließe sich mit Hilfe der Lokalzeichen eine Erklärung anbieten, auf welche Weise oder anhand welcher Prinzipien die mit den Lokalzeichen versehenen Nervenerregungen die Wahrnehmung unterschiedlich lokalisierter Lichtpunkte oder Objekte zur Folge hätten.

---

<sup>64</sup>„Es war indessen auch keineswegs unsere Absicht, aus jenen Lokalzeichen die Fähigkeit der Seele, *Raum überhaupt* anzuschauen, oder ihre Nöthigung abzuleiten, das Empfundene in diese Anschauung aufzunehmen. Wir setzen vielmehr voraus, dass es in der Natur der Seele Motive gibt, um deren willen sie einer räumlichen Anschauungsform nicht nur *fähig ist*, sondern auch zu ihrer Anwendung auf den Inhalt der Empfindungen *gedrängt* wird; und weder jene Fähigkeit noch diese Nöthigung suchten wir aus den vorausgesetzten physiologischen Verhältnissen jener Lokalzeichen zu erklären“ (Lotze, 1852, S. 334f.).

Allerdings könne, so Lotze, die bisher dargestellte Theorie der Lokalzeichen höchstens eine Erklärung dafür bieten, wie unterschiedliche Lichtpunkte auf einer Fläche unterschiedlich lokalisiert werden könnten. Die bisherigen Ausführungen Lotzes bezogen sich ja darauf, dass verschiedene Stellen der Retina mit verschiedenen Lokalzeichen assoziiert sind, d.h. die Lokalzeichen können lediglich als ‘Kodes’ für die ‘Koordinaten der Retina’ genutzt werden. Wie kommt es nun zur Tiefenwahrnehmung? Laut Lotze müsse diese durch Erfahrungen erlernt werden:

Die Vorstellung nun, daß es so etwas wie eine dritte Dimension des Raumes überhaupt gebe, kann nicht von selbst, sondern nur durch die Erfahrung entstehen, welche wir machen, wenn wir uns durch die gesehenen Objecte hindurch bewegen. (Lotze, 1881, S. 34)

An diesem Punkt werden die Erläuterungen Lotzes äußerst kryptisch, man könnte seine Ausführungen jedoch so lesen, dass wahrgenommene Räumlichkeit dadurch entstehe, dass in einem ersten Schritt (wie bei Müller) mittels der Lokalzeichen lediglich ‘Flächen’ wahrgenommen würden. Wenn wir uns nun jedoch durch diese ‘gesehenen Flächen’ „hindurch bewegen“ würden (falls man dieser Wendung einen Sinn zuschreiben kann) und diese somit ständig ihre Form veränderten, wäre der Wahrnehmende in der Lage, aufgrund verschiedener geometrischer Relationen, die bei diesen Formveränderungen erhalten blieben, die sich verändernden Flächen als sich in geringerer oder größerer Entfernung zueinander befindende *Körper* zu ‘interpretieren’. Der Sehende, so Lotze, gelange dazu,

durch seine Bewegungen im Raume, durch die Verschiebung und die veränderliche Parallaxe der Bilder, die während seines Fortschreitens entsteht, durch ihre gegenseitigen momentane Deckungen, durch das Ausfallen einiger aus dem Sehfeld und den Eintritt neuer, sich eine Vorstellung von der Tiefe des Raums und der Lagerung der Objekte zu verschaffen. (Lotze, 1852, S. 419)

Die (nicht erklärbaren) Tätigkeiten des Bewusstseins scheinen demnach von Lotze allein zur Erklärung des Übergangs von den mit den Lokalzeichen versehenen Nerveneregerungen zu den von uns gesehenen ‘Bildern’ der Objekte herangezogen zu werden, nicht jedoch zur Erklärung der Wahrnehmung unterschiedlich im Raum lokalisierter Objekte, denn diese sei nur durch über Bewegung vermittelte *Erfahrungen des Wahrnehmenden* verständlich. In diesem Zusammenhang greift Lotze dann auch auf sämtliche der bisher in dieser historischen Darstellung erwähnten Hilfsmittel der Tiefenwahrnehmung zurück (so nennt er u.a. die Konvergenz, die Akkommodation, die Verdeckung, die ‘tatsächliche’ und ‘erscheinende Größe’ eines Objektes sowie *image blur* und *aerial*

*perspective*), behandelt sie jedoch nicht ausführlich, sondern zählt sie mehr oder weniger auf, ohne sie zu erläutern oder in seine allgemeinen theoretischen Vorstellungen zu integrieren. So heißt es etwa: „Ist uns die Natur und damit auch die wahre Größe eines Gegenstandes von früher her bekannt, so leiten wir aus seiner scheinbaren Größe die Weite seiner Entfernung ab“ (Lotze, 1852, S. 439). Oder, bezüglich der *aerial perspective*: „Lichtstarke und deutliche Bilder gehören im Allgemeinen grösserer Nähe an; was uns trüber erscheint oder mit ineinander verschwimmenden Grenzen seiner kleinsten Theilchen, das versetzen wir in weitere Ferne“ (Lotze, 1852, S. 440). Bezüglich der Konvergenz scheint Lotze eher der Auffassung des Descartes zuzustimmen, dass anhand dieser Entfernung gleichsam ‘berechnet’ werden könne:

Für nahe Gegenstände erregt die grössere oder geringere Convergenz unserer Augenaxen, durch welche wir den Blick auf sie fixieren, sehr merkbare Gefühle, obgleich keine solchen, dass aus ihnen genauere Grössenbestimmungen über die Differenz zweier Entfernungen nach der Tiefe des Raumes zu sich ableiten liessen. (Lotze, 1852, S. 441)

Durch dieses mehr oder weniger bloße Aufzählen und das Fehlen weitergehender Erläuterungen – etwa was es heißen soll, dass Entfernung ‘abgeleitet’ werde oder wir etwas in die Ferne ‘versetzen’ – fällt es schwer, ein klares Bild dieser Hilfsmittel in Lotzes Wahrnehmungstheorie zu gewinnen. Es scheint aber auch hier der Fall zu sein, dass die Beschreibungen der Hilfsmittel eine Wahrnehmung bereits voraussetzen; die ‘scheinbare Größe’ eines Objekts muss wahrgenommen werden, um als Hilfsmittel der Tiefenwahrnehmung dienen zu können, ebenso muss im Fall der *aerial perspective* zunächst etwas wahrgenommen werden, was „trüber erscheint“.

### Zusammenfassung

Die Lokalzeichentheorie übte eine nicht zu unterschätzende Wirkung auf die Entwicklung wahrnehmungstheoretischer Ansätze im späten 19. und frühen 20. Jahrhundert aus: Wurden Fragen der Raum- bzw. Tiefenwahrnehmung behandelt, bezog man sich in der Regel in irgendeiner Weise auf Lotzes Theorie. Helmholtz griff die Idee der Lokalzeichen in leicht veränderter Form in seiner als nächstes zu behandelnden Zeichentheorie wieder auf. Für ihn stand außer Frage, dass es nur gelinge, eine „Erklärung der räumlichen Gesichtswahrnehmungen ... zu geben, wenn man mit Lotze annimmt, dass den Empfindungen der räumlich verschieden gelagerten Nervenfasern gewisse Verschiedenheiten, *Localzeichen* anhaften“ (Helmholtz, 1896/2002d, S. 394). Auch Wilhelm Wundt hatte an der Tragweite der Lotzeschen Theorie keinen Zweifel:

Das größte Verdienst um die Analyse der Sinneswahrnehmungen hat sich von psychologischer Seite Lotze erworben. ... Durch Lotze war insofern

ein erster Schritt der Vermittlung [zwischen Physiologie und ‘Seelenkunde’] geschehen, als er dem aus psychologischen Gründen Geforderten eine anatomisch-physiologische Basis zu verschaffen suchte, so kam er auf sein System der Lokalzeichen. (Wundt, 1862, S. 12)

Aber nicht nur im deutschsprachigen Raum, auch über dessen Grenzen hinaus gelangte die Lokalzeichentheorie zu einiger Bedeutung und konnte die weitere Theoriebildung im Bereich der Raumwahrnehmung und insbesondere auch der taktilen Wahrnehmung wesentlich beeinflussen; William James (1890/2007) etwa widmete ein eigenes Unterkapitel seiner „*Principles of Psychology*“ der Lokalzeichentheorie Lotzes.

Zusammenfassend lässt sich feststellen, dass Lotze im Rahmen seiner Lokalzeichentheorie einen großen Aufwand betreibt, um Tiefenwahrnehmung unter Bezugnahme auf die physiologischen Kenntnisse seiner Zeit verständlich zu machen. Lotze sah sich mit dem Problem konfrontiert, dass die räumlichen Verhältnisse bei Erregung unterschiedlich lokalisierter Netzhautstellen *als solche* nicht durch die physiologische Aktivität der zugehörigen Nervenfasern einfach zum Bewusstsein ‘übertragen’ werden könnten. Die räumlichen Verhältnisse der erregten Netzhautstellen müssten daher laut Lotze gleichsam ‘kodiert’ werden, in sogenannte Lokalzeichen, bestimmten und für jeden Punkt der Retina spezifischen Erregungen, aus denen sie dann durch das Bewusstsein wieder ‘rekonstruiert’ werden könnten. Lotze ließe sich aus diesem Grund mit einem gewissen Recht auch als ein Vordenker der heutigen komputationalen Wahrnehmungstheorien bezeichnen. Das Konzept der Lokalzeichen wird von Lotze ausführlich und detailliert erläutert, wie das Bewusstsein aus den mit den Lokalzeichen behafteten Nervenregungen die räumlichen Verhältnisse ‘rekonstruiere’, wie also die Lokalzeichen gleichsam als Zeichen für diese ‘interpretiert’ werden, darauf gibt Lotze keine Antwort. Er verweist darauf, dass man gewisse Tatsachen, wie eben jene, *dass* das Bewusstsein bestimmte Nervenaktivitäten räumlich ‘interpretiere’, einfach anerkennen müsse und nicht weiter erklären könne, was letztlich dazu führt, dass in seiner Theorie eine unüberbrückbare Kluft zwischen den mit den Lokalzeichen behafteten Nervenregungen und dem Sehen externer Objekte bestehen bleibt.<sup>65</sup> Gleichzeitig spricht sich Lotze an mehreren Stellen dafür aus, dass Tiefenwahrnehmung über Erfahrungen des Wahrnehmenden erlernt

---

<sup>65</sup>Diese Kritik an der Lotzeschen Theorie der Lokalzeichen wurde bereits von Johannes von Kries vorgebracht, der in den von ihm verfassten Zusätzen des dritten Bandes von Helmholtz’ „*Handbuch der Physiologischen Optik*“ schreibt:

Denn wenn, wie soeben erwähnt, die Theorie Lotzes gerade, indem sie die Raumschauung als etwas neben den Empfindungen und ihren Lokalzeichen von vornherein Gegebenes betrachtete, für eine Entwicklung der Lokalisation im empiristischen Sinne die ausreichenden Grundlagen bot, so werden wir die Entwicklung einer räumlich geordneten Wahrnehmung allein aus den in den Lokalzeichen gegebenen unräumlichen Verhältnissen durchaus unverständlich finden müssen. (Helmholtz, 1910/2003a, S. 522)

werden müsse und scheint diesbezüglich ähnliche Vorstellungen zu haben wie Müller, nämlich dass der Wahrnehmende zunächst nur ‘flächenhafte Bilder’ sehen würde. Auf welche Weise es über Bewegungen des Wahrnehmenden dann dazu komme, dass dieser sich „eine Vorstellung von der Tiefe des Raums und der Lagerung der Objekte“ (Lotze, 1852, S. 419) verschaffe, wird jedoch nicht deutlich. Lotze bedient sich schließlich auch der Annahme verschiedener Hilfsmittel der Tiefenwahrnehmung, um den Übergang zwischen den durch die Lokalzeichen vermittelten ‘flächenhaften Bildern’ der Objekte und den als unterschiedlich im Raum lokalisiert wahrgenommenen Objekten zu erklären, die entsprechenden, äußerst knappen Beschreibungen dieser Hilfsmittel lassen aber nicht erkennen, welche genauen Vorstellungen Lotze mit diesen verbindet.

## 2.6. Die Theorie der unbewussten Schlüsse: Hermann von Helmholtz

Dieser Abschnitt nimmt den größten Teil der historischen Darstellung des Konzepts der Tiefencues ein und aus diesem Grunde sollen zunächst einige rechtfertigende Bemerkungen vorangestellt werden. Bei Helmholtz’ wahrnehmungstheoretischem Schaffen handelt es sich erstens um eine der komplexesten und umfangreichsten Arbeiten in diesem Gebiet, da es den Versuch darstellt, eine an Kant angelehnte Erkenntnistheorie mit empirischen Ergebnissen aus Physiologie und Psychologie in Einklang zu bringen. Zweitens lassen sich viele der Helmholtzschen Überlegungen gleichsam als ‘Blaupause’ für das *cue*-Konzept der aktuellen Wahrnehmungspsychologie ansehen, so dass eine umfangreiche Darstellung sinnvoll erschien. Helmholtz’ Theorie visueller Wahrnehmung war bis in die achtziger Jahre des 20. Jahrhunderts der dominierende wahrnehmungspsychologische Ansatz, auch der moderne sogenannte *komputationale Ansatz* der Wahrnehmung lässt sich als generell an von Helmholtz vorgebrachten Ideen orientiert kennzeichnen. Drittens ist der Umfang dieses Abschnittes dem Fakt geschuldet, dass es Helmholtz seinen Lesern an vielen Stellen schlicht nicht einfach macht. So werden etwa zwischen verschiedenen Schriften (aber auch innerhalb ein- und derselben) Begriffe auf deutlich unterschiedliche Weise und in deutlich verschiedenen Kontexten verwendet oder bestimmte, für ein Verständnis der theoretischen Überlegungen notwendige Sachverhalte von Helmholtz schlicht ausgeklammert, so dass es notwendig erschien, an gewissen Stellen dieses Abschnitts noch stärker als bisher auf Zitate zurückzugreifen, um dem Leser die Möglichkeit zu geben, die Stimmigkeit der vorgeschlagenen Lesart(en) zu überprüfen. Darüber hinaus haben die wahrnehmungstheoretischen Vorstellungen Helmholtzens im Laufe seines Lebens einige markante Änderungen erfahren, so dass von einer zusammenfassenden Darstellung *der* Wahrnehmungstheorie von Helmholtz

zugunsten einer eher chronologisch orientierten Vorgehensweise abgesehen wurde.<sup>66</sup> Zunächst sollen einige biographische und allgemeine Bemerkungen vorangestellt werden, die dem Leser den Zugang zu Helmholtz' Arbeiten und dem theoretischen Kontext, in dem diese anzusiedeln sind, erleichtern sollen.

### 2.6.1. Biographische und allgemeine Vorbemerkungen

Hermann Ludwig Ferdinand von Helmholtz (1821-1894) war ohne Zweifel eine die wissenschaftliche Öffentlichkeit seiner Zeit bestimmende Persönlichkeit; nicht selten findet man die Charakterisierung Helmholtzens als „Universalgenie“ (Krüger, 1994) oder „*homo universalis*“ (Meyering, 1989, S. 179) des 19. Jahrhunderts. Er veröffentlichte in akademischen Kreisen hoch angesehene und wirkungsmächtige Arbeiten in so unterschiedlichen Fachgebieten wie der theoretischen Physik (in den Bereichen Hydro- und Elektrodynamik), der Akustik, der Optik, der Meteorologie, der Erkenntnistheorie und der Physiologie. Gleichzeitig war es Zeit seines Lebens Helmholtz' Anliegen, über populärwissenschaftliche Vorträge und Schriften die stetig zunehmenden neuen Erkenntnisse aus den Naturwissenschaften einem breiten Publikum zugänglich zu machen und so das naturwissenschaftliche Denken im allgemeinen Denken seiner Zeit zu verankern (Mausfeld, 1994). Helmholtz entwickelte zudem viele neuartige Messgeräte und experimentelle Vorrichtungen, etwa den Augenspiegel zur Beobachtung des Augenhintergrundes, und gilt nicht zuletzt, neben Wundt und Fechner, als einer der Wegbereiter oder gar 'Begründer' der experimentellen Psychologie (Boring, 1942).

Eigentlich ausgebildeter und praktizierender Mediziner (Helmholtz promovierte bei Johannes Müller und war anschließend fünf Jahre als Militärarzt in Potsdam tätig), galt sein hauptsächliches Interesse jedoch bereits in jungen Jahren der Physik; er sei „wesentlich durch äußere zwingende Umstände“ zur Medizin und zur Physiologie geführt worden (Koenigsberger, 1902-1903). Zu Beginn seiner akademischen Karriere war Helmholtz Lehrstuhlinhaber für Physiologie und Anatomie in Königsberg und in Bonn, daran anschließend Professor für Physiologie in Heidelberg (von 1858 bis 1863 war Wilhelm Wundt dort sein Assistent), bis er schließlich 1871 zum Ordinarius für Physik an die Universität Berlin berufen wurde, wo unter anderem Max Planck und Heinrich Hertz bei ihm studierten und Helmholtz bis zu seinem Tode wirkte. Im Jahre 1873 erhielt er den Orden *Pour le Mérite für Wissenschaften und Künste*, 1882 wurde er in den

---

<sup>66</sup>Erdmann (1921, S. 5) schreibt in diesem Zusammenhang:

Ganz leicht ist das Verständnis von Helmholtz' Wahrnehmungstheorie nicht zu erlangen. Sie hat sich ... von philosophischen Grundgedanken aus entwickelt, [die] durch eine so verwickelte wie originale Synthese ineinander verschlungen [sind] und in ihrem einheitlichen Zusammenhang aus den beiden Hauptwerken sowie zahlreichen verschieden orientierten Abhandlungen, Vorträgen und Reden herausgelesen werden [müssen].

erblichen Adelsstand erhoben und 1888 erster Präsident der von ihm mitgegründeten „Physikalisch-Technischen-Reichsanstalt“ in Charlottenburg.<sup>67</sup>

Im Gegensatz zu der bei vielen Naturwissenschaftlern seiner Zeit anzutreffenden „Verwerfung aller Philosophie“ (Erdmann, 1921) aufgrund deren vermeintlich spekulativer Natur war Helmholtz der Überzeugung, dass sich die Naturwissenschaften und die Philosophie, insbesondere die Erkenntnistheorie, gegenseitig bereichern sollten (siehe etwa Helmholtz, 1896/1998, S. 149). In einer Zeit der „Tyrannis der Naturwissenschaften“ (Riehl, 1904, S. 47), in der sich die naturwissenschaftliche Forschung „von der Philosophie als von etwas gänzlich Abstrusem und Unwissenschaftlichem abgewandt hatte“ (Schlick, 1922/2012, S. 476), lässt sich Helmholtz’ nachdrückliche Bekräftigung einer Berechtigung und Bedeutung der Philosophie für die Naturwissenschaften als großes Verdienst ansehen; für Schlick (1922/2012, S.475) war Helmholtz gar „schlecht-hin ein Philosoph“.<sup>68</sup> Helmholtz hielt es für eine der wesentlichen Aufgaben seiner Zeit, die Metaphysik, insbesondere apriorische metaphysische Spekulationen, wie er sie in der Naturphilosophie Schellings und Hegels zu finden glaubte, mit Hilfe naturwissenschaftlicher Erkenntnisse für immer aus der Philosophie zu verbannen (Erdmann, 1921; Heidelberger, 1995). Helmholtz’ philosophisches Interesse galt im Wesentlichen der Epistemologie (insbesondere der Kantischen Erkenntnistheorie fühlte er sich verbunden); die Neigung zu erkenntnistheoretischen Fragen, so schreibt er später in seinen Erinnerungen, schien ihm selbst von Jugend auf durch Gespräche seines Fichte verehrenden Vaters mit Anhängern Kants und Hegels eingepägt (Helmholtz, 1896/2002c).

Die weiteren Ausführungen in diesem Abschnitt werden verdeutlichen, dass es durchgängig Helmholtz’ Anliegen war, in seinen wahrnehmungspsychologischen Schriften Er-

---

<sup>67</sup>Die immer noch maßgebliche Biographie über Helmholtz ist das dreibändige Werk von Koenigsberger (1902-1903). Für einen kompakten Überblick des Lebens und Wirkens von Helmholtz siehe die Gedächtnisrede zum Tode Helmholtz’ von du Bois-Reymond (1912).

<sup>68</sup>Über das Verhältnis zwischen den Naturwissenschaften und der Philosophie zu seiner Zeit stellt Helmholtz (1896/2002d) fest:

Weiss man nicht allgemein, dass Naturforscher und Philosophen gegenwärtig nicht gerade gute Freunde sind, wenigstens in ihren wissenschaftlichen Arbeiten? Weiss man nicht, dass zwischen beiden lange Zeit hindurch ein erbitterter Streit geführt worden ist, der neuerdings zwar aufgehört zu haben scheint, aber jedenfalls nicht deshalb, weil die eine Partei die andere überzeugt hätte, sondern weil jede daran verzweifelte, die andere zu überzeugen? (S. 87)

In verschiedenen privaten Korrespondenzen zeigt sich allerdings auch ein gewisses Ressentiment Helmholtz’, insbesondere gegenüber der akademischen Philosophie seiner Zeit. In einem Brief an den Mathematiker Lipschitz stellt er etwa fest, dass es sich bei akademischen Philosophen meistens um „impotente Bücherwürmer“ handle, „die nie ein neues Wissen erzeugt haben, also auch gar keine Ahnung davon haben, wie es dabei zugeht“ (zitiert nach Schiemann, 2014, S. 199). In einer weiteren, in Koenigsberger (1902-1903, Bd. 3, S. 248) wiedergegebenen, privaten Korrespondenz zwischen Helmholtz und dem Physiologen Adolf Fick heißt es, der Philosophie sei „nur wieder aufzuhelfen ... wenn sie sich mit Ernst und Eifer der Untersuchung der Erkenntnisprozesse und der wissenschaftlichen Methoden zuwendet“, im für die Wissenschaft besten Falle sei es nötig, „einen der Philosophie zugewendeten Naturforscher zum Philosophen zu berufen“.

kenntnisse aus der Physiologie und Psychologie mit prononciert erkenntnistheoretischen Erwägungen in *einer* Theorie zu verbinden. Helmholtz selbst sah seine eigene erkenntnistheoretische Arbeit als an Kants in der „Kritik der reinen Vernunft“ durchgeführten Kritizismus anschließend (Helmholtz, 1896/2002a, 1910/2003a, 1896/1998), in Helmholtz' Schriften lässt sich jedoch eine deutliche theoretische Ausrichtung dahingehend feststellen, epistemologische Fragen stets in Verbindung mit empirischen Erkenntnissen bezüglich des Aufbaus und der Funktion der menschlichen Sinnesorgane zu behandeln und so gewissermassen auf *empirischer* Basis zu erörtern. Diese Haltung scheint für den naturwissenschaftlichen Impetus der Psychologie bzw. Physiologie des mittleren und späten 19. Jahrhunderts kennzeichnend; die als rein theoretisch angesehene Behandlung des Erkenntnisproblems seitens Kant (oder allgemein der Philosophie) sollte durch Ergebnisse der Sinnesphysiologie auf einen 'sicheren', d.h. im Wesentlichen *empirischen* Boden gestellt werden.

Die neben Fragen der Farbwahrnehmung am stärksten und heftigsten diskutierte Frage im Bereich der visuellen Wahrnehmungspsychologie des 19. Jahrhunderts war die durch Überlegungen Kants angeregte, ob es sich bei dem wahrgenommenen Raum um eine physikalisch ausgedehnte, empirische Entität handle, oder, wie Kant meinte, um eine „notwendige Vorstellung *a priori*, die all unseren äußeren Anschauungen zum Grunde liegt“ (Kant, 1781/1998, S. 98), d.h. eine notwendige Bedingung *a priori* der Möglichkeit, überhaupt äußere Dinge sinnlich zu erfahren und vorzustellen. Der Raum, so Kant, sei kein empirischer Begriff, keine Eigenschaft der äußeren Dinge oder ein Verhältnis dieser untereinander, das von äußeren Erfahrungen abgezogen werden könne, sondern die Form, in der uns alle über die Sinne wahrgenommenen Dinge notwendigerweise erscheinen; die „subjektive Bedingung der Sinnlichkeit, unter der allein uns äußere Anschauung möglich ist“ (Kant, 1781/1998, S. 101).<sup>69</sup> Diese Frage wurde dann in der psychologischen Literatur der Zeit zunehmend zu der (wesentlich anderen) Frage umgewandelt, ob die Fähigkeit zur Raumwahrnehmung auf einem 'angeborenen', so genannten *primitiven Raumgefühl* beruhe, oder allein durch Erfahrung erworben werde. Die damit zusammenhängende Debatte, die teils äußerst heftig geführt wurde und nicht selten in persönlichen Animositäten ausartete (am bekanntesten ist sicher die äußerst persönlich anmutende Fehde zwischen Helmholtz und Ewald Hering), ging als sogenannte *Empirismus-Nativismus-Kontroverse* in die Psychologiegeschichte ein.<sup>70</sup>

An die Kantischen Überlegungen angelehnt, war es Helmholtz' dezidierte Ansicht, dass Wahrnehmung aufzufassen sei als eine vermittelte Wirkung, die ein Objekt auf die

---

<sup>69</sup>Aufgrund der Ausrichtung dieser Arbeit kann die Position Kants zum Begriff des Raumes – von der zugehörigen Argumentation einmal ganz abgesehen – hier nur schemenhaft dargestellt werden, für weitergehende Erläuterungen siehe etwa Vaihinger (1922).

<sup>70</sup>Für eine kompakte Darstellung der Debatte zwischen Helmholtz und Hering siehe Heidelberger (1993), Kalkofen (1992) sowie Turner (1993). Eine ausführliche und äußerst lesenswerte Darstellung bietet die Monographie von Turner (1994).

Sinne habe, wobei diese Wirkung ebenso vom Wirkenden, wie von der Natur dessen, auf das eingewirkt werde, abhängig sei. Laut Helmholtz könne eine Aufklärung des Wahrnehmungsproblems und dessen erkenntnistheoretischer Implikationen jedoch nur über eine psychophysiologische Untersuchung nach naturwissenschaftlichen Methoden angegangen werden; das ursprünglich erkenntnistheoretische Wahrnehmungsproblem wird so zu einem *psychophysiologischen*.<sup>71</sup> Psychologisch sei eine solche Untersuchung, da die Wahrnehmung äusserer Gegenstände ein Akt des Vorstellungsvermögens sei, der mit Bewusstsein einhergehe und somit eine psychische Tätigkeit darstelle. Bewusstsein habe aber eine physiologische Basis, das Gehirn sei als das „körperliche Organ unseres Bewusstseins“ (Helmholtz, 1896/2002a, S. 90) anzusehen. Die epistemologische Wichtigkeit einer solchen psychophysiologischen Untersuchung könne man gar nicht überschätzen, denn, so Helmholtz, „sinnliche Wahrnehmung liefert ja am Ende unmittelbar oder mittelbar den Stoff zu allem menschlichen Wissen“ (Helmholtz, 1896/2002a, S. 267). Ein wesentlicher und revolutionärer Punkt in den Kantischen Analysen ist die so genannte *transzendente Methode*, über die, so das Hauptanliegen der „Kritik der reinen Vernunft“, gezeigt werden soll, wie synthetische Urteile *a priori* möglich seien, wobei Kant transzendental bestimmt als „alle Erkenntnis, die sich nicht sowohl mit Gegenständen, sondern mit unserer Erkenntnisart von Gegenständen, sofern diese *a priori* möglich sein soll, beschäftigt“ (Kant, 1781/1998, S. 25). Helmholtz schlägt diesen aus Platzgründen hier nicht weiter auszuführenden methodischen Weg oder einen vergleichbaren nicht ein, sondern wählt einen *empirischen*, auf physiologische Erkenntnisse über die Sinnesorgane des Menschen beruhenden Zugang, den Kant selbst für eine Behandlung des Erkenntnisproblems für vollkommen unangemessen hielt.<sup>72</sup>

Ein weiteres auffälliges Merkmal der Arbeiten Helmholtzens ist die Tendenz, die methodische Zugangsweise der Physik – der zu dieser Zeit sicher erfolgreichsten Naturwissenschaft – auf möglichst viele Bereiche, inklusive dem der ‘mentalen Geschehnisse’ auszudehnen, was sich bei Helmholtz oft in einer als naturalistisch zu bezeichnenden Einstellung gegenüber diesen äußert (Hatfield, 1990). Für Helmholtz besteht keinerlei Zweifel, dass mentale wie physikalische Prozesse dem Kausalgesetz unterworfen und somit durch Gesetzmäßigkeiten beschreibbar seien, die wiederum empirisch begründet werden könnten.<sup>73</sup> Gesetzmäßigkeit überhaupt müsse jedoch eine notwendige Voraus-

<sup>71</sup>Die Helmholtzsche Position weist dadurch eine gewisse Ähnlichkeit mit der sogenannten *naturalisierten Erkenntnistheorie* auf, wie sie etwa von Quine (2003) vertreten wird.

<sup>72</sup>Erdmann (1921, S. 18) spricht in diesem Zusammenhang auch von einer „anthropologischen Umdeutung des Kantischen Kritizismus“, Riehl (1904, S. 11) von einer „physiologischen Auffassung der kritischen Philosophie“. Zu weitergehenden Erörterungen über das Verhältnis zwischen der Kantischen Philosophie und Helmholtz’ eigenen Arbeiten siehe Erdmann (1921), Fullinwider (1990), Riehl (1904), Schiemann (2014) sowie Schlick (1922/2012).

<sup>73</sup>Für Kant stellt das Kausalgesetz – grob vereinfacht gesprochen – eine *notwendige Bedingung a priori* für Erkenntnis überhaupt dar und ist dahingehend der Kantischen Auffassung von Zeit und Raum nicht unähnlich. Sinnliches Erleben geschehe im Menschen *a priori* in den Kategorien von Ur-

setzung für Erkenntnis sein, sie sei eine „Bedingung der Begreifbarkeit“ (Helmholtz, 1896/2002b, S. 243) schlechthin. Wenn man als Forscher tätig werden möchte, so müsse man auf die „Gesetzmäßigkeit alles Geschehens“ *vertrauen*, denn „Vertrauen in die Gesetzmäßigkeit ist also zugleich Vertrauen auf die Begreifbarkeit der Naturerscheinungen“ (Helmholtz, 1896/2002b, S. 243), zu denen laut Helmholtz eben auch mentale Geschehnisse gehörten.<sup>74</sup> Da es Helmholtz (zumindest in seinen späten Jahren) wie Kant unmöglich schien, ein ‘Ding an sich’, also ein auf die Sinne Wirkendes und verschiedene Wahrnehmungen Hervorrufendes, *unabhängig* von diesen Wahrnehmungen positiv bestimmen zu können, lasse sich das Kausalgesetz im Kontext der Wahrnehmung folgendermaßen formulieren:

Die einzige Voraussetzung, welche wir festhalten, ist die des Causalgesetzes, dass nämlich die mit dem Charakter der Wahrnehmung in uns zu Stande kommenden Vorstellungen nach festen Gesetzen zu Stande kommen, so dass, wenn verschiedene Wahrnehmungen sich uns aufdrängen, wir berechtigt sind, daraus auf Verschiedenheit der realen Bedingungen zu schliessen, unter denen sie sich gebildet haben. Übrigens wissen wir über diese Bedingungen selbst, über das eigentlich Reale, was den Erscheinungen zu Grunde liegt, nichts; alle Meinungen, die wir sonst darüber hegen möchten, sind nur als mehr oder minder wahrscheinliche Hypothesen zu betrachten. Die vor-

---

sache und Wirkung, das Kausalgesetz sei „ein notwendiges Gesetz unserer Sinnlichkeit, mithin eine formale Bedingung aller Wahrnehmungen“ (Kant, 1781/1998, S. 296). Helmholtz drückt es folgendermaßen aus: „Also führt uns die Untersuchung der Sinneswahrnehmungen auch noch zu der schon von Kant gefundenen Erkenntnis: dass der Satz ‚Keine Wirkung ohne Ursache‘ ein vor aller Erfahrung gegebenes Gesetz unseres Denkens sei“ (Helmholtz, 1896/2002d, S. 116). Trotz aller auf den ersten Blick äußerlichen Übereinstimmungen der Kantischen und Helmholtzschen Auffassung des Kausalgesetzes gibt es bedeutsame Unterschiede zwischen diesen, im Wesentlichen die Modifikation des Kantischen *a priori* im starken Sinn in einen zunehmend empiristischen Gedankengang bei Helmholtz. Der an diesem Aspekt der Helmholtzschen Theorie interessierte Leser sei auf Fullinwider (1990), den äußerst instruktiven Aufsatz von Erdmann (1921, S. 6-17) sowie auf Riehl (1904) verwiesen, der insbesondere auf die vom Früh- zum Spätwerk deutlichen Veränderungen der Helmholtzschen Auffassung des Kausalgesetzes eingeht. Allgemein lässt sich feststellen, dass Helmholtz in seinen Schriften zuweilen eine größere Übereinstimmung seiner erkenntnistheoretischen Ansichten mit denen Kants annahm, als sich tatsächlich feststellen lässt (siehe dazu Schlick, 1998). Erdmann (1921, S. 41) spricht sich gar dafür aus, dass Helmholtz nie ein Kantianer gewesen sei (siehe dazu aber auch Fullinwider, 1990). Hervorzuheben ist noch, dass Helmholtz häufig Begriffe aus der Kritischen Philosophie Kants in deutlich anderer Weise als dieser selbst verwendet (vgl. Riehl, 1904; Schlick, 1922/2012), was die Lektüre der Helmholtzschen Schriften zusätzlich erschwert.

<sup>74</sup>An dieser Stelle wird die Helmholtzsche Abweichung von der Kantischen Konzeption des Kausalgesetzes beispielhaft deutlich. An Stelle des Kantischen *a priori* steht bei Helmholtz ein *Vertrauen* in die Gesetzmäßigkeit aus rein praktischen Gründen. Besonders in seinen späteren Schriften stellt das Kausalgesetz für Helmholtz lediglich eine *Hypothese* dar, auf die wir eben *vertrauen* müssten, um überhaupt wissenschaftlich arbeiten zu können. Es sei lediglich ein *regulatives* Prinzip unseres Denkens, dessen Anwendbarkeit durch den Erfolg gerechtfertigt werde. Das Kausalgesetz lasse sich weder durch Vernunft noch Erfahrung begründen, es sei allein *praktisch* begründet, wohingegen es bei Kant *konstitutiv* für Erfahrung überhaupt ist.

angestellte Voraussetzung dagegen ist das Grundgesetz unseres Denkens.  
(Helmholtz, 1896/2002b, S. 401f.)

Ein beachtenswerter Punkt in diesem Zusammenhang ist, dass Helmholtz es an verschiedenen Stellen seiner wahrnehmungspsychologischen Arbeiten explizit vermeidet, sich auf einen metaphysischen Realismus festlegen zu wollen, der eine Wirklichkeit unabhängig von einem erlebenden Bewusstseins schlicht voraussetzen würde.<sup>75</sup> Allerdings wird gleichzeitig, wie sich im weiteren Verlauf dieses Abschnittes zeigen wird, fast durchgängig zumindest die Sprechweise der realistischen Hypothese von Helmholtz verwendet. Dies führt zuweilen zu deutlichen Spannungen zwischen verschiedenen Texten, aber auch innerhalb dieser. Möglicherweise sah Helmholtz die realistische Position als die ‘wahrscheinlichste Hypothese’ über das „eigentlich Reale“ an, in der überwiegenden Mehrzahl der Helmholtzschen Texte lässt sich der mögliche hypothetische Charakter dieser Position jedoch nicht deutlich ausmachen.

Helmholtz’ im Kontext visueller Wahrnehmungspsychologie einflussreichstes und auch relativ früh seinen Ruhm begründendes Werk ist das zwischen 1856 und 1867 in drei Bänden erschienene „Handbuch der physiologischen Optik“, welches seinen Versuch darstellt, eine systematische und umfassende Ordnung in das bis zu diesem Zeitpunkt durch weitgehend unzusammenhängende Einzelbeobachtungen gekennzeichnete Forschungsgebiet der visuellen Wahrnehmung herzustellen.<sup>76</sup> Auch über den europäischen Kontinent hinaus erntete Helmholtz’ Werk große Bewunderung. So befand etwa William James (1890/2007, S. 278), dass es sich bei dem *Handbuch* um „one of the four or five greatest monuments of human genius in the scientific line“ handle.

Die Trisektion des *Handbuchs* in die Teile „Dioptrik des Auges“, „Lehre von den Gesichtsempfindungen“ und „Lehre von den Gesichtswahrnehmungen“ spiegelt dabei Helmholtz’ theoretische Überzeugung wider, dass Wahrnehmung als ein Prozess anzusehen sei, der sich wesentlich in drei verschiedene Stufen unterteilen lasse – eine Auffassung, die seit Descartes’ Theorie weithin Anerkennung gefunden hatte. Der erste Band des Handbuchs ist, dieser Stufentheorie folgend, dem physikalischen Reiz Licht,

<sup>75</sup>So äußert Helmholtz etwa: „Ich kann nicht umhin, selbst den extremsten subjektiven Idealismus [etwa Fichtes Lehre vom *absoluten Ich*] als eine mögliche und in sich konsequente Hypothese anzusehen“ (zitiert nach Erdmann, 1921, S. 8).

<sup>76</sup>Emil du Bois-Reymond (1912) sprach in seiner Helmholtz-Gedächtnisrede von 1912 folgendermaßen über diesen aus seiner Sicht mehr als gelungenen Versuch:

In einem umfangreichen, einheitlichen, doch auf das Feinste gegliederten Werke, seinem ‚Handbuche der physiologischen Optik‘ stellte er diesen Zweig der Physiologie systematisch und literaturgeschichtlich in größter Vollständigkeit dar, von den mathematischen Anfangsgründen der geometrischen Optik bis zu den letzten erkenntnistheoretischen und ästhetischen Gesichtspunkten. Man kann ohne Übertreibung sagen, daß keine wissenschaftliche Literatur irgendeiner Nation ein Buch besitzt, welches diesem an die Seite gestellt werden kann. (S. 536)

der Anatomie des Auges sowie den optischen Gesetzmäßigkeiten gewidmet, denen das Licht bei Eintritt in das Auge unterworfen ist. Die Empfindungen, welche durch Reizung der Sehnerven und der Übertragung dieser Nervenerrregung ausgelöst würden, sind Gegenstand des zweiten Bandes. Der dritte, von Helmholtz selbst oft als ‘genuin psychologisch’ bezeichnete Band befasst sich schließlich mit der ‘Interpretation’ dieser Empfindungen, welche die Wahrnehmung ausgedehnter, im Raum lokalisierter und als unabhängig vom Wahrnehmenden existierend angesehener Objekte zur Folge habe.

Die von Helmholtz vorgenommene Unterscheidung zwischen Empfindung und Wahrnehmung ähnelt der Reidschen Konzeption, ist aber bei Helmholtz – wie auch schon bei Müller und Lotze – deutlich physiologisch ausgerichtet:

*Empfindungen* nennen wir die Eindrücke auf unsere Sinne, insofern sie uns nur als Zustände unseres Körpers (speciell unserer Nervenapparate) zum Bewusstsein kommen; *Wahrnehmungen*, insofern wir aus ihnen uns die Vorstellung äusserer Objecte bilden. (Helmholtz, 1896/2002a, S. 297)

Diese Beschreibung weist im Kontext visueller Wahrnehmung ähnliche Schwierigkeiten auf, die im Abschnitt über Müller bereits diskutiert wurden, insbesondere, dass im Erleben diese Empfindungen als „Zustände unseres Körpers“ überhaupt nicht auffindbar sind, also uns gerade *nicht* „zum Bewusstsein kommen“. Schon unter Helmholtz’ Zeitgenossen kamen Bedenken auf, ob der Begriff der ‘visuellen Empfindung’ nicht ein von Wahrnehmungstheoretikern eingeführtes, abstraktes „Kunstprodukt der Wahrnehmungsanalyse“ (Erdmann, 1921, S. 44) sei.<sup>77</sup> Leider wird zudem der Empfindungsbegriff von Helmholtz häufig in inkonsistenter Weise verwendet, was dem Leser die Lektüre und das Verständnis an vielen Stellen erschwert. So schwankt Helmholtz (ähnlich wie Müller) auffallend oft zwischen der Charakterisierung der Empfindungen als einerseits Wirkungen *in unseren Organen*, d.h. als *Erregungen* der Sinnesorgane, und andererseits als uns *bewussten Zuständen unseres Körpers*, d.h. als *psychischen Data*, ein theoretisch durchaus bedeutsamer Unterschied (vgl. Schlick, 1998, S. 201).

Die Empfindungen würden dann in einem nächsten Schritt ‘genutzt’, „um uns aus ihnen Vorstellungen über die Existenz, die Form und die Lage äußerer Objekte zu bilden“ (Helmholtz, 1910/2003a, S. 1). Diese Vorstellungen werden von Helmholtz als Wahrnehmungen bzw. im speziellen Fall des Sehens als *Gesichtswahrnehmungen* bezeichnet. Der primäre Fokus der wahrnehmungstheoretischen Erörterungen Helmholtzens lässt sich in der Frage zusammenfassen, nach welchen Prinzipien aus Nervenerrregungen bzw.

<sup>77</sup>So bemerkt Riehl (1879, S. 195) ironisch, dass nach dieser Erklärung allein der Physiologe Empfindungen habe,

denn er nur *weiss*, dass sie zunächst als Zustände der Sinne aufzufassen seien, aber auch er weiss dies nur als Ergebnis seines Nachdenkens in abstracto, während sie unmittelbar, in concreto, für ihn, wie für Jedermann, *Theile* der directen *Wahrnehmung* sind.

den mit diesen einhergehenden Empfindungen entsprechende Wahrnehmungen hervor-  
gehen. Helmholtz' wesentliches Ergebnis in diesem Zusammenhang ist, wie er selbst in  
einer Tischrede zu seinem 70. Geburtstag 1891 rekapituliert, dass „die Sinnesempfin-  
dungen nur Zeichen für die Beschaffenheit der Aussenwelt sind, deren Deutung durch  
Erfahrung gelernt werden muss“ (Helmholtz, 1896/2002c, S. 17). Dieser in der Lite-  
ratur auch als *Helmholtz' Zeichentheorie* anzutreffende Ansatz soll zusammen mit der  
sogenannten *Theorie der unbewussten Schlüsse* von Helmholtz im Zentrum dieses Ab-  
schnitts stehen. Beide Theorien stehen in unmittelbarer Verbindung zueinander, wie  
im Laufe dieses Abschnittes deutlich werden wird.

Die nach Wissen des Autors erste, wenngleich knappe Ausformulierung der Zeichen-  
theorie lässt sich in Helmholtz' Habilitationsvortrag von 1852 finden, in der – wie schon  
bei Berkeley und Reid – die Analogie zur Sprache verwendet wird:

Licht- und Farbenempfindungen sind nur Symbole für Verhältnisse der  
Wirklichkeit; sie haben mit den letzteren ebenso wenig und ebenso viel  
Aehnlichkeit oder Beziehung, als der Name eines Menschen, oder der Schrift-  
zug für den Namen mit dem Menschen selbst. Sie benachrichtigen uns durch  
die Gleichheit oder Ungleichheit ihrer Erscheinung davon, ob wir es mit den-  
selben oder anderen Gegenständen und Eigenschaften der Wirklichkeit zu  
thun haben. Weiter leisten sie aber auch nichts. Ueber die wirkliche Na-  
tur der durch sie bezeichneten äusseren Verhältnisse erfahren wir durch  
sie ebenso wenig wie aus den Namen über die unbekanntenen Menschen und  
Städte. Der wesentliche Unterschied zwischen der Symbolik der menschli-  
chen Sprache und dieser Symbolik unserer Sinnesnerven scheint mir der zu  
sein, dass jene ein Erzeugniss der Willkür, letztere uns von der Natur selbst,  
welche unseren Körper in der bestimmten Weise aufgebaut hat, mitgegeben  
ist. (Helmholtz, 1883/2003b, S. 608)

Es lassen sich viele der bereits dargestellten Aspekte der Helmholtzschen Auffassung  
von Wahrnehmung und Erkenntnis in diesem kurzen Abschnitt finden, so etwa die an  
Kant angelehnte Unmöglichkeit der Erkenntnis der „wirklichen Natur“ oder die spe-  
zifische Auffassung des Kausalgesetzes im Wahrnehmungszusammenhang (Gleichheit  
oder Ungleichheit der Empfindungen ist zurückzuführen auf Gleichheit oder Ungleich-  
heit der sie auslösenden Gegenstände).

Besonderes Gewicht legte Helmholtz in seiner wahrnehmungspsychologischen Arbeit  
auf das zu seiner Zeit in der psychologischen, physiologischen und philosophischen Li-  
teratur virulente Problem der Raum- bzw. Tiefenwahrnehmung. Helmholtz hat diesem  
Problembereich in seinen Publikationen insgesamt mehr Seiten gewidmet, als jedem  
anderen singulären Thema (Hatfield, 1990), was möglicherweise auch auf das intensive

Studium der Auseinandersetzung Kants mit diesem Thema zurückzuführen ist. Helmholtz formuliert in seinen speziell der Wahrnehmung von Tiefe und Distanz gewidmeten Ausführungen ebenfalls die Vorstellung, dass diese nur über bestimmte Hilfsmittel möglich sei, die in der aktuellen Wahrnehmungspsychologie als Tiefencues bezeichnet werden. Aufgrund der Ausrichtung dieser Arbeit soll dieser Aspekt in der Darstellung besonders berücksichtigt werden. Um der in der Literatur nicht selten vernachlässigten Tatsache gerecht zu werden, dass Helmholtz' wahrnehmungstheoretische Ansichten im Laufe seines Lebens einigen, auch teilweise deutlichen Veränderungen unterworfen waren, soll im Folgenden die Entwicklungsgeschichte dieser, ausgehend von der ersten großen Publikation zu diesem Thema, dem 1855 gehaltenen Vortrag „Über das Sehen“, über das Hauptwerk, dem „Handbuch der physiologischen Optik“, bis zur letzten, stark epistemologisch orientierten Ausarbeitung, dem 1878 gehaltenen Vortrag „Die Tatsachen in der Wahrnehmung“, nachgezeichnet werden.

### 2.6.2. Der Vortrag „Über das Sehen“ (1855)

Der 1855 in Königsberg zu Ehren Kants gehaltene und 1884 erstmals veröffentlichte Vortrag stellt die erste umfassende Darstellung der frühen Helmholtzschen Sehtheorie sowie deren epistemologischer Implikationen dar und gilt zudem als eines der maßgeblichen Dokumente, die zur Entwicklung des Neukantianismus in Deutschland beigetragen haben (Heidelberger, 1993, 1995; Schiemann, 2014). Anhand dieses Vortrages wird unter anderem exemplarisch deutlich werden, wie stark sich Helmholtz selbst in der Tradition Kants sah und meinte, an das in der „Kritik der reinen Vernunft“ dargelegte Unterfangen anzuschließen und dieses mit naturwissenschaftlicher Ausrichtung fortzuführen. Als Thema seines Vortrages wählte er die „Lehre von den sinnlichen Wahrnehmungen des Menschen“, da sich hier „Philosophie und Naturwissenschaften am nächsten berühren“ (Helmholtz, 1896/2002d, S. 90) würden.

In diesem frühen Vortrag – Helmholtz war gerade 33 Jahre alt – lässt sich Helmholtzens Auffassung von visueller Wahrnehmung als eines dreistufigen Prozesses bereits gut ausmachen. Er beginnt mit detaillierten Ausführungen zum damaligen Kenntnisstand der Physiologie der Sehorgane (des „optischen Apparates“, wie Helmholtz sich auch ausdrückt), sowie den optischen Gesetzen, denen das von Gegenständen reflektierte und ins Auge projizierte Licht unterworfen ist. Zusammenfassend sei in dieser Hinsicht der optische Apparat mit einer *camera obscura* zu vergleichen; eine Analogie die auch in der heutigen Literatur gelegentlich anzutreffen ist.<sup>78</sup> Die Projektion des Lichts auf die Retina führe dazu, „dass auf der Fläche der Netzhaut ein optisches Bild entworfen wird, wie es auch in jeder Camera obscura geschieht“ (Helmholtz, 1896/2002d, S.

---

<sup>78</sup>„Das Auge ist ein von der Natur gebildetes optisches Instrument, eine natürliche Camera obscura“ (Helmholtz, 1896/2002d, S. 90).

94). Durch Einwirkung des Lichts auf die Netzhaut würden anschließend Empfindungen hervorgerufen, die „durch die im Sehnerven zusammengefassten Nervenfasern der Netzhaut dem Gehirn, als dem körperlichen Organe des Bewusstseins, zugeführt werden“ (Helmholtz, 1896/2002d, S. 90). Dies sei ein rein physiologischer Vorgang, der empirisch feststellbaren, festen Gesetzen unterliege. Verschiedene Stellen dieses Vortrages deuten darauf hin, dass Helmholtz 1855 der Auffassung zu sein scheint, dass das Auge als Organ *selbst* ‘empfinde’ und diese Empfindungen dann gleichsam an das Gehirn ‘weitergeleitet’ würden.<sup>79</sup> Helmholtz bezieht sich in diesem Zusammenhang ausdrücklich auf Müllers „Gesetz der spezifischen Sinnesenergien“ und bekräftigt an dieser Stelle seine vermeinte erkenntnistheoretische Nähe zu Kant durch die Feststellung, dass durch Müllers Lehre die Physiologie auf empirischem Wege ‘nachgewiesen’ habe, was „Kant schon früher für die Vorstellungen des menschlichen Geistes überhaupt nachzuweisen“ suchte, nämlich „dass die Art unserer Wahrnehmungen ebenso sehr durch die Natur unserer Sinne, wie durch die äusseren Objecte bedingt sei“ (Helmholtz, 1896/2002d, S. 99).

Das Haben von Empfindungen, von Helmholtz als rein physiologischer Zustand charakterisiert, stelle für sich aber noch keine Wahrnehmung im eigentlichen Sinne dar, Wahrnehmung bestehe in einem „*Verständnis* der Lichtempfindung“. Von Wahrnehmung sei nur dann zu sprechen, „insofern wir *durch sie* [die Empfindungen] zur Kenntnis der Gegenstände der Außenwelt gelangen“ (Helmholtz, 1896/2002d, S. 100).<sup>80</sup> Zur Erläuterung dieser dritten Stufe des Wahrnehmungsprozesses geht Helmholtz zunächst auf das Phänomen der sogenannten ‘Sinnestäuschungen’ ein, um anschließend in allgemeinerer Weise über das ‘Deuten der Empfindungen’ zu sprechen.

Das Hervorrufen von Empfindungen geschehe nach festen und unabänderlichen Gesetzen, die aufgrund der Natur der Sinnesorgane schlicht gegeben seien. So führe etwa eine Reizung der mit dem Sehorgan assoziierten Sinnesnerven stets zu einer Lichtempfindung, selbst wenn es sich bei der Reizung um eine mechanische Einwirkung auf das Auge handle. Dies könne nun auch zu ‘Sinnestäuschungen’ führen, allerdings unterstreicht Helmholtz (man erinnere sich an ähnliche Ausführungen des Descartes), dass es in einem solchen Falle nicht das Sinnesorgan sei, dass den Wahrnehmenden ‘täusche’, da es eben nur den mitgegebenen festen Gesetzen unterliege, sondern *der Wahrnehmende* täusche sich im *Verständnis* der hervorgerufenen Empfindungen.<sup>81</sup>

<sup>79</sup> „Dass die Netzhaut, welche im Auge das optische Bild empfängt ... ein empfindlicher Teil unseres Nervensystems ist, und dass durch die Einwirkung des Lichtes als eines äusseren Reizes, *in ihr* [Hervorhebung von mir] Lichtempfindung hervorgerufen wird“ (Helmholtz, 1896/2002d, S. 94).

<sup>80</sup> Dieses Zitat kann als eines der zahlreichen Beispiele für den in den Vorbemerkung angesprochenen Punkt gelten, dass sich Helmholtz im Wesentlichen des Sprachgebrauchs eines metaphysischen Realismus bedient, obwohl er sich eigentlich auf diesen nicht festlegen wollte.

<sup>81</sup> „Das Sinnesorgan täuscht uns dabei nicht, es wirkt in keiner Weise regelwidrig, im Gegentheil, es wirkt nach seinen festen und unabänderlichen Gesetzen und kann gar nicht anders wirken“ (Helmholtz, 1896/2002d, S. 100). In diesem Zusammenhang zeigt sich ebenfalls ein Problem des Konzepts der

Im Kontext der Auffassung von Wahrnehmung als ‘Deuten’ oder ‘Verstehen’ von Empfindungen spricht Helmholtz davon, dass wir dabei auf das Vorhandensein äußerer Objekte „schließen“, wir diesbezügliche „Urteile fassen“, oder auch, dass „unsere Vorstellung ... Schlussfolgerung[en]“ vornehme (Helmholtz, 1896/2002d, S. 101f.). An dieser Stelle werden Helmholtz’ Erläuterungen zunehmend vage, man könnte sie jedoch so lesen, dass sich durch wiederholte Erfahrung bestimmte Schlussfolgerungen, die „sich Millionen Male als richtig erwies[en]“ (Helmholtz, 1896/2002d, S. 101) hätten, auf *sämtliche* Fälle mit ‘hinreichender Ähnlichkeit’ angewendet würden, was in gewissen Fällen eben auch zu ‘falschen Schlüssen’, den Sinnestäuschungen, führen könne. Unklar bleibt dabei jedoch, was man im Kontext visueller Wahrnehmung unter Schlüssen, die sich als ‘richtig’ erwiesen hätten, verstehen soll.<sup>82</sup> Es zeige sich aber auch, dass bei den ‘Sinnestäuschungen’ – Helmholtz zieht nun das Beispiel des im Wasser gekrümmt erscheinenden Stabes heran – der Wahrnehmende unzählige Male die *praktische* Erfahrung machen könne, dass es sich um eine solche handle (im Falle des gekrümmten Stabes etwa durch Betasten), dennoch werde sich an dem Verständnis der Empfindungen, also der Wahrnehmung, laut Helmholtz nichts ändern:

Wir haben nämlich unsere Vorstellung auf einem Fehlschlusse ertappt; wir haben auch wissenschaftlich eingesehen, wie sie dazu gekommen ist, und wie sich die Sache eigentlich verhält. Nun sollte unsere Vorstellung doch billiger Weise ein Einsehen haben. Wir werden gestehen müssen, nein. Sie trägt sich wie der ungelehrigste aller Schüler, sie lässt die arme Lehrerin Wissenschaft reden, so lange sie will und bleibt hartnäckig bei ihrer Aussage stehen. Sie verschwinden niemals dadurch, dass wir ihren Mechanismus einsehen, sondern bleiben in ungestörter Kraft bestehen. (Helmholtz, 1896/2002d, S. 101f.)

Im Fall der ‘Sinnestäuschungen’ werde also gleichsam sowohl gegen die Erfahrung, als auch gegen besseres Wissen ‘gesehen’, ein Punkt, der deutlich gegen die von Helmholtz vorgebrachte allgemeine Auffassung, dass Wahrnehmung wesentlich erlernt werden müsse, zu sprechen scheint, aber nicht weiter diskutiert wird. Andererseits gebe es laut Helmholtz vielfältige Situationen, etwa die sogenannten *Scheinbewegungen*, an denen sich zeige, „wie schnell eine veränderte Einübung in der Deutung der Sinneswahr-

---

‘visuellen Empfindungen’. Helmholtz Beispiel der mechanischen Einwirkung auf das Auge, die eine Lichtempfindung zu Folge habe, soll als Beispiel einer ‘Sinnestäuschung’ dienen. Seinen Ausführungen gemäß handelt es sich gleichzeitig auch nicht um eine Täuschung, da die Sinnesorgane selbst sich aufgrund ihrer von der Natur festgelegten Funktionsweise nicht ‘täuschen’ können. Falls Helmholtz hier meinen sollte, dass wir uns als Wahrnehmende ‘täuschen’ würden, wenn wir die zugehörige Empfindung als *Licht* ‘interpretieren’, ist unklar, was dann mit einer Lichtempfindung gemeint sein könnte.

<sup>82</sup>Es wird sich zeigen, dass Helmholtz’ in seinen späteren Schriften zumindest andeutet, dass ‘richtige Schlüsse’ jene seien, die sich als in *praktischer* Hinsicht erfolgreich herausgestellt haben.

nehmungen eintreten kann“ (Helmholtz, 1896/2002d, S. 107).<sup>83</sup> Eine weitergehende Erläuterung oder Diskussion, worin sich diese beiden offenbar fundamental verschiedenen Fälle unterscheiden, liefert Helmholtz nicht, so dass das Phänomen der Sinnestäuschungen im Rahmen des von Helmholtz vorgebrachten Ansatzes unverständlich bleibt. Wir werden auf diesen Punkt an verschiedenen Stellen in diesem Abschnitt zurückkommen, da dieser ein ernstzunehmendes Problem der Helmholtzschen Theorie zu sein scheint.

Helmholtz spricht an den wesentlichen Stellen des Vortrags davon, dass ‘die Vorstellung’ urteile und *nicht* wir selbst, wobei nicht weiter erläutert wird, was mit dem Begriff der Vorstellung gemeint sein könnte.<sup>84</sup> Die Deutung der Empfindungen unterliege im Gegensatz zu üblichen logischen Schlussfolgerungen nicht unserem eigenen Willen und Wissen und sei der Introspektion unzugänglich. Ein weiteres wesentliches Merkmal, welches das Zustandekommen von Wahrnehmungen anhand von Schlussfolgerungen von üblichen Schlussfolgerungen unterscheide, sei der Umstand, dass diese mit außerordentlicher Schnelligkeit geschehen würden – man erinnere sich an ähnliche Ausführungen von Alhazen und Descartes.<sup>85</sup>

Ein das Verständnis der Helmholtzschen Auffassung bezüglich des Zustandekommens der Wahrnehmung erschwerender Umstand besteht darin, dass er an vielen Stellen seiner Beschreibungen zwischen einer Beschreibung der Wahrnehmung als eines dem Bewusstsein nicht zugänglichen Schlusses und als einer auf wiederholten gemeinsamen Auftretens von ‘Vorstellungen’ basierenden assoziativen Verknüpfung im Sinne Berkeleys zu schwanken scheint, wobei die letztere Beschreibung häufiger verwendet wird:

Dass es ein nicht mit Selbstbewusstsein vollzogener Schluss sei, darüber sind wir einig. Vielmehr hat er mehr den Charakter eines mechanisch eingeübten,

---

<sup>83</sup>Man beachte, dass Helmholtz hier davon spricht, dass „Sinneswahrnehmungen“ und *nicht* Empfindungen gedeutet werden. Es zeigt sich hier einerseits eine inkonsistente Verwendung des Begriffs der Wahrnehmung, sowie die an mehreren Stellen in diesem Vortrag auffindbare Auffassung, dass die Sinne *selbst* empfinden oder gar ‘wahrnehmen’ würden.

<sup>84</sup> In späteren Schriften, etwa dem als nächstes zu behandelnden „Handbuch der physiologischen Optik“, führt Helmholtz den Begriff der Vorstellung ein als „Erinnerungsbild von Gesichtsobjekten“ (Helmholtz, 1910/2003a, S. 11), eine Formulierung, die sich nicht gut mit dem Vortrag von 1855 in Übereinstimmung bringen lässt.

<sup>85</sup>„Ich habe bisher immer gesagt, die Vorstellung in uns urtheile, schliesse, überlege u.s.w., habe mich aber wohl gehütet zu sagen, *wir* urtheilen, schliessen, überlegen; denn ich habe schon anerkannt, dass diese Acte ohne unser Wissen vor sich gehen und auch nicht durch unseren Willen und unsere bessere Ueberzeugung abgeändert werden können. Dürfen wir denn nun, was hier geschieht, dieses Denken ohne Selbstbewusstsein und nicht unterworfen der Controlle der selbstbewussten Intelligenz, wirklich als Prozesse des Denkens bezeichnen? Dazu kommt, dass die Genauigkeit und Sicherheit in der Construction unserer Gesichtsbilder so gross, so augenblicklich und unzweifelhaft ist, wie wir sie unseren Schlussfolgerungen zuzutrauen eigentlich nicht geneigt sind. ... Und was die Schnelligkeit der Schlussfolgerungen betrifft, so sind diese auch da, wo wir absolute Zuverlässigkeit erreichen können, wie in mathematischen Folgerungen und Rechnungen, so schwerfällig, umständlich und langsam, dass sie mit den blitzschnellen Auffassungen unseres Auges nicht im Entferntesten verglichen werden können“ (Helmholtz, 1896/2002d, S. 110f.).

der in die Reihe der unwillkürlichen Ideenverbindungen eingetreten ist, wie solche zu entstehen pflegen, wenn zwei Vorstellungen sehr häufig mit einander verbunden vorgekommen sind. (Helmholtz, 1896/2002d, S. 112).

Wir hätten, so Helmholtz weiter, gelernt zu sehen, indem bestimmte ‘Vorstellungen’ (wobei Helmholtz auch hier offen lässt, was damit gemeint ist) auf rein mechanische Weise verknüpft werden, Wahrnehmung sei aber gleichzeitig auch „mehr als eine mechanisch eingeübte Ideenverbindung“ (Helmholtz, 1896/2002d, S. 115), sondern beinhalte auch und wesentlich eine Schlussfolgerung. Das beständige Schwanken zwischen diesen beiden Polen der Beschreibung der Wahrnehmung als einerseits ‘rein assoziativer Verknüpfung’ und andererseits als Schlussfolgerung, ist kennzeichnend für diesen Vortrag.

Man könnte die Helmholtzschen Ausführungen jedoch auch so lesen, dass es sich beim Schließen und der rein assoziativen Verknüpfung um *zwei fundamental verschiedene* Prozesse handle, die *beide* beim Zustandekommen der Wahrnehmung zum Tragen kämen. So könnte die Verbindung zwischen bestimmten Empfindungen und bestimmten Eigenschaften der wahrgenommenen Objekte, etwa die Form oder Oberflächenbeschaffenheit, eine rein assoziative, durch wiederholte Erfahrungen mit Empfindungen aus Tast- und Sehsinn zustande gekommene sein, es müsse aber zusätzlich *auch* zu einem Schluss kommen, da wir die als Objekte gedeuteten Empfindungen üblicherweise als Gegenstände einer Außenwelt wahrnehmen. Es sei laut Helmholtz ausgeschlossen, dass sich eine assoziative Verknüpfung zwischen bestimmten Empfindungen und der Vorstellung eines Gegenstandes der Außenwelt bilden könne, da wir keinen unmittelbaren Zugang zu den Gegenständen der Außenwelt hätten, sondern nur zu den *Wirkungen* dieser auf unsere Sinnesorgane. Um überhaupt „aus der Welt der Empfindungen unserer Nerven ... in die Welt der Wirklichkeit“ (Helmholtz, 1896/2002d, S. 115) zu gelangen, bedarf es eines Schlusses und *allein* ein Schluss könne dies leisten. Dieser Schluss bestehe darin, dass wir die Gegenwart von Objekten der Außenwelt als Ursache der Wirkungen auf unsere Sinnesorgane *voraussetzen* müssten, da keine Wirkung ohne Ursache sei, wie Helmholtz mit Bezug auf das Kausalgesetz feststellt, welches „ein vor aller Erfahrung gegebenes Gesetz unseres Denkens“ (Helmholtz, 1896/2002d, S. 116) sei. Genauere Erläuterungen diesbezüglich lassen sich nicht finden, es scheint Helmholtz große Schwierigkeiten bereitet zu haben, die Natur dieses Prozesses genauer zu bestimmen.

So bleibt die von Helmholtz selbst angedeutete Frage bestehen, wie man sich eine dem Bewusstsein nicht zugängliche Schlussfolgerung vorzustellen habe, die aufgrund des vorausgesetzten Kausalgesetzes dazu führe, dass Empfindungen *notwendigerweise* als Wirkungen von Gegenständen einer Außenwelt gedeutet werden *müssten*. Wie im Abschnitt über Alhazen, der Wahrnehmung als einem deduktiven Schlussfolgern gleich auffasst, bereits angesprochen, wird der Begriff der Schlussfolgerung üblicherweise in

einer Weise verwendet, so dass damit ein *bewusster geistiger Akt* bezeichnet wird, der dem diesen Vollziehenden erlaubt, über Anwendung logischer Regeln aus gewissen Prämissen eine Konklusion abzuleiten. Dabei sind die logischen Regeln, die Prämissen und die Konklusion sprachliche und/oder formale Gebilde. Der von Helmholtz behauptete Schlussfolgerungsprozess scheint vollkommen anderer Natur zu sein. Um in Helmholtz' möglicher Metaphorik zu sprechen, könnte man sagen, dass im Fall der Wahrnehmung die 'Prämissen' die Empfindungen (physiologisch zu charakterisierende Nervenzustände) seien und die Konklusion in einer Wahrnehmung außerhalb von mir existierender Gegenstände mit bestimmten Eigenschaften bestehe. Der behauptete notwendige Übergang von Empfindung zu Wahrnehmung soll jedoch durch ein gleichsam 'sprachliches' Gebilde gesichert sein, das Kausalgesetz. Ohne weitergehende Erläuterungen ist jedoch nicht einzusehen, welche Verbindung zwischen diesen Elementen bestehen soll.

Möchte man Helmholtz' Ausführungen wohlwollend begegnen, könnte man vermuten, dass in Ermangelung eines angemesseneren Begriffes und des zu dieser Zeit vergleichsweise begrenzten Wissens um Aufbau und Funktion des Hirnes, Helmholtz den Begriff des Schlussfolgerns hier lediglich als *Analogie* für einen rein physiologischen und nicht kognitiven Prozess verwenden möchte. Gegen eine solche Auslegung spricht allerdings erstens, dass Helmholtz an gewissen Stellen sehr viel Wert darauf legt, diesen Prozess explizit als Schluss des Wahrnehmenden selbst zu bezeichnen und zweitens an keiner Stelle von Helmholtz deutlich gemacht wird, dass es sich hierbei um eine Analogie handeln soll, sich also der rein physiologische Prozess so beschreiben ließe, *als ob* eine 'Schlussfolgerung' durchgeführt werde.<sup>86</sup> Das von Helmholtz behauptete Zustandekommen der Wahrnehmung externer Objekte als Ergebnis einer Schlussfolgerung bleibt somit in diesem Vortrag weitgehend undurchsichtig.

Das Argument, das über das Kausalgesetz die Notwendigkeit der Deutung der Empfindungen als Objekte einer Außenwelt rechtfertigen soll und das Helmholtz auch in späteren Schriften wieder anführen wird, ist bei eingehender Betrachtung zudem nicht unproblematisch. So scheint dieses Argument doch entweder vorauszusetzen, dass die Empfindungen, die dann als Gegenstände einer Außenwelt gedeutet werden, in der Tat die Wirkung (vorausgesetzter) Objekte einer vom Wahrnehmenden unabhängigen Außenwelt seien oder es zumindest keine andere Möglichkeit gebe, als was die Empfindungen ursächlich gedeutet werden könnten. Ohne weitergehende Voraussetzungen ist es nicht zwingend, dass wir aufgrund unseres Habens von Empfindungen und der von Helmholtz angenommenen Gültigkeit des Kausalgesetzes, dass also keine Wirkung ohne Ursache sei, wir „die Gegenwart äusserer Objecte als *Ursache* unserer Nervenerrregung voraussetzen“ *müssen*. Des Weiteren gibt dieses Argument nicht Auskunft

---

<sup>86</sup>„Auf welche Weise sind wir denn nun zuerst aus der Welt der Empfindungen unserer Nerven hingekommen? Offenbar nur durch einen Schluss; wir müssen die Gegenwart äusserer Objecte als *Ursache* unserer Nervenerrregung voraussetzen“ (Helmholtz, 1896/2002d, S. 115).

darüber, *weshalb* von den Empfindungen gerade auf Objekte einer Außenwelt geschlossen werden soll, dahingestellt, ob diese nun die Empfindungen bewirken oder nicht. Die angenommene Gültigkeit des Kausalgesetzes impliziert zwar, dass die Empfindungen eine Ursache haben müssen, es lässt sich aber ohne Weiteres weder daraus schließen, dass eine von unserem Bewusstsein unabhängige Außenwelt mit darin existierenden Objekten besteht oder wir diese voraussetzen müssen, noch die Ursache unserer Empfindungen in eben diesen Objekten der Außenwelt besteht oder die Empfindungen *als* Objekte einer unabhängig existierenden Außenwelt gedeutet werden müssten. Schließlich scheint der Status des Kausalgesetzes in diesem Argument selbst nicht klar zu sein. Die Empfindungen als zu deutende Zeichen *müssten* laut Helmholtz aufgrund des Kausalgesetzes vom Wahrnehmenden als von einer unabhängig bestehenden Außenwelt kausal Bewirktes interpretiert werden, was einen ‘unkantischen’ Gebrauch des Kausalgesetzes darstellt, da sich die Kategorie der Kausalität im Sinne Kants *allein* auf Erscheinungen bezieht, nicht jedoch auf das ‘Ding an sich’ oder mögliche Beziehungen zwischen diesem und dem wahrnehmenden Subjekt. In der Kantischen Lehre ist durch die Kategorie der Kausalität *a priori* gleichsam festgelegt, dass sich dem Wahrnehmenden die Inhalte der Wahrnehmung stets nur als in kausaler Beziehung zueinander stehend darbieten können. In diesem Sinne würde das Kantische Kausalgesetz, als Prinzip der Möglichkeit von Erfahrung überhaupt, lediglich implizieren, dass der sich mit Wahrnehmung befassende *Wissenschaftler* diese als bestimmt durch vorangehende Ursachen auffasst. In Helmholtz’ Theorie scheint allerdings über das Kausalgesetz eine Beziehung zwischen den Objekten einer unabhängig existierenden Außenwelt und den Wahrnehmungen eines Subjekts hergestellt werden zu sollen. Nicht allein die Inhalte unserer Wahrnehmung sollen uns notwendigerweise als in kausaler Beziehung zueinander stehend erscheinen, sondern der Helmholtzschen Auffassung des Kausalgesetzes zufolge müssen die Empfindungen notwendigerweise als von einem unabhängig vom Wahrnehmenden kausal Bewirktes interpretiert werden. Bei Helmholtz bezieht sich das Kausalgesetz auf eine unabhängig vom Wahrnehmenden existierende Außenwelt mit in kausaler Beziehung zueinander stehenden Objekten, der der Wahrnehmende mit seinen Sinnesorganen angehören muss, damit die Objekte kausal auf diese einwirken und die Empfindungen hervorbringen können, die dann wiederum über das Kausalgesetz als von diesen unabhängigen Objekten bewirkt gedeutet werden müssten. Die Helmholtzsche Konzeption scheint hier zur Legitimation der realistischen Grundhaltung Helmholtzens zu dienen, ohne dass diese Grundhaltung selbst eine Begründung findet oder thematisiert wird.

Bezüglich der Wahrnehmung von Tiefe und Entfernung seien es ebenfalls Urteile oder Schlussfolgerungen, die diese ermöglichen sollen. Die wahrgenommene Lokalisation eines Gegenstandes sei darauf zurückzuführen, dass anhand der durch Licht gereizten

Stelle der Netzhaut die *Richtung* beurteilt werde, in der sich die diese Reizung verursachenden Gegenstände befänden.<sup>87</sup> Helmholtz scheint hier eine der Lotzeschen Lokalzeichentheorie ähnliche Vorstellung vor Augen zu haben, ohne diese explizit zu nennen; eine Vermutung, die sich in späteren Schriften bestätigen wird, in denen Helmholtz in diesem Zusammenhang auch explizit von Lokalzeichen spricht.<sup>88</sup> Die Lokalzeichen sollen bei Helmholtz (wie bei Lotze auch) lediglich dazu dienen, die Lokalisation der wahrgenommenen Objekte gleichsam ‘auf einer Ebene’ verständlich zu machen. Zur Wahrnehmung von Tiefe seien zusätzliche „Hilfsmittel für die Beurtheilung der Tiefendimension“ (Helmholtz, 1896/2002d, S. 104) vonnöten, die es uns erlaubten, „die Raumverhältnisse der uns umgebenden Gegenstände fortdauernd [zu] konstruieren“ (Helmholtz, 1896/2002d, S. 105).

Helmholtz nennt in diesem Vortrag drei solcher Hilfsmittel, nämlich die perspektivische Darstellung in der Kunst, die Konvergenz (aufgefasst im Berkeleyschen Sinne, d.h. als zugehörige Muskelbewegungen der Augen) und die sogenannte *Disparität der Netzhautbilder* bei binokularer Betrachtung, ein Mittel der Tiefenwahrnehmung, das in der bisherigen Darstellung noch keine Erwähnung fand. Mit diesem Begriff wird der Umstand bezeichnet, dass in einer festen Sehsituation durch den Abstand der beiden Augen zueinander differierende Lichtmuster (oder ‘Netzhautbilder’, wenn man diese betrachten würde) auf die beiden Retinae projiziert werden, die aus rein geometrischer Perspektive bestimmte Relationen zueinander aufweisen. In Lehrbüchern zur Wahrnehmungspsychologie (etwa Coren et al., 1999) wird zur Verdeutlichung häufig das Beispiel herangezogen, dass man einen Finger nahe an das Gesicht halten und diesen Finger dann abwechselnd mit dem einen und dem anderen Auge betrachten solle; eine Verdeutlichung, die sich jedoch auf die Wahrnehmung selbst und nicht das retinale Lichtmuster bezieht. Die binokulare Disparität wird auch in der modernen wahrnehmungspsychologischen Literatur als Wahrnehmung von Tiefe ermöglichender Tiefencue

---

<sup>87</sup>„Nach der Stelle unserer Netzhaut, in welcher Lichtempfindung angeregt wird, beurtheilen wir also, in welcher Richtung die verschiedenen hellen Gegenstände, die uns umgeben, sich befinden.“ (Helmholtz, 1896/2002d, S. 102).

<sup>88</sup>Zum Einfluss der Lokalzeichentheorie von Lotze auf die Helmholtzsche Theorie siehe Lenoir (1993).

angesehen.<sup>89</sup> Wir werden auf die binokulare Disparität im Abschnitt über Helmholtz' „Handbuch der physiologischen Optik“ ausführlicher eingehen.

Es ist anzunehmen, dass Helmholtz' Theorie zu dieser Zeit wesentlich mehr Hilfsmittel als die von ihm tatsächlich aufgezählten vorsah (etwa die Akkommodation), er aber aufgrund des populärwissenschaftlichen Charakters dieses Vortrages auf eine extensive Aufzählung und Untergliederung, wie er sie dann später im „Handbuch der Physiologischen Optik“ vornehmen sollte, verzichtet. Aber selbst anhand der im Vortrag von 1855 eher kurz gehaltenen Beschreibung der Hilfsmittel wird deutlich, dass es sich bei diesen um Konzepte aus deutlich unterschiedlichen Beschreibungsebenen handelt. Die Konvergenz lässt sich eher einer physiologischen Beschreibungsebene zuordnen, die binokulare Disparität eher einer physikalischen oder geometrischen Beschreibungsebene (da sie sich aus den optischen Gesetzen ergibt) und bei der perspektivischen Darstellung in der Kunst schließlich fällt die Einordnung ungemein schwerer. Dass es sich bei der linearperspektivischen Darstellung um ein solches Hilfsmittel handle, sieht Helmholtz darin begründet, dass uns in entsprechend gemalten Bildern die dargestellte Szene räumlich erscheint; eine Beschreibung, die allein unserer Wahrnehmung zu entnehmen ist. Was die gemeinsame Verbindung dieser Hilfsmittel sein soll – außer dass sie allesamt Hilfsmittel zur Beurteilung der Tiefe seien sollen – und auf welche Weise sie zur Beurteilung der Tiefe oder Entfernung beitragen sollen, bleibt unklar, Helmholtz geht in diesem Vortrag auf beide Punkte nicht weiter ein. Insbesondere wird nicht ersichtlich, ob Helmholtz diese Hilfsmittel als Empfindungen auffassen möchte (was insbesondere bei der perspektivischen Darstellung nicht verständlich wäre) oder als etwas anderes, möglicherweise zwischen Empfindung und Wahrnehmung Anzuesiedelndes. Er äußert sich allerdings in einer aufschlussreichen Passage etwas genauer zur binokularen Disparität:

Nun hat der Mensch zwei Augen, welche fortdauernd die Welt von zwei verschiedenen Standpunkten aus betrachten, welche also auch fortdauernd zwei verschiedene perspectivische Ansichten dem Bewusstsein zur Begutachtung darbieten, so oft sie einen körperlich ausgedehnten Gegenstand wahrnehmen. ... So construieren wir uns also die Raumverhältnisse der uns umge-

---

<sup>89</sup>Dass im 19. Jahrhundert die binokulare Disparität als ein Hilfsmittel der Tiefenwahrnehmung angesehen wird, ist zu großen Teilen auf die Erfindung des Stereoskops durch Charles Wheatstone zurückzuführen. Das Stereoskop ermöglicht es, zwei Bilder so zu präsentieren, dass das von einem Bild reflektierte Licht nur auf eines der beiden Augen projiziert wird, oder anschaulich gesprochen, dem einen Auge nur das eine und dem anderen nur das andere Bild 'dargeboten' wird. Handelt es sich bei den beiden dargebotenen Bildern um Fotografien derselben Szene, die aus zwei verschiedenen Positionen aufgenommen wurden, die gerade den Positionen der beiden Augen entsprechen würden, wenn eine Person die entsprechende Szene sähe, so ist mit Betrachtung dieser Bilder durch ein Stereoskop ein eindringlicher räumlicher Eindruck verbunden. Stereoskopische Fotografien wurden im ausgehenden 19. und beginnenden 20. Jahrhundert geradezu zu einem Massenphänomen. Für weitergehende Erläuterungen siehe etwa Arditti (1986) oder die Originalarbeit von Wheatstone (1879/2011).

benden Gegenstände fortdauernd aus zwei verschiedenen perspectivischen Ansichten, welche unsere beiden Augen uns von ihnen liefern. (Helmholtz, 1896/2002d, S. 104f.)

Auffällig an dieser Beschreibung ist das von Helmholtz verwendete mentale Vokabular, das ohne weitere Erläuterungen jedoch äußerst vage scheint, so dass die Funktionsweise dieses Hilfsmittels nicht wirklich deutlich wird und zudem nicht in die bisher von Helmholtz dargestellte Vorstellung von Wahrnehmung, als über assoziative Verbindungen und Schlussfolgerungen zustande kommend, zu passen scheint. Helmholtz spricht davon, dass dem Bewusstsein „zwei verschiedene perspectivische Ansichten“ zur ‘Begutachtung’ dargeboten werden und der Wahrnehmende schließlich ‘Raumverhältnisse konstruiere’. Was mag das heißen? ‘Sieht’ das Bewusstsein die retinalen Lichtmuster so wie wir, wenn wir sie betrachten könnten, als *Bilder*? Wenn ja, wie ist das ‘Sehen des Bewusstseins’ dann zu erklären? Helmholtz’ Ausführungen scheinen zu implizieren, dass das Bewusstsein selbst oder etwas ‘im Bewusstsein’, einem Beobachter gleich, das auf die Netzhaut projizierte Lichtmuster in einer Weise ‘sehen’ könne, wie wir es sehen würden, wenn wir es sehen könnten. Aber welche Verbindung soll dann zwischen diesem ‘Sehen des Bewusstseins’ und *unserem* Sehen bestehen? Der Wahrnehmende selbst sieht die retinalen Lichtmuster nicht, er sieht externe Objekte, also kann das ‘Sehen des Bewusstseins’ nicht das vom Wahrnehmenden Gesehene sein. Warum spricht Helmholtz dennoch im gleichen Absatz davon, dass *wir* etwas aus diesen „perspectivischen Ansichten“ ‘konstruieren’ würden? Wir sehen diese nicht und können somit auch nichts aus ihnen ‘konstruieren’. Helmholtz’ Beschreibungen der Funktionsweise des Hilfsmittels der binokularen Disparität scheint ein gutes Beispiel für eine mit dem Begriff des Sehens verbundene konzeptuelle Verwirrung zu sein, die uns auch schon im Zusammenhang mit dem Hilfsmittel der ‘retinalen Größe’ begegnet ist. Das Bewusstsein, insofern damit nicht der Wahrnehmende selbst gemeint ist, kann nichts sehen oder ‘begutachten’, allein der Wahrnehmende sieht oder begutachtet etwas; was er sieht sind jedoch Objekte und keine „zwei verschiedenen perspectivischen Ansichten“. Die von Helmholtz offenbar gemeinten unterschiedlichen retinalen Lichtmuster kann der Wahrnehmende nicht sehen.

Innerhalb des Vortrags wird auch im Fall der Tiefen- und Entfernungswahrnehmung über die angesprochenen Hilfsmittel ein Schwanken Helmholtz’ zwischen der Beschreibung derselben als eines Urteils und einer rein assoziativen Verknüpfung (etwa zwischen Empfindungen des Tast- und des Sehsinnes) deutlich. So heißt es beispielsweise an einer Stelle, die deutlich an ähnliche Ausführungen Berkeleys erinnert und sich auf die der Konvergenz zugehörigen Augen- oder Augenmuskelbewegungen als Hilfsmittel bezieht:

Unser ganzes Leben hindurch haben wir in millionenfacher Wiederholung erfahren, dass, wenn in gewissen Nervenfasern unserer beiden Augen, bei

einer gewissen Stellung derselben, ein Gegenstand Lichtempfindung erregte, wir den Arm so weit ausstrecken mussten, oder eine bestimmte Zahl von Schritten gehen mussten, um ihn zu erreichen. Dadurch ist denn die unwillkürliche Verbindung zwischen dem bestimmten Gesichtseindruck und der Entfernung und Richtung, in welcher der Gegenstand zu suchen ist, hergestellt. (Helmholtz, 1896/2002d, S. 113)

Die mit dieser Vorstellung verbundenen Schwierigkeiten sind dementsprechend dieselben, die schon im Abschnitt über Berkeley angeführt wurden, aber dennoch erneut in aller Kürze in Erinnerung gerufen werden sollen. Erstens wird hiermit das Problem lediglich verschoben. Visuell wahrgenommene Entfernung soll durch Entfernungswahrnehmung über den Tastsinn erklärt werden, die vorausgesetzt werden muss. Die Wahrnehmung von Entfernung über den Tastsinn, etwa durch Ausstrecken des Armes, setzt aber bereits einen Wahrnehmenden voraus, der sich als sich in einem Raum mit anderen Gegenständen befindend auffasst. Zweitens bleibt in dieser Konzeption ungewiss, wie der Zusammenhang zwischen der etwa durch Ausstrecken des Armes ‘gefühlten Entfernung’ und dem *Sehen* von Entfernung gedacht werden soll. Drittens müssten, wenn man von einer rein assoziativen Verknüpfung ausgeht, *sämtliche möglichen* Empfindungen des Tastsinnes, die zu einer Tiefenwahrnehmung über den Tastsinn führen, mit den entsprechenden Empfindungen des Sehsinnes wiederholt assoziiert werden, um zu erklären, wie es möglich ist, dass wir *stets* Objekte in bestimmter Entfernung visuell wahrnehmen. Viertens ließe sich so nicht erklären, auf welche Weise wir die Entfernung von Objekten wahrnehmen, die wir unmöglich über den Tastsinn wahrgenommen haben können, etwa die des Mondes. Zur Vermeidung der beiden letztgenannten Probleme könnte man anführen, dass sich das Bewusstsein hier möglicherweise einer ‘abstrakten Regel’ bedienen könnte, die es erlauben würde, in Fällen, in denen noch keine assoziativen Verbindungen vorliegen, Empfindungen als ‘ähnlich’ zu anderen, bei denen solche Verbindungen bereits bestehen, zu klassifizieren und so zu einer Wahrnehmung räumlicher Verhältnisse per ‘Ähnlichkeit’ zu kommen. Es wird sich zeigen, dass Helmholtz in seinen nachfolgenden Schriften einen in diese Richtung weisenden Schritt gehen wird, allerdings werden sich auf diese Weise auch neue theoretische Probleme einstellen.

Insgesamt bietet Helmholtz’ theoretischer Ansatz zur Tiefenwahrnehmung – wenn man die vergleichsweise kurzen Ausführungen von 1855 so nennen möchte – ein durchaus heterogenes Bild. Die Wahrnehmung von Tiefe soll ebenfalls auf Urteilen beruhen, werde durch Einübung erlernt und sei durch Erfahrung veränderbar.<sup>90</sup> Die von ihm vorgeschlagenen Hilfsmittel scheinen differenten Beschreibungsebenen zu entstammen, die Funktionsweise dieser Hilfsmittel bleibt weitestgehend unklar. Im Fall der binoku-

---

<sup>90</sup>„Was die Beurtheilung der Entfernung durch die Augen betrifft, so können wir wohl ebenfalls nicht zweifeln, dass diese durch Einübung angelernt sei“ (Helmholtz, 1896/2002d, S. 114).

laren Disparität scheint Helmholtz eine aktive Tätigkeit des Bewusstseins (und/oder des Wahrnehmenden) nahelegen zu wollen, im Fall der Konvergenz soll es sich um bloße assoziative Verknüpfungen zwischen Empfindungen des Tast- und des Sehannes handeln. Insbesondere in Helmholtz' Beschreibung der Funktionsweise der binokularen Disparität wird offenbar, dass entweder ein Homunkulus vorausgesetzt werden muss, dessen 'Sehen' des retinalen Lichtmusters dann letztlich unerklärt bleibt, oder aber vorausgesetzt werden muss, dass der Wahrnehmende selbst die retinalen Lichtmuster *als Bilder* sieht. Unklar bleibt auch, welchen Status die von Helmholtz angesprochenen Hilfsmittel haben sollen. In Helmholtz' allgemeiner Konzeption sollte es sich auch hier um Empfindungen handeln, im Fall der perspektivischen Darstellung in der Kunst ist das jedoch nicht einleuchtend. Wir können zwar wahrnehmen, dass etwas perspektivisch dargestellt ist und uns räumlich erscheint, wie man sich jedoch eine diesbezügliche 'Empfindung' vorzustellen habe, ist nicht auszumachen.

In seinen weiteren Schriften wird Helmholtz diese theoretischen Ansätze stetig erweitern und modifizieren, der nächste Abschnitt soll die wohl umfassendste und systematischste Darstellung der Helmholtzschen Wahrnehmungstheorie zum Gegenstand haben, in der sich zudem markante Veränderungen im Vergleich zum Vortrag von 1855 zeigen werden.

### 2.6.3. Das „Handbuch der physiologischen Optik“

In Helmholtz' in der Wissenschaftsgeschichte am ehesten mit seinem Namen verbundenen Werk, dem dreibändigen „Handbuch der physiologischen Optik“ (im Folgenden nur noch als *Handbuch* bezeichnet), werden die in dem Vortrag von 1855 dargelegten Elemente seines wahrnehmungstheoretischen Ansatzes genauer erläutert und an verschiedenen Punkten modifiziert.<sup>91</sup> Der dritte, oft als 'genuin psychologisch' bezeichnete Band des *Handbuchs*, der allein der „Lehre von den Gesichtswahrnehmungen“ vorbehalten ist und Helmholtz laut Koenigsberger (1902-1903) am meisten Kopfzerbrechen bereitet habe, beginnt mit §26 „Von den Wahrnehmungen im allgemeinen“, der als klassische Darlegung von Helmholtzens *Theorie der unbewussten Schlüsse* gelten darf.

Es wird von Helmholtz zunächst eine neue Begriffsunterscheidung eingeführt, nämlich zwischen *Anschaung*, *Vorstellung* und *Perzeption*. Der Begriff der *Anschaung* bezeichne die „von den sinnlichen Empfindungen begleitete Wahrnehmung“, eine *Vorstellung* sei ein „Erinnerungsbild von Gesichtobjekten, welches von keinen gegenwärtigen sinnlichen Empfindungen begleitet ist“ und eine *Perzeption* eine „Anschaung,

<sup>91</sup>Wenn in dieser Arbeit von *dem* „Handbuch der physiologischen Optik“ gesprochen wird, so ist damit stets die zweite bzw. dritte Auflage desselben gemeint, die wesentliche Veränderungen gegenüber der ersten Auflage aufweist. Der Unterschied zwischen der zweiten und dritten Auflage besteht in von Johannes von Kries geschriebenen Zusätzen zu einzelnen Kapiteln, der eigentliche Helmholtzsche Text wurde unverändert übernommen.

in der nichts enthalten ist, was nicht aus den unmittelbar gegenwärtigen sinnlichen Empfindungen hervorgeht, also eine Anschauung, wie sie auch ohne alle Erinnerung an früher Erfahrenes sich bilden könnte“ (Helmholtz, 1910/2003a, S. 11).<sup>92</sup> Dem aufmerksamen Leser mag an dieser Stelle nicht entgangen sein, dass der Begriff der Vorstellung hier auf gänzlich andere Weise verwendet wird, als es noch in dem Vortrag von 1855 der Fall war, in dem Helmholtz den Prozess des Schlussfolgerns der Vorstellung zugeschrieben hatte (vgl. Fußnote 84 auf Seite 86).

Welchen Zweck diese begriffliche Unterscheidung innerhalb der Helmholtzschen Theorie erfüllen soll, bleibt undeutlich; ein Eindruck, der dadurch verstärkt wird, dass sich Helmholtz selbst an vielen Stellen nicht an diese Unterscheidung zu halten scheint bzw. diese Begriffe zumindest in inkonsistenter Weise verwendet.<sup>93</sup> Insbesondere erscheint der Begriff der Anschauung im Rahmen dieser Unterscheidung schwierig, da er im Kontext der Helmholtzschen Theorie mit dem der Wahrnehmung *prima facie* identisch zu sein scheint, diese neue Begriffsunterscheidung allerdings nahe legt, dass es auch Wahrnehmung *ohne* sinnliche Empfindungen geben könne, was wiederum in Spannung zu Helmholtz' allgemeiner Auffassung von Wahrnehmung als gedeuteten Empfindungen stehen würde. Möglicherweise sollte die Einführung dieser Begriffe dazu dienen, dem sich zu dieser Zeit herauskristallisierenden Streit zwischen empiristischen und nativistischen Wahrnehmungstheorien eine gemeinsame begriffliche Basis zur Verfügung zu stellen. Eine Anschauung könne laut Helmholtz aus einer Verbindung von „Vorstellung und Perzeption in den verschiedensten Verhältnissen“ (Helmholtz, 1910/2003a, S. 11) bestehen und daher sowohl auf Erfahrungen als auch auf ‘angeborenen Eigenheiten’ des Sehens beruhen, wobei Helmholtz mittels der Formulierung „in den verschiedensten Verhältnissen“ offen lässt, ob bei Anschauungen *stets* Perzeption und Vorstellung gemeinsam auftreten, oder ob es auch Fälle von Anschauungen geben könne, in denen nur eins der beiden auftritt.

---

<sup>92</sup>Der so von Helmholtz eingeführte Begriff der Anschauung weist starke Ähnlichkeit mit dem von Helmholtz häufig verwendeten Begriff der *Apperzeption* auf. Dieser wurde von Leibniz in seinen „Neuen Abhandlungen über den menschlichen Verstand“ von 1704 zuerst eingeführt (Eisler, 1904) und dann – mit sich oft stark wandelnder Bedeutung – von Kant, Fichte, Herbart und auch Wundt benutzt. Ein nützlicher Überblick der verschiedenen Verwendungsweisen dieses Terminus lässt sich etwa bei Lange (1906) finden.

<sup>93</sup>So heißt es etwa an einer anderen Stelle im *Handbuch*: „Der Inbegriff aller dieser möglichen Empfindungen in eine Gesamtvorstellung zusammengefaßt, ist unsere *Vorstellung* von dem Körper, welche wir *Wahrnehmung* nennen, solange sie durch gegenwärtige Empfindungen unterstützt ist, *Erinnerungsbild*, wenn sie das nicht ist“ (Helmholtz, 1910/2003a, S. 433). Was in dem oben zitierten §26 als Anschauung bezeichnet wird, heißt hier nun Wahrnehmung und der Begriff der Vorstellung ist hier ein ganz anderer, viel allgemeinerer. Auch schon im ersten Satz des den dritten Band des *Handbuchs* einleitenden §26 heißt es: „Wir benutzen die Empfindungen, welche Licht in unserem Sehnervenapparate erregt, um uns aus ihnen Vorstellungen über die Existenz, die Form und die Lage äußerer Objekte zu bilden. Dergleichen Vorstellungen nennen wir *Gesichtswahrnehmungen*“ (Helmholtz, 1910/2003a, S. 1).

## Die Schlußtheorie

Der Großteil des „Wahrnehmungen im allgemeinen“ behandelnde §26 des *Handbuchs* ist der Natur des Deutens der Empfindungen in Form eines ‘Schlussaktes’ gewidmet. Beim Lesen der übrigen Teile des dritten Bandes des *Handbuchs*, die dann Einzelthe-men wie etwa der wahrgenommenen Lokalisation von Gegenständen gewidmet sind, entsteht jedoch eine auffällige Spannung zu diesem Eingangsparagrafen, da Helmholtz in den späteren Abschnitten diese einzelnen Phänomene im Wesentlichen über rein assoziative Verbindungen verständlich machen möchte. Im bereits dargestellten Vortrag von 1855 sprach sich Helmholtz, nach der vorgeschlagenen Lesart, für *zwei* Prozesse aus, die *beide* zum Zustandekommen von Wahrnehmung, im Sinne des Sehens externer Objekte mit bestimmten Eigenschaften wie Form etc., notwendig seien. Dass Empfindungen überhaupt als Gegenstände einer Außenwelt gedeutet werden, beruhe auf dem Kausalgesetz, die genauen Bestimmungen des Wahrgenommenen seien hingegen auf rein assoziative Verknüpfungen zurückzuführen. Die wohl markanteste Modifikation der Helmholtzschen Wahrnehmungstheorie besteht nun darin, dass, wie sich zeigen wird, Helmholtz im *Handbuch* nun nicht mehr den Schlussprozess und den der assoziativen Verknüpfung als *zwei distinkte*, für Wahrnehmung notwendige Prozesse ansieht, sondern sie für *identisch* hält.

Zunächst beschreibt Helmholtz das Deuten der Empfindungen als Zeichen für Objekte einer Außenwelt nicht mehr – wie noch 1855 – selbst als einen Akt des Schließens, sondern es sei lediglich *vergleichbar* mit diesem, und zwar hinsichtlich des Ergebnisses:

Die psychischen Tätigkeiten, durch welche wir zu dem Urteile kommen, daß ein bestimmtes Objekt von bestimmter Beschaffenheit an einem bestimmten Orte außer uns vorhanden sei, sind im allgemeinen nicht bewußte Tätigkeiten, sondern unbewußte. Sie sind in ihrem Resultate einem Schlusse gleich, insofern wir aus der beobachteten Wirkung auf unsere Sinne die Vorstellung von einer Ursache dieser Wirkung gewinnen, während wir in der Tat doch immer nur die Nervenerregungen, also die Wirkungen wahrnehmen können, niemals die äußeren Objekte. (Helmholtz, 1910/2003a, S. 5f.)

Diese Formulierung deutet stärker als die im Vortrag von 1855 verwendeten Beschreibungen darauf hin, dass Helmholtz den Begriff des Schlusses im Kontext seiner Wahrnehmungstheorie eher metaphorisch verstanden wissen möchte. Darüberhinaus beziehe sich der Schlussakt nun nicht mehr allein darauf, dass die Empfindungen als Gegenstände einer Außenwelt gedeutet werden, sondern auch die wahrgenommenen Eigenschaften dieser Gegenstände wie Lage im Raum, Form etc. soll auf einer „einem Schlusse gleichen“ Tätigkeit beruhen. Visuelle Wahrnehmung sei demnach *in Gänze* auf einen

solchen Prozess zurückzuführen. Das Deuten der Empfindungen unterscheide sich dabei von einem üblichen logischen Schluss dadurch, dass es kein bewusster, sondern ein ‘unbewusster Prozess’ sei. Helmholtz verwendet in diesem Zusammenhang nun auch erstmals den Begriff der *unbewussten Schlüsse*: „Indessen mag es erlaubt sein, die psychischen Akte der gewöhnlichen Wahrnehmung als *unbewußte Schlüsse* zu bezeichnen“ (Helmholtz, 1910/2003a, S. 6). Die von Helmholtz an allen wesentlichen Stellen gebrauchte Wendung, dass die logischen Schlüsse und die unbewussten Schlüsse des Wahrnehmenden in ihren Resultaten gleich seien, scheint allerdings verwirrend, da sich doch beide *gerade* hinsichtlich ihrer Resultate zu unterscheiden scheinen. Das Ergebnis eines üblichen logischen Schlusses ist ein sprachliches Gebilde, wohingegen das Ergebnis des unbewussten Schlusses des Wahrnehmenden eine Wahrnehmung sein soll.

In den etwa zur Zeit zwischen der Veröffentlichung der ersten und zweiten Auflage des *Handbuchs* von Helmholtz gehaltenen Vorlesungen „Die neueren Fortschritte in der Theorie des Sehens“ von 1868, die man als konzise Darstellung des *Handbuchs* lesen kann und daher an verschiedenen Stellen dieses Abschnitts zur Erläuterung herangezogen werden sollen, geht Helmholtz näher auf den Unterschied zwischen den üblichen logischen Schlüssen und den unbewussten Schlüssen des Wahrnehmenden ein. Dieser Unterschied sei lediglich ein ‘oberflächlicher’, der darin bestehe, dass sich die logischen Schlüsse sprachlich ausdrücken lassen, die unbewussten Schlüsse hingegen nicht. Der bei beiden involvierte psychische Prozess, „die Art der Thätigkeit selbst“ (Helmholtz, 1896/2002a, S. 361), sei jedoch *derselbe*.

Der Unterschied zwischen den Schlüssen der Logiker und den Inductionsschlüssen, deren Resultat in den durch die Sinnesempfindungen gewonnen Anschauungen der Aussenwelt zu Tage kommt, scheint mir in der That nur ein äusserlicher zu sein, und hauptsächlich darin zu bestehen, dass erstere des Ausdrucks in Worten fähig sind, letztere nicht, weil bei ihnen statt der Worte nur die Empfindungen und Erinnerungsbilder der Empfindungen eintreten. Eben darin, dass diese sich nicht in Worten beschreiben lassen, liegt auch die grosse Schwierigkeit, von diesem ganzen Gebiete der Geistesoperationen überhaupt zu reden. (Helmholtz, 1896/2002a, S. 358)<sup>94</sup>

Dass sich die unbewussten Schlüsse des Wahrnehmenden diesem gleichsam zwingend aufdrängten, selbst gegen besseres Wissen, wie es etwa im Fall der ‘Sinnestäuschungen’ geschehe, sei ebenfalls kein wesentliches Unterscheidungsmerkmal zwischen diesen und den üblichen logischen Schlüssen. Der einzige mit Bewusstheit und Willkür ausgeführte Prozess beim logischen Schlussfolgern (etwa in der Form eines Syllogismus) bestehe

<sup>94</sup>Helmholtz verwendet im *Handbuch* die Begriffe „Induktionsschluss“ und „unbewusster Schluss“ im Wahrnehmungskontext synonym.

laut Helmholtz darin, dass „wir das Material für seine Vordersätze vollständig herbeischaffen“ (Helmholtz, 1896/2002a, S. 361) könnten, also, grob gesagt, in der Lage seien, die einzelnen Prämissen uns selbst sprachlich zu explizieren. Sobald dies geschehen sei, dränge sich uns der Schluss auch beim logischen Schlussfolgern ‘in zwingender Weise’ auf. Die weiteren allgemeinen Ausführungen zum Prozess des Schlussfolgerns lassen deutlich werden, dass Helmholtz hier eine von John Stuart Mill inspirierte Theorie des logischen Schlussfolgerns vertritt, auf die er an verschiedenen Stellen auch explizit verweist (etwa Helmholtz, 1910/2003a, S. 23) und die daher zwecks der Darstellung des Konzepts der unbewussten Schlüsse in aller Kürze erläutert werden soll (für eine ausführlichere Darstellung siehe etwa Skorupski, 1994).

Mill behauptet in seiner 1843 veröffentlichten Schrift „*System of Logic*“, dass die epistemologisch, psychologisch und ontogenetisch fundamentale Art des Schlussfolgerns nicht die Deduktion, also der Schluss vom Allgemeinen auf das Besondere, sondern die *Induktion*, eine einfache Verallgemeinerung von spezifischen empirischen Erfahrungen, sei (Skorupski, 1994).<sup>95</sup> Induktion sei die Tendenz, anhand einer begrenzten Anzahl von Beobachtungen zu verallgemeinern, dass das für die beobachteten Fälle Geltende nicht nur auf diese, sondern auf sämtliche Fälle ‘vergleichbarer Art’ zutrefe und so in einer allgemeinen Aussage zusammengefasst werden könne. Wenn sich etwa eine Person in einer bestimmten Situation bei Kontakt mit einem bestimmten heißen Gegenstand die Finger verbrenne, so würde sie diese Erfahrung dahingehend verallgemeinern oder aufgrund dieser Erfahrung ‘schließen’, dass man sich an *allen* heißen Sachen die Finger verbrenne und diese in Zukunft meiden. Mill versucht in diesem Zusammenhang aufzuzeigen, welche Rolle sogenannte *Allaussagen* in unserem Denken und Erkennen spielen und kommt zu dem Ergebnis, dass diese einen rein *psychologischen* und nicht epistemologischen Wert hätten. Allaussagen seien lediglich ‘psychologisch nützlich’, da sich so die einzelnen Erfahrungen leichter merken und anderen Menschen mitteilen ließen, epistemologisch seien sie durchaus verzichtbar. Im etwa oft angeführten klassischen Beispiel eines Syllogismus:

- (1) „Alle Menschen sind sterblich.“
- (2) „Sokrates ist ein Mensch.“
- (3) „Sokrates ist sterblich.“

setze laut Mill der Obersatz (1) die Wahrheit der Konklusion (3) bereits voraus, denn (1) sei schließlich nur dann wahr, wenn auch (3) wahr wäre. Mill, und Helmholtz schließt sich dem in einer an diesem Beispiel angelehnten Passage im *Handbuch* an, folgert daraus, dass ein solcher Schluss uns „in der Tat nichts Neues, was wir nicht schon gewußt

<sup>95</sup>„All experience begins with individual cases, and proceeds from them to generals“ (Mill, 1882, S. 201).

haben, ehe wir ihn aufgestellt haben“ lehre, sondern vielmehr ein „Herumgehen im Zirkel“ sei (Helmholtz, 1910/2003a, S. 23).<sup>96</sup> In Wahrheit verhalte es sich laut Mill und Helmholtz bei dieser Schlussfolgerung so, dass von Menschen bisher ausnahmslos *beobachtet* wurde, dass kein Mensch über ein gewisses Alter hinaus gelebt habe und diese *empirische Erfahrung* würde nun einfach in dem allgemeinen Satz (1) zusammengefasst und könne so gleichsam zu Bewusstsein gebracht und kommuniziert werden. Die Basis einer solchen allgemeinen Aussage bildeten jedoch die einzelnen diesbezüglichen Erfahrungen. Zu der Überzeugung, dass Sokrates sterben werde, hätten wir auch „unmittelbar, ohne in unserem Bewußtsein den allgemeinen Satz zu bilden“ durch *Verallgemeinerung* des bereits Erfahrenen kommen können, „indem wir seinen Fall mit allen und bekannten früheren verglichen hätten“ (Helmholtz, 1910/2003a, S. 23), also durch ein induktives Vorgehen.<sup>97</sup>

Helmholtz argumentiert im *Handbuch* auf dieser Basis weiter, dass es sich bei sämtlichen Deduktionen, insofern diese als Prämissen Aussagen enthalten, die auf Erfahrung beruhen, eigentlich um ein induktives Vorgehen handle, welches wir uns lediglich über sprachliche Explikation der Prämissen gleichsam bewusst machen würden:

Bei den eigentlich sogenannten und mit Bewußtsein vollzogenen Schlüssen, wenn sie sich nicht auf Gebote, sondern auf Erfahrungssätze stützen, tun wir also in der Tat nichts anderes, als daß wir mit Überlegung und sorgfältiger Prüfung diejenigen Schritte der induktiven Verallgemeinerung unserer Erfahrungen wiederholen, welche schon vorher in schnellerer Weise ohne bewußte Reflexion ausgeführt waren. (Helmholtz, 1910/2003a, S. 24)

In diesem Zusammenhang spricht Helmholtz (etwas unglücklich) nun auch von einem *Induktionsschluss* oder auch *Analogieschluss*. Ein solcher beruhe demnach (ganz im Sinne einer empiristischen Theorie) auf der bisher unzählige Male erfahrenen *empirischen* Verbindung zwischen zwei Sachverhalten. Bei Auftreten eines neuen, bisher nicht beobachteten Falles, finde eine Art ‘Vergleich’ desselben mit den bisher erfahrenen Fällen statt und es werde anschließend einfach diese Assoziation ‘aktiviert’, *ohne bewusste Reflexion*. Ebenso verhalte es sich laut Helmholtz nun auch beim Prozess des unbewussten Schliessens durch den Wahrnehmenden, der der Deutung der Empfindungen als durch Objekte einer Außenwelt verursacht, zugrundeliege. Wir hätten, etwa durch wie-

<sup>96</sup>In den Worten von Mill (1882, S. 238): „No reasoning from generals to particulars can, as such, prove any thing: since from a general principle we cannot infer any particulars, but those which the principle itself assumes as known“.

<sup>97</sup>„All inference is from particulars to particulars. General propositions are merely registers of such inferences already made, and short formulae for making more. The major premise of a syllogism consequently is a formula of this description and the conclusion ist not an inference drawn *from* the formula, but an inference drawn *according* to the formula, the real logical antecedent, or premise, being the particular facts from which the general proposition was collected by induction“ (Mill, 1882, S. 240).

derholtes Ausstrecken des Armes in eine gewisse Richtung, hinreichend oft durch die zugehörigen Empfindungen aus Tast- und Sehsinn erfahren, dass bei Erregung einer gewissen Netzhautstelle ein Objekt in einer bestimmten Richtung lokalisiert sei und aufgrund dieser Erfahrungen habe sich eine assoziative Verknüpfung zwischen diesen bestimmten Empfindungen des Sehens und denen des Tastsinnes ausgebildet. Diese Verknüpfung werde nun bei Auftreten eines ‘hinreichend ähnlichen Falles’ von Empfindungen des Sehens einfach ‘aktiviert’. Dabei handle es sich um einen nicht bewussten, rein assoziativen Prozess ohne aktive Reflexion im üblichen Sinne, den Helmholtz dennoch – mit Verweis auf die Ausführungen Mills – als ein Schlussfolgern bezeichnen möchte:

Wir haben nun genau denselben Fall bei unseren Sinneswahrnehmungen. Wenn wir Erregung in denjenigen Nervenapparaten gefühlt haben, deren periphere Enden an der rechten Seite beider Netzhäute liegen, so haben wir in millionenfach wiederholten Erfahrungen unseres ganzen Lebens gefunden, daß ein leuchtender Gegenstand nach unserer linken Seite hin vor uns lag. Wir mußten die Hand nach links hin erheben, um das Licht zu verdecken, oder das leuchtende Objekt zu ergreifen, oder uns nach links hin bewegen, um uns ihm zu nähern. Wenn also in diesen Fällen kein eigentlich bewußter Schluß vorliegt, so ist doch die wesentliche und ursprüngliche Arbeit eines solchen vollzogen, und das Resultat desselben erreicht, aber freilich nur durch die unbewußten Vorgänge der Assoziation von Vorstellungen, die im dunklen Hintergrunde unseres Gedächtnisses vor sich geht, und deren Resultate sich daher auch unserem Bewußtsein aufdrängen. (Helmholtz, 1910/2003a, S.24)

Dass wir diesen Prozess nun im Unterschied zu den logischen Schlussfolgerungen in Mills Sinn nicht durch Ausformulierung der Prämissen in unser Bewusstsein überführen könnten, sei einzig in einer Unmöglichkeit der sprachlichen Ausdrückbarkeit unserer Empfindungen begründet, da „wir gar nicht näher bezeichnen können, was in uns vorgegangen ist, wenn wir eine Empfindung in einer bestimmten Nervenfasern hatten, und wodurch diese zu unterscheiden ist von entsprechenden Empfindungen in anderen Nervenfasern“ (Helmholtz, 1910/2003a, S. 25).<sup>98</sup> Von diesem, laut Helmholtz oberflächlichen Unterschied abgesehen seien die üblichen logischen Schlussfolgerungen mit den unbewussten Schlüssen des Wahrnehmenden im Prinzip *identisch*.

Helmholtz scheint nun mit dieser Konzeption etwas durchaus Bemerkenswertes zu gelingen. Indem sowohl die sensorische Wahrnehmung als auch das logische Denken oder

<sup>98</sup>Diese Stelle weist erneut auf die Schwierigkeit des Konzepts der ‘visuellen Empfindungen’ hin. Im paradigmatischen Fall einer Empfindung, dem Schmerz, lässt sich relativ gut beschreiben, wodurch sich etwa ein Kopfschmerz von Schmerz in der Hand unterscheidet.

Schlussfolgern von Helmholtz im Wesentlichen auf ein- und denselben psychischen Prozess, nämlich *Assoziation*, zurückgeführt werden soll, hebt er praktisch die seit Kant in der Philosophie bestehende Trennung zwischen Sinnlichkeit und Verstandestätigkeit als *fundamental verschiedener* Vermögen auf. Ein kühner und bewundernswerter Versuch, denn das gesamte mentale Erleben ließe sich so auf einen einzigen Prozess zurückführen, oder, wie Hatfield (1990, S. 204) es prägnant ausdrückt: „Mental life is unified“. Allerdings birgt diese Vorstellung der unbewussten Schlüsse des Wahrnehmenden als (passive) assoziative Verknüpfungen auch neue theoretische Probleme

Zunächst ist festzuhalten, dass Helmholtz in seiner Auffassung der Wahrnehmung als Ergebnis unbewusster Schlüsse den Begriff der Wahrnehmung mehrdeutig verwendet. Einerseits spricht Helmholtz von Wahrnehmung im Sinne des *Sehens* von etwas, andererseits spricht er, gerade im Kontext der Erläuterung seiner Schlusstheorie, von Wahrnehmung im Sinne von *Wahrnehmungsurteilen*, etwa *dass* sich irgendein Objekt oder Licht links oder rechts neben uns befinde. Die Aussage, dass das *Sehen von etwas* selbst das Ergebnis einer Schlussfolgerung sei, ist jedoch, wie bereits erwähnt, aus einer sprachanalytischen Perspektive unverständlich. Das Ergebnis einer Schlussfolgerung ist stets, *dass* dieses oder jenes der Fall ist; etwa *dass* Sokrates sterblich ist oder *dass* „ein leuchtender Gegenstand nach unserer linken Seite hin vor uns lag“. Eine Wahrnehmung im Sinne des Sehens von etwas kann jedoch ebensowenig das Ergebnis einer Schlussfolgerung sein wie das Werfen eines Balls oder das Gewinnen eines Wettlaufs (Hacker, 1995). Bewegt sich jemand zielstrebig auf etwas zu, was wir vorher versteckt haben, so hat es zwar Sinn zu sagen, *dass* er das von uns Versteckte wohl gesehen hat und es hat ebenfalls Sinn, dies als eine Schlussfolgerung, die auf dem von der Person gezeigten Verhalten beruht, anzusehen. Welchen Sinn es jedoch haben soll, das Sehen des von uns Versteckten durch die Person als das Ergebnis einer Schlussfolgerung anzusehen, ist unklar. Bestenfalls könnte man aus dieser Perspektive Helmholtz zugestehen, dass seine Schlusstheorie eine Theorie über das Zustandekommen von *Wahrnehmungsurteilen* darstellt.<sup>99</sup> Helmholtz scheint aber das Sehen selbst erklären zu wollen. Möchte man auch hier Helmholtz wohlwollend begegnen und seine Ausführungen bezüglich der Wahrnehmung als Ergebnis einer Schlussfolgerung als rein metaphorisch auffas-

---

<sup>99</sup>Hacker (1995, S. 208) bemerkt aus dieser sprachanalytischen Perspektive, dass selbst bezogen auf Wahrnehmungsurteile Helmholtz' Ansatz fehlerhaft sei, da er „not one confusion but a whole skein of misunderstandings“ enthalte. Eine Schlussfolgerung sei kein 'mentaler Prozess', weder ein bewusster, noch ein unbewusster, sondern eine *logische Operation*, so wie etwa die Division eine mathematische Operation sei. Konzepte wie 'Prozess', 'Aktivität' oder 'Zustand', die einer *physikalischen* Beschreibungsebene entstammen, ließen sich nicht ohne Weiteres auf den Bereich des Mentalen anwenden; „a mental state is not a physical one save for being mental“. Zudem könnten die von Helmholtz als Prämissen angesehenen Empfindungen überhaupt keine solchen sein, da sie laut Helmholtz' Angabe einfach gegeben und daher weder wahr noch falsch seien. Für eine ausführliche Diskussion dieser und weiterer Argumente siehe Hacker (1995).

sen, stellt sich jedoch die Frage, wie diese Metaphorik dazu beitragen kann, das Sehen verständlicher zu machen.

Selbst wenn man von diesem Punkt absieht, scheint Helmholtz' Ansatz nicht konsistent zu sein. So ist etwa unklar, was Helmholtz glaubt, wir „in millionenfach wiederholten Erfahrungen unseres ganzen Lebens gefunden“ zu haben. In seinem Beispiel (s.o.) bilde sich eine assoziative Verbindung zwischen den Erregungen bestimmter Nerven, die mit bestimmten Netzhautstellen verbunden seien, und bestimmten Empfindungen des Tastsinnes. Beide stellen in seiner Konzeption gleichsam 'Prämissen' für die durchzuführende Schlussfolgerung dar, die zum Sehen externer Objekte führe. Im selben Abschnitt erläutert er dies aber dann dadurch, dass wir „die Hand nach links hin erheben, um das Licht zu verdecken, oder das leuchtende Objekt zu ergreifen, oder uns nach links hin bewegen, um uns ihm zu nähern“. Erregungen bestimmter Nerven sind jedoch weder Licht, geschweige denn ein leuchtendes 'Objekt'. Hier zeigt sich erneut die bereits mehrfach angesprochene Problematik des Empfindungsbegriffes bei Helmholtz. Erregungen bestimmter Nerven lassen sich physiologisch feststellen und sind vermutlich eine *Bedingung* dafür, dass wir sehen, nur kann der Wahrnehmende diese Nervenerregungen selbst nicht empfinden. Der Wahrnehmende 'empfindet' aber auch kein Licht und erst recht kein leuchtendes *Objekt*, dem er sich nähern kann. Bestenfalls ließe sich sagen, dass der Wahrnehmende Licht oder leuchtende Objekte *sieht*, aber in dieser Beschreibung ist schon das enthalten, was Helmholtz mit seiner Theorie erklären möchte, nämlich Wahrnehmung. Besonders deutlich wird dies in der von Helmholtz gebrauchten Wendung, dass man sich dem 'empfundenen leuchtenden Objekt' *nähern* könne, was bei aller sprachlichen Unschärfe doch offensichtlich voraussetzt, dass das Objekt vom Wahrnehmenden bereits als *außerhalb* von ihm aufgefasst werden muss, was in Helmholtz' Konzeption doch aber erst das Ergebnis der Schlussfolgerung sei. Wie an anderer Stelle bereits bemerkt, wird zudem nicht verständlich, auf welche Weise es durch bloße assoziative Verknüpfung von Empfindungen des Sehens und Empfindungen des Tastsinn zum *Sehen* eines Objektes kommen soll.

Darüberhinaus scheint nun das Konzept der *Ähnlichkeit* oder der *Regularität* für Helmholtz' Theorie eine entscheidende Rolle zu spielen. Ein unbewusster Schluss, etwa von einer bestimmten Erregung einer Netzhautstelle auf einen sich in einer bestimmten Richtung außerhalb des Wahrnehmenden befindenden Gegenstand, soll bei Vorliegen *hinreichender Ähnlichkeit* gleichsam automatisiert vollzogen werden. Doch was kann hinreichende Ähnlichkeit im Falle 'visueller Empfindungen' bedeuten? Versetzt man sich, Helmholtz' Theorie folgend, in die Lage eines 'noch lernenden Wahrnehmenden', so erscheint das Erlernen der Wahrnehmung ohne weitere Voraussetzungen extrem schwierig zu sein. Da Helmholtz 'visuelle Empfindungen' nicht weiter charakterisiert oder sie als bloße Nervenerregungen auffasst, bleibt die Frage, wie, durch wen und

nach welchen ‘Kriterien’ eine Empfindung als zu einer anderen Empfindung ‘ähnlich’ klassifiziert wird. Der von Helmholtz postulierte Prozess scheint unter anderem das Ausbilden einer Art ‘Regel’ seitens des Wahrnehmenden (oder des Bewusstseins) zu erfordern. Im Zusammenhang mit dem von Helmholtz zumeist angeführten Beispiel der Lokalisation von Objekten über assoziative Verbindungen zwischen Empfindungen aus Seh- und Tastsinn stellt sich zudem die Frage, warum es gerade zu einer assoziativen Verknüpfung *dieser* beiden ‘Empfindungstypen’ kommt, da sich der Wahrnehmende in Helmholtz’ Konzeption normalerweise in einem Strom von Empfindungen unterschiedlichster Qualität aus den unterschiedlichsten Sinnen befinde. Helmholtz geht in Teilen auf diese Punkte näher ein, indem er in §26 des *Handbuchs* Wahrnehmen als dem Experimentieren gleich beschreibt.

### Wahrnehmen als Experimentieren

Das Zustandekommen einer solchen, für einen Schluss per Ähnlichkeit notwendigen, assoziativen Verknüpfung zwischen Empfindungen des Tast- und denen des Sehsinnes geschehe laut Helmholtz nicht lediglich auf passive Weise, sondern beruhe auf einer *aktiven Tätigkeit* des Wahrnehmenden selbst, die mit der naturwissenschaftlichen Methode des kontrollierten Experimentes zu vergleichen sei:

Die Deutung unserer Sinnesempfindungen beruht auf dem Experiment und nicht auf blosser Beobachtung äusseren Geschehens. Das Experiment lehrt uns, dass die Verbindung zwischen zwei Vorgängen in jedem von uns gewählten beliebigen Augenblicke bestehe, unter übrigens von uns beliebig abgeänderten Verhältnissen. Die Zusammengehörigkeit der beiden Vorgänge bewährt sich dabei als unmittelbar constant in der Zeit, da wir sie in jedem beliebigen Augenblicke prüfen können. Blosser Beobachtung gewährt uns kaum je dieselbe Sicherheit der Kenntniss, trotz noch so häufiger Wiederholung unter vielfach veränderten Umständen. Denn sie lehrt uns wohl, dass die Vorgänge, um deren Zusammengehörigkeit es sich handelt, oft, oder bisher immer, zusammen eingetreten sind, nicht aber, dass sie zu jeder beliebigen von uns gewählten Zeit eintreten. (Helmholtz, 1896/2002a, S. 355)

Überträgt man diese allgemeine Beschreibung auf den Fall der Wahrnehmung, so lernen wir laut Helmholtz die Empfindungen als Zeichen für Gegenstände einer Außenwelt zu deuten, indem wir zunächst über *absichtliche* und *willentlich durchgeführte* Handlungen überprüfen, ob bestimmte Bewegungen unseres Körpers – etwa das Ausstrecken eines Armes oder bestimmte Bewegungen der Augen – und die damit einhergehenden Empfindungen konstant mit bestimmten Empfindungen des Sehsinnes einhergingen.

Anhand wiederholter ‘Prüfung’ dieses durch unser Handeln herbeigeführten gemeinsamen Auftretens der bestimmten Empfindungen aus Tast- und Sehsinn lernten wir die *Regelmäßigkeit* dieses gemeinsamen Auftretens:

Aber wir lernen sie [die zu deutenden Zeichen] lesen, indem wir sie mit dem Erfolge unserer Bewegungen und den Veränderungen, die wir selbst durch diese in der Aussenwelt hervorbringen, vergleichen. (Helmholtz, 1896/2002a, S. 354)

Von der größten Wichtigkeit endlich für die Festigkeit unserer Überzeugung von der Richtigkeit unserer sinnlichen Wahrnehmung sind die Prüfungen, welche wir mittels der willkürlichen Bewegungen unseres Körpers anstellen. (Helmholtz, 1910/2003a, S. 26)

Auch die Wahrnehmung räumlicher Verhältnisse werde auf diese Weise erlernt. Der Wahrnehmende generiere gleichsam ‘Hypothesen’, welche ‘visuellen Empfindungen’ sich bei welchen willentlich ausgeführten Bewegungen einstellten, die dann *in actu* getestet würden. Im Falle der Augenbewegungen verhalte es sich beispielsweise so, dass eine willentlich hervorgebrachte Bewegung der Augen eine Veränderung der mit den motorischen Nerven assoziierten Empfindungen *und gleichzeitig* eine Veränderung der mit den Sehnerven assoziierten ‘visuellen Empfindungen’ nach sich ziehe.<sup>100</sup> Die Aufgabe des Wahrnehmenden bestehe nun darin, eine *Trennung* zwischen jenen Empfindungen des Sehens vorzunehmen, die allein auf die willentlichen Augenbewegungen zurückgingen, und jenen, die allein auf Veränderungen der (angenommenen) Objekte der Außenwelt beruhten, um eine allgemeine Regelmäßigkeit feststellen zu können. Diese Trennung werde dadurch möglich, dass dem Wahrnehmenden die willentlich ausgeführten Bewegungen mitsamt den daraus resultierenden Empfindungen durch das oben dargestellte ‘Experimentieren’ bewusst und zugänglich seien. Der Wahrnehmende werde so in die Lage versetzt, bei auftretenden Empfindungsveränderungen jenen Teil der Veränderungen, der auf seine willentlichen Bewegungen zurückgehe, von dem Teil, der auf Veränderungen der Gegenstände der Außenwelt zurückgehe, voneinander zu scheiden.

Durch diese Konzeption des Wahrnehmenden als eines gleichsam nach naturwissenschaftlichen Methoden Vorgehenden kommt es zu einer deutlichen Anreicherung der

---

<sup>100</sup>Helmholtz spricht bezüglich der mit willentlich herbeigeführten Augenbewegungen einhergehenden Empfindungen auch davon, dass wir „den *Grad der Innervation*, die wir den Augenmuskeln zufließen lassen“ (Helmholtz, 1910/2003a, S. 433) fühlten. Der empfundene Grad der Innervation sei dabei von einer bloß passiv empfundenen Kontraktion der Augenmuskeln zu unterscheiden, da er aus der „Intention unseres Willens“ bestehe, die sich bei willkürlichen Bewegungen „immer nur auf die Erreichung eines direkt und deutlich wahrnehmbaren äußeren Erfolges bezieht“ (Helmholtz, 1910/2003a, S. 54).

theoretischen Ausrichtung von Helmholtz' Zeichentheorie, die nun einen stark *handlungsorientierten* Zug aufweist:

Unsere Vorstellungen von den Dingen können gar nichts anderes sein, als Symbole, natürlich gegebene Zeichen für die Dinge, welche wir zur *Regelung unserer Bewegungen und Handlungen* benutzen lernen. Wenn wir jene Symbole richtig zu lesen gelernt haben, so sind wir imstande, mit ihrer Hilfe *unsere Handlungen* so einzurichten, daß dieselben den gewünschten Erfolg haben, d.h. *daß die erwarteten neuen Sinnesempfindungen eintreten* [sämtliche Hervorhebungen von mir]. (Helmholtz, 1910/2003a, S. 18)

Dem willentlichen, intentionalen und gleichsam 'experimentellen' Herbeiführen der Empfindungsveränderungen und dem Erlernen einer diesbezüglichen Regelmäßigkeit durch den Wahrnehmenden kommt der Erklärung der Wahrnehmung aus empiristischer Sicht offenbar eine enorm große, ja geradezu konstituierende Bedeutung zu, da aus dieser Annahme eines *aktiv tätigen* Wahrnehmenden offenbar sämtliche Wahrnehmungsleistungen theoretisch abgeleitet werden sollen.<sup>101</sup>

An dieser Stelle unterstreicht Helmholtz die fundamentale Wichtigkeit des Kausalgesetzes, das nun auch eine andere Interpretation erfährt. Im Rahmen der 'Induktionsschlüsse' soll das Kausalgesetz begründen, dass die Empfindungen bzw. Empfindungsveränderungen vom Wahrnehmenden als *eindeutig bestimmt* durch vorangehende Ursachen (nämlich u.a. durch von ihm willentlich hervorgebrachte Bewegungen) aufgefasst werden und so über diese durch das Kausalgesetz garantierte Gesetzmäßigkeit ein induktives Vorgehen überhaupt erst möglich wird. Ohne die zugrundeliegende Annahme des Wahrnehmenden, dass jede Wirkung eine Ursache habe, sei laut Helmholtz der dargestellte Prozess des hypothesengeleiteten Experimentierens, der sowohl für die Wahrnehmung als auch für die Naturwissenschaften selbst konstituierend sei, überhaupt nicht denkbar:

Der eigentliche letzte Grund, durch welchen alle unsere bewußt vollzogenen Induktionen überzeugende Kraft erhalten, ist das Kausalgesetz. ... Solange wir nun aber auf bloße Beobachtung ... beschränkt sind, welche ohne unser Zutun von selbst eintreten, ohne Experimente anstellen zu können, bei denen wir den Komplex der Ursachen verändern, gewinnen wir schwer die Überzeugung, dass wir alle Bedingungen, welche auf den Erfolg Einfluß haben können, wirklich schon ermittelt haben. (Helmholtz, 1896/2002a, S. 26f.)

<sup>101</sup> „Zunächst erfahren wir unmittelbar nur, dass wir durch die Willensimpulse Veränderungen hervorbringen, die wir durch den Tastsinn und Gesichtssinn wahrnehmen“ (Helmholtz, 1896/2002a, S. 356). „Das Wesentliche bei diesem Verfahren ist eben das Prinzip des Experimentierens. Wir verändern einen Teil der Bedingungen, unter denen das Objekt wahrgenommen wird, aus eigenem Antrieb und eigener Machtvollkommenheit“ (Helmholtz, 1910/2003a, S. 28).

Wenn – so könnte man Helmholtz' Vorstellungen paraphrasieren – der aus empiristischer Sicht Wahrnehmung Lernende oder der Naturwissenschaftler nicht davon ausgehen würde, dass ein Phänomen *immer* eine Ursache haben muss und Phänomen und Ursache durch eine allgemeine Regel miteinander verbunden sind, so wäre das Aufstellen und Überprüfen von Gesetz- oder Regelmäßigkeit über experimentelles Vorgehen von vornherein zum Scheitern verurteilt. Im Wahrnehmungskontext würde das bedeuten, dass – im Beispiel der Augenbewegungen – die laut Helmholtz zu vollziehende Trennung von Empfindungsveränderungen, die auf Aktivitäten des Wahrnehmenden zurückgingen, und Empfindungsveränderungen, die auf Veränderungen der Außenwelt zurückgingen, überhaupt nicht möglich wäre, wenn der Wahrnehmende nicht davon ausginge, dass er durch willentliche Bewegung seiner Augen *stets* Veränderungen der Empfindungen des Sehsinnes herbeiführen könnte.

Das Kausalgesetz, und hier zeichnet sich nun eine merkbare Abweichung Helmholtzens vom Kantischen *a priori* ab, müsse vom Naturwissenschaftler wie vom Wahrnehmenden *vorausgesetzt* werden; es sei „nichts anderes als die Forderung, alles begreifen zu wollen“ (Helmholtz, 1910/2003a, S. 31). Neben dieser das experimentelle Vorgehen überhaupt rechtfertigenden Funktion des Kausalgesetzes hat es in der im *Handbuch* dargestellten Theorie eine weitere Funktion; es soll, wie im Vortrag von 1855, ebenfalls rechtfertigen, dass der Wahrnehmende die als nicht von seinen willentlich herbeigeführten Bewegungen stammend identifizierten Empfindungen als *Veränderungen einer unabhängigen Außenwelt* deutet. Da die mit dieser Argumentation verbundenen Probleme, insbesondere die Frage nach dem Status des Kausalgesetzes, bereits im Abschnitt über den Vortrag von 1855 besprochen wurden, soll an dieser Stelle auf eine Wiederholung verzichtet werden. Festzuhalten ist jedoch, dass im *Handbuch* das Kausalgesetz innerhalb der Helmholtzschen Theorie nun *zwei* verschiedene Funktionen hat.

Es ist durchaus denkbar, dass diese Konzeption des Entstehens assoziativer Verknüpfungen zwischen Empfindungen des Tast- und Sehsinns über hypothesengeleitetes Experimentieren nach naturwissenschaftlichem Vorbild in der Lage wäre, einige der im letzten Abschnitt dargestellten Probleme (zumindest teilweise) zu vermeiden. Allerdings verschiebt sich dadurch auch die Ausrichtung der Helmholtzschen Schlusstheorie auf deutliche Weise. Ein als hypothesengeleitetes Testen oder Experimentieren verstandenes Handeln der Art, wie Helmholtz es beschreibt, ist nun selbst nicht mehr über bloße Assoziation als *den* vereinheitlichenden psychischen Prozess erklärbar, den Helmholtz sowohl bezüglich der sinnlichen Wahrnehmung als auch bezüglich des logischen Denkens so stark gemacht hatte.<sup>102</sup> Für eine solche Konzeption müssten nun weitaus stärkere Annahmen über den Wahrnehmenden gemacht werden; so müsste er etwa die

<sup>102</sup>„Die *einzig*e psychische Tätigkeit, die dazu gefordert wird, ist die gesetzmäßig wiederkehrende Assoziation zweier Vorstellungen, die schon oft miteinander verbunden gewesen sind“ (Helmholtz, 1910/2003a, S. 434).

Fähigkeit haben, intentional zu handeln, Erwartungen über räumliche Verhältnisse zu generieren und diese Erwartungen mit den tatsächlich aufgetretenen zu vergleichen.

An einer Stelle im Vortrag von 1868 (siehe aber auch Helmholtz, 1910/2003a, S. 29) unternimmt Helmholtz den Versuch einer genaueren Darstellung, wie es aus empiristischer Sicht zu einem Erlernen der Trennung der Empfindungen komme, die einerseits auf eigene willentliche Bewegungen seitens des Wahrnehmenden und andererseits auf Veränderungen in der Außenwelt zurückgingen:

Zunächst erfahren wir unmittelbar nur, dass wir durch die Willensimpulse Veränderungen hervorbringen, die wir durch den Tastsinn und Gesichtssinn wahrnehmen. Die meisten dieser Aenderungen, die wir willkürlich hervorbringen, sind nur Raumänderungen, d.h. Bewegungen; es können freilich auch andere, Aenderungen, die an den Dingen selbst, bewirkt werden. Können wir nun die Bewegungen unserer Hände und Augen als Raumänderungen erkennen, ohne dies vorher zu wissen, und von anderen Aenderungen, welche die Eigenschaften der Dinge betreffen, unterscheiden? Ich glaube, ja! ... Eine Augenbewegung, die eine Verschiebung des Netzhautbildes auf der Netzhaut hervorbringt, bringt bei gleicher Wiederholung dieselbe Reihe von Veränderungen hervor, welches auch der Inhalt des Gesichtsfeldes sein mag; sie bewirkt, dass die Eindrücke, welche bisher die Localzeichen  $a_0, a_1, a_2, a_3$  hatten, die neuen Localzeichen  $b_0, b_1, b_2, b_3$  bekommen; und dies kann stets in gleicher Weise geschehen, welches auch die Qualitäten dieser Eindrücke sein mögen. Dadurch sind diese Veränderungen charakterisiert von der eigenthümlichen Art, welche wir eben Raumveränderungen nennen. (Helmholtz, 1896/2002a, S. 356)

Helmholtz spricht an dieser Stelle über den Umstand, dass Veränderungen des retinalen Lichtmusters verschiedene, sich nicht gegenseitig ausschließende Ursachen haben können – vorausgesetzt, dass es eine unabhängig vom Wahrnehmenden existierende Außenwelt mit Licht reflektierenden Objekten gibt. Entweder der Wahrnehmende verändert seine Sehrichtung relativ zu den Objekten (das was Helmholtz „Raumänderungen“ nennt) oder die Objekte verändern ihre Lage relativ zum Wahrnehmenden, oder beides. Helmholtz' Auffassung ist es nun offenbar, dass der Wahrnehmende eine Unterscheidung zwischen diesen beiden Ursachen der Veränderung allein durch ein Experimentieren der oben dargestellten Art erlernen könne, wobei er explizit auf das Konzept der Lokalzeichen von Lotze zurückgreift.

Der Wahrnehmende könne mit Hilfe dieser Lokalzeichen lernen zu unterscheiden, ob Empfindungsveränderungen auf Bewegungen seines eigenen Körpers oder auf Veränderungen in der Außenwelt zurückgingen, da sich bei einer bestimmten Augenbewegung stets dieselbe Veränderung der Lokalzeichen einstelle, gleichgültig, welches Objekt der

Wahrnehmende betrachte („welches auch der Inhalt des Gesichtsfeldes sein mag“). Bei einer bestimmten Augenbewegung verschiebe sich das ‘Lokalzeichenmuster’ stets auf die gleiche Weise und der Wahrnehmende könne durch wiederholtes Ausführen derselben Augenbewegung unter völlig verschiedenen Sehbedingungen dieses stets gleich bleibende Muster der Verschiebung erfahren und gewissermassen ‘herausfiltern’ und so als allein auf die von ihm initiierte Bewegung zurückführen. Würden nun bei Ausführung dieser bestimmten Augenbewegung andere Lokalzeichenveränderungen auftreten, dann könnten diese nicht auf die vom Wahrnehmenden willentlich hervorgebrachten Bewegungen zurückgehen, sondern müssten eine andere Ursache haben.

Helmholtz erläutert anschließend ausführlicher, wie es aufgrund dieser vom Wahrnehmenden zu erlernenden Unterscheidung zwischen den beiden möglichen Ursachen der Empfindungsveränderungen zur Wahrnehmung – oder besser Setzung – von unabhängig vom Wahrnehmenden existierenden Objekten komme. All jene Empfindungsveränderungen, die nicht auf die Bewegungen des Wahrnehmenden zurückführbar seien, müssten laut Helmholtz unter Voraussetzung der Gültigkeit des Kausalgesetzes ebenfalls eine Ursache haben. Da sie nun aber Veränderungen von Empfindungen, d.h. Veränderungen von Nervenprozessen seien, könnten sie laut Helmholtz selbst keine unabhängig vom Wahrnehmenden bestehenden Eigenschaften externer Gegenstände sein. Die Zeichentheorie besagt doch gerade, dass sämtliche

Anschauungen und Vorstellungen *Wirkungen* [sind], welche die angeschauten und vorgestellten Objekte auf unser Nervensystem und unser Bewusstsein hervorgebracht haben. Jede Wirkung hängt ihrer Natur nach ganz notwendig ab, sowohl von der Natur des Wirkenden, als von der desjenigen, auf welches gewirkt wird. (Helmholtz, 1910/2003a, S. 18)

Helmholtz argumentiert nun mit eher vagen Begriffen weiter, dass dem Bewusstsein eine gewisse ‘Tendenz’ innewohne, seine ‘Aufmerksamkeit’ einseitig auf die Beobachtung jener *nicht* konstanten, nicht auf die vom Wahrnehmenden initiierten Bewegungen zurückführbaren Empfindungsveränderungen ‘auszurichten’ und so von der eigentlichen Natur dieser Empfindungen als *Wirkungen* von Objekten auf die Sinnesorgane ‘abzusehen’ und sie stattdessen als *objektive Eigenschaften* von Gegenständen einer Außenwelt zu deuten, wie Helmholtz in einer lyrisch anmutenden Passage ausführt:

Das natürliche Bewußtsein, welches ganz im Interesse der Beobachtung der Außenwelt aufgeht, und wenig Veranlassung hat, seine Aufmerksamkeit dem neben dem bunten Wechsel der äußeren Objekte immer unverändert erscheinenden Ich zuzuwenden, pflegt nicht zu beachten, daß die *Eigenschaften* der betrachteten und betasteten Objekte Wirkungen derselben teils auf andere Naturkörper, hauptsächlich aber auf unsere Sinne sind. Indem nun

so ganz abgesehen wird von unserem Nervensystem und unserem Empfindungsvermögen, als dem gleichbleibenden Reagens, auf welches die Wirkung ausgeübt wird, und die Verschiedenheit der Wirkung nur als Verschiedenheit des Objekts, von dem sie ausgeht, beachtet wird, kann die Wirkung auch nicht mehr als Wirkung anerkannt werden (denn jede Wirkung muß Wirkung auf etwas anderes sein), sondern sie wird als Eigenschaft des Körpers objektiv hingestellt, und nur als ihm angehörig betrachtet. (Helmholtz, 1910/2003a, S. 29)

Es ist offensichtlich, dass Helmholtz hier mehrere, im Rahmen seiner empiristischen Theorie nicht unkritische Annahmen machen muss. Erstens ist nicht klar, wie der Wahrnehmende diese Fähigkeit des ‘Herausfilterns’ der beiden möglichen Ursachen der Empfindungsveränderungen erlernt haben soll. Der Verweis auf bloße Assoziation ist hier keinesfalls ausreichend, darüber hinausgehende Annahmen über bestimmte diesbezügliche ‘kognitive Fähigkeiten’ scheinen hier erforderlich. Zweitens setzt diese Konzeption eine unabhängig vom wahrnehmenden existierende Aussenwelt schon voraus, in der es sich nicht ständig verändernde Gegenstände geben muss, die sich zudem nicht in stetiger Bewegung befinden dürfen, damit die behauptete invariante Verschiebung der Lokalzeichenmuster stattfinden und vom Wahrnehmenden erfahren werden kann. Drittens ist auch hier wieder fraglich, warum es keine andere Möglichkeit für das ‘natürliche Bewusstsein’ geben soll, als die Empfindungen als externe Gegenstände zu deuten.

### Die Hilfsmittel zur Wahrnehmung von Tiefe

Eine zusätzliche Schwierigkeit für das Verständnis der im *Handbuch* ausgeführten Theorie ergibt sich dadurch, dass an mehreren Stellen die Konzeption einer assoziativen Verknüpfung zwischen Empfindungen des Tast- und Sehsinnes als einer notwendigen Bedingung für das Erlernen räumlicher Wahrnehmung aufgegeben wird, indem Helmholtz diese auch ohne jegliche Beteiligung von Empfindungen aus dem Tastsinn zu erklären sucht, nämlich über die von ihm so genannten Hilfsmittel der Tiefenwahrnehmung (bzw. über das Konzept der Tiefencues), wobei Helmholtz der binokularen Disparität eine zentrale Rolle einräumt:

Was die Abbildung der Raumverhältnisse betrifft, so geschieht eine solche allerdings an den peripherischen Nervenenden im Auge ... in einem gewissen Grade, aber doch nur *in beschränkter Weise*, da das Auge nur perspektivische Flächenabbildungen gibt. Ein direktes Bild einer nach drei Dimensionen ausgedehnten Raumgröße gibt weder das Auge noch die Hand. Erst

durch die Vergleichung der Bilder beider Augen ... kommt die Vorstellung von Körpern zustande. (Helmholtz, 1910/2003a, S. 21f)

Die Wahrnehmung räumlich ausgedehnter *Körper* komme, so könnte man die obige Passage lesen, *allein* durch eine „Vergleichung der Bilder beider Augen“, die für sich genommen nur „perspektivische Flächenabbildungen“ seien, zustande; eine Formulierung, die uns bereits im Vortrag von 1855 begegnet ist. Es lässt nicht mit Sicherheit entscheiden, ob Helmholtz der Ansicht ist, dass der Wahrnehmende die ‘Flächenabbildungen’ zunächst selbst wahrnehme und dann über diese das räumlich ausgedehnte Objekt gleichsam ‘konstruiere’, Hatfield (1990) etwa spricht sich gegen eine solche, zweistufige Interpretation aus. Helmholtz spricht an verschiedenen Stellen vom „flächenhaften Gesichtsfeld“ (Helmholtz, 1910/2003a, S. 330) oder der „Fläche des Sehfeldes“ (Helmholtz, 1910/2003a, S. 235), schränkt dann aber ein, dass uns die Gegenstände nicht erscheinen als „an oder auf einer Fläche verteilt, sondern nur *wie an einer Fläche*, in flächenartiger Anordnung, in einer nach zwei Dimensionen unterschiedenen Anordnung“ (Helmholtz, 1910/2003a, S. 132), wobei der angesprochene Unterschied zwischen „an einer Fläche“ und „wie an einer Fläche“ verteilt keineswegs deutlich wird. Es scheint jedoch notwendig zu sein, dass diese ‘Flächenabbildungen’ irgendeiner Instanz zugänglich sein *müssen*, ansonsten wäre die behauptete „Vergleichung der Bilder beider Augen“ vollends unverständlich. Gleichzeitig müsste diese Instanz über das Wissen verfügen, dass es sich bei den beiden ‘Flächenabbildungen’ um Abbildungen *derselben* Szene mit *denselben* Objekten handelt. Wir wollen uns nun genauer den Hilfsmitteln der Tiefenwahrnehmung im *Handbuch* zuwenden.

In dem oft auch als ‘technischer Teil’ bezeichneten Abschnitt des *Handbuchs*, in denen Helmholtz verschiedene Einzelthemen wie Augenbewegungen abhandelt und jeweils zugehörige empirische Befunde vorstellt, wird in §30 auch das Thema der Wahrnehmung von Tiefe näher ausgeführt. Wie schon im Vortrag von 1855 unterstreicht Helmholtz auch hier, dass eine Kenntnis der „Tiefendimension“ *nur* über bestimmte Hilfsmittel (gelegentlich auch als „Momente“ oder „Motive“ der Tiefenwahrnehmung bezeichnet) möglich sei. Helmholtz teilt diese Hilfsmittel in zwei verschiedene Klassen ein, nämlich Hilfsmittel, die auf „Erfahrung über die besondere Natur der gesehenen Objekte“ (Helmholtz, 1910/2003a, S. 236) beruhen und Hilfsmittel, die, unabhängig von Erfahrungen, auf der Anwendung ‘allgemeiner Regelmäßigkeiten’ bezüglich bestimmter Empfindungen beruhen. Helmholtz selbst kennzeichnet diese beiden Klassen nicht mittels eigener Termini, die von ihm vorgenommene Unterteilung fand allerdings ihren Niederschlag in der etwa bis zur Mitte des 20. Jahrhunderts in der Literatur vorherr-

schenden begrifflichen Unterscheidung zwischen sogenannten *primären* und *sekundären* Tiefenkriterien (siehe Boring, 1942).<sup>103</sup>

Zu jenen Hilfsmitteln, die allein auf der Anwendung allgemeiner Regelmäßigkeiten hinsichtlich bestimmter Empfindungen beruhen (die primären Tiefenkriterien) zählt Helmholtz die Akkomodation, die Konvergenz, die binokulare Disparität und die Bewegungsparallaxe. Der Begriff der Bewegungsparallaxe bezieht sich darauf, dass bei Bewegung des Wahrnehmenden unterschiedlich weit entfernte Objekte als unterschiedlich schnell am Wahrnehmenden ‘vorbeigleitend’ erscheinen, eine Beobachtung, die sich etwa leicht in einem fahrenden Zug sitzend machen lässt. Zur zweiten Klasse (den sekundären Tiefenkriterien) zählt Helmholtz all jene Hilfsmittel, die bei monokularer Betrachtung ohne Bewegung des Kopfes bzw. der Augen dennoch helfen würden, räumliche Relationen von Objekten wahrzunehmen, die so weit entfernt sind, dass Akkommodationsunterschiede des Auges für den Wahrnehmenden nicht mehr merkbar sind. Dazu zählt Helmholtz die Kenntnis über die Größe und Form der Objekte in Verbindung mit der ‘erscheinenden Größe’, die atmosphärische Trübung (*aerial perspective*), die perspektivische Darstellung in der Kunst und die Verdeckung (Helmholtz, 1910/2003a, S. 236). Beide Klassen an Hilfsmitteln unterschieden sich darüber hinaus hinsichtlich ihrer ‘Güte’ als Hilfsmittel zur Tiefenwahrnehmung, eine Unterscheidung die uns schon des Öfteren begegnet ist. Die sekundären Tiefenkriterien gäben „nur *Vorstellungen des Abstandes*“, wohingegen die primären Hilfsmittel eine „wirkliche *Wahrnehmung des Abstandes*“ erlaubten (Helmholtz, 1910/2003a, S. 236). Wie ist dieser Unterschied in Helmholtz’ theoretischem Ansatz zu verstehen und auf welche Weise werden die Hilfsmittel zur Tiefenwahrnehmung laut Helmholtz genutzt?

Folgt man dem von Helmholtz eingeführten Begriffsapparat und der Logik seiner Zeichentheorie, ist man etwas ratlos, was dieser von Helmholtz als wesentlich herausgestellte Unterschied zwischen einer ‘Vorstellung des Abstandes’ und einer ‘wirklichen Wahrnehmung’ desselben bedeuten mag. Falls Helmholtz den Begriff der ‘wirklichen Wahrnehmung’ im Sinne einer ‘Übereinstimmung’ zwischen wahrgenommener Entfernung und tatsächlicher Entfernung verstanden wissen möchte, scheint diese Auffassung mit der generellen Ausrichtung seiner Theorie nicht kompatibel. Die Zeichentheorie besagt doch, dass es sich bei Wahrnehmungen um *Wirkungen*, welche sowohl von dem, was auf unsere Sinne wirkt, als auch von den Sinnesorganen selbst, auf die gewirkt wird, abhängig sind, so dass das Prädikat „wirklich“ im Bezug auf die wahrgenommene Entfernung eines Objektes keine solche Bedeutung tragen kann.<sup>104</sup> Zudem besteht

---

<sup>103</sup>Wie wir in Abschnitt 3.1 dieser Arbeit sehen werden, konnte sich diese Unterteilung der Tiefencues in der aktuellen wahrnehmungspsychologischen Literatur nicht durchsetzen.

<sup>104</sup>„Daraus geht nun hervor, daß in Wahrheit die *Eigenschaften* der Naturobjekte, trotz dieses Namens, gar nichts dem einzelnen Objekte an und für sich eigens zukommen, sondern immer eine Beziehung zu einem zweiten Objekte (einschließlich unserer Sinnesorgane) bezeichnen“ (Helmholtz, 1910/2003a, S. 20).

die Frage, ob Helmholtz an dieser Stelle durch die Art der Formulierung zusätzlich einen Unterschied zwischen den verschiedenen Klassen der Hilfsmittel deutlich machen möchte, der sich darauf bezieht, dass mit Hilfe der primären Tiefenkriterien ein Abstand wahrgenommen, hingegen vermöge der sekundären Tiefenkriterien ein solcher nur ‘vorgestellt’ werde.<sup>105</sup>

Wie eingangs dargestellt wurde, führt Helmholtz in §26 des *Handbuchs* den Begriff der Vorstellung ein als „Erinnerungsbild von Gesichtsobjekten, welches von keinen gegenwärtigen sinnlichen Empfindungen begleitet ist“ (Helmholtz, 1910/2003a, S. 11), als Wahrnehmungen werden von Helmholtz die über Empfindungen gewonnenen „Vorstellungen über die Existenz, die Form und die Lage äußerer Objekte“ (Helmholtz, 1910/2003a, S. 1) bezeichnet.<sup>106</sup> Möglicherweise zielt Helmholtz hier darauf ab, dass bei den primären Tiefenkriterien *stets* auf Empfindungen zurückgegriffen werde und es daher zu einer Wahrnehmung des Abstandes komme, wohingegen im Fall der sekundären Tiefenkriterien keine (oder vielleicht in einem viel geringerem Maße) Empfindungen ‘genutzt’ würden, was dann in einer bloßen Vorstellung des Abstandes resultiere.

Eine weitere, mit dieser Lesart zusammenhängende Möglichkeit wäre, dass Helmholtz den Unterschied zwischen den beiden Klassen der Tiefenkriterien in der ‘Funktionsweise’ oder der Beteiligung *qualitativ* verschiedener ‘Prozesse’ begründet sieht. Im Falle der sekundären Tiefenkriterien, deren Nutzung, der oben vorgeschlagenen Lesart folgend, in nicht geringem Maße von Erinnerungen an spezifische Erfahrungen mit den gesehenen Objekten abhängig ist, könnte es sich möglicherweise um einen aktiven, bewussten Prozess seitens des Wahrnehmenden handeln. So könnte beispielhaft die Nutzung des sekundären Tiefenkriteriums der ‘erscheinenden’ oder ‘retinalen Größe’ eine *aktive* Schätzung der Entfernung eines bekannten Objektes *durch* den Wahrnehmenden beinhalten. Die folgende Passage aus dem *Handbuch* könnte man zumindest in dieser Weise interpretieren:

Derselbe Gegenstand aus verschiedener Entfernung gesehen gibt verschiedenen große Netzhautbilder und erscheint unter verschiedenen Gesichtswinkeln. Je entfernter er ist, desto kleiner der Gesichtswinkel, unter dem er erscheint. *Wie* also die Astronomen aus der Messung der verschiedenen Gesichtswinkel, unter denen uns Sonne und Mond erscheinen, die Änderungen in der Entfernung dieser Gestirne *berechnen* können, *so können wir* aus dem Gesichtswinkel, oder was dem entspricht, aus der Größe des Netzhautbildchens eines gesehenen Gegenstandes von bekannter Größe die Entfernung

<sup>105</sup>Man erinnere sich, dass sich bei den Ausführungen des Descartes zur Tiefen- bzw. Distanzwahrnehmung eine ähnliche Unterscheidung finden ließ, nämlich die zwischen einer *gesehenen* Distanz und einer etwa bei einem Objekt mit uns bekannter Größe über Imagination *vorgestellten* Distanz.

<sup>106</sup>Es wurde bereits darauf hingewiesen, dass Helmholtz diese Begriffe in inkonsistenter Weise im *Handbuch* verwendet.

*schätzen*, in der er sich von uns befindet [sämtliche Hervorhebungen von mir]. (Helmholtz, 1910/2003a, S. 236)

Die „Größe des Netzhautbildchens“ eines Gegenstandes werde demnach in Verbindung mit der Kenntnis der (bekannten) Größe des Gegenstandes genutzt, um die Entfernung desselben zu schätzen, da die Größe des ‘Netzhautbildes’ eines bestimmten Gegenstandes sich umgekehrt proportional zur Entfernung dieses Gegenstandes vom Wahrnehmenden verhalte.

Es bieten sich nun verschiedene Interpretationen dieser Passage an. Nimmt man Helmholtz beim Wort, dann müsste man diese so auslegen, dass *wir* als Wahrnehmende *aktiv*, auf Basis der „Größe des Netzhautbildchens“ und dem Wissen über die ‘tatsächliche Größe’ des Gegenstandes, die Entfernung (wie ein Astronom) schätzten. Da wir als Wahrnehmende allerdings nicht unsere Netzhaut sehen und daher keinen Zugang zu den „Netzhautbildchen“ und deren Größe haben, wäre die von Helmholtz angeführte Analogie des Astronomen irreführend, da es sich um prinzipiell unterschiedliche Situationen handelt. Eine solche Analogie wäre nur sinnvoll unter der zusätzlichen Annahme, dass sich hinter unseren Retinae ein ‘kleiner Astronom’ oder Homunkulus befindet, der das auf die Retinae projizierte Lichtmuster *als ein Bild* mit einer bestimmten Größe *sehen* kann und auf dieser Basis und mit Hilfe geometrischer Berechnungen die Entfernung des Objektes schätzt. Und selbst wenn man diese Voraussetzung machen würde, wäre, wie bereits mehrfach erwähnt, immer noch nicht einzusehen, wie es aufgrund der Schätzung durch den Homunkulus zu *meiner* Wahrnehmung eines sich von mir entfernten Gegenstandes kommt. Eine den möglichen metaphorischen Charakter der Helmholtzschen Ausführungen stärker hervorhebende Interpretation dieser Passage wäre, dass es sich bei dieser ‘Entfernungsschätzung’ um eine Art im Gehirn ‘implementierten Mechanismus’ handelt, der ohne jede bewusste oder aktive Beteiligung seitens des Wahrnehmenden abläuft, sich aber so beschreiben lässt, *als ob* solche ‘Berechnungen’ angestellt würden. Aber selbst dann wäre (neben des fraglichen ‘als-ob-Charakters’ der Erklärung) immer noch unklar, welcher Zusammenhang zwischen möglichen ‘Berechnungen’ durch das Gehirn und meiner Wahrnehmung eines entfernten Gegenstandes bestehen soll. Zudem müsste dieser abstrakte ‘Mechanismus’ in der Lage sein, zu ‘erkennen’, dass das in Frage stehende retinale Lichtmuster das *Bild eines Objektes* ist.

Die Interpretation der Helmholtzschen Auffassung der Funktionsweise der beiden Klassen der Tiefencues wird dadurch erschwert, dass Helmholtz an anderen Stellen die ‘Nutzung’ der sekundären Tiefenkriterien als rein auf Erinnerungen an bereits bestehende Assoziationen basierend beschreibt:

Das gegenwärtige Bild ruft in uns wach die Erinnerung an alles, was in früheren Gesichtsbildern Ähnliches sich gefunden hat, und auch an alles,

was von sonstigen Erfahrungen mit diesen früheren Gesichtsbildern regelmäßig verbunden war, also z.B. die Anzahl von Schritten, die wir haben machen müssen, um an einen Menschen heranzukommen, dessen Erscheinung im Gesichtsfelde eine gewisse Größe gehabt hatte usw. Diese Art der Assoziation der Vorstellungen geschieht nicht bewußt und nicht willkürlich, sondern wie durch eine blinde Naturgewalt ... und was wir daher vermittels dieser auf die gesammelten Erfahrungen sich stützenden Ideenassoziationen den gegenwärtigen Empfindungen hinzufügen, erscheint ebensogut, wie letztere, uns ohne Willkür und ohne bewußte Tätigkeit von unserer Seite als unmittelbar gegeben, also als unmittelbare Wahrnehmung, während es doch nur zu den Vorstellungen zu rechnen ist. (Helmholtz, 1910/2003a, S. 245)

Diese Passage könnte man nun so auslegen, dass offenbar vor Ausführung (oder anstelle) eines unbewussten Schlusses von den bestehenden Empfindungen auf ein Objekt in der Außenwelt mit bestimmter Entfernung etc., den Empfindungen bestimmte Erfahrungen und damit verbundene Ideenassoziationen auf nicht bewusste Weise ‘hinzugefügt’ würden, was anschließend „wie durch blinde Naturgewalt“ zu einer von einer Wahrnehmung nicht zu unterscheidenden Gegebenheit eines Objektes der Außenwelt mit bestimmter Entfernung etc. führe, die Helmholtz ausdrücklich als Vorstellung bezeichnet und von einer Wahrnehmung abgrenzt. Liest man den Begriff der Vorstellung hier so, wie Helmholtz ihn selbst einführt, nämlich als „Erinnerungsbild von Gesichtobjekten, welches von keinen gegenwärtigen sinnlichen Empfindungen begleitet ist“ (Helmholtz, 1910/2003a, S.11), so würde dies bedeuten, dass der Wahrnehmende in einem solchen Fall über Assoziation bildhaft gleichsam seine Erinnerungen ‘sieht’, dieses ‘Sehen’ aber nicht von einer üblichen Wahrnehmung zu unterscheiden sei. Auch diese Beschreibung ist nicht wirklich verständlich. Helmholtz spricht zunächst von einem ‘gegenwärtigen Bild’, was darauf hindeutet, dass wir als Wahrnehmende ein solches in irgendeiner Weise *sehen* müssen. Dann spricht er davon, dass aufgrund der durch dieses Bild ‘wachgerufenen Erinnerungsbilder’ den *Empfindungen* etwas ‘hinzugefügt’ werde, das er als Vorstellungen kennzeichnet. Wie jedoch den Empfindungen, die hier als ‘gegenwärtige Bilder’ aufgefasst werden, eine Vorstellung ‘hinzugefügt’ werden kann und wie dies schließlich ein *einheitliches* Erleben eines Objektes der Außenwelt zur Folge haben soll, ist nicht verständlich.

Problematisch an beiden angeführten Darstellungen Helmholtz’ zur Funktionsweise der sekundären Tiefenkriterien ist, dass in beiden ein Sehen in gewisser Weise schon vorausgesetzt werden muss. Um die Größe des ‘Netzhautbildes’ in Verbindung mit der Kenntnis der Größe eines bekannten Objektes als Hilfsmittel überhaupt nutzen zu können, muss der Wahrnehmende (oder ein Homunkulus oder ein ‘abstrakter Mecha-

nismus') erstens das 'Netzhautbild' sehen können und zweitens dieses als Bild eines *bestimmten Objektes* sehen können. Wie die bekannte Größe eines Objektes (oder einer abstrakten Objektklasse) in Verbindung mit der 'retinalen Größe' als Hilfsmittel genutzt werden soll, ohne dass das 'Netzhautbild' als ein Bild eben dieses Objektes gesehen wird, ist nicht einzusehen. Der Verweis auf Empfindungen scheint hier – von dem fraglichen Charakter dieses Konzeptes abgesehen – nicht zu genügen. In der zweiten Helmholtzschen Beschreibung stellt sich das Problem so dar, dass das 'gegenwärtige Bild' doch überhaupt nur Erinnerungen an ähnliche Fälle auslösen kann, wenn die Empfindungen bereits als ein bestimmtes Objekt der Außenwelt gedeutet wurden, um so diese Situation als zu einer bestimmten, bereits erfahrenen Situation 'ähnlich' klassifizieren zu können. Am deutlichsten tritt dieser Punkt an einer Stelle im *Handbuch* zutage, in der Helmholtz näher auf das sekundäre Tiefenkriterium der Verdeckung eingeht und argumentiert, dass wir in einer bestimmten Sehsituation ein Objekt *A* als weiter entfernt wahrnehmen würden als ein anderes Objekt *B*, weil Objekt *A* von *B* *verdeckt* werde.

Zur Kenntnis der Größe kommt ferner in sehr vielen Fällen die Kenntnis der Form der Objekte, namentlich in solchen Fällen, wo das eine zum Teil vom anderen gedeckt wird. Wenn wir z.B. in der Entfernung zwei Hügel *sehen* [Hervorhebung von mir], von denen der eine mit seiner Basis sich vor den anderen vorschiebt und den letzteren zum Teil verdeckt, so schließen wir daraus unmittelbar, daß der deckende vor dem gedeckten liegt. (Helmholtz, 1910/2003a, S. 237)

Um jedoch feststellen zu können, dass ein Objekt *A* von einem Objekt *B* verdeckt wird, scheint es doch notwendig, erstens *A als A* und *B als B* und unterschieden von *A* bereits wahrgenommen zu haben. Andernfalls scheint es nicht sinnvoll zu sein, davon zu sprechen, dass etwas verdeckt werde. Falls Helmholtz hier meinen sollte, dass sich eine Verdeckung in den 'Netzhautbildern' zeige, steht man erneut vor dem Problem, dass man einen Homunkulus annehmen müsste, der die retinalen Lichtmuster als Bilder sich teilweise verdeckender Objekte sehen kann.

Hinsichtlich der primären Tiefenkriterien, die eine „wirkliche *Wahrnehmung* des Abstandes“ (Helmholtz, 1910/2003a, S. 236) erlauben sollen, behandelt Helmholtz lediglich die Konvergenz und die binokulare Disparität ausführlicher. Bezüglich der Konvergenz lassen sich nur eher knappe Erläuterungen im *Handbuch* finden, die sich im Wesentlichen mit den entsprechenden Erläuterungen im Vortrag von 1855 decken, so dass an dieser Stelle auf eine erneute ausführliche Darstellung dieses Hilfsmittels und der mit der Helmholtzschen Auffassung verbundenen Probleme verzichtet werden soll. Die Funktionsweise dieses Hilfsmittels, das soll erwähnt werden, lässt sich jedoch besser

in die grundsätzliche Logik der Helmholtzschen Wahrnehmungstheorie einer durch Erfahrung entstandenen assoziativen Verknüpfung zwischen Empfindungen des Tast- und Sehannes integrieren, als etwa die Verdeckung oder die nun zu behandelnde binokulare Disparität.

Die binokulare Disparität wird von Helmholtz, wie schon im Vortrag von 1855, als eine „Vergleichung verschiedener Netzhautbilder“ charakterisiert, die als Hilfsmittel zur Tiefenwahrnehmung zudem „außerordentlich genau“ (Helmholtz, 1910/2003a, S. 254) sei. Die bei binokularer Fixation einer bestimmten Szene durch den Abstand der beiden Augen zueinander sich ergebenden Unterschiede des retinalen Lichtmusters kämen laut Helmholtz dem Wahrnehmenden „als solche ... nicht zum Bewusstsein“, sondern lediglich die „Unterschiede in der Tiefendimension, die von jenen Unterschieden abhängen“ (Helmholtz, 1910/2003a, S. 254). Helmholtz scheint an dieser Stelle auszuschließen, dass der Wahrnehmende die beiden unterschiedlichen ‘Netzhautbilder’ selbst in irgendeiner Weise wahrnehme, was ihm zu Bewusstsein komme, seien damit zusammenhängende „Unterschiede in der Tiefendimension“. Um dieser Beschreibung näher nachgehen zu können, soll in aller Kürze etwas genauer auf die übliche (geometrische) Charakterisierung der binokularen Disparität als Tiefencue eingegangen werden.

Man stelle sich vor, man fixiere mit beiden Augen einen bestimmten Punkt  $F$  eines Objektes, das sich in bestimmter Entfernung befinde (vgl. Abbildung 2.2 auf der nächsten Seite, die diese Situation geometrisch vereinfacht darstellt). Das von diesem Punkt reflektierte Licht breitet sich in alle Richtungen aus, wir wollen im Folgenden lediglich die beiden sogenannten *Hauptstrahlen* betrachten, also jene beiden Lichtstrahlen, die durch den jeweiligen Mittelpunkt der Linse des Auges verlaufen und nicht gebrochen werden (vgl. Hecht, 2009). Die vom fixierten Punkt  $F$  reflektierten Hauptstrahlen werden auf *korrespondierende* Stellen  $f_l, f_r$  der beiden Retinae projiziert, nämlich jeweils in den Bereich der Fovea. Man stelle sich nun einen weiteren Punkt  $A$  vor, dessen Entfernung geringer ist als die Entfernung des fixierten Punktes  $F$ . Betrachtet man nun die von  $A$  reflektierten Hauptstrahlen, so zeigt sich, dass diese auf *nicht korrespondierende* Netzhautstellen  $a_l, a_r$  projiziert werden. Anschaulich gesprochen liegt  $a_l$  ‘links’ von  $f_l$ ,  $a_r$  hingegen ‘rechts’ von  $f_r$ . Nun stelle man sich vor, dass bei Fixation von  $F$  die Entfernung des Punktes  $A$  zum Wahrnehmenden vergrößert werde, so dass dieser nicht vor, sondern hinter  $F$  liege (ein solcher Punkt ist in Abbildung 2.2 mit  $A'$  bezeichnet). Auch hier zeigt sich geometrisch, dass die von  $A'$  reflektierten Hauptstrahlen ebenfalls auf nicht korrespondierende Netzhautstellen  $a'_l, a'_r$  projiziert werden. Vergleicht man diese Situation mit der des Punktes  $A$ , so sieht man, dass nun allerdings (wieder anschaulich gesprochen)  $a'_l$  ‘rechts’ von  $f_l$  und  $a'_r$  ‘links’ von  $f_r$  liegt. Anhand Abbildung 2.2 macht man sich folgende Punkte leicht klar: Befindet sich in der dargestellten einfachen Situation bei Fixation von  $F$  der Punkt  $A$  vor  $F$ , so wird  $a_l$  stets ‘links’ von  $f_l$

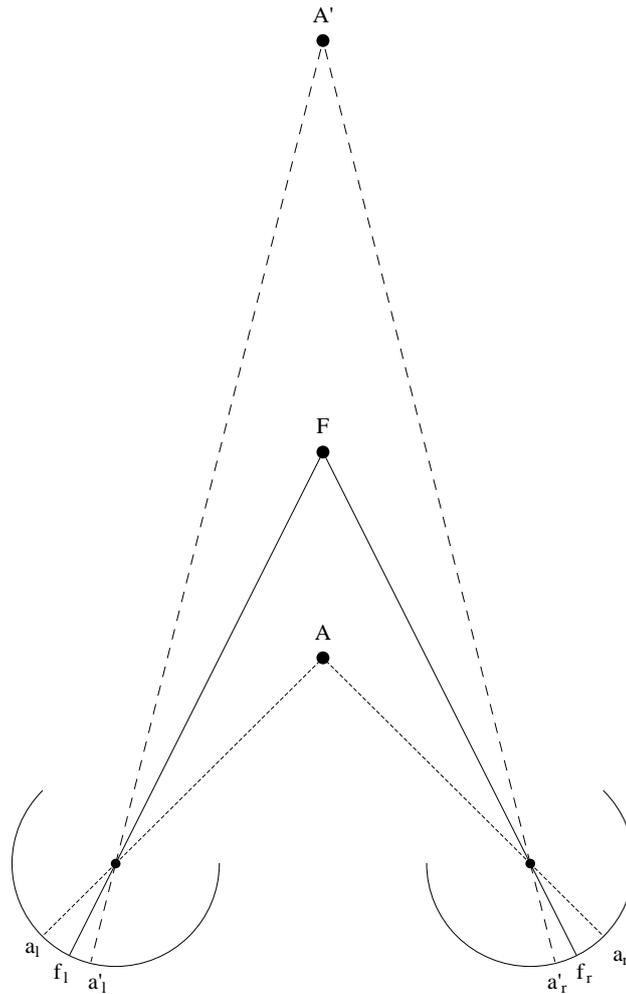


Abbildung 2.2. Geometrische Veranschaulichung der binokularen Disparität. Erläuterungen im Text.

und  $a_r$  stets ‘rechts’ von  $f_r$  liegen. Befindet sich (wie im Fall von  $A'$ ) der Punkt *hinter*  $F$ , so wird  $a'_l$  stets ‘rechts’ von  $f_l$  und  $a'_r$  stets ‘links’ von  $f_r$  liegen. Die Größe der horizontalen Abweichung zwischen den Punkten  $a_l$  und  $f_l$  bzw.  $a_r$  und  $f_r$  auf der Retina ist zudem abhängig von der Entfernung zwischen dem fixierten Punkt  $F$  und Punkt  $A$ . Anschaulich gesprochen bewegen sich die Punkte  $a_l$  und  $a_r$  auf die jeweiligen im Bereich der Fovea liegenden Punkte  $f_l$  und  $f_r$  zu, wenn der Abstand zwischen  $F$  und  $A$  geringer wird und sie ‘entfernen’ sich von diesen Punkten, wenn der Abstand zwischen  $F$  und  $A$  wächst, wobei die Zu- bzw. Abnahme der Entfernung dieser Punkte auf der Retina in Abhängigkeit von der Distanz zwischen  $A$  und  $F$  unterschiedlich ausfällt, je nachdem, ob sich  $A$  vor oder hinter  $F$  befindet.<sup>107</sup> Man beachte, dass es sich hier um eine geometrische Abstraktion der Sehsituation handelt.

Helmholtz scheint davon auszugehen, dass diese *geometrischen Eigenschaften* in irgendeiner Weise vom Wahrnehmenden ‘genutzt’ werden könnten, indem ihm diese

<sup>107</sup>Für eine ausführliche Darstellung der geometrischen Eigenschaften solcher vereinfachten Sehsituationen siehe etwa Arditti (1986).

durch ‘Vergleichung der Netzhautbilder’ in einer konkreten Situation als „Unterschiede in der Tiefendimension“ zu Bewusstsein kämen. Weitergehende Erläuterungen lassen sich leider nicht finden, Helmholtz geht ohne Weiteres zur Darstellung zugehöriger empirischer Befunde über. Der grundsätzlichen Logik der Helmholtzschen Theorie folgend wäre denkbar, dass diese ‘Vergleichung’ über assoziative Verknüpfungen bestimmter Empfindungen geschehe, allerdings ist es nicht auszumachen, welche Empfindungen in diesem Fall verknüpft werden sollen. Laut der oben zitierten Passage ist ausgeschlossen, dass die beiden unterschiedlichen retinalen Lichtmuster selbst vom Wahrnehmenden ‘empfunden’ werden, da sie „als solche ... nicht zum Bewusstsein kommen“. Daher, so müsste man schlussfolgern, können sie auch nicht mit Empfindungen verknüpft werden. Dass die „Unterschiede in der Tiefendimension“ selbst von Helmholtz als Empfindungen angesehen werden, scheint in seiner Konzeption jedoch ebenfalls nicht plausibel zu sein, da Tiefe nicht empfunden werden könne, sondern erst im Rahmen einer Wahrnehmung ‘konstruiert’ werde. Aufgrund des deutlich mentalistisch geprägten Begriffs der ‘Vergleichung’ wäre auch möglich, dass Helmholtz hier eine Vorstellung vor Augen hat, die bereits im Zusammenhang mit dem sekundären Tiefenkriterium der ‘retinalen Größe’ erläutert wurde. So wie im Fall der ‘retinalen Größe’ der Wahrnehmende ‘wie ein Astronom’ die Entfernung ‘berechnen’ könne, könnte es Helmholtz Ansicht sein, dass auch im Fall der binokularen Disparität die Entfernung aus den unterschiedlichen retinalen Lichtmustern ‘berechnet’ oder ‘geschätzt’ wird. Da der Wahrnehmende selbst aber keinen Zugang zu den Lichtmustern auf seinen Retinae hat, müsste man diese Tätigkeit nun wieder einem Homunkulus zuordnen oder als rein metaphorische Beschreibung eines abstrakten ‘Gehirnmechanismus’ ansehen, was wiederum die Annahme beinhalten würde, dass das Gehirn ‘Berechnungen’ vornehmen könne. In beiden Fällen würde sich jedoch erneut die Frage stellen, wie das Ergebnis einer ‘Vergleichung’ oder einer ‘Berechnung’ der Tiefe durch einen Homunkulus oder einen abstrakten Mechanismus zur Wahrnehmung unterschiedlich weit entfernter Objekte führen soll.

Anzumerken ist noch, dass Helmholtz’ Behandlung der ‘Sinnestäuschungen’ im *Handbuch* keinerlei Veränderungen im Vergleich zum Vortrag von 1855 erfahren hat. Die Sinnestäuschungen blieben selbst gegen ‘besseres Wissen’ und gegen wiederholte Erfahrungen, dass es sich um solche handle, bestehen, da sich eine assoziative Verknüpfung gebildet habe, so dass dem Wahrnehmenden über den entsprechenden Analogieschluss wie „durch fremde Naturkraft“ die ‘falsche Wahrnehmung’ gewissermassen aufgezwungen werde.<sup>108</sup> Wie in der Darstellung des Helmholtzschen Vortrags von 1855 schon

---

<sup>108</sup> „Daß wir nun, nachdem wir den physiologischen Ursprung und Zusammenhang der Sinnestäuschungen kennen gelernt haben, doch die Täuschung trotz unserer besseren Einsicht nicht los werden können, rührt eben davon her, daß die Induktion durch eine unbewußte und unwillkürliche Tätigkeit des Gedächtnisses gebildet ist, die eben deshalb unserem Bewußtsein als eine fremde, zwingende Naturkraft erscheint“ (Helmholtz, 1910/2003a, S. 26).

erwähnt wurde, lassen sich die ‘Sinnestäuschungen’ nicht mit der Helmholtzschen empiristischen Konzeption von Wahrnehmung als eines auf Erfahrung basierenden Prozesses in Übereinstimmung bringen. Wenn Wahrnehmung über hypothesengeleitetes Experimentieren durch den Wahrnehmenden *erlernt* werden müsse, so sollte doch die Möglichkeit bestehen, diese ‘Täuschungen’ durch Erfahrungen gleichsam zu ‘korrigieren’. Üblicherweise sei es laut Helmholtz (1896/2002a, S. 329) auch so, „dass wir die Anschauungen des Auges durch die des Tastsinnes, wo es angeht, fortdauernd kontrollieren und corrigieren und dabei die Aussagen des letzteren immer als die entscheidenden betrachten“. Bei den ‘Täuschungen’ (man denke an den ins Wasser getauchten Stab) ist dies jedoch nicht der Fall, was Helmholtz selbst auch zugibt, aber nicht weiter problematisiert. Die ‘Sinnestäuschungen’ bleiben somit auch im *Handbuch* ein wunder Punkt der Helmholtzschen Theorie.

Wir wollen uns nun einer weiteren, ‘späten’ Ausarbeitung der wahrnehmungstheoretischen Ansichten Helmholtzens zuwenden, dem 1878 gehaltenen Vortrag „Die Tatsachen in der Wahrnehmung“, in der insbesondere erkenntnistheoretische Erwägungen im Vordergrund stehen, die nicht zuletzt für Helmholtz’ Auffassung des Status der Naturwissenschaften von Relevanz sind und großen Einfluss auf namhafte Physiker wie etwa Heinrich Hertz ausübten.<sup>109</sup>

#### 2.6.4. Der Vortrag „Die Tatsachen in der Wahrnehmung“ (1878)

Die zur Stiftungsfeier der Berliner Universität 1878 von Helmholtz in seiner Eigenschaft als Rektor gehaltene Rede schließt in den wesentlichen Punkten an die Ausführungen im *Handbuch* an, hier aber mit einem deutlich erkenntnistheoretisch und weniger technisch ausgerichteten Einschlag. Schlick (1998, S. 197) gilt diese Rede als die „gehaltvollste von Helmholtz’ erkenntnistheoretischen Arbeiten“, da sie „die vollkommenste, geschlossenste Darstellung seiner Erkenntnislehre“ enthalte, weshalb sie auch eine separate Behandlung erfahren soll. Ein weiterer Grund, diesen Vortrag zum Gegenstand eines eigenen Abschnittes zu machen, besteht darin, dass an einigen Stellen des Vortrages die Helmholtzsche Theorie auch nennenswerte Modifikationen bzw. Erweiterungen erfährt. Da allerdings die Hilfsmittel zur Tiefenwahrnehmung oder die Tiefencues, die ja im Zentrum dieser Arbeit stehen sollen, keine besondere Erwähnung in diesem Vortrag finden, sondern es sich eher um eine allgemeine Darstellung der Helmholtzschen Wahrnehmungs- bzw. Erkenntnistheorie handelt, soll die Behandlung vergleichsweise kurz gehalten werden.

---

<sup>109</sup>Siehe etwa Hertz (1894) sowie Schieman (2008) für eine Darstellung des Einflusses des Helmholtzschen Denkens auf Hertz.

Was beim Lesen dieses Vortrages zunächst deutlich auffällt ist, dass Helmholtz sich erstens an vielen Stellen explizit auf Kant bezieht und zweitens stärker als in den bisher vorgestellten Schriften mit dem von Kant in der „Kritik der reinen Vernunft“ entwickelten und bereitgestellten Begriffsapparat hantiert, die Kantische Terminologie jedoch in zuweilen vollkommen anderer Bedeutung gebraucht.<sup>110</sup> Gleich zu Beginn formuliert Helmholtz jene Fragen, die seiner Ansicht nach im Fokus der Erkenntnistheorie stehen und daher in seinem Vortrag behandelt werden sollen, nämlich: „Was ist Wahrheit in unserem Anschauen und Denken? In welchem Sinne entsprechen unsere Vorstellungen der Wirklichkeit?“ (Helmholtz, 1896/1998, S. 149). Mit Verweis auf Müllers Theorie der spezifischen Sinnesenergien stellt Helmholtz dann – in annähernd gleicher Wortwahl wie im *Handbuch* – seine Zeichentheorie dar, die Antworten auf diese Fragen geben soll. Die Empfindungen seien durch äußere Ursachen hervorgebrachte Wirkungen auf die Sinnesorgane, die daher sowohl von diesen äußeren Ursachen, als auch von der Eigentümlichkeit der Sinnesorgane abhängig seien.<sup>111</sup> Von größter Wichtigkeit für die aufgeworfenen erkenntnistheoretischen Fragen sei jedoch, dass (unter ansonsten gleichen Umständen) eine gleiche äußere Einwirkung auf ein bestimmtes Sinnesorgan *stets* zu einer gleichen Qualität der Empfindung führe und daher ungleiche Empfindungsqualitäten auf ungleichen äußeren Einwirkungen beruhten. Unsere Empfindungen könnten daher auch als Zeichen für diese äußeren Einwirkungen aufgefasst werden:

Insofern die Qualität unserer Empfindung uns von der Eigentümlichkeit der äußeren Einwirkung, durch welche sie erregt ist, eine Nachricht gibt, kann sie als ein *Zeichen* derselben gelten, aber nicht als ein Abbild. ... Ein Zeichen aber braucht gar keine Art der Ähnlichkeit mit dem zu haben, dessen Zeichen es ist. Die Beziehung beschränkt sich darauf, daß das gleiche Objekt, unter gleichen Umständen zur Wirkung kommend, das gleiche Zeichen hervorruft, und daß also ungleiche Zeichen immer ungleicher Einwirkung entsprechen. (Helmholtz, 1896/1998, S. 153)<sup>112</sup>

Was können diese Zeichen, die als solche keinerlei Ähnlichkeit mit dem, das sie bezeichnen sollen, aufweisen müssten, aus erkenntnistheoretischer Perspektive über das ‘Verhältnis zwischen unseren Vorstellungen und der Wirklichkeit’ aussagen? Da gleiche Empfindungen durch gleiche äußere Einwirkung und verschiedene Empfindungen

<sup>110</sup>Einige Beispiele hierfür sind insbesondere die Begriffe „*a priori*“, „transzendental“ und „Anschauungsform“. Da dieser Punkt für die Ausrichtung der vorliegenden Arbeit nicht wesentlich ist, soll auf eine ausführliche Diskussion verzichtet werden. Der interessierte Leser sei auf Riehl (1904) sowie Schlick (1922/2012, 1998) verwiesen.

<sup>111</sup>Auffällig ist auch in diesem Vortrag, dass Helmholtz *stets* zwischen einer Beschreibung der Empfindungen als etwas, das uns ‘zu Bewusstsein kommt’ und etwas, das „in unseren Organen hervorgebracht“ wird, schwankt (etwa Helmholtz, 1896/1998, S. 153).

<sup>112</sup>Mathematisch ausgedrückt soll die Zeichenrelation in Helmholtz’ Sinne einer Zuordnung im Sinne einer Abbildung genügen.

durch verschiedene äußere Einwirkung hervorgerufen würden, kämen wir über die Zeichen zu einer „Abbildung der *Gesetzmäßigkeit* in den Vorgängen der wirklichen Welt“ (Helmholtz, 1896/1998, S. 153):

Da Gleiches in unserer Empfindungswelt durch gleiche Zeichen angezeigt wird, so wird der naturgesetzlichen Folge gleicher Wirkungen auf gleiche Ursachen, auch eine ebenso regelmäßige Folge im Gebiete unserer Empfindungen entsprechen. ... Wenn also unsere Sinnesempfindungen in ihrer Qualität auch nur *Zeichen* sind, deren besondere Art ganz von unserer Organisation abhängt, so sind sie doch nicht als leerer Schein zu verwerfen, sondern sie sind eben Zeichen von *Etwas*, sei es etwas Bestehendem oder Geschehendem, und was das Wichtigste ist, das *Gesetz* dieses Geschehens können sie uns abbilden. (Helmholtz, 1896/1998, S. 154.)

Falls, so ein Beispiel Helmholtz', eine bestimmte Sorte von Beeren beim Reifen stets rotes Pigment und Zucker ausbildete, so würden wir im Kontakt mit Beeren dieser Art auch stets die Qualitäten 'rot' und 'süß' zusammen in unseren Empfindungen auffinden und auf diese Weise würde sich die Gesetzmäßigkeit des Geschehens in der Außenwelt, und allein diese, in unseren Empfindungen widerspiegeln. Die Qualitäten der Empfindungen selbst seien laut Helmholtz, unter Bezugnahme auf die Kantische Erkenntnistheorie, „bloße Formen der Anschauung“ (Helmholtz, 1896/1998, S. 154), die daher nicht der 'Wirklichkeit' entsprechen müssten.<sup>113</sup>

Im Verlauf des Vortrages widmet sich Helmholtz dann auch dem Problem der Tiefen- oder Raumwahrnehmung, hier ebenfalls von einem deutlich erkenntnistheoretischen Standpunkt aus. Ganz ähnlich den entsprechenden Ausführungen im *Handbuch* spricht Helmholtz auch hier dem willentlichen Experimentieren über motorische Impulse und Bewegungsveränderungen durch den Wahrnehmenden eine konstituierende Rolle zu. Zudem möchte Helmholtz zeigen, dass sich sämtliche wahrgenommenen räumlichen Bestimmungen *allein* aus diesem Experimentieren mit motorischen Impulsen und Bewegungsveränderungen ableiten ließen und daher, ganz im Sinne einer empiristischen Theorie der Raumwahrnehmung, keine zusätzlichen Voraussetzungen diesbezüglich gemacht werden müssten. Interessant in diesem Zusammenhang ist nun, dass Helmholtz unter Raumwahrnehmung *sämtliche* 'empfundenen Verhältnisse' versteht, die durch motorische Impulse seitens des Wahrnehmenden Veränderungen unterworfen sind. Da laut Helmholtz sämtliche Empfindungen des Sehsinns auf motorischen Impulsen beruhten, seien mithin *sämtliche* Empfindungen des Sehsinns in diesem Sinn *räumlich*

---

<sup>113</sup>Wie Schlick (1998) richtig bemerkt, erweckt die Darstellung Helmholtzens den Eindruck, dass Kant Empfindungsqualitäten als Formen der Anschauung, ähnlich Zeit und Raum, angesehen hätte, was allerdings nicht der Fall ist, da für Kant Empfindungsqualitäten *Inhalte* von Vorstellungen darstellen.

*bestimmt.* In diesem Argumentationsgang versucht Helmholtz nun auch, die Kantische Auffassung des Raumes als einer *Form der Anschauung* gleichsam auf physiologischer Basis zu rechtfertigen, ein Weg, der mit der Kantischen Transzendentalphilosophie freilich nicht mehr viel zu tun hat:

Wenn wir nun Impulse solcher Art geben (den Blick wenden, die Hände bewegen, hin- und hergehen), so finden wir, daß die gewissen Qualitätskreisen angehörigen Empfindungen (nämlich, die auf räumliche Objekte bezüglichen) dadurch geändert werden können; andere psychische Zustände, deren wir uns bewusst sind, Erinnerungen, Absichten, Wünsche, Stimmungen, durchaus nicht. Dadurch ist in unmittelbarer Wahrnehmung ein durchgreifender Unterschied zwischen den ersteren und letzteren gesetzt. Wenn wir also dasjenige Verhältnis, welches wir durch unsere Willensimpulse unmittelbar ändern, dessen Art uns übrigens noch ganz unbekannt sein könnte, ein *räumliches* nennen wollen, so treten die Wahrnehmungen *psychischer* Tätigkeiten gar nicht in ein solches ein; wohl aber müssen alle Empfindungen der äußeren Sinne unter irgend welcher Art der Innervation vor sich gehen, d.h. räumlich bestimmt sein. Demnach wird uns der Raum auch sinnlich erscheinen, behaftet mit den Qualitäten unserer Bewegungsempfindungen. Die Raumschauung würde also in diesem Sinne eine subjektive *Anschauungsform* sein, wie die Empfindungsqualitäten Rot, Süß, Kalt (Helmholtz, 1896/1998, S. 155).<sup>114</sup>

Zur Wahrnehmung von dauerhaft bestehenden Objekten in einer als vom Wahrnehmenden unabhängig erlebten Außenwelt komme es über die unbewussten Schlüsse, die Helmholtz in diesem Vortrag nur noch als Induktionsschlüsse bezeichnet. Befände sich der Wahrnehmende „einer Umgebung von ruhenden Objekten gegenüber“ (Helmholtz, 1896/1998, S. 156), so könne er erstens erkennen, dass es sich um ruhende Objekte handle, da ohne motorischen Impuls seinerseits die Empfindungen und die mit diesen einhergehenden räumlichen Bestimmungen unverändert blieben. Zweitens sei der Wahrnehmende in der Lage, durch einen solchen motorischen Impuls die Empfindungen zu verändern und „durch Nachlaß oder den zugehörigen Gegenimpuls“ (Helmholtz, 1896/1998, S. 156) wieder in den ursprünglichen Empfindungszustand zurückzukehren. Der Wahrnehmende könne so aus allen möglichen ‘Empfindungskonstellationen’, die in dieser gegebenen Situation durch seine willentlichen Impulse hervorgerufen werden könnten (Helmholtz nennt dies auch *Präsentabilien*) zu jeder Zeit jede einzelne dieser ‘Konstellationen’ durch den entsprechenden motorischen Impuls hervorrufen und in jedem Moment, in dem er den entsprechenden motorischen Impuls gibt, diese spezifische

<sup>114</sup>Auch hier wird der Begriff der Anschauungsform von Helmholtz in einem völlig anderen Sinn verwendet als bei Kant.

Konstellation ‘beobachten’. Aus diesen Einzelbeobachtungen schließe er nun ‘per Induktion’, dass er auch in jedem anderen Moment, in dem er diesen motorischen Impuls hätte geben *können*, die entsprechenden ‘Beobachtungen’ gemacht haben würde. Dadurch „erscheint ihm jedes Einzelne aus dieser Gruppe der Präsentabilien als *bestehend in jedem Augenblick*“ (Helmholtz, 1896/1998, S. 156) und so könne laut Helmholtz die „Vorstellung von einem *dauernden Bestehen von Verschiedenem gleichzeitig nebeneinander* gewonnen werden“ (Helmholtz, 1896/1998, S. 157). Wenn sich nun für dieselben motorischen Impulse und Bewegungsveränderungen auch *unterschiedliche* Präsentabilien ergäben, so könne der Wahrnehmende die durch seine eigenen motorischen Impulse sich ergebenden Veränderungen der ‘Empfindungskonstellation’ von jenen Veränderungen trennen, die nicht Folge seiner motorischen Impulse und auch nicht durch diese rückgängig zu machen seien. Der Wahrnehmende würde dann diesen nicht durch seine motorischen Impulse veränderbaren Teil der ‘Empfindungskonstellation’ als *von ihm unabhängige Objekte* ansehen, sie träten ihm „als ein Gegebenes, ein *objectum* entgegen“ (Helmholtz, 1896/1998, S. 157).

Bei jeder dieser durch den Wahrnehmenden hervorgebrachten willentlichen Veränderungen der Empfindungen handle es sich um ein Experiment, das auf *aktiver und intentionaler* Tätigkeit des Wahrnehmenden beruhe und dazu diene, eine *Gesetzmäßigkeit* zwischen den willentlichen motorischen Impulsen und den damit einhergehenden Empfindungsveränderungen herzustellen und beständig zu prüfen.<sup>115</sup> Dass ein Experimentieren dieser Art überhaupt zur Feststellung einer Gesetzmäßigkeit zwischen motorischen Impulsen und Empfindungsveränderungen geeignet ist, möchte Helmholtz über das Kausalgesetz verständlich machen.

Die Ausführungen Helmholtz’ zum Kausalgesetz in diesem Vortrag sind mannigfaltiger als in seinen anderen Schriften; das Kausalgesetz nimmt zudem nun eine wesentlich wichtigere, man könnte sogar sagen, *die zentrale* Position innerhalb der Helmholtz’schen Theorie ein. Dabei lassen sich mehrere Funktionen des Kausalgesetzes innerhalb Helmholtz’ Theorie unterscheiden: Zum Einen stellt das Kausalgesetz die allgemeine Grundlage für sämtliche ‘Induktionsschlüsse’ des Wahrnehmenden dar, da laut Helmholtz ohne die *Annahme* einer „allgemeinen Gesetzmäßigkeit alles Geschehenden“ ein solches Vorgehen keine vernünftige Basis hätte. Damit zusammenhängend ist das Kausalgesetz zweitens unabdinglich für ein willentliches Experimentieren seitens des Wahrnehmenden, da nur durch dieses sichergestellt werde, dass bei ansonsten konstanten

<sup>115</sup> „Wir haben, wie wir gesehen, nicht nur wechselnde Sinneseindrücke, die über uns kommen, ohne daß wir etwas dazu tun, sondern wir beobachten unter *fortdauernder eigener Tätigkeit*, und gelangen dadurch zur Kenntnis des *Bestehens* eines gesetzlichen Verhältnisses zwischen unseren Innervationen und dem Präsentwerden der verschiedenen Eindrücke aus dem Kreise der zeitweiligen Präsentabilien. Jede unserer willkürlichen Bewegungen, durch die wir die Erscheinungsweise der Objekte abändern, ist als ein Experiment zu betrachten, durch welches wir prüfen, ob wir das gesetzliche Verhalten der vorliegenden Erscheinung ... richtig aufgefaßt haben“ (Helmholtz, 1896/1998, S. 166).

Bedingungen *dieselben* motorischen Impulse stets zu *denselben* Empfindungsveränderungen führten. Schließlich soll durch das Kausalgesetz gleichsam verbürgt werden, dass im Sinne der Zeichentheorie identische Empfindungen stets durch identische äußere Einwirkungen hervorgebracht werden, dessen Zeichen sie seien.

Damit weicht Helmholtz trotz der sprachlichen Übereinstimmung – auch Helmholtz (1896/1998, S. 171) nennt das Kausalgesetz ein „a priori gegebenes, ein transzendentes Gesetz“ – deutlich von der Kantischen Auffassung der Kausalität ab. Wie bereits erwähnt, stellt für Kant das Kausalgesetz eine Bedingung der Möglichkeit von Erfahrung überhaupt dar und ist in diesem Sinne *a priori*, ein „notwendiges Gesetz unserer Sinnlichkeit, mithin eine *formale* Bedingung aller Wahrnehmungen“ (Kant, 1781/1998, S. 296). Bei Kant erstreckt sich der Geltungsbereich des Kausalgesetzes allein auf die Erscheinungen, wohingegen Helmholtz' Argument, dass Empfindungen als Zeichen für die „Gesetzmäßigkeit in den Vorgängen der wirklichen Welt“ zu betrachten seien (da gleiche Zeichen stets auf gleicher äußerer Einwirkung und ungleiche Zeichen stets auf ungleicher äußerer Einwirkung beruhten) einen anderen Geltungsbereich für das Kausalgesetz aufweist, nämlich als gültig für die Beziehung zwischen den Vorgängen in einer unabhängigen Außenwelt und dem wahrnehmenden Subjekt. Durch das so verstandene Kausalgesetz wird somit eine Beziehung zwischen dieser und dem wahrnehmenden Subjekt hergestellt und es hat zudem die bereits an anderer Stelle erwähnte Funktion, die Annahme einer unabhängigen Außenwelt nachträglich zu rechtfertigen, ohne diese Annahme selbst zu thematisieren oder zu begründen.

Darüber hinaus kennzeichnet Helmholtz das Kausalgesetz nun als „ein regulatives Prinzip unseres Denkens“, welches im Grunde nur Ausdruck eines „Vertrauens auf die *vollkommene Begreifbarkeit* der Welt“ (Helmholtz, 1896/1998, S. 171) und daher nicht ‚beweisbar‘, sondern allein durch den Erfolg seiner Anwendbarkeit gerechtfertigt sei:

Jeder Induktionsschluß stützt sich auf das Vertrauen, daß ein bisher beobachtetes gesetzliches Verhalten sich in allen noch nicht zur Beobachtung gekommenen Fällen bewähren werde. Es ist dies ein Vertrauen auf die Gesetzmäßigkeit allen Geschehens. ... Setzen wir aber voraus, ... daß wir ein letztes Unveränderliches als *Ursache* der beobachteten Veränderungen werden hinstellen können, so nennen wir das regulative Prinzip unseres Denkens, das uns dazu treibt, das *Kausalgesetz*. (Helmholtz, 1896/1998, S. 171)

Helmholtz bezeichnet das Kausalgesetz in einem von Kant abweichenden Sinne als *a priori*, da es sich nicht aus der Erfahrung ableiten lasse, sondern auf die Gültigkeit des Kausalgesetzes *vertraut* werden müsse, damit Erfahrungen über induktives Vorgehen gemacht werden könnten.<sup>116</sup> Dadurch kommt es zu einer „Hypothetisierung der Kau-

<sup>116</sup>„Das Kausalgesetz ist wirklich ein a priori gegebenes, transzendentes Gesetz. Ein Beweis des Kausalgesetzes aus der Erfahrung ist nicht möglich; denn die ersten Schritte der Erfahrung sind

salität“ (Schiemann, 2008, S. 38), das Kausalgesetz nimmt die Form einer empirisch nicht abschließend verifizierten und auch nicht abschließend verifizierbaren Aussage an, auf deren Gültigkeit jedoch vertraut werden müsse, damit Erfahrungen (und auch wissenschaftliche Erkenntnisse) überhaupt möglich seien.<sup>117</sup>

Über die Gültigkeit des Kausalgesetzes, in die man vertrauen müsse, kommt Helmholtz dann zu einer Antwort auf die von ihm zu Beginn des Vortrags gestellte, seiner Ansicht nach zentrale erkenntnistheoretische Frage, in welchem Sinne unsere Vorstellungen der ‘Wirklichkeit’ entsprächen. Alles, was wir laut Helmholtz erreichen könnten, wäre, durch die Wahrnehmungen zu einer „Kenntnis der gesetzlichen Ordnung im Reiche des Wirklichen“ (Helmholtz, 1896/1998, S. 170) zu gelangen, da sich diese, und nur diese, über das Zeichensystem unserer Empfindungen darstelle. Das *Gesetzmäßige* selbst, d.h. die sich in den Zeichen ausdrückenden regelmäßigen Beziehungen zwischen dem, was auf unsere Sinne wirke, und unseren Empfindungen, sei die *einzig*e ‘Tatsache in der Wahrnehmung’. Die sachlichen Bestimmungen unserer Wahrnehmungen, die Eigenschaften der von uns als einer unabhängigen Außenwelt angehörig wahrgenommenen Objekte als Ergebnis der Deutung der Zeichen seien mehr oder weniger akzidentiell, da das Deuten der Zeichen vom Wahrnehmenden über Induktion nach pragmatischen Gesichtspunkten gelernt werden müsse und – etwa im Falle der ‘Sinnestäuschungen’ – auch ‘fehlerhaft’ sein könne. Die sich im Zeichensystem der Empfindungen ausdrückende gesetzmäßige Beziehung zwischen dem auf unsere Sinne Wirkenden, das „hinter dem Wechsel ursprünglich Bleibende und Bestehende“ (Helmholtz, 1896/1998, S. 169) könne vom Wahrnehmenden hingegen unverfälscht erfahren werden unter Voraussetzung der Gültigkeit des Kausalgesetzes. Dadurch kommt nun auch dem Begriff des Wirklichen bei Helmholtz eine ganz bestimmte Bedeutung zu. Es soll nämlich genau nur diese sich in den Empfindungen ausdrückende Gesetzmäßigkeit des Wirkenden auf den Wahrnehmenden bezeichnen und nicht die sachlichen Bestimmungen oder Eigenschaften dieses Wirkenden.<sup>118</sup>

Wir haben in unserer Sprache eine sehr glückliche Bezeichnung für dieses, was hinter dem Wechsel der Erscheinungen stehend auf uns einwirkt, nämlich: ‚das Wirkliche‘. Hierin ist nur das Wirken ausgesagt; es fehlt die Nebenbeziehung auf die Substanz, welche der Begriff des Reellen, d.h. des Sachlichen einschließt. (Helmholtz, 1896/1998, S. 170)

---

nicht möglich, wie wir gesehen haben, ohne die Anwendung von Induktionsschlüssen, d.h. ohne das Kausalgesetz“ (Helmholtz, 1896/1998, S. 172).

<sup>117</sup>Zu den Auswirkungen der sich wandelnden Auffassung des Kausalitätsgesetzes auf Helmholtz’ Wissenschaftsauffassung siehe Schiemann (1994, 2008).

<sup>118</sup>Über die in diesem Zusammenhang wichtige aber in dieser Arbeit nicht weiter zu behandelnde Verbindung zwischen diesem Aspekt der Helmholtz’schen Wahrnehmungstheorie und dessen Naturauffassung als gesetzmäßiger Veränderungen von *Kräften* siehe Heimann (1974) sowie Schiemann (2014).

Die Zeichen bekommen dadurch innerhalb der Helmholtzschen Theorie eine doppelte Funktion: Einerseits sollen sie das *Ergebnis* einer kausalen Beziehung zwischen unabhängiger Außenwelt und dem wahrnehmenden Subjekt sein, andererseits sollen sie als Zeichen auf die Gesetzmäßigkeit des auf den Wahrnehmenden Wirkenden verweisen, so dass dieser über die Zeichen zu einer Kenntnis dieser Gesetzmäßigkeit gelangen kann.

### 2.6.5. Zusammenfassung

Die Helmholtzsche Zeichentheorie markiert einen Höhepunkt der Theoriebildung im Bereich der visuellen Wahrnehmung. Helmholtz greift die zentralen Aspekte der bisher in diesem historischen Teil dargestellten Theorien auf und versucht, diese auf originelle Weise in einen gemeinsamen und einheitlichen theoretischen Rahmen zu integrieren. Reids Unterscheidung zwischen Wahrnehmung und Empfindung spielt dabei ebenso wie Müllers Theorie der spezifischen Sinnesenergien und Lotzes Theorie der Lokalzeichen eine eminent wichtige Rolle. Die Auffassung der Wahrnehmung als Ergebnis einer Schlussfolgerung, wie sie sich bereits bei Alhazen zeigte, wird in Helmholtz' Konzept der 'Induktionsschlüsse' verbunden mit Berkeleys Ansatz, dass Wahrnehmung von Tiefe auf assoziativen Verknüpfungen beruhe. Die assoziative Verknüpfung von Erfahrungen stellt bei Helmholtz – in Verbindung mit dem Kausalgesetz, dessen Gültigkeit diese assoziativen Verknüpfungen ermögliche – das zentrale Element dar, über das sowohl das Denken, als auch die Wahrnehmung erklärt werden soll. Gleichzeitig versucht Helmholtz über seine Wahrnehmungstheorie an Kant angelehnte erkenntnistheoretische Erwägungen *empirisch* zu rechtfertigen. Die Empfindungen seien Zeichen für die Vorgänge des auf den Wahrnehmenden Wirkenden, da sie, laut der Theorie der spezifischen Sinnesenergien, die für Helmholtz eine *empirische Tatsache* darstellt, mit diesem keinerlei Ähnlichkeit aufweisen müssten, sondern durch die Natur der Sinnesorgane mitbestimmt seien. In der Gesetzmäßigkeit dieser Zeichen bilde sich jedoch die Gesetzmäßigkeit des auf den Wahrnehmenden Wirkenden ab und der Wahrnehmende könne so zu einer Kenntnis dieser Gesetzmäßigkeit, und nur dieser, kommen. Auf die verschiedenen theoretischen Probleme der unterschiedlichen Ausarbeitungen der Helmholtzschen Theorie wurde in den vorangehenden Abschnitten ausführlich eingegangen, es sollen dennoch kurz einige dieser Schwierigkeiten, insbesondere die mit den Hilfsmitteln der Tiefenwahrnehmung verbundenen, abschließend nochmals hervorgehoben werden.

Wahrnehmung müsse laut Helmholtz *erlernt* werden, ein solches Lernen sei nur über aktives und willentliches Experimentieren durch den Wahrnehmenden sowie die voraussetzende Gültigkeit des Kausalgesetzes möglich. Er müsse – ganz nach Vorbild der naturwissenschaftlichen experimentellen Methode – gleichsam Hypothesen generieren und diese durch selbst herbeigeführte Bedingungsvariationen *in actu* testen. Den

Ausgangspunkt der Helmholtzschen Theorie visueller Wahrnehmung stellt die Erregung der optischen Nerven durch Licht dar, diese Erregung führe zu Empfindungen. Das Konzept der ‘visuellen Empfindung’ ist – dies wurde nun mehrfach erläutert – überaus problematisch, nicht nur, weil Helmholtz ständig zwischen der Beschreibung einer solchen als eines rein physiologischen Phänomens (die Nervenregungen *selbst* seien Empfindungen bzw. die Sinnesorgane würden ‘empfinden’) und eines *psychischen* Phänomens (Empfindungen seien erst dann gegeben, wenn die Nervenregungen ‘zu Bewusstsein kommen’) schwankt. Beide Beschreibungen sind mit Schwierigkeiten verbunden. Was es heißen soll, dass die Sinnesorgane (etwa das Auge) selbst ‘empfinden’ würden oder dass Nervenregungen als solche Empfindungen seien, ist schlicht unverständlich. Man erinnere sich wieder an den als paradigmatischen Fall einer Empfindung angesehenen Schmerz. Wenn ich Schmerzen in meiner Hand habe, so ist es nicht meine Hand, die Schmerzen empfindet, sondern ich und was ich empfinde, ist Schmerz und keine Nervenregung in der Hand. Die Erregung bestimmter Nerven ist vermutlich eine kausale Bedingung für das Schmerzempfinden *einer Person*, aber die Nervenregungen selbst sind keine Empfindungen. Bezüglich der zweiten Beschreibung stellt sich das Problem, dass uns diese ‘visuellen Empfindungen’ offenbar gerade nicht ‘zu Bewusstsein kommen’. Wir ‘empfinden’ nichts visuell, sondern wir *sehen Objekte*. Man kann sich zwar in der *Beschreibung* des Gesehenen in eine künstliche Situation versetzen und etwa versuchen, bloße Helligkeitsunterschiede ohne Objektbezug in dem Gesehenen zu beschreiben, aber dies ist eine Beschreibung unserer *Wahrnehmung* und keiner ‘Empfindungen’, die Wahrnehmung laut Helmholtz erst ermöglichen sollen. Helmholtz selbst beschreibt auch nie näher, was eine ‘visuelle Empfindung’ sei und bemerkt darüber hinaus an mehreren Stellen, dass wir keinen bewussten Zugang zu diesen hätten. Von Empfindungen zu reden, die uns nicht bewusst seien, ist aber nicht sinnvoll. Das wesentliche Kennzeichen von Empfindungen ist doch, dass wir sie *haben*. Es ist insgesamt unklar, wie sich das Konzept der Empfindungen auf den visuellen Bereich übertragen lassen soll. Die ‘visuellen Empfindungen’ scheinen in Helmholtz’ Theorie eine Art ‘Zwischenstufe’ zwischen physiologischer Nervenregung und Wahrnehmung im eigentlichen Sinn darzustellen, werden aber nicht verständlich.

In Helmholtz’ Theorie ist das Konzept der ‘visuellen Empfindungen’ jedoch essentiell, soll es doch aus diesen über die ‘unbewussten Schlüsse’ zu visueller Wahrnehmung kommen. In frühen Ausarbeitungen seiner Theorie möchte Helmholtz den Übergang von Empfindungen zur Wahrnehmung äußerer Objekte allein über das Kausalgesetz rechtfertigen, was aus den im Abschnitt über den Vortrag von 1855 ausgeführten Gründen problematisch ist. Ab der Zeit des *Handbuchs* schlägt Helmholtz dann eine kompliziertere, aber auch etwas undurchsichtigere Konzeption vor. Der Wahrnehmende könne durch aktives Experimentieren (verbunden mit der vorauszusetzenden Gültigkeit des

Kausalgesetzes) lernen, von ihm willentlich hervorgebrachte Empfindungsveränderungen von nicht durch ihn hervorgebrachten zu unterscheiden. Diese als nicht von ihm hervorgebracht identifizierten Empfindungsveränderungen würden dann ebenfalls über das Kausalgesetz und einer gewissen ‘Tendenz des Bewusstseins’ als Gegenstände einer äußeren Welt aufgefasst. In beiden Fällen scheint Helmholtz die Existenz einer unabhängig vom Wahrnehmenden existierenden Aussenwelt voraussetzen zu müssen, die er lediglich als Hypothese betrachten wollte. Im Fall des aktiv experimentierenden Wahrnehmenden muss Helmholtz zudem wesentlich mehr ‘kognitive Fähigkeiten’ voraussetzen, als die bloße Fähigkeit zur assoziativen Verknüpfung. Insgesamt scheint – nimmt man Helmholtz’ Ausführungen wörtlich – die Helmholtzsche Auffassung von Wahrnehmung als Ergebnis einer Schlussfolgerung nicht geeignet, unser *Sehen* externer Objekte verständlich zu machen. Das Ergebnis einer Schlussfolgerung ist stets (Helmholtz’ eigene Beispiele weisen sogar darauf hin), *dass* etwas der Fall ist, etwa *dass* sich ein Objekt in bestimmter Entfernung befinde. Was aber durch die Schlusstheorie erklärt werden soll ist, das *Sehen* von Objekten. Der von Helmholtz behauptete Zusammenhang zwischen Empfindungskonstellationen und Kausalgesetz als möglichen ‘Prämissen’ eines solchen Schlusses und der Wahrnehmung externer Gegenstände als ‘Konklusion’ ist unverständlich. Begegnet man Helmholtz’ Ausführungen wohlwollend, indem man sie so auslegt, dass Helmholtz die Wahrnehmung so beschreibt, *als ob* sie das Ergebnis einer Schlussfolgerung sei, bleibt man etwas ratlos zurück, was mit solch einer ‘als-ob-Erklärung’ gewonnen ist. Ein bloß metaphorischer Bezug zwischen dem Sehen von Gegenständen und dem Ziehen von Schlüssen macht unser Sehen nicht begreifbarer, es muss deutlich werden, auf welche Weise die behaupteten Schlussfolgerungen unser phänomenales Erleben konstituieren. Zudem deuten die Texte nicht darauf hin, dass Helmholtz die ‘Induktionsschlüsse der Wahrnehmung’ als bloße Metapher auffassen möchte. Die sogenannten ‘Sinnestäuschungen’ stellen ebenfalls ein ernstzunehmendes Problem in Helmholtz’ Theorie dar, da sie – wie mehrfach erläutert – sich nur schlecht in Übereinstimmung bringen lassen mit Helmholtz’ allgemeiner Auffassung, dass Wahrnehmung wesentlich *erlernt* werde.

Wenden wir uns nun den Hilfsmitteln der Tiefenwahrnehmung im Rahmen der Helmholtzschen Theorie zu. Wie in allen bisher dargestellten Theorien geht auch Helmholtz davon aus, dass Tiefen- oder Entfernungswahrnehmung die ‘Nutzung’ besonderer Hilfsmittel erfordere, da es sich bei dem auf die beiden Retinae projizierten Lichtmuster nur um „perspektivische Flächenabbildungen“ der Objekte handle. Das *Handbuch* enthält die bisher umfangreichste Auflistung dieser Hilfsmittel, wobei viele der aufgelisteten Elemente – etwa die Konvergenz, die Akkommodation, die Verdeckung oder die ‘Größe des Netzhautbildes’ eines Objekts – früheren theoretischen Ansätzen entnommen, an-

dere hingegen, in diesem Sinne, neu sind, insbesondere die binokulare Disparität, der Helmholtz eine besondere Bedeutung bei der Tiefenwahrnehmung zumisst.

Auch bei Helmholtz lässt sich zunächst feststellen, dass diese eine durchaus heterogene Gruppe bilden. Die eher physiologisch charakterisierten Hilfsmittel Akkommodation und Konvergenz lassen sich dabei noch am besten in Helmholtz' allgemeines Schema der assoziativen Verknüpfung zwischen 'Empfindungen des Sehsinns' und Empfindungen des Tastsinns integrieren, es ist aber auch hier, wie bei Berkeley, festzuhalten, dass so das Problem der Tiefenwahrnehmung auf den Tastsinn verschoben wird. Eine Raumwahrnehmung über den Tastsinn muss hier schon vorliegen, damit die der Konvergenz und der Akkommodation zugehörigen Empfindungen mit dieser verknüpft werden können. Wie es zu einer Wahrnehmung der Entfernung oder Tiefe durch den Tastsinn komme (im Rahmen einer empiristischen Theorie müsste diese ja auch erlernt werden) erläutert Helmholtz nicht eigens, manche seiner Ausführungen deuten jedoch, ebenfalls wie bei Berkeley, darauf hin, dass diese etwa durch das Ausstrecken des Armes oder simples 'Zählen' der Schritte geschehe, die benötigt werden, um ein Objekt zu betasten. In diesen Fällen muss der Wahrnehmende sich jedoch selbst bereits als *in einem Raum* mit anderen *Objekten* sich befindend auffassen, also Raumwahrnehmung in gewisser Weise schon gegeben sein. Zudem kann Helmholtz nicht verständlich machen, wie es aufgrund der durch den Tastsinn gleichsam 'gefühlten Entfernung' zum *Sehen* unterschiedlich weit entfernter Objekte kommt.

Andere, aber bereits in den anderen theoretischen Ansätzen aufgezeigte Probleme zeigen sich im Fall der binokularen Disparität, dem Hilfsmittel der 'Größe des Netzhautbildes' eines Objekts und der Verdeckung. Damit diese die ihnen von Helmholtz zugeschriebene Funktion erfüllen können, muss eine personenähnliche Instanz (ein Homunkulus) vorausgesetzt werden, die Zugang zu den retinalen Lichtmustern *als Bilder* der äußeren Objekte hat, um etwa diese als Bilder sich verdeckender Objekte auffassen, sie miteinander 'vergleichen' oder in diesen die 'retinale Größe' eines Objektes identifizieren zu können. Der vorauszusetzende Homunkulus muss gleichsam die retinalen Lichtmuster als Bilder 'sehen' können, der Wahrnehmende selbst sieht die retinalen Lichtmuster nicht und kann sie nicht sehen (oder gar 'empfinden'). Die Beschreibung der binokularen Disparität im *Handbuch* könnte man zwar auch so auslegen, dass 'abstrakte Berechnungen' anhand des retinalen Lichtmusters durchgeführt werden, um „Unterschiede in der Tiefendimension“ zu 'ermitteln', aber auch in dieser Lesart muss eine Instanz vorausgesetzt werden, die diese 'Berechnungen' anhand bestimmter Kriterien vornimmt. Zudem müsste, wie ebenfalls bei den anderen theoretischen Ansätzen bereits bemerkt, nun erstens das 'Sehen' des Homunkulus erklärt und zweitens eine Verbindung zwischen dem 'Sehen' des Homunkulus (oder den vorgenommenen 'Berechnungen') und dem Sehen des Wahrnehmenden hergestellt werden. Diese Hilfsmit-

tel zur Tiefenwahrnehmung scheinen in Helmholtz' Theorie weder den Empfindungen, noch den Wahrnehmungen zuzuordnen zu sein. Sie können nicht bloße 'visuelle Empfindungen' sein, da sie auf bereits Wahrgenommenes bezogen sind, gleichzeitig kann es sich in Helmholtz' Ansatz bei diesen auch nicht um Wahrnehmungen handeln, da die Hilfsmittel eine Wahrnehmung erst ermöglichen sollen.

Die Helmholtzsche Wahrnehmungstheorie kann als klassische Formulierung des sogenannten *Inferenzparadigmas* gelten, das bis heute die Wahrnehmungspsychologie dominiert (Ben-Zeev, 1987), wenn auch in anderer Ausgestaltung. Verschiedene Wahrnehmungstheoretiker, etwa Gregory (1998, 2002), knüpfen explizit an die Helmholtzsche Theorie an, der Mainstream der aktuellen visuellen Wahrnehmungspsychologie lässt sich allerdings als eine Verbindung aus der prinzipiellen Ausrichtung der Helmholtzschen Theorie und dem sogenannten *komputationalen Ansatz* der Wahrnehmung kennzeichnen. Da der komputationale Ansatz auch in aktuellen Formulierungen des Konzepts der Tiefencues eine eminent wichtige Rolle spielt, soll dieser, den historischen Teil dieser Arbeit abschließend, im Folgenden Gegenstand der Behandlung sein.

## 2.7. Der Komputationale Ansatz: David Marr

Verschiedene theoretische und technische Entwicklungen wie etwa die Informationstheorie von Shannon und Weaver (1949/1962), die Kybernetik von Wiener (1948/1963), die theoretischen Arbeiten zur Architektur eines Digitalcomputers von von Neumann (1945/1993), die theoretischen Arbeiten zur 'Künstlichen Intelligenz' von Turing (1950) sowie der Bau immer leistungsstärkerer Großrechner führte in der Mitte des letzten Jahrhunderts zum Aufkommen eines neuen Paradigmas (und Menschenbildes) in der psychologischen Theoriebildung, welches bis heute die Kognitionspsychologie in Gänze bestimmt. Die zentrale These dieses *Informationsverarbeitungsparadigmas* lautet, dass mentale Aktivitäten in derselben Relation zum Gehirn stünden, wie Programme zu den Computern, auf denen sie ablaufen. 'Geistige Prozesse' könnten aufgefasst werden als Programme, die auf einer speziellen biologischen Maschine 'ablaufen', dem Gehirn. Ulric Neisser, eine Schlüsselfigur in der sogenannten *kognitiven Wende* der Psychologie formulierte die Situation des Psychologen im Rahmen dieses Paradigmas wie folgt: „The task of a psychologist trying to understand human cognition is analogous to that of a man trying to discover how a computer has been programmed“ (Neisser, 1967/2014, S. 6). Wie stark diese Analogie zwischen Geist und Gehirn auf der einen und Computer und Programm auf der anderen Seite aufzufassen sein soll, ist ein umstrittener Punkt. Vertreter der von Searle (1980) als 'starke künstliche Intelligenz' (*strong AI*) bezeichneten Auffassung halten die Ähnlichkeiten für so umfassend, dass sie angemessen programmierten, sich 'erfolgreich verhaltenden' Maschinen in der Tat mentale Aktivi-

täten zuschreiben wollen, Vertreter der ‘schwachen künstlichen Intelligenz’ (*weak AI*) behaupten hingegen, dass in einem bestimmt umgrenzten Kontext solche Maschinen lediglich mentale Aktivitäten *simulieren* würden, es aber nicht gerechtfertigt wäre, ihnen diese tatsächlich zuzuschreiben.<sup>119</sup>

Die wohl elaborierteste Formulierung einer Theorie visueller Wahrnehmung im Rahmen des Informationsverarbeitungsparadigmas stammt von dem einflussreichen Neuroinformatiker David Marr (1945-1980), der 1980 an Leukämie verstarb, noch bevor er sein Hauptwerk „*Vision*“ vollenden konnte. Für Marr (1982, S. 3) ist Wahrnehmung „first and foremost, an information-processing task“, Sehen müsse als Ergebnis einer Informationsverarbeitung durch das ‘visuelle System’ aufgefasst werden, welches im Wesentlichen aus den beiden Retinae, den optischen Nerven und dem Gehirn als zentraler informationsverarbeitender Komponente bestehe. Um die Funktionsweise eines informationsverarbeitenden Systems (wie des visuellen Systems) angemessen verstehen zu können, müsse es auf drei qualitativ verschiedenen Ebenen analysiert werden, nämlich (i) der komputationalen Ebene, (ii) der Ebene der Repräsentation und des Algorithmus und (iii) der Ebene der Implementierung. Bezüglich der komputationalen Ebene müsse gefragt werden, was das System überhaupt leisten, d.h. welches Problem der Informationsverarbeitung gelöst werden soll. Man müsse klären, „what is the goal of the computation, why it is appropriate, and what is the logic of the strategy by which it is carried out“ (Marr, 1982, S. 25). Beschreibungen auf dieser Ebene seien abstrakt und liegen in der Regel in der Form einer formal-mathematischen Beschreibung der zu lösenden Aufgabe vor. Als simples Beispiel führt Marr eine übliche Registrierkasse an. Was diese im einfachsten Fall leisten müsse, sei die Addition von natürlichen Zahlen. Auf der komputationalen Ebene ließe sich die Funktionsweise einer Registrierkasse durch eine Abbildung der Menge  $\mathbb{N}^2$  in die Menge  $\mathbb{N}$  kennzeichnen, welche je zwei Zahlen ihre Summe zuordnet. Auf der Ebene der Repräsentation und des Algorithmus sei zu klären, wie die auf der komputationalen Ebene gegebene Beschreibung implementiert werden könne, d.h. wie das System die formulierte Aufgabe tatsächlich ‘löse’. Insbesondere sei zu klären, welche Repräsentation für den Input und Output des Informationsverarbeitungsprozesses verwendet wird und welche Algorithmen auf diese Repräsentationen angewendet werden sollen, um von Input zu Output zu gelangen. Im Fall der Registrierkasse könne die Repräsentation etwa durch arabische Ziffern geschehen und der Algorithmus ließe sich durch das jedem aus dem Mathematikunterricht aus der Schule bekannten Verfahren des ‘Addierens mit Übertrag’ kennzeichnen. Der Algorithmus ist offensichtlich abhängig von der Art der Repräsentation; hätte man statt der arabischen Ziffern römische Zahlzeichen oder das Binärsystem gewählt, müsste auch der Algorith-

---

<sup>119</sup>Für einen Überblick über diese beiden Positionen siehe Searle (1980) sowie die sich an diesen Aufsatz anschließenden Kommentare.

mus ein anderer sein. Auf der Ebene der Implementierung schließlich sei zu klären, wie die Repräsentation und der Algorithmus *physikalisch realisiert* seien, es müsse gefragt werden nach der ‘Hardware’ des Systems, der Natur der dieses konstituierenden Teile und deren physikalischer Funktionsweise.

Wie soll der komputationale Ansatz zu einem besseren Verständnis visueller Wahrnehmung beitragen? In Marrs Arbeiten lässt sich eine gewisse Ambivalenz bezüglich der Frage feststellen, wozu der komputationale Ansatz genau dienen soll. Einerseits entsteht der Eindruck, dass Marr eine Theorie oder ein Modell entwickeln möchte, welche das (menschliche) Sehen verständlich machen soll (siehe etwa die Einleitung in Marr, 1982). Andererseits könnte man auch den Eindruck bekommen, dass Marr einen Beitrag zur Entwicklung künstlicher Sehsysteme leisten möchte, die in einem bestimmten Kontext ‘erfolgreiches Verhalten’ zeigen sollen. Die Beschreibungen Marrs erwecken aufgrund der verwendeten, deutlich mentalistisch geprägten Begriffe jedoch den Eindruck, dass es ihm um mehr als nur ‘erfolgreiches Verhalten’ geht:

What does it *mean*, to see. The plain man’s answer (and Aristotle’s, too) would be, to *know* what is where by looking. In other words, vision is the process of *discovering* what is present in the world, and where it is [sämtliche Hervorhebungen von mir]. (Marr, 1982, S. 3)

Die vom visuellen System zu ‘lösende’ Aufgabe sei laut Marr (1982, S. 3) „to extract from images the various aspects of the world that are useful to us“, oder, in anderer Formulierung „to reliably derive properties of the world from images of it“ (S. 23).<sup>120</sup> Der Input dieses Informationsverarbeitungsprozesses seien die auf die Retinae projizierten Lichtmuster, der Output sei eine *explizite symbolische Beschreibung* dessen, „what is present in the world, and where it is“ (Marr, 1982, S. 3). Marrs Ansatz verzichtet vollkommen auf das Konzept der Empfindungen, er beschreibt Wahrnehmung als die Verarbeitung abstrakter, in den retinalen Lichtmustern enthaltener ‘Informationen’, deren Ergebnis eine explizite symbolische Beschreibung der Umwelt sein soll. Jeder Informationsverarbeitungsprozess lasse sich laut Marr als Transformation von Repräsentationen kennzeichnen, komputationale Prozesse erforderten stets ein symbolisches Medium, in dem sich diese vollziehen könnten. Das visuelle System operiere mit ‘internen Symbolen’ oder ‘internen Repräsentationen’, die verschiedene Eigenschaften des retinalen Lichtmusters repräsentierten, um durch sukzessive Transformation dieser zu einer symbolischen Beschreibung der Form, Orientierung und Distanz von Objekten der Umwelt zu gelangen.

---

<sup>120</sup>Aus epistemologischer Perspektive lässt sich Marr demnach als Realist kennzeichnen, er geht davon aus, dass eine unabhängig vom Wahrnehmenden existierende Umwelt und dieser zugehörige Objekte existieren. Insgesamt spielen epistemologische Fragen im Vergleich etwa zu Helmholtz’ Theorie bei Marr keine besondere Rolle; eine Haltung, die sich auch in der aktuellen Wahrnehmungspsychologie zeigt.

Dieser Prozess vollziehe sich in drei unterscheidbaren Stufen. Zunächst würden die im retinalen Lichtmuster enthaltenen Informationen in einem sogenannten *retinal image* repräsentiert. Dieses lasse sich als eine Wertematrix auffassen, welche die punktuellen absorbierten Lichtintensitäten des retinalen Lichtmusters repräsentiere. Physikalisch realisiert sei diese Repräsentation durch das mit der Absorption durch die Photorezeptoren einhergehende neuronale Aktivitätsmuster. Aus diesem *retinal image* ‘konstruiere’ das visuelle System anschließend den sogenannten *primal sketch*, welcher bestimmte Intensitätsunterschiede und deren geometrische Organisation im *retinal image* expliziere. Ein beträchtlicher Teil der Analysen von Marr ist den diesem Prozess zugrundeliegenden, aufwendigen Algorithmen gewidmet, auf die jedoch nicht weiter eingegangen werden soll, da hier der allgemeine konzeptuelle Rahmen der Theorie Marrs im Vordergrund der Darstellung stehen soll, für die eine genaue Kenntnis dieser Algorithmen nicht vonnöten ist.<sup>121</sup> Dieser *primal sketch* repräsentiere die Umgrenzungen oder Oberflächen der Objekte der Umwelt. Der Übergang vom *retinal image* zum *primal sketch* sei auf komputationaler Ebene nur mit Hilfe gewisser Annahmen über die Umwelt möglich, etwa dass Objekte sich in ihrer Form nicht veränderten und bestimmte Regelmäßigkeiten bezüglich ihrer Oberfläche aufwiesen. Diese Annahmen (oder *constraints*) müssten in das visuelle System implementiert sein. Anschließend werde aus dem *primal sketch* der sogenannte  $2\frac{1}{2}$ -*D-sketch* ‘konstruiert’, der eine beobachterzentrierte („viewer-centered“) oder standpunktabhängige symbolische Beschreibung der Oberflächenorientierung, Konturen, Entfernung und Tiefe der Objektoberflächen der Umwelt sei. Abbildung 2.3 auf der nächsten Seite zeigt beispielhaft eine Darstellung eines solchen  $2\frac{1}{2}$ -*D-sketches* aus Marr (1982).

Der  $2\frac{1}{2}$ -*D-sketch* werde schließlich transformiert in eine standpunktunabhängige, ‘objektzentrierte’ („object-centered“) dreidimensionale Repräsentation der Objekte der Umwelt, indem (vereinfacht gesprochen) die Elemente des  $2\frac{1}{2}$ -*D-sketches* mit einem im visuellen System ‘gespeicherten Katalog’ an dreidimensionalen Modellen ‘abgeglichen’ werde. Die theoretische Erläuterung des Übergangs vom  $2\frac{1}{2}$ -*D-sketch* zur dreidimensionalen Repräsentation fällt im Vergleich zu den anderen behaupteten ‘Informationsverarbeitungsstufen’ allerdings eher fragmentarisch aus.<sup>122</sup> Marrs Analysen bezüglich der ‘Konstruktion’ des *primal sketch* aus dem *retinal image* sind voller Ausführlichkeit und technischer Brillanz, je weiter man jedoch auf den von Marrs Theorie postulierten Stufen des Wahrnehmungsprozesses voranschreitet – insbesondere beim Übergang vom  $2\frac{1}{2}$ -*D-sketch* zu einer expliziten symbolischen Beschreibung dreidimensionaler Objekte

<sup>121</sup>Der interessierte Leser sei auf Marr (1980, 1982) sowie Marr und Poggio (1979) oder auf die kompakte Darstellung von Poggio (1981) verwiesen.

<sup>122</sup>„Unfortunately, this is an aspect of the Marr ... model that has not been developed in much detail“ (Pinker, 1984, S. 22).

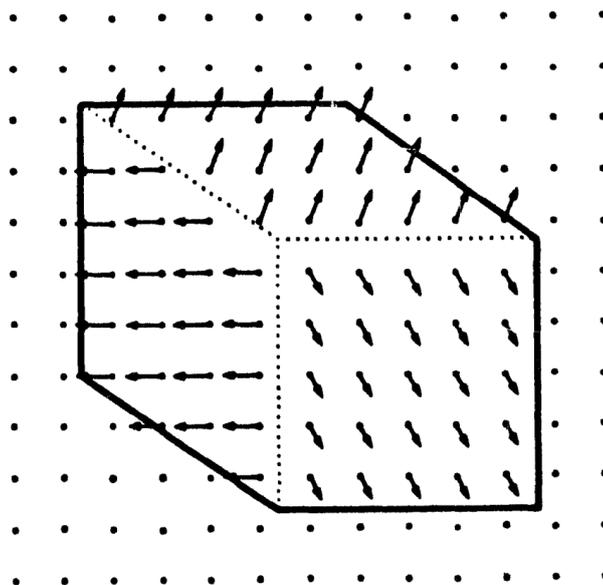


Abbildung 2.3. Illustration eines  $2\frac{1}{2}D$ -sketches eines Quaders vor einem zum Betrachter senkrechten Hintergrund. Die Oberflächenorientierung relativ zum Betrachter werde in diesem Beispiel symbolisch durch Pfeile repräsentiert. Die gepunktete Linie repräsentiere diskontinuierliche Übergänge der Oberflächenorientierung („ridges“), die durchgezogene Linie repräsentiere diskontinuierliche Intensitätsveränderungen (die Kanten des Quaders). Abbildung aus Marr (1982).

und deren räumlicher Relationen – umso stärker schwindet diese Ausführlichkeit und Detailliertheit der Darstellung, wie Marr (1982) an mehreren Stellen selbst eingesteht.

In Marrs Ansatz stellt die Tiefenwahrnehmung kein eigenständiges oder besonders hervorzuhebendes Problem dar, sie ist in den gesamten Informationsverarbeitungsprozess vom *retinal image* zur symbolischen Beschreibung der Umwelt eingebettet. Marr spricht sehr selten explizit von Tiefencues (eine Ausnahme stellt etwa Marr, 1980, S. 208 dar, wo von „*perspective cues*“ gesprochen wird), verschiedene in der aktuellen Literatur als Tiefencues bezeichnete Elemente tauchen jedoch an mehreren Stellen in Marrs Arbeiten als ‘Informationsquellen’ für den visuellen Verarbeitungsprozess auf. Insbesondere bei der Konstruktion des  $2\frac{1}{2}D$ -sketches würden viele dieser ‘Informationsquellen’ gemeinsam genutzt, in dem die ‘Informationen’ aus diesen zunächst in verschiedenen unabhängig voneinander arbeitenden ‘Modulen’ getrennt gewonnen und anschließend ‘kombiniert’ würden.<sup>123</sup> Wir werden im nächsten Kapitel sehen, dass diese Ansicht die Grundlage für die heutigen Modelle der sogenannten *cue integration* darstellt. Der Grad der Detailliertheit, mit der Marr die einzelnen Tiefencues behandelt, variiert stark. Im Fall der binokularen Disparität etwa geht Marr (1982) in aller Ausführlichkeit auf das damit verbundene komputationale Problem und entsprechende Algorithmen ein, die Verdeckung wird hingegen lediglich als ‘Informationsquelle’ erwähnt.

<sup>123</sup> „To make the most efficient use of different and often complementary channels of information deriving from stereopsis, from motion, from contours, from texture, from colour, from shading, they need to be combined in some way“ (Marr, 1980, S. 208).

Bemerkenswert ist, dass Marr an mehreren Stellen der Beschreibung der ‘höheren Verarbeitungsprozesse’ auf ein Vokabular zurückgreift, das Ähnlichkeiten zu Helmholtz’ Formulierung der Theorie der unbewussten Schlüsse aufweist. Im Zusammenhang mit der Verdeckung als ‘Informationsquelle’ wird etwa davon gesprochen, dass das Gehirn, um diese ‘Informationen’ nutzen zu können, bestimmte Annahmen über die Umwelt machen müsse, „that we unconsciously employ“ (Marr, 1982, S. 219). Zudem wird visuelle Wahrnehmung an mehreren Stellen von Marr im Kern als das Ergebnis einer *Schlussfolgerung* beschrieben: „The true heart of visual perception is the inference from the structure of an image about the structure of the real world outside“ (Marr, 1982, S. 68). Was Helmholtz mit dem Begriff der Schlussfolgerung umschreibt, wird bei Marr dadurch konkretisiert, dass komputationale Prozesse als Schlussfolgerungen zu charakterisieren seien.<sup>124</sup> Ein Unterschied zwischen Helmholtz’ und Marrs Theorie besteht allerdings darin, dass Helmholtz der Erfahrung einen immensen Stellenwert einräumt. Marr geht nicht explizit darauf ein, auf welche Weise das Gehirn zu diesen für den Prozess des Schlussfolgerns notwendigen Annahmen gelangt sei, es lässt sich jedoch vermuten, dass Marr diese als in das visuelle System gleichsam ‘eingebaut’ auffassen wollte (Johnson-Laird, 1988).

Eine aktuell stark verbreitete Spielart des komputationalen Ansatzes ist der sogenannte *Bayes-Ansatz* der visuellen Wahrnehmung, in dessen Rahmen Wahrnehmung als ein probabilistischer Prozess angesehen wird (etwa Knill & Richards, 1996; Kersten & Schrater, 2002; Kersten & Yuille, 2003; Mamassian, Landy & Maloney, 2002). Die diesem Ansatz zugehörigen Vorstellungen lassen sich in Marrs Klassifikationsschema der Ebene des Algorithmus und der Repräsentation zuordnen. Die Funktionsweise des visuellen Systems wird in diesem Ansatz so beschrieben, als ob dieses das Bayes-Theorem ‘verwenden’ würde, um für ein vorliegendes retinales Lichtmuster zugehörige bedingte Wahrscheinlichkeiten für verschiedene mögliche dreidimensionale Szenen zu ‘berechnen’, die dieses Lichtmuster verursacht haben könnten. Vereinfacht dargestellt wird angenommen, dass für ein bestimmtes vorliegendes retinales Lichtmuster  $L$  die mit *allen möglichen* dreidimensionalen Szenen  $S_i$  verbundenen bedingten Wahrscheinlichkeiten  $P(S_i | L)$  ‘berechnet’ werden, die sich nach dem Bayes-Theorem folgendermaßen

---

<sup>124</sup>Fodor und Pylyshyn (1981, S. 140), ebenfalls Vertreter des Informationsverarbeitungsparadigmas, argumentieren ebenfalls, dass sich eine jegliche Schlussfolgerung als ein komputationaler Prozess auffassen lasse und daher jede Theorie visueller Wahrnehmung, die diese als ein Ergebnis einer Schlussfolgerung betrachte, eine komputationale Theorie sei: „The psychological mechanism of inference is the transformation of mental representations, it follows that perception is in relevant aspects a computational process“. Ähnlich äußert sich Palmer (1999, S. 71): „The computer analogy is quite compatible with the inferential analogy of constructivism because making inferences is, in effect, what computers do when they execute programs“.

ergeben:

$$P(S_i | L) = \frac{P(L | S_i) \cdot P(S_i)}{P(L)}, \quad (2.3)$$

wobei  $P(L | S_i)$  die bedingte Wahrscheinlichkeit des Vorliegens des Lichtmusters  $L$  unter der dreidimensionalen Szene  $S_i$ ,  $P(S_i)$  die Wahrscheinlichkeit des Vorliegens von  $S_i$  und  $P(L)$  die Wahrscheinlichkeit des Vorliegens des Lichtmusters  $L$  bezeichne. Was dann bei Vorliegen eines bestimmten retinalen Lichtmusters wahrgenommen werde, sei jene dreidimensionale Szene, deren bedingte Wahrscheinlichkeit  $P(S_i | L)$  am größten ist. Es ist offensichtlich, dass diese bedingten Wahrscheinlichkeiten nur dann berechnet werden können, wenn sämtliche Größen auf der rechten Seite der Gleichung (2.3) *bekannt* sind. Im Rahmen des Bayes-Ansatzes der visuellen Wahrnehmung wird üblicherweise angenommen, dass das visuelle System *a priori* über ein solches Wissen verfüge (Knill & Richards, 1996), dieses Wissen aber auch durch Erfahrungen Veränderungen unterworfen sei (Knill, 2007b).

Der Bayes-Ansatz lässt sich noch stärker in Beziehung zur Helmholtzschen Theorie der unbewussten Schlüsse bringen, da in diesem explizit davon gesprochen wird, dass das visuelle System über die Ermittlung der bedingten Wahrscheinlichkeiten  $P(S_i | L)$  ‘schlussfolgere’, welches die wahrscheinlichste ‘Interpretation’ des vorliegenden retinalen Lichtmusters sei.<sup>125</sup> Es liegen zahlreiche experimentelle Untersuchungen vor, bei denen sich das Verhalten der Versuchspersonen sehr gut mit Hilfe des Bayes-Modells *beschreiben* lässt. Ein Problem besteht jedoch darin, dass über die Größe der Elemente auf der rechten Seite der Gleichung (2.3) vor der experimentellen Untersuchung in der Regel keine allgemeinen Aussagen getroffen werden, sondern diese anhand der empirischen Ergebnisse erst geschätzt werden. Dieses Vorgehen birgt die Gefahr, dass sich möglicherweise *jedes* empirische Ergebnis durch das Bayes-Modell beschreiben lässt, solange diese Größen *ad hoc* gewählt werden. Der Bayes-Ansatz wäre dann letztlich nicht mehr als eine Methode der Beschreibung erhaltener Daten unter einem bestimmten formalen Modell, hätte aber so gut wie keine prädiktive Kraft (Colombo & Seriès, 2012).

Eine Schwierigkeit, die sich beim Lesen der Texte von Vertretern des komputationalen Ansatzes einstellt ist, dass häufig die Ebene einer rein formal-mathematischen Beschreibung oder Modellierung verlassen zu werden scheint, indem darauf verwiesen wird, dass das visuelle System oder das Gehirn mit ‘internen Repräsentationen’ operiere und die in der mathematischen Beschreibung enthaltenen ‘Berechnungen’ tatsächlich ‘durchführe’. So lässt sich etwa auch der Verbrennungsvorgang in einem Motor oder

---

<sup>125</sup>So etwa bei Kersten und Yuille (2003, S. 150): „The visual cortex arrives at a simple and unambiguous interpretation of the data from the retinal image.“

der Verdauungsvorgang des Menschen durch Algorithmen *beschreiben* und *simulieren*, man hätte in diesem Fall aber sicher Zweifel an der Beschreibung, dass der Motor oder der Magen ein Problem der Informationsverarbeitung zu ‘lösen’ habe, etwas ‘intern repräsentiere’ und sich bestimmter Algorithmen ‘bediene’. Von besonderer Wichtigkeit für die im nächsten Kapitel vorzunehmenden Analysen aktueller Formulierungen des Konzepts der Tiefencues, in denen wesentlich auf den komputationalen Ansatz zurückgegriffen wird, ist, dass Marr den *gesamten* Prozess der Informationsverarbeitung dem visuellen System bzw. dem Gehirn zuschreibt. Das Gehirn ‘analysiere’ und verarbeite die im *retinal image* enthaltenen ‘Informationen’, es operiere mit ‘internen Symbolen’ oder ‘internen Repräsentationen’ und es ‘konstruiere’ schließlich die explizite dreidimensionale Beschreibung der Umwelt.<sup>126</sup> Oder, um den Begriff der Schlussfolgerung wieder aufzugreifen, das Gehirn selbst nehme „the inference from the structure of an image about the structure of the real world outside“ vor. Ähnliche Bemerkungen lassen sich auch bezüglich vieler Vertreter des Bayes-Ansatzes machen, deren Beschreibungen den Eindruck erwecken, dass das Gehirn Wahrscheinlichkeiten ‘berechne’ und anhand dieser Wahrscheinlichkeiten ein Perzept ‘konstruiere’ oder ‘generiere’. Die Frage ist, ob dies einen sinnvollen konzeptuellen Rahmen visueller Wahrnehmung darstellt, der dazu beitragen kann, Wahrnehmung verständlich zu machen.<sup>127</sup> In diesem Zusammenhang ist das abstrakte Konzept der ‘Informationen’, die vom Gehirn verarbeitet würden, sowie das damit zusammenhängende Konzept der ‘internen Repräsentationen’, mit denen das Gehirn ‘operiere’, näher zu beleuchten. Nicht zuletzt besteht auch die Frage, insofern es sich hier nicht nur um eine formal-mathematische Beschreibung handeln sollte, in welchem Zusammenhang das behauptete Operieren mit ‘internen Repräsentationen’ und die ‘Konstruktion’ einer symbolischen Beschreibung der Umwelt *durch das Gehirn* mit dem *Sehen des Wahrnehmenden* stehen soll. Da die im nächsten Kapitel näher zu behandelnden aktuellen Formulierungen der Tiefencues dem konzeptuellen Rahmen des komputationalen Ansatzes zuzuordnen sind, soll eine weitere Diskussion an dieser Stelle unterdrückt und auf das nächste Kapitel verschoben werden.

---

<sup>126</sup>Um nur einige Beispiele zu nennen, welche Fähigkeiten Marr dem Gehirn zuschreibt: „Our brains usually receive similar images of a scene“ (Marr, 1982, S. 100), „the brain is capable of measuring the disparity and using it to create the sensation of depth“ (Marr, 1982, S. 100), „the brain must construct three-dimensional representations of objects and of the space they occupy“ (Marr, 1982, S. 326).

<sup>127</sup>Man erinnere sich, dass ähnliche Fragen in den bereits dargestellten Theorien schon mehrfach auftauchten, insbesondere bei Müller, Lotze und Helmholtz, die dem ‘Sinnesapparat’, dem Bewusstsein oder der Vorstellung ähnliche Fähigkeiten zuschreiben (etwa einen „Vergleich der Netzhautbildchen“ vorzunehmen).



### 3 | Der Begriff des Tiefencues in der heutigen Wahrnehmungsforschung

„Scientific psychology, as the well known saying goes, having first lost its soul, later its consciousness, seems finally to lose its mind altogether.“

---

*Feigl (1958/1967, S. 1)*

Das *cue*-Konzept ist – die in der Einleitung dieser Arbeit angeführten Zahlen zur Auftretenshäufigkeit dieses Terminus in einschlägigen wissenschaftlichen Fachartikeln machen dies deutlich – aus der aktuellen wahrnehmungspsychologischen Literatur nicht mehr wegzudenken. In diesem Kapitel sollen, auf den im vorangegangenen Kapitel anhand der Entwicklung des Konzepts der Tiefencues herausgearbeiteten Probleme aufbauend, aktuelle Verwendungsweisen des *cue*-Konzepts aus eher systematischer Perspektive beleuchtet werden. Das *cue*-Konzept ist historisch aus der Analyse der Tiefenwahrnehmung erwachsen und später dann auch auf weitere Bereiche der Wahrnehmungsforschung angewendet worden. Da nicht nur die Verwendungslogik, sondern auch die mit dieser Logik verbundenen theoretischen Probleme sich schon in aller Deutlichkeit in der Analyse der Verwendungsweisen des Terminus „Tiefencue“ aufzeigen lassen, ist es gerechtfertigt, im Folgenden die Diskussion auf den theoretischen Status des Konzepts der Tiefencues zu beschränken.

In diesem Kapitel soll zunächst in Abschnitt 3.1 dargestellt werden, auf welche Weise dieser Terminus in der aktuellen Literatur zur Tiefenwahrnehmung verwendet wird. Zu diesem Zweck sollen aktuelle Standardlehrbücher, Monographien zur Tiefenwahrnehmung sowie Fachartikel, die sich explizit dem Thema der Tiefenwahrnehmung widmen und auf Tiefencues als Erklärungshilfen zurückgreifen, herangezogen werden. Anschließend ist zu erörtern, ob sich aus den unterschiedlichen Verwendungsweisen dieses Terminus ein gemeinsamer Bedeutungskern herauschälen lässt. Darüber hinaus wird die Frage zu untersuchen sein, welche impliziten Annahmen jenen Verwendungsweisen innewohnen. Wie sich zeigen wird, spielen in aktuellen Charakterisierungen der Tiefencues die Termini „Information“ und „interne Repräsentation“ eine zentrale Rolle.

Diesen beiden Konzepten und deren Verwendungsweisen im Rahmen des *cue*-Konzepts wird in Abschnitt 3.2 näher nachgegangen werden. Dabei soll auch hier untersucht werden, mit welchen impliziten Annahmen diese Konzepte im Rahmen der aktuellen Wahrnehmungspsychologie verbunden sind. In Abschnitt 3.3 soll schließlich aufgrund der herausgearbeiteten Erkenntnisse dafür argumentiert werden, dass das Konzept der Tiefencues aufgrund einer deutlichen Nähe zum sogenannten *Homunkulus-Trugschluss* logische Mängel aufweist.

### 3.1. Aktuelle Charakterisierungen des Konzepts der Tiefencues

Wie zu Ende des letzten Kapitels angedeutet wurde, lässt sich der allgemeine theoretische Rahmen, in dem visuelle Wahrnehmung aktuell üblicherweise behandelt wird, als eine Variante des Inferenzparadigmas mit spezifisch komputationaler Ausrichtung auffassen. Das von den Objekten der Außenwelt reflektierte und ins Auge gelangende Licht, das retinale Lichtmuster, stelle den ‘visuellen Input’ dar, der im Vergleich zum ‘Output’, der Wahrnehmung räumlich ausgedehnter Objekte bestimmter Form, Farbe und Lage im Raum, wesentlich *unterbestimmt* sei.<sup>1</sup> Es bestehe gleichsam eine ‘Kluft’ zwischen dem visuellen Reiz und unserer Wahrnehmung. Die Überbrückung dieser Kluft soll theoretisch dadurch verständlich gemacht werden, dass Wahrnehmung als abstrakte ‘Problemlösetätigkeit’ des visuellen Systems (oder des Gehirns) aufgefasst wird. Diese wird je nach Autor als ein Prozess der Informationsverarbeitung im Sinne Marrs, als ‘Interpretation’ des retinalen Lichtmusters oder auch als ein Prozess des ‘Schlussfolgerns’ (nicht selten mit Verweis auf die Helmholtzsche Theorie) beschrieben; häufig werden diese Beschreibungen synonym verwendet.

Eine Behandlung der Tiefencues wird in der aktuellen Fachliteratur üblicherweise mit der Frage motiviert, wie es zu erklären sei, dass wir räumlich wahrnehmen, obwohl die auf die Retinae projizierten Lichtmuster (die meisten Autoren sprechen von einem *retinal image*) lediglich ‘zweidimensional’ seien. Zur Beantwortung dieser Frage wird dann unmittelbar auf die Tiefencues verwiesen; folgende, Fachartikeln entnommene Zitate mögen dies exemplarisch verdeutlichen:

An enduring question in the study of visual perception is how we manage to perceive depth ... from retinal images that are essentially flat, having only two dimensions. One answer is that retinal images contain cues to depth, such as retinal disparity, movement parallax, superimposition and familiar size. (O’Shea et al., 1994, S. 1595)

<sup>1</sup>Dies wird gelegentlich auch als „*poverty of the stimulus thesis*“ bezeichnet (Richeimer, 2006).

The fundamental problem in depth perception is due to the geometry of perspective projection, which reduces the three-dimensional (3D) coordinates of the visual scene to the 2D coordinates of the retinal images. The third dimension of space has to be inferred from the 2D images. The visual system uses several sources of information – depth cues such as disparity, perspective, and motion parallax – to estimate the layout of the 3D scene. (Hillis et al., 2004, S. 967)

The human visual system expends much effort generating a stable three-dimensional percept from two-dimensional retinal images. The visual system can estimate depth from many monocular depth cues, such as perspective, texture-gradient, and occlusion for static images; and motion parallax for moving ones. (Troscianko et al., 1991, S. 1923)

Bemerkenswert ist, wie verschieden die Akzentsetzungen dieser drei Beschreibungen sind. Bei O’Shea et al. (1994) soll das *cue*-Konzept erklären, wie wir es anstellen, Tiefe *wahrzunehmen*. Laut der Ausführungen von Hillis et al. (2004) sollen die Tiefencues zur Erklärung beitragen, auf welche Weise *das visuelle System* die räumlichen Verhältnisse der Umgebung *schätzt*, indem es anhand dieser *cues* diesbezügliche ‘Schlussfolgerungen’ ziehe. Bei Troscianko et al. (1991) schließlich soll über das *cue*-Konzept verständlich gemacht werden, wie das visuelle System über eine *Schätzung der Tiefe* ein dreidimensionales, stabiles Perzept ‘*generiert*’. Diese drei einführenden Beispiele mögen schon andeuten, dass in der aktuellen Literatur das Konzept der Tiefencues und dessen Funktion innerhalb der theoretischen Vorstellungen zur Tiefenwahrnehmung mit nicht deckungsgleichen Vorstellungen verbunden ist – abgesehen von der generellen Vorstellung, dass die Tiefencues die Wahrnehmung von Tiefe ermöglichen sollen. Eine ähnliche Feststellung wurde bereits bezüglich der im letzten Kapitel dargestellten Theorien getroffen, in erster Linie in Bezug auf die Frage, ob die Tiefencues Wahrnehmung oder Wahrnehmungsurteile verständlich machen sollen.

Die Fachartikeln zugrundeliegenden theoretischen Überzeugungen werden in der Regel durch Wendungen umschrieben, wie etwa, dass das visuelle System sensorischen Input ‘interpretiere’, zu ‘Schlussfolgerungen’ über die Umwelt gelange, ‘Tiefe generiere’, ‘Berechnungen anstelle’, ‘Tiefe ermittle’ oder schlicht Tiefencues ‘nutze’. Viele Arbeiten scheinen eher dem komputationalen Ansatz verpflichtet, was sich unter anderem dadurch ausdrückt, dass bestimmte Algorithmen für die Lösung des behandelten Problems entwickelt und dargestellt werden. Gleichzeitig wird Wahrnehmung in deutlich intentionalen oder geistige Leistungen bezeichnenden Begriffen, wie dem der Interpretation, beschrieben, die sich aber auf das visuelle System oder das Gehirn selbst beziehen sollen. Da es sich bei Fachartikeln in der Regel um kompakt abgefasste, empirische Forschungsberichte im Rahmen sehr spezieller wahrnehmungspsychologischer Fragestellun-

gen handelt, ist es durchaus verständlich, dass hier auf breitere theoretisch-konzeptuelle Überlegungen verzichtet wird. Das Konzept der Tiefencues scheint in Fachartikeln im Wesentlichen als Allgemeingut der *scientific community* behandelt zu werden, welches nicht immer wieder eigens eingeführt werden müsse.

Es ist daher sinnvoll, nach allgemeineren Charakterisierungen dieses Konzepts in der Literatur zu suchen, etwa in Lehrbuchtexten, Übersichtsartikeln oder Monographien zur Tiefenwahrnehmung, in denen eher auch Raum für theoretisch-konzeptuelle Überlegungen gegeben wäre. Aber auch hier wird in der Regel auf eine kritische Diskussion der Tiefencues verzichtet. Konsultiert man aktuelle Lehrbücher der Wahrnehmungspsychologie oder Übersichtsartikel zur Tiefenwahrnehmung, findet man in der Regel sehr knappe und vage Charakterisierungen dieses Konzeptes. Die Vorstellungen, die mit dem Terminus „Tiefencue“ verwoben sind, deuten sich in den folgenden Wendungen an: Tiefencues seien „information about the third dimension (depth) of visual space“ (Wolfe et al., 2009, S. 136), „depth information“ (Palmer, 1999, S. 202), „information in the retinal image that is correlated with depth in the scene“ (Goldstein, 1999, S. 215), „features of visual stimuli that prompt the perception of depth or distance“ (Coren et al., 1999, S. 303), „a source of information regarding depth“ (Levine, 2000, S. 297) oder etwas „which provides information about depth“ (Yantis, 2014, S. 190). In dem dreibändigen Kompendium von Howard (2012), welches ausschließlich der Tiefenwahrnehmung gewidmet ist und als das Standardwerk für diesen Bereich angesehen wird, wird der Begriff eingeführt als „information about depth arising from a specified stimulus feature“ (Howard, 2012, Bd. 1, S. 4), an anderer Stelle als „perceptually equivalent sources of information“ (Howard, 2012, Bd.1, S. 170). In einer anderen Publikation des gleichen Autors (Howard, 2002, S. 78) werden die Tiefencues umschrieben als „sources of information, [which] contribute to the perception of depth“.

Versucht man, möglichst viele Gemeinsamkeiten dieser Charakterisierungen zu finden, so lässt sich zunächst feststellen, dass Tiefencues allgemein als Informationen (oder Informationsquellen) über Tiefe beschrieben werden. In der Regel werden Tiefencues als bestimmte Eigenschaften des ‘visuellen Reizes’ angesehen, die in irgendeiner Form mit den räumlichen Verhältnissen der Außenwelt zusammenhängen und dadurch a) Informationen über Tiefe darstellen, b) mit ‘Tiefe in der Szene’ korreliert sind oder c) zur Wahrnehmung oder ‘Schätzung’ von Tiefe beitragen sollen. Bei vielen Autoren lässt sich zudem das bereits bemerkte Schwanken zwischen der Beschreibung der Tiefencues als Informationen zur *Wahrnehmung von Tiefe* und als Information, die genutzt wird, um *Urteile* bezüglich der wahrgenommenen Tiefe zu fällen, finden.<sup>2</sup> Das Konzept der

---

<sup>2</sup>Howard (2002, S. 78) etwa spricht von den Tiefencues als „sources of information, [which] contribute to the perception of depth“, an anderer Stelle (Howard, 2012, Bd.1, S. 170) hingegen heißt es „judgements of the relative depth between two objects may be based on binocular disparity, perspective, accommodation, and object overlap“.

Tiefencues wird in der aktuellen Literatur fast durchweg über den Terminus der „Information“ charakterisiert. Ein Großteil der ‘theoretischen Last’ dieses Konzeptes scheint an diesem Begriff zu hängen, so dass eine nähergehende Erörterung notwendig scheint. Wir werden diesem Begriff und dessen Funktion im Kontext wahrnehmungstheoretischer Ansätze in diesem Kapitel, insbesondere in Abschnitt 3.2.1, weiter nachgehen, zunächst sollen weitere, in Zusammenhang mit dem Konzept der Tiefencues stehende Beschreibungen aus der wahrnehmungspsychologischen Literatur zusammengestellt werden.

### Klassifizierungsschemata

Den meist knappen Charakterisierungen der Tiefencues als ‘Informationen über Tiefe’ folgt in der Regel eine Auflistung von *Beispielen*, die schließlich in eine Klassifizierung mündet; ein Vorgehen, das uns in der Darstellung von Helmholtz’ *Handbuch* bereits begegnet ist. Auch diese Klassifizierungen sind nicht einheitlich, es lassen sich aber auch hier weitgehend Gemeinsamkeiten feststellen. Das am häufigsten anzutreffende Klassifikationsschema zerlegt die Tiefencues in Hinblick auf gewisse Sehbedingungen in Teilbereiche: Bestimmte Tiefencues, etwa binokulare Disparität oder Konvergenz, erforderten das Sehen mit zwei Augen, andere, etwa Verdeckung oder Texturgradient, könnten auch allein bei monokularer Betrachtung als Information genutzt werden. Wieder andere Tiefencues wie die Bewegungsparallaxe stünden erst bei zusätzlicher Bewegung des Wahrnehmenden oder der betrachteten Objekte zur Verfügung.

So unterscheidet Levine (2000) die Tiefencues in *monokulare* und *binokulare* Tiefencues (zu denen er die binokulare Disparität und die Konvergenz zählt), wobei die monokularen Tiefencues wiederum unterschieden werden in sogenannte *pictorial cues*, *kinetic cues* (Bewegungsparallaxe) und die Akkommodation. Goldstein (1999) differenziert zwischen *oculomotor cues* (Konvergenz und Akkommodation), *pictorial cues*, durch Bewegung induzierte Tiefencues und binokulare Disparität. Foley und Matlin (2010) unterscheiden zwischen monokularen Tiefencues, binokularer Disparität und *eye muscle cues* (Konvergenz und Akkommodation), Coren et al. (1999) allein zwischen *pictorial depth cues* und *physiological depth cues*.

Howard (2012) bietet die wohl umfangreichste Auflistung und differenzierteste Klassifizierung. Es werden insgesamt 20 verschiedene *cues* für Tiefe aufgeführt, welche zunächst in monokulare und binokulare Tiefencues unterteilt werden. Die monokularen *cues* werden anschließend weiter unterschieden in *statische* und *dynamische* Tiefencues. Unter die statischen Tiefencues werden u.a. die Akkommodation und die Verdeckung gezählt, ein dynamischer monokularer Tiefencue sei etwa die Bewegungsparallaxe. Ab-

bildung 3.1 auf der nächsten Seite veranschaulicht die von Howard (2012) vorgeschlagene Klassifizierung der Tiefencues.<sup>3</sup>

Es gibt vielfältige Diskussionen darüber, ob es sich hierbei um eine Auflistung ‘voneinander unabhängiger’ Tiefencues handle oder nicht. So argumentieren etwa Cutting und Vishton (1995) dafür, dass viele der in Abbildung 3.1 aufgelisteten Tiefencues Konglomerate sich bereits in dieser Klassifizierung befindender Tiefencues seien und daher nicht eigenständig aufgelistet werden müssten. Nachvollziehbar ist das etwa im Fall des von Howard (2012) als *linear perspective* bezeichneten Tiefencues, der offensichtlich ein Konglomerat verschiedener ebenfalls aufgelisteter Tiefencues (etwa Texturgradient und Verdeckung) darstellt. Auffallend an diesen Klassifizierungen ist die Heterogenität der aufgelisteten Elemente, eine Feststellung, die bereits mehrfach im Rahmen der Darstellung der historischen Entwicklung des Konzepts der Tiefencues getroffen wurde. Manche der aufgelisteten Tiefencues scheinen auf Eigenschaften der Umwelt zu verweisen (etwa die in Abbildung 3.1 unter *aerial effects* und *lighting* aufgelisteten), andere auf physiologische Korrelate des Sehens (etwa Konvergenz und Akkommodation) und wiederum andere auf Eigenschaften des *retinal image* (etwa *image blur* und *image size*).

Bei vielen der aufgelisteten Tiefencues ist zudem nicht gut ersichtlich, worauf sich diese beziehen sollen, da sie in Wahrnehmung bereits voraussetzenden Begriffen umschrieben zu sein scheinen. Ein gutes Beispiel hierfür ist der Tiefencue der Verdeckung, der im Abschnitt über Helmholtz’ Theorie bereits ausführlich zur Sprache kam. Im Bezug auf das retinale Lichtmuster von sich verdeckenden Objekten (oder im Fall der ‘retinalen Größe’ von der ‘Größe des Bildes des Objektes auf der Retina’) zu sprechen, scheint in gewisser Weise irreführend zu sein. Eine solche Beschreibung ist der *Wahrnehmung* entnommen, im Fall der Verdeckung ist sie an die Wahrnehmung zweier voneinander verschiedener und verschieden im Raum situierter Objekte gebunden, von denen eins das andere *verdeckt*. Man nehme etwa die Ausführungen von Goldstein (1999) zum Tiefencue der Verdeckung:

If one object partially covers another object, as the houses in the foreground of Figure 8.1 cover the houses in the background, the object that is partially covered must be at a greater distance than the object that is covering it.  
(S. 215f.)

Diese Beschreibung scheint vollkommen richtig. Wenn eine Person von einem bestimmten Standpunkt aus feststellt, dass sich zwei Objekte teilweise verdecken, dann muss, das liegt im Sinn des Ausdrucks „verdecken“, die Person davon ausgehen, dass das eine

---

<sup>3</sup>Da ein genaues Verständnis jener aufgelisteten Tiefencues, die im Rahmen des vorangegangenen Kapitels dieser Arbeit *nicht* erörtert wurden, für die weitere Diskussion nicht notwendig ist, soll auf eine ausführliche Darstellung dieser verzichtet werden. Der interessierte Leser sei auf Howard (2012) verwiesen.

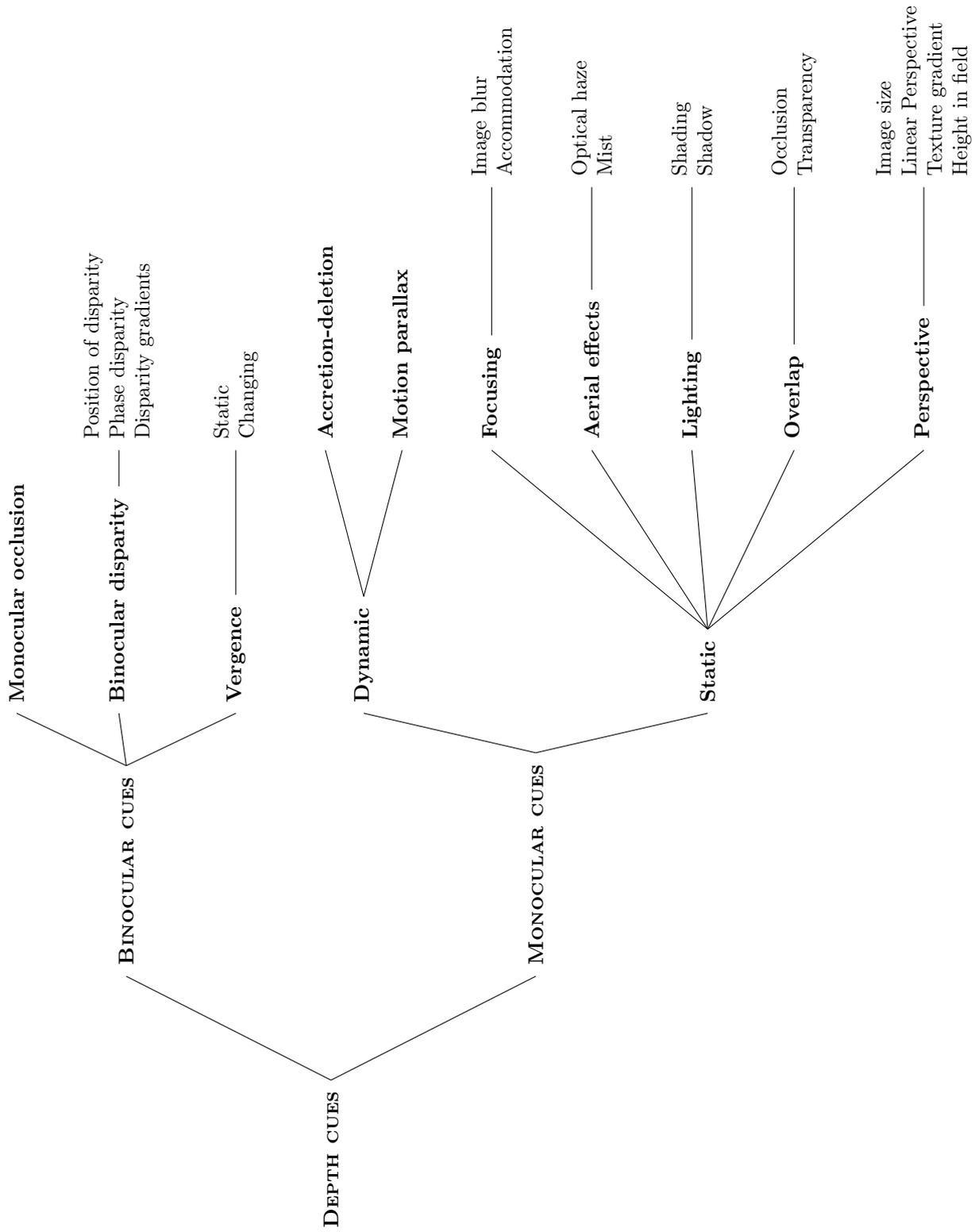


Abbildung 3.1. Klassifizierung der Tiefencues nach Howard (2012).

Objekt weiter entfernt ist als das andere. Eine derartige Feststellung setzt aber doch bereits ein ‘fertiges Perzept’ voraus, es wird davon ausgegangen, dass der Wahrnehmungsprozess in gewisser Weise bereits abgeschlossen ist. Die Person nimmt zwei in ihrer Sehrichtung befindliche Objekte *als* hintereinander lokalisiert *wahr*. Nun soll aber eine notwendige Voraussetzung für die Wahrnehmung verdeckter Objekte das Vorliegen des Tiefencues ‘Verdeckung’ sein. Soll dieser Tiefencue im Perzept, in der Wahrnehmung selbst ‘enthalten’ sein? Das würde in einen logischen Zirkel führen. Oder steht hinter dieser Beschreibung eine Vorstellung der Art, dass der Tiefencue der Verdeckung aus dem retinalen Lichtmuster ‘extrahiert’ wird und so ‘Anlass’ zur Wahrnehmung zweier unterschiedlich weit entfernter Objekte gibt? Wie kann in diesem Kontext die Beschreibung dieses Tiefencues sinnvoll interpretiert werden?

Die Erläuterungen zur Verdeckung als Tiefencue werden in der Regel durch *ein Bild* illustriert, das eine Szene zeigt, in der ein Objekt ein anderes teilweise verdeckt. Ein solches Bild kann nicht dazu dienen, das Perzept zu illustrieren, das wäre logisch nicht belastbar. Es kann aber auch nicht ein ‘retinales Bild’ illustrieren, auf der Retina ist lediglich ein Lichtmuster, eine Energieverteilung aufzufinden. Das retinale Lichtmuster würde allein dann zu einem Bild im üblichen Sinne werden, wenn jemand die Retina betrachten würde und sich so auf der Retina des Betrachters ein Lichtmuster bildete, das dem der betrachteten Retina ähnlich ist. Auch diese Interpretation würde jedoch keine Antwort auf die Frage nach dem Sinn der obigen Beschreibung der Verdeckung als Tiefencue geben. Eine harmlosere Interpretation wäre, dass der Tiefencue der Verdeckung nicht unsere Wahrnehmung erklären soll, sondern unsere Wahrnehmungsurteile. Wenn ich etwa in einer bestimmten Wahrnehmungssituation gefragt werde, ob zwei Objekte unterschiedlich weit von mir entfernt seien, kann ich womöglich meine Aufmerksamkeit stärker auf bestimmte Aspekte meiner Wahrnehmung richten und *feststellen*, dass ich sich verdeckende Objekte *sehe* und dies dann als ein ‘Hinweis’ zur Rechtfertigung meines Urteils ansehen. Beim Lesen der entsprechenden Literatur entsteht jedoch nicht der Eindruck, dass allein dies gemeint sei. Die Tiefencues sollen die *unmittelbare Wahrnehmung* von Tiefe verständlich machen. Die Beschreibungen vieler Tiefencues – insbesondere jene, die sich auf bestimmte Eigenschaften des *retinal image* beziehen sollen – scheinen jedoch mit dem Problem belastet, dass ihnen keine sinnvolle Interpretation gegeben werden kann, die eben dies verständlich macht.

### Weitere Merkmale der Tiefencues

Häufig werden den Tiefencues neben den den Klassifizierungen zugrundeliegenden Kategorien (z. B. monokular / binokular) noch weitere ‘Merkmale’ zugeschrieben. Howard (2012, Bd. 3, S. 147) spricht diesbezüglich von „objective attributes“, die den Tiefencues zukämen. So ließen sich diese auch dahingehend unterscheiden, welcher ‘Art’ die

von ihnen bereitgestellten Informationen seien. Manche Tiefencues, meist als *quantitative* oder *metric depth cues* bezeichnet, würden ‘metrische Informationen’ über die räumlichen Relationen der Umwelt zur Verfügung stellen, andere hingegen, sogenannte *qualitative* oder *ordinal depth cues*, lediglich ‘ordinale Informationen’ (Bertamini, Martinovic & Wuerger, 2008; Burge et al., 2005; Goldstein, 2010; Mather, 2009; Wolfe et al., 2009).<sup>4</sup> Beim Tiefencue der Verdeckung handle es sich beispielsweise um einen ordinalen Tiefencue, da sich diesem nur die Information ‘entnehmen’ lasse, *dass* ein Objekt weiter entfernt sei als ein anderes, aber nicht, *wie weit* diese beiden Objekte entfernt seien (Cutting & Vishton, 1995). Im Falle metrischer Tiefencues, etwa der Konvergenz, soll sich diesen direkt ‘entnehmen’ lassen, wie weit die Objekte der Umwelt entfernt seien. Manche Autoren gehen sogar so weit, zu behaupten, dass manche Tiefencues die *exakte* Entfernung eines Objektes angäben: „Moreover, these cues are the only ones ... that can tell us the *exact* distance to an object“ (Wolfe et al., 2009, S. 146). Die Vorstellung, dass sich etwa im Fall der Konvergenz aus der Stellung der beiden Augen die Entfernung eines fixierten Objektes über Anwendung trigonometrischer Funktionen *berechnen* lasse, geht dabei offenbar auf die im letzten Kapitel dargestellten Überlegungen des Descartes bezüglich einer ‘natürlichen Geometrie’ zurück.

Fragt man sich nun, auf welche Beschreibungsebene die Unterscheidung zwischen metrischen und ordinalen Tiefencues zu beziehen ist, so stellt man zunächst fest, dass dazu eigentlich nicht mehr gesagt wird, als dass sich diese Informationen in den entsprechenden Tiefencues ‘verbergen’ oder in diesen ‘enthalten’ ist. Die in den wahrnehmungspsychologischen Darstellungen mehr oder weniger explizit aufzufindenden Vorstellungen scheinen davon auszugehen, dass auf komputationaler Ebene eine geometrische ‘Berechnung’ stattfindet, deren Resultat schließlich eine Größe, ein Größenverhältnis oder zumindest eine nach der Größe gegebene Anordnung sei. Es sollen physikalisch realisierte Prozesse ablaufen, die vom Ergebnis her *ähnlich* einer von einem Menschen angestellten geometrischen Berechnung seien. Gleichzeitig wird anhand der Beschreibungen nicht deutlich, ob die Tiefencues zur Wahrnehmung von Tiefe oder zur Wahrnehmung von Entfernungen dienen sollen. Die Wahrnehmung von Tiefe geht nicht unmittelbar mit der Wahrnehmung von Entfernung einher, Entfernungen werden *geschätzt* und diese Schätzung ist in der Regel ein neuer, bewusster Akt. Wenn wir über unsere Wahrnehmung befragt werden, so können wir *Urteile* darüber fällen oder Schätzungen abgeben, wie weit in einer festgelegten Einheit ein Objekt von uns wohl entfernt ist, und diese Urteile können mehr oder minder genau sein. Die Beschreibungen der Tiefencues deuten jedoch darauf hin, dass das visuelle System oder das Gehirn Zugang zu den für die Berechnung notwendigen Größen oder ‘Informationen’ habe,

---

<sup>4</sup>„Most of the monocular depth cues are qualitative; that is, they indicate that one object is closer than another, but they do not give a reliable indication of how much closer“ (Goldstein, 2010, S. 51).

diese Berechnungen vornehme und schließlich anhand der berechneten Entfernung, um eine Wendung von Troscianko et al. (1991) wieder aufzugreifen, ein ‘dreidimensionales Perzept generiere’, das, so soll diese Beschreibung vermutlich verstanden werden, mit unserer Wahrnehmung identisch sei. Häufig wird auch davon gesprochen, dass die ‘berechneten Distanzen’ im visuellen System oder im Gehirn ‘intern repräsentiert’ würden. Was damit gemeint sein könnte, soll ebenso in diesem Kapitel näher beleuchtet werden.

Tiefencues ließen sich auch dahingehend unterscheiden, in welchem Intervall von Entfernungen sie ‘effektiv’ oder ‘wirksam’ seien. So sei etwa die Akkommodation als Tiefencue – man erinnere sich an ganz ähnliche Formulierungen der im historischen Teil dargestellten Theorien – nur bei eher geringen Entfernungen wirksam, da sich physiologisch ab einer gewissen Entfernung des fixierten Objektes keine wesentlichen Veränderungen der Linsenkrümmung mehr feststellen ließen. Ähnliches gelte auch für die binokulare Disparität und die Konvergenz, die Verdeckung sei hingegen über ein weit breiteres Intervall an Entfernungen eine ‘Information über Tiefe’. Zudem ließen sich Tiefencues hinsichtlich ihrer ‘Genauigkeit’ kennzeichnen, eine Charakterisierung, die ebenfalls in den im letzten Kapitel dargestellten Theorien fast durchgängig aufzufinden ist. Howard (2012, Bd. 3, S. 148) versteht unter der Genauigkeit eines Tiefencues eine *Zahl*, nämlich „the signed difference between the perceived depth created by the cue and the actual depth in the stimulus“.

Ein weiteres Merkmal der Tiefencues sei deren Validität (Howard, 2012) oder Reliabilität (Epstein, 1995). Manche Tiefencues seien gewissermassen ‘zuverlässigere Informationen’ bzw. ‘zuverlässigere Informationsquellen’ über Tiefe als andere.<sup>5</sup> Ein „fully valid depth cue“ liege dann vor, wenn dieser „only as a function of a change in the distance or relative depth of the distal stimuli“ variere (Howard, 2012, Bd.3, S. 148). Binokulare Disparität oder Konvergenz seien etwa vergleichsweise valide Tiefencues, da im Falle der Konvergenz die Stellung der beiden Augen zueinander allein von der Entfernung des fixierten Objektes abhängt – zumindest in einem bestimmten Intervall

---

<sup>5</sup>Im Gegensatz zu anderen Bereichen der Psychologie, etwa der psychologischen Diagnostik, scheinen im Rahmen des *cue*-Konzeptes die Begriffe „valide“ und „reliabel“ synonym gebraucht zu werden, es lässt sich zumindest keine unterschiedliche Verwendungsweise erkennen. Das Konzept einer (ökologischen) Validität von Tiefencues geht zurück auf Egon Brunswik, der darunter eine quantitative und objektive Erfassung der ‘Zuverlässigkeit’ der Tiefencues als potentielle Indikatoren der räumlichen Verhältnisse der Umwelt eines Lebewesens verstehen wollte. Für Brunswik (1956) ist Wahrnehmung prinzipiell *probabilistischer Natur*; die wahrgenommenen räumlichen Verhältnisse seien gleichsam als die aufgrund der momentan vorliegenden Tiefencues ‘wahrscheinlichste Repräsentation’ der räumlichen Verhältnisse der Umwelt anzusehen. Die Validität eines *cues* soll die ‘Stärke der Beziehung’ zwischen den räumlichen Verhältnissen der Umwelt und den Tiefencues, die Indikatoren für diese sein sollen, erfassen. Brunswik selbst verwendete den Korrelationskoeffizienten als Maß für die ‘Stärke’ dieser Beziehung und fand insgesamt eher gering ausgeprägte Validitäten. Für eine ausführlichere Darstellung der dabei verwendeten Methode siehe Brunswik und Kamiya (1953/2001) sowie Brunswik (1956). Allgemeine Darstellungen von Brunswiks wahrnehmungstheoretischen Auffassungen lassen sich in Gordon (1997) sowie Postman und Tolman (1958) finden. Eine zeitgenössische psychophysikalische Methode der Erfassung der Validität von Tiefencues wird von Cutting und Vishton (1995) vorgeschlagen.

an Entfernungen. Die Bewegungsparallaxe sei hingegen ein wenig valider Tiefencue, da diese sowohl durch unterschiedliche Entfernungen von Objekten der Umwelt bei Bewegung des Betrachters als auch durch unterschiedliche Bewegung von Objekten bei gleichbleibender Entfernung vom Betrachter ‘hervorgerufen’ werden könnte. Die Reliabilität oder Validität eines Tiefencues sei zudem keine absolute, stets gleich bleibende Größe, sondern kontext- und erfahrungsabhängig (Atkins, Fiser & Jacobs, 2001); so verringere sich etwa die Reliabilität der Konvergenz mit zunehmender Distanz des fixierten Objektes vom Betrachter. Vollkommen ‘zuverlässig’ oder valide sei für sich genommen keiner der Tiefencues, sie seien lediglich als „imperfect indicators of the distal configuration“ (Epstein, 1995, S. 11) anzusehen, die prinzipiell mehrdeutig seien (Jacobs, 2002).

### Die Integration von Tiefencues

Um der Tatsache gerecht zu werden, dass wir uns trotz der imperfekten ‘Zuverlässigkeit’ der einzelnen Tiefencues im Alltag ohne größere Probleme sicher und zielgerichtet durch unsere Umwelt bewegen und mit Objekten interagieren können, werden im Rahmen des *cue*-Konzeptes nun weitere Annahmen gemacht. Die erste Annahme lautet, dass in Alltagssituationen stets *mehrere* der Tiefencues als ‘Informationen’ oder ‘Informationsquellen über Tiefe’ zur Verfügung stünden. Weiterhin wird angenommen, dass die vielfältigen verfügbaren Tiefencues im Rahmen der Informationsverarbeitung ‘interagieren’ könnten und in einem auch *cue integration* oder *cue combination* genannten Prozess nach bestimmten Regeln ‘kombiniert’ würden, um auf diese Weise eine einheitliche Wahrnehmung der räumlichen Verhältnisse zu ermöglichen bzw. zu ‘erzeugen’ (siehe etwa Atkins et al., 2001; Jacobs, 2002; MacKenzie, Murray & Wilcox, 2008; Knill, 2007a; Landy, Maloney, Johnston & Young, 1995; Wismeijer, Erkelens, van Ee & Wexler, 2010).<sup>6</sup> Es gibt eine Unmenge an Literatur zum Thema der Interaktion und Integration der Tiefencues, der sich ganz unterschiedliche Beschreibungen entnehmen lassen.

Im Fall der metrischen Tiefencues wird der Integrationsprozess häufig im Rahmen eines formal-mathematischen Modells beschrieben, in der die ‘quantitativen Informationen’ beispielsweise linear kombiniert werden.<sup>7</sup> Weiterhin könnten auch sogenannte

---

<sup>6</sup>Wie bereits an anderer Stelle bemerkt, findet man in der Literatur unterschiedliche Auffassungen dazu, was das Ergebnis eines solchen Integrationsprozesses sei. Bei manchen Autoren (etwa Burge et al., 2005; Howard, 2012) wird das Ergebnis als eine *Wahrnehmung* bezeichnet, bei anderen (etwa Mamassian & Landy, 1998; MacKenzie et al., 2008; Muller, Brenner & Smeets, 2009) als eine *quantitative Schätzung* der räumlichen Verhältnisse der Umwelt. In der Regel lassen sich in einer Arbeit beide Auffassungen finden.

<sup>7</sup>Es soll hier darauf verzichtet werden, auf die unterschiedlichen Modellvorstellungen der Interaktion oder Integration der Tiefencues näher einzugehen, da eine Kenntnis dieser für die weitere Diskussion nicht notwendig ist. Einen Überblick über die verschiedenen Ansätze gibt Howard (2012, Bd. 3), eine

*cue-Konflikte* (etwa Landy et al., 1995; Knill, 2007c) auftreten, falls verschiedene Tiefencues sich gegenseitig ‘widersprechende Informationen’ lieferten.<sup>8</sup> Üblicherweise wird der Prozess der *cue*-Integration als ein sich in zwei Stufen vollziehender aufgefasst:

First, depth estimates based on individual cues are derived. Second, a weighted combination of these estimates is calculated and used as the observer’s composite depth percept; the cue weights are based on the relative reliabilities of the cues in the current visual context. (Jacobs, 2002, S. 345)<sup>9</sup>

Die Beschreibungen legen insgesamt nahe, dass es sich bei den ‘Tiefenschätzungen’ bezüglich einzelner Tiefencues (häufig auch als *depth values* bezeichnet), den Reliabilitäten und dem Ergebnis der Integration um *quantitative Größen* handeln soll. Die Auffassungen, auf welche Weise das visuelle System zu den ‘Tiefenwerten’ kommt, um mit diesen weiter zu operieren, variieren stark, sind jedoch häufig intentional gefärbt. Die Tiefencues würden diesbezüglich ‘interpretiert’ (Seydell et al., 2010) oder ‘bewertet’ (Bertamini et al., 2008), die Tiefenwerte ‘geschätzt’ (Hillis et al., 2004; MacKenzie et al., 2008; van Mierlo, Brenner & Smeets, 2007) oder im Rahmen einer Schlussfolgerung ‘abgeleitet’ bzw. ‘deduziert’ (Allison & Howard, 2000; Jacobs, 2002; Knill, 2007a; Svarverud, Gilson & Glennerster, 2010). Wie bereits erwähnt, wird bezüglich der Frage nach dem Agenten dieser Aktivitäten in den meisten Fällen auf das visuelle System (etwa Burge et al., 2005; MacKenzie et al., 2008; Wismeijer et al., 2010), das Gehirn (etwa Knill, 2007a; Seydell et al., 2010) oder auch auf spezifische Neurone des Gehirns (Nadler, Angelaki & DeAngelis, 2008) referiert. Üblicherweise lassen sich auch bezüglich dieses Punktes mehrere dieser Beschreibungen innerhalb eines Textes auffinden. Eine Ermittlung oder Schätzung der ‘Tiefenwerte’ anhand des *retinal image* erfordere (man erinnere sich an die Ausführungen von Marr, 1982) die Annahme zusätzlicher *constraints* der Umwelt. Beispielsweise erfordere eine ‘Interpretation’ des Tiefencues der Verdeckung *als* Information über Tiefe neben der Annahme, dass sich Licht geradlinig ausbreite, weitere Annahmen, etwa dass (unbelebte) Objekte der Umwelt üblicherweise eine konstante Form und regelmäßige Oberflächen aufwiesen und üblicherweise nicht ausführlichere Darstellung lässt sich in Clark und Yuille (1990), Jacobs (1998) sowie Knill (2007a) finden.

<sup>8</sup>Ein häufig angeführtes Beispiel für das Vorliegen eines solchen ‘*cue*-Konfliktes’ ist die Betrachtung von Photographien, Bildern oder Filmen, bei denen verschiedene monokulare Tiefencues wie etwa Verdeckung und Linearperspektive darauf ‘hindeuten’ würden, dass sich die dargestellten Objekte in unterschiedlicher Entfernung befänden, andere Tiefencues wie etwa Konvergenz würden hingegen auf Objekte in identischer Entfernung ‘hinweisen’ (Kitazaki, Kobiki & Maloney, 2008). Wir werden den interessanten Fall der Wahrnehmung von Bildern ausführlicher in Abschnitt 4.1 behandeln.

<sup>9</sup>Anhand dieser Beschreibung lässt sich wiederum exemplarisch ablesen, dass in Fachartikeln zum Thema der Tiefencues ein relevanter Punkt, nämlich auf welche Weise es über diese Tiefencues zur *Wahrnehmung* räumlicher Relationen kommen soll, schlicht übergangen wird. In dem angeführten Zitat, das über eine formal-mathematische Beschreibung hinausgeht, wird einfach davon gesprochen, dass (quantitative) Entfernungsschätzungen *als* „the observer’s composite depth percept“ ‘benutzt’ würden.

transparent seien Diese *constraints* seien entweder in das visuelle System ‘eingebaut’ oder müssten ‘erlernt’ werden (Knill, 2007b).

Zusammenfassend sind die meisten Diskussionen der Tiefencues deutlich von komputationalen Vorstellungen inspiriert. Das *retinal image* werde unter Verwendung bestimmter *constraints* hinsichtlich der in diesem enthaltenen ‘Informationen über Tiefe’ verarbeitet, zu jedem Tiefencue werde ein zugehöriger ‘Tiefenwert’ berechnet oder geschätzt, auch im Fall der ordinalen Tiefencues. Bei diesen ‘Tiefenwerten’ soll es sich offenbar um Zahlen handeln. Auf diese Zahlen würden anschließend mathematische Operationen angewendet, in die u.a. die verschiedenen Reliabilitäten der einzelnen Tiefencues (ebenfalls Zahlen) eingingen. Bei dem ‘Output’ dieses Verarbeitungsprozesses sollte es sich daher ebenfalls um Zahlen handeln. Viele Autoren (etwa Howard, 2012; Landy et al., 1995) sprechen diesbezüglich auch von einer sogenannten *Tiefenkarte* (*depth map*), welche für jeden Punkt im visuellen Feld einen zugehörigen ‘Tiefenwert’ angebe, der sich aus der – je nach Modell unterschiedlichen – Kombination der einzelnen ‘Tiefenwerte’ der *cues* ergebe.<sup>10</sup> In der Regel wird diese Tiefenkarte als eine ‘interne Repräsentation’ der räumlichen Verhältnisse der Umwelt beschrieben. Gleichzeitig werden, wie auch schon bei Marr (1982), in den Beschreibungen dieser Prozesse intentional aufgeladene Begriffe verwendet, die sich auf das Gehirn oder das visuelle System beziehen. Das Gehirn ‘erhalte Informationen’ und ‘werte’ diese aus, es ‘mache verschiedene Annahmen’, könne ‘lernen’, etwas ‘schätzen’ und ‘generiere’ schließlich ein dreidimensionales Perzept. Tiefencues würden von diesem ‘interpretiert’ und ‘reinterpretiert’, die ‘Tiefenwerte’ seien das Ergebnis von ‘Schlussfolgerungen’ des Gehirns oder des visuellen Systems, es ‘verwende verschiedene Strategien’ und ‘rekonstruiere eine dreidimensionale Szene’.<sup>11</sup> Es ist auffällig, dass in den aktuellen Beschreibungen

---

<sup>10</sup>„We will speak of a ‘depth map’ as if it consisted of an array of measured distances across the visual field“ (Landy et al., 1995, S. 392).

<sup>11</sup>„To provide useful information about the physical environment, the visual system must generate a reasonably accurate three-dimensional percept from optical information in two 2D retinal images“ (Burge et al., 2005, S. 534). „By efficiently integrating information from all available cues, the brain can derive more accurate and robust estimates of three-dimensional space“ (Knill, 2007a, S. 189). „The simplest explanation for this behavior is that the visual system infers the 3D orientations and shapes of figures. It assumes that a large percentage of ellipses in the world have an aspect ratio equal to 1“ (Knill, 2007b, S. 13). „The visual system uses multiple complementary sources of information (cues) to estimate properties of interest“ (Muller et al., 2009, S. 1). „One of the biggest puzzles in perception is how the brain reliably and accurately estimates properties of the world from ambiguous sensory information. ... By incorporating prior knowledge of these regularities into perceptual computations, the brain can resolve much of the apparent ambiguity“ (Seydell et al., 2010, S. 1). „Overall, our results demonstrate a remarkable flexibility within the visual system to change the prior models it uses for the interpretation of pictorial depth cues“ (Seydell et al., 2010, S. 18). „The visual system uses different strategies depending on the congruency and reliability of available depth cues“ (Wismeijer et al., 2010, S. 1). „In normal circumstances there are multiple sources of stimulus information about aspects of the environment (such as the distance to an object). In such conditions it is typically found that the perceptual contribution depends upon the ‘confidence’ the nervous system attaches to the information“ (Tresilian, Mon-Williams & Kelly, 1999, S. 39). „Binocular disparities and motion parallax provide

der Tiefencues solche, *geistige Leistungen* bezeichnende Begriffe immer wieder auftauchen, die an ähnliche Formulierungen aus den im vorangehenden Kapitel dargestellten früheren wahrnehmungstheoretischen Ansätzen erinnern. Die aktuellen Ansichten verstehen und erklären die Wahrnehmung von Tiefe durch einen im Prinzip mathematisch modellierbaren Prozess einer komplexen Integration von Tiefencues, welche gleichsam die ‘Datenbasis’ für diesen Informationsverarbeitungsprozess darstellen. Das Ergebnis dieses Verarbeitungsprozesses wird als *Wahrnehmung von Tiefe* charakterisiert, die Tiefencues sollen die Wahrnehmung von Tiefe ermöglichen und über einen Prozess der Informationsverarbeitung verständlich machen.

Diese Beschreibungen provozieren jedoch Fragen. Welche Bedeutung könnte die Beschreibung haben, dass das visuelle System oder das Gehirn ‘Informationen verarbeite’, etwas ‘intern repräsentiere’, etwas ‘interpretiere’ oder ein ‘Perzept generiere’? Und nicht zuletzt: Wie soll die Verbindung zwischen diesen, dem Gehirn oder dem visuellen System zugeschriebenen Aktivitäten und *meiner Wahrnehmung* gedacht werden? Es soll an dieser Stelle unterstrichen werden, wie zentral die Konzepte ‘Information’, ‘Informationsverarbeitung’ und ‘Repräsentation’ für die in der Regel stark an komputationalen Vorstellungen orientierten Untersuchungen zum Konzept der Tiefencues sind. Nach dieser Vorstellung lägen ‘Informationen über Tiefe’ (u.a.) im *retinal image* vor, die mit Hilfe eines komplexen Integrationsprozesses zu einer ‘internen Repräsentation’ der Tiefenstruktur der betrachteten Szene führe, welche letztlich die Wahrnehmung von Tiefe verständlich machen soll. Es scheint daher ein Großteil der theoretischen Last des Konzepts der Tiefencues auf den abstrakten Konzepten der Information und der ‘internen Repräsentation’ zu liegen. Dies soll zum Anlass genommen werden, diese beiden für das *cue*-Konzept offenbar zentralen Konzepte näher zu erörtern, um so zu einem möglicherweise besseren Verständnis des Konzepts der Tiefencues und den diesem zugrundeliegenden impliziten theoretischen Annahmen zu gelangen.

## 3.2. Weiterführende Erörterungen

### 3.2.1. Das Informationskonzept im Kontext der Tiefencues

Ein sich durch sämtliche Beschreibungen der Tiefencues durchziehender Begriff ist der der *Information*. Wie im letzten Abschnitt dargestellt, werden die Tiefencues entweder selbst als Information über Tiefe angesehen oder als Informationsquellen, die Informationen über Tiefe ‘bereitstellen’ sollen. Wir wollen in diesem Abschnitt nun der Frage

---

information that allows a 3D reconstruction of the scene to be computed“ (Svarverud et al., 2010, S. 1).

nachgehen, welche Bedeutung der Informationsbegriff im Kontext der Tiefencues haben könnte.

In einer groben Annäherung lässt sich Information im alltäglichen Sprachgebrauch als *kommuniziertes Wissen* umschreiben, etwa wenn man davon spricht, dass man durch das Lesen einer Zeitung Informationen erhalte. Im Kontext der Natur- und Geisteswissenschaften zeigt sich allerdings, dass sich eine Unmenge an heterogenen Charakterisierungen des Konzepts der Information finden lassen, die an den Bezugsrahmen der entsprechenden wissenschaftlichen Disziplin und den in dieser untersuchten Phänomenen gebunden zu sein scheint (Capurro & Hjørland, 2003). Verschiedene Autoren äußern sich aus diesem Grund eher skeptisch, dass diesen verschiedenen Charakterisierungen in den Einzelwissenschaften auch nur eine rudimentäre gemeinsame Auffassung des Informationskonzepts zugrunde liegt:

The notion of information has been taken to characterize a measure of physical organization (or decrease in entropy), a pattern of communication between source and receiver, a form of control and feedback, the probability of a message being transmitted over a communication channel, the content of a cognitive state, the meaning of a linguistic form, or the reduction of uncertainty. These concepts of information are defined in various theories such as physics, thermodynamics, communication theory, cybernetics, statistical information theory, psychology, inductive logic, and so on. There seems to be no unique idea of information upon which these various concepts converge and hence no proprietary theory of information. (Bogdan, 1994, zitiert nach Capurro & Hjørland, 2003, S. 356)

Eine ähnliche Feststellung trifft auch Dennett (1989):

The term ‘information’ is commonly used as a mass noun, as if information were some kind of substance that could be moved, stored, chopped up ... but in fact we have no sound mutually recognized understanding of what information is and what kind of parcels it should be measured in. (S. 205)

Einer der strittigsten und breite philosophische Diskussionen auslösenden Punkte in diesem Zusammenhang ist die für die weitere Diskussion nicht unwichtige Frage, ob sich das Konzept der Information auf ein Phänomen bezieht, welches bestimmte ‘mentale Leistungen’ wie *Verstehen* voraussetzt, oder ob es auch sinnvoll ist, *unabhängig* von Subjekten, die zu solchen mentalen Leistungen fähig sind, von Informationen zu sprechen. Die meisten Charakterisierungen des Informationskonzepts lassen sich zwischen den beiden Polen ‘Subjektabhängigkeit vs. Subjektunabhängigkeit’ einordnen. Im Folgenden sollen kurz zwei diesen Polen entsprechende Informationskonzepte skizziert

werden, nämlich das *semantische Informationskonzept* und das *informationstheoretische Informationskonzept*. Für eine detaillierte Darstellung der Vielfältigkeit des Informationsbegriffes, insbesondere in den Wissenschaften, siehe Capurro und Hjørland (2003).

Verwendet man den Begriff der Information in einem semantischen Sinn, so ist dies der alltäglichen Verwendung dieses Begriffs sehr nahe, etwa wenn man davon spricht, dass man beim Lesen einer Zeitung viele Informationen erhalten habe oder man noch Informationen benötige, um sich in einer fremden Umgebung zurechtzufinden. Information im semantischen Sinn ist ein durchaus vages Konzept (insbesondere wenn man Verwendungsweisen wie ‘falsche Informationen’ als nicht unsinnig ansieht), es scheint aber zwingend an die Verwendung eines *Symbolsystems* (etwa einer Sprache) und ein von allen am Informationsaustausch beteiligten Subjekten geteiltes Wissen über die meist per Konvention festgelegten Regeln der Verwendung dieses Symbolsystems (der Syntax) sowie die Bedeutung der Symbole oder deren Kombinationen (der Semantik) gebunden. Informationen im semantischen Sinne lassen sich als *Aussagen* oder *Propositionen* charakterisieren, d.h. sprachliche Gebilde, denen man einen Wahrheitswert zuordnen kann.<sup>12</sup> Information in diesem Sinn „is addressed to human minds and is received by human minds“ (Machlup, 1983, zitiert nach Capurro & Hjørland, 2003, S. 357) und lässt sich daher nicht naturalisieren, d.h. auf bloße physikalische Zustände reduzieren. Es bedarf mindestens eines Subjektes, das diese physikalischen Zustände *als* (mögliche) Informationen über etwas *interpretiert*.

In der Informationstheorie hat der Begriff der Information eine formale, auf Wahrscheinlichkeitstheoretischen Konzepten beruhende, quantitative und deutlich vom semantischen Sinn abgegrenzte Bedeutung. In der für die Informationstheorie wegweisenden Arbeit von Shannon und Weaver (1949/1962) wird ausdrücklich darauf hingewiesen, dass Information im informationstheoretischen Sinn *nicht* mit Bedeutung verwechselt werden dürfe („*information must not be confused with meaning*“, S. 8). In der Informationstheorie werden *abstrakte Kommunikationssysteme* betrachtet. Be-

---

<sup>12</sup>Diese Charakterisierung ist einerseits sehr allgemein, und andererseits auch zu eng. Sie ist zu allgemein, da dieser zufolge *jede* Aussage gleichzeitig eine Information im semantischen Sinn wäre. Bennett und Hacker (2001) etwa kennzeichnen allein *wahre* Aussagen bzw. eine Menge an wahren Aussagen als Information im semantischen Sinn, um den Begriff der Information von dem der Aussage abzuheben, allerdings lässt eine solche Kennzeichnung dann alltägliche Verwendungsweisen wie ‘Fehlinformation’ oder ‘falsche Information’ nicht zu. Zudem würde es sich nach dieser Charakterisierung auch bei logischen Tautologien (z.B. „ $a = a$ “) um Informationen handeln, die man im üblichen Sprachgebrauch jedoch nicht als Informationen bezeichnet. Für unsere Zwecke ist dieser Punkt nicht von Relevanz. Die Charakterisierung ist andererseits zu eng, da sie vieles ausschließt, was wir alltäglich als Informationen bezeichnen. Fragt man etwa, an einer Kreuzung stehend, welcher der Wege zu einem bestimmten Ziel führe und der Befragte *zeigt* auf einen der Wege, spricht man üblicherweise auch davon, dass man entsprechende Informationen erhalten habe. Auch dieser Punkt wird für unsere Betrachtungen jedoch nicht entscheidend sein, entscheidend ist allerdings auch hier, dass der Fragende die Semantik der Geste *versteht*.

standteile solcher Kommunikationssysteme ist eine *Nachrichtenquelle*, welche abstrakte Nachrichten produziert, die an einen *Empfänger* weitergeleitet werden. Im einfachsten Fall – den wir im Folgenden betrachten wollen – produziert (oder ‘wählt’) die Nachrichtenquelle genau eine Nachricht aus einer endlichen Menge  $X = \{x_1, x_2, \dots, x_i, \dots, x_n\}$  von  $n$  unterschiedlichen Nachrichten.<sup>13</sup> Mögliche Bedeutungsgehalte dieser abstrakten Nachrichten sind irrelevant, vom informationstheoretischen Standpunkt aus werden etwa der gesamte Text des „Hamlet“, die Zeichenkette „ASTC“ oder das Symbol „!“ als äquivalente potentielle Nachrichten betrachtet (Shannon & Weaver, 1949/1962). Weiterhin wird angenommen, dass die Wahl einer Nachricht aus  $X$  mit einer bestimmten Wahrscheinlichkeit  $p_i$  verbunden ist, so dass die Größe  $p_i$  die Wahrscheinlichkeit dafür angibt, dass Nachricht  $x_i$  aus  $X$  gewählt wird (wobei die Summe aller Wahrscheinlichkeiten  $p_i$  Eins ergeben muss). Es wird nun festgelegt, dass die durch eine produzierte Nachricht übertragbare *Menge* an Informationen sich umgekehrt proportional verhält zur Wahrscheinlichkeit, mit der die Nachricht produziert wird. Die Informationsmenge (oder *self-information* in der Terminologie von Shannon & Weaver, 1949/1962) einer Nachricht  $x_i$  wird meist durch  $I(x_i)$  bezeichnet und in der Regel auf folgende Weise definiert (vgl. Coombs et al., 1970):

$$I(x_i) := -\log_2(p_i). \quad (3.1)$$

Die Informationsmenge oder der Informationsgehalt  $I(x_i)$  einer Nachricht  $x_i$  ist eine nicht-negative *Zahl*. Die Wahl des Logarithmus zur Basis 2 ist dabei arbiträr, die gewählte Basis des Logarithmus bestimmt gleichsam die Einheit, in der die Informationsmenge angegeben wird.<sup>14</sup> Die Einheit bei Verwendung des Logarithmus zur Basis 2 ist das sogenannte *bit*, eine Abkürzung für *binary digit*. Eine aus einer Menge von Nachrichten mit einer Auftretenswahrscheinlichkeit von 0.5 gewählte Nachricht enthält dieser Definition zufolge eine Informationsmenge von einem *bit*, eine Nachricht mit einer Auftretenswahrscheinlichkeit von 0.1 enthält ca. 4.32 *bit* an Informationen. Wie bereits erwähnt, ist der Inhalt oder eine mögliche Bedeutung der Nachricht für die Informationsmenge irrelevant, die Informationsmenge ist allein eine Funktion der

---

<sup>13</sup>Es handelt sich im Folgenden um eine sehr vereinfachte Darstellung des Informationskonzepts im informationstheoretischen Sinn, in der insbesondere nicht auf die zugehörige wahrscheinlichkeitstheoretische Fundierung eingegangen werden soll. Für ausführlichere Darstellungen siehe etwa Ash (1990) oder Yeung (2002). Für eine lohnenswerte Einführung in die Informationstheorie siehe Coombs, Dawes und Tversky (1970, Kap. 10).

<sup>14</sup>Dass die Informationsmenge  $I(x_i)$  als eine Funktion des Logarithmus der Wahrscheinlichkeit  $p_i$  definiert wird, ist indes nicht arbiträr. Es lässt sich zeigen, dass der Logarithmus die einzige Funktion ist, die bestimmte geforderte Eigenschaften erfüllt. Auf diesen Punkt soll hier jedoch nicht näher eingegangen werden, siehe dazu Coombs et al. (1970, Kap. 10).

Auftretenswahrscheinlichkeit der Nachricht. Je kleiner die Auftretenswahrscheinlichkeit einer Nachricht, desto größer ist ihre Informationsmenge und *vice versa*.

Gleichzeitig lässt sich leicht sehen, dass die Informationsmenge einer Nachricht *kontextgebunden* ist, da die Auftretenswahrscheinlichkeit einer Nachricht nicht durch die Nachricht selbst bestimmt ist, sondern vom Kontext der Wahrscheinlichkeitsverteilung auf den möglichen Nachrichten abhängt. Die Informationsmenge der Nachricht, die aus dem gesamten Text des „Hamlet“ besteht, kann in verschiedenen Kontexten identisch, größer oder kleiner sein als die Informationsmenge der Nachricht „!“, sie ist allein gegeben durch die Auftretenswahrscheinlichkeit dieser Nachricht in dem gegebenen Kontext.

Betrachtet man die durch eine Nachrichtenquelle zu übertragende *erwartete* Informationsmenge der Menge aller Nachrichten, so ergibt sich diese durch die mit den entsprechenden Wahrscheinlichkeiten gewichtete Summe der Informationsmenge aller Nachrichten:

$$\frac{1}{n} \sum_{i=1}^n I(x_i) = - \sum_{i=1}^n p_i \log_2(p_i). \quad (3.2)$$

Diese auch als *Unsicherheit* oder *Entropie* der Nachrichtenquelle bezeichnete Größe spielt in der Informationstheorie eine zentrale Rolle und wird in der Regel mit  $U(X)$  oder  $H(X)$  bezeichnet. Die Entropie wird üblicherweise ebenfalls in *bit* angegeben und lässt sich anschaulich auch als ein Maß der Unsicherheit auffassen, welche der  $n$  Nachrichten von der Nachrichtenquelle gewählt wird. Die Entropie einer Nachrichtenquelle mit  $n$  möglichen zu wählenden Nachrichten ist eine Funktion *sämtlicher* Wahrscheinlichkeiten  $p_i$  ( $1 \leq i \leq n$ ) und ist dann am größten, wenn für alle  $i$   $p_i = 1/n$  gilt, d.h. wenn die Auftretenswahrscheinlichkeiten aller  $n$  potentiell zu wählenden Nachrichten identisch sind. Im Fall identischer Auftretenswahrscheinlichkeiten ist die Entropie zudem identisch der Informationsmenge einer jeden Nachricht aus der Menge  $X = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ , nämlich gerade  $\log_2(n)$  *bits*, da

$$I(x_i) = -\log_2(p_i) = -\log_2\left(\frac{1}{n}\right) = \log_2(n) \quad (3.3)$$

und

$$\begin{aligned} U(X) &= - \sum_{i=1}^n p_i \log_2(p_i) = - \sum_{i=1}^n \frac{1}{n} \log_2\left(\frac{1}{n}\right) \\ &= -\frac{n}{n} \log_2\left(\frac{1}{n}\right) = -\log_2\left(\frac{1}{n}\right) = \log_2(n). \end{aligned} \quad (3.4)$$

Falls eine Nachrichtenquelle stets dieselbe Nachricht produziert, so ist die zugehörige Entropie gleich Null, es besteht gleichsam keinerlei Unsicherheit darüber, welche Nachricht gewählt werden wird, in diesem Fall liefert die Produktion dieser Nachricht auch keinerlei Information in diesem speziellen Sinn.

Leider wird in der Literatur zur Informationstheorie der Begriff der Information häufig sowohl für die Informationsmenge *einer* Nachricht als auch für die Entropie, die sich auf die Menge *aller* Nachrichten bezieht, verwendet.<sup>15</sup> Für die weitere Diskussion wird dieser Punkt jedoch nicht von Relevanz sein. Festzuhalten ist, dass das Konzept der Information in der Informationstheorie – sei damit nun der Informationsgehalt einer Nachricht oder die Entropie bezeichnet – *keinerlei* Berührungspunkte mit einem möglichen *semantischen Gehalt* der Nachrichten aufweist. In beiden Fällen handelt es sich um *statistische Größen*, die allein durch die Auftretenswahrscheinlichkeiten  $p_i$  der Nachrichten bestimmt sind und nicht durch eine mögliche inhaltliche Bedeutung dieser Nachrichten *für* einen Empfänger. Eine Nachricht ist im Kontext der Informationstheorie (quantitativ) umso informativer, je seltener sie von einer Nachrichtenquelle aus einer Menge von Nachrichten ausgewählt wird, egal um was es sich für Nachrichten handelt. Eine Nachrichtenquelle ist umso informativer, je höher der erwartete Informationsgehalt der von dieser Nachrichtenquelle zu produzierenden Nachrichten ist, d.h. je höher ihre Entropie ist.

Lässt sich anhand dieser Gegenüberstellung ausmachen, in welchem dieser Sinne der Begriff der Information im Rahmen des *cue*-Ansatzes (eher) verwendet wird? Innerhalb der aktuellen kognitionspsychologischen bzw. wahrnehmungspsychologischen Fachliteratur lassen sich selten explizite Hinweise finden, wie der Begriff der Information aufgefasst werden soll. Epstein (1995, S. 9) stellt diesbezüglich fest: „Cognitive psychologists ... did not provide a defined sense of information“. Er scheint in der Regel als abstrakter Begriff verwendet zu werden, was das Verständnis der fundamentalen Rolle dieses Konzepts innerhalb der aktuellen Kognitionspsychologie, die das Gehirn im Wesentlichen als ein *Informationsverarbeitungssystem* auffassen möchte, erschwert.

Howard (2012) spricht davon, dass ‘Information über Tiefe’ auf fünf verschiedenen Ebenen ‘definiert’ werden könne („depth information ... may be defined at any of the following five levels“ , Bd. 3, S. 147), nämlich u.a. auf Ebene des ‘distalen Stimulus’, d.h. der dreidimensionalen Szene selbst, auf Ebene des *retinal image* und auf Ebene ‘zentraler oder globaler neuronaler Verarbeitung’. Es ist nicht ersichtlich, wie hier der Begriff der Information als übergreifendes Konzept zu verstehen sein soll. In all die-

---

<sup>15</sup>Auch bei Shannon und Weaver (1949/1962) ist die Verwendungsweise nicht durchgängig klar, insbesondere hinsichtlich der Frage, ob sich der Begriff der Information auf die Nachrichtenquelle, eine Nachricht, die Menge aller Nachrichten oder den Empfänger der Nachrichten bezieht, was in verschiedenen Passagen zu teilweise paradox anmutenden Charakterisierungen des Informationskonzeptes führt. Siehe dazu auch Cole (1993).

sen Fällen handelt es sich offensichtlich nicht um das unserem alltäglichen Gebrauch nahe, semantische Informationskonzept. Andererseits scheint dieser Verwendung des Informationsbegriffes auch nicht das rein informationstheoretische Konzept der Information nach Shannon und Weaver (1949/1962) zugrunde zu liegen, da das *cue*-Konzept einen bestimmten *semantischen Gehalt* der Informationen impliziert, es soll sich ja um Informationen *über Tiefe* handeln.<sup>16</sup> Die mit den Tiefencues als Information verbundenen ‘Tiefenwerte’ sollen zwar Zahlen sein, aber Zahlen, die eine ganz bestimmte *Bedeutung* haben sollen, sie sollen u.a. Entfernungen von Objekten spezifizieren. Diese Zahlen lassen sich nicht im Sinne der Informationstheorie als *Informationsmenge* auffassen. Sie sollen *für* das visuelle System die Entfernung der Objekte der Umwelt untereinander und zum Wahrnehmenden spezifizieren und das visuelle System oder Gehirn soll unter Verwendung dieser Werte ein dreidimensionales Perzept ‘generieren’ oder ‘konstruieren’. Alles, was dem Gehirn jedoch zur ‘Verfügung’ stehen soll, seien die durch das retinale Lichtmuster ausgelösten *neuronalen Aktivitätsmuster*, in denen diese Informationen gleichsam ‘kodiert’ vorlägen. Die Vorstellung, diese als ‘Informationen über Tiefe’ für das Gehirn anzusehen, d.h. als Informationen mit semantischem Gehalt, scheint voraussetzen, dass das Gehirn über ein *Symbolsystem* mit zugehöriger Semantik ‘verfügt’, um diese neuronalen Aktivitätsmuster *als Symbole* (oder Zeichen oder ‘Information’) *über Tiefe* aufzufassen, oder, um eine häufig verwendete Umschreibung aufzugreifen, die genau diese Voraussetzung deutlich macht, zu ‘interpretieren’. Es ist schwer einzusehen, wie sich Muskelaktivitäten, Energieverteilungen auf der Retina oder damit einhergehende neuronale Aktivität *per se* als ‘Informationen über Tiefe’ auffassen lassen sollen. Hinter der Beschreibung der Tiefencues als ‘Informationen über Tiefe’ scheint eine zentrale, aber nicht explizite Voraussetzung zu stehen, nämlich, dass das Gehirn einer gewissen Semantik fähig wäre. Im der historischen Entwicklung des Konzepts der Tiefencues gewidmeten Teil dieser Arbeit wurde mehrfach darauf hingewiesen, dass in den dargestellten Theorien (etwa insbesondere bei Reid, Müller, Lotze

---

<sup>16</sup>Es ist nicht uninteressant, an dieser Stelle darauf hinzuweisen, dass in der Kognitionspsychologie selbst die Nützlichkeit des informationstheoretischen Informationskonzeptes deutlich in Frage gestellt worden ist. Uttal (1981) etwa stellt fest, dass sich Theorien und empirische Untersuchungen, die sich explizit des informationstheoretischen Rahmens bedienen (etwa Attneave, 1954), als wenig fruchtbar für die Kognitionspsychologie herausgestellt haben:

[T]he hopeful promises expressed in Attneave’s work that the quantitative measures of information quantity that had proven so successful in the engineering sciences could be simply and directly applied to psychological processes has been largely unfilled. (Uttal, 1981, S. 117)

Ähnliche Bemerkungen lassen sich bei Luce (2003) finden:

Although Shannon’s information theory is alive and well in a number of fields, after an initial fad in psychology during the 1950s and 1960s it no longer is much of a factor, beyond the word *bit*, in psychological theory. (S. 1)

und Helmholtz) im Kontext der Tiefencues Aussagen aufzufinden sind, die kaum anders zu verstehen sind, als dass das retinale Lichtmuster von irgendeiner Instanz (der Seele, der Vorstellung oder dem Bewusstsein) *als Bild* wahrgenommen wird. Eine ähnliche, wenngleich in abstraktere Begriffe eingekleidete Voraussetzung scheint in aktuellen Beschreibungen des *cue*-Konzeptes vorzuliegen, wenn etwa davon gesprochen wird, dass das *retinal image* 'Informationen über Tiefe' enthalte, welche das Gehirn 'verarbeiten' könne. Gleichzeitig kann man diesen Punkt in den Beschreibungen nicht eindeutig 'festnageln', u.a. durch den immer wieder auftauchenden Verweis auf das Konzept der 'internen Repräsentation', das in unmittelbarer Verbindung zu dem der Information zu stehen scheint. Die 'Tiefenkarte', ein Ergebnis des komputationalen Prozesses, *repräsentiere* die Informationen über die räumlichen Verhältnisse der Umwelt und diese Repräsentation soll unsere Wahrnehmung von Tiefe und Entfernung verständlich machen. Um diese Beschreibung angemessen verstehen zu können und möglicherweise auch ein besseres Verständnis des Informationskonzeptes im Kontext der Tiefencues zu erlangen, scheint es nötig, sich auch näher mit diesem Konzept auseinanderzusetzen.

### 3.2.2. Das Konzept der 'internen Repräsentation von Tiefe'

Leider zeigen sich hier zunächst ähnliche Probleme wie beim Konzept der Information. Der Begriff der 'internen' oder 'mentalenen Repräsentationen' ist in der aktuellen wahrnehmungspsychologischen und kognitionspsychologischen Literatur ebenso wie der Informationsbegriff allgegenwärtig, es sind aber, insofern eine Auseinandersetzung mit diesem Begriff überhaupt stattfindet, wenig einheitliche und klare Vorstellungen mit diesem verbunden.<sup>17</sup> Die Annahme interner Repräsentationen scheint – Ähnliches lässt sich bezüglich des Informationsbegriffs feststellen – gleichsam als eine Art 'Axiom' der aktuellen Kognitionspsychologie angesehen zu werden, das sich unmittelbar daraus ergibt, dass man geistige Prozesse als Prozesse der Informationsverarbeitung auffassen möchte (Ramsey, 2007).<sup>18</sup>

In einer ersten Annäherung ist es auch hier lohnenswert, sich übliche Verwendungsweisen des Terminus „Repräsentation“ vor Augen zu führen. Man spricht etwa im Zusammenhang mit Bildern oder Fotografien davon, dass ein Portrait eine bestimmte Person *repräsentiere*. Das Portrait selbst könnte man dann auch als „das Repräsentierende“ oder „die Repräsentation“ und die dargestellte Person als „das Repräsentierte“ bezeichnen. In diesem Kontext scheint unter anderem das Konzept der *Ähnlichkeit* zwi-

---

<sup>17</sup>„Anyone who has attempted to read the literature related to cognitive representation quickly becomes confused – and with good reason. The field is obtuse, poorly defined, and embarrassingly disorganized. ... *What is a representation?* This is a question few psychologists have ever asked and even fewer have made any serious attempt to answer“ (Palmer, 1978, S. 259).

<sup>18</sup>So schreibt Ullmann (1980, S. 333): „Central to contemporary cognitive science is the notion that mental processes involve computations defined over internal representations“.

schen Repräsentierendem und Repräsentierten eine wichtige Rolle zu spielen – ohne auf den Punkt, was Ähnlichkeit hier genau bedeuten mag, näher eingehen zu wollen (siehe dazu etwa Goodman, 1976). Wir sprechen aber auch bei Landkarten davon, dass etwa Punkte auf dieser Karte Städte repräsentieren, die Farbe dieser Punkte die Größe oder Einwohnerzahl dieser Städte repräsentiert und die Zeichenkombination neben einem solchen Punkt (etwa „HALLE“) die Stadt selbst. Der Begriff der Repräsentation ist eng mit dem des Symbols und des Zeichens verwandt; man spricht im Beispiel der Landkarte auch davon, dass die Elemente einer Karte etwas *symbolisieren* oder *bezeichnen*. Im Fall der Landkarte spielt das Konzept der Ähnlichkeit zwischen dem Repräsentierenden und dem Repräsentierten (etwa die Farbe der Punkte als Repräsentation der Größe der Stadt) offenbar keine Rolle. All diesen Beispielen gemeinsam ist, dass ein ‘Etwas’, etwa ein Punkt, eine Zeichenkette oder eine Farbe, gleichsam für etwas anderes, etwa eine Stadt oder deren Einwohnerzahl, ‘stehen’ soll – womit freilich noch nicht wirklich viel gesagt ist, da der zu erklärende Begriff hier einfach durch den nicht weniger erklärungsbedürftigen des ‘Einstehens von etwas für etwas anderes’ ersetzt wurde. Palmer (1978), ein Vertreter des komputationalen Ansatzes visueller Wahrnehmung, umschreibt eine Repräsentation auf ähnliche Weise: „A representation is, first and foremost, something that stands for something else“ (S. 262).

Eine Repräsentation ( $X$ ) scheint demnach stets eine Repräsentation von etwas zu sein, das von dieser Repräsentation *verschieden* ist ( $Y$ ). Wir sprechen nicht davon, dass etwas sich selbst repräsentiere oder sich selbst symbolisiere. Dieses, von  $X$  verschiedene  $Y$ , für das  $X$  als Repräsentation ‘stehen’ soll, könnte man auch als den *Inhalt* oder die *Bedeutung* der Repräsentation  $X$  bezeichnen.<sup>19</sup> In dem oben angeführten Beispiel der Landkarte wäre demnach die Bedeutung der Farbe der Punkte auf dieser Karte die Einwohnerzahl der durch die Punkte repräsentierten Städte. Den oben angeführten Beispielen lässt sich darüberhinaus entnehmen, dass die Eigenschaft von  $X$ , eine Repräsentation von  $Y$  zu sein, nicht zwingend durch die *intrinsischen Eigenschaften* von  $X$  selbst bestimmt wird. Im Beispiel der Landkarte könnte man auch Dreiecke statt Punkte (oder beliebige andere Symbole) zur Repräsentation von Städten und statt der Farbe die Größe dieser Dreiecke (oder beliebige andere Symbole) als Repräsentation der Einwohnerzahl verwenden. Entscheidend ist offenbar, dass  $X$  als Repräsentation für  $Y$  benutzt wird, und zwar im Rahmen eines impliziten oder expliziten *sozialen Aktes*, in dem die Regeln der Verwendung von  $X$  als Repräsentation für  $Y$  festgelegt werden, bzw. sich durch den *Gebrauch* von  $X$  als Repräsentation für  $Y$  bestimmen. Im Fall der Landkarte werden diese Regeln etwa durch die entsprechende Legende expliziert. Eine Repräsentation scheint also nicht nur eine Repräsentation von etwas zu sein, sondern

<sup>19</sup> „That which stands in for something else in this way is a *representation*; that which it stands for is its *content*; and its standing in for that content is *representing* it“ (Haugeland, 1991, S. 62).

auch stets eine Repräsentation von etwas *für jemanden*, der um die Verwendung von  $X$  als Repräsentation von  $Y$  weiß, *durch jemanden*, der  $X$  als Repräsentation von  $Y$  *verwendet*. Wir sprechen erst dann davon, dass  $X$  ein  $Y$  repräsentiert, wenn innerhalb eines bestimmten sozialen Rahmens  $X$  als Repräsentation von  $Y$  *verwendet* werden soll, d.h. wenn es bestimmte implizite oder explizite Übereinkünfte darüber gibt, *dass*  $X$  als Repräsentation von  $Y$  verwendet wird und bestimmte implizite oder explizite Regeln der Verwendung von  $X$  als Repräsentation von  $Y$  vorliegen, die bekannt sein müssen, damit  $X$  als Repräsentation von  $Y$  verwendet werden kann. Das Konzept der Repräsentation ist demnach auch mit dem der *Intentionalität* verbunden, wir sprechen nicht davon, dass etwas repräsentiert wird, wenn keine Absicht zur Verwendung von etwas als Repräsentation vorliegt.<sup>20</sup> Damit  $X$  als Repräsentation für  $Y$  in dem eben dargestellten Sinn verwendet werden kann, scheint es zudem erforderlich zu sein, dass ein gewisses Wissen über das zu repräsentierende  $Y$  *unabhängig* von der repräsentationalen Verwendung von  $X$  für  $Y$  vorhanden ist.  $X$  kann nicht als Repräsentation für  $Y$  verwendet werden, wenn keinerlei Wissen über  $Y$  vorliegt, das darüber hinausgeht, dass  $X$   $Y$  repräsentieren soll. Wir wollen zur besseren Unterscheidung verschiedener Möglichkeiten, den Begriff der Repräsentation zu verwenden, die eben dargestellte Verwendungsweise auch als *semantischen Repräsentationsbegriff* bezeichnen.

Wird im Rahmen des modernen *cue*-Ansatzes mit ‘Tiefe repräsentierenden Tiefenkarten’ oder allgemeiner mit ‘mental Repräsentationen’ auf den semantischen Repräsentationsbegriff referiert? Das Konzept interner oder mentaler Repräsentationen ist ein zentrales Element des sogenannten *Repräsentationalismus*, einer Strömung in der Philosophie des Geistes, welche in seiner aktuellen Formulierung in enger Verbindung zu einer komputationalen Auffassung des Geistes steht. Einer der berühmtesten Befürworter dieses Ansatzes ist Jerry Fodor (etwa Fodor, 1975, 1981, 1983, 1987), der diesen für „the *only* available theory of mental processes that isn’t *known* to be false“ (Fodor, 1987, S. 20) hält. Komputationale Prozesse seien als *symbolmanipulierende Prozesse* zu kennzeichnen, daher erforderten sie *notwendigerweise* ein Medium oder

---

<sup>20</sup>Eines der wohl berühmtesten Beispiele zur Verdeutlichung dieses Punktes lässt sich in dem Aufsatz „Brains in a vat“ von Hilary Putnam (1981) finden: Man stelle sich eine Ameise vor, die sich im Sand fortbewegt und durch diese Bewegungen eine Spur hinterlässt, die *für den menschlichen Betrachter* wie ein gelungenes Portrait von Winston Churchill aussieht. Handelt es sich in diesem Fall um eine Repräsentation von Winston Churchill? Für die Ameise offensichtlich nicht, diese „has never seen Churchill, or even a picture of Churchill, and it had no intention of depicting Churchill“ (Putnam, 1981, S. 1). Die von der Ameise gezogene Spur ist *für sich genommen* keine Repräsentation. Eine Abwandlung dieses Beispiels gibt Ramsey (2007, S. 23), das noch eindringlicher ist:

The proverbial driftwood washed up on an uninhabited beach does not, intuitively, represent anything, even if it happens to spell out the word ‘UNINHABITED BEACH’ or is arranged in a way that maps a course to a nearby lake. However, if someone were to come along and use the driftwood as a type of map, then it would indeed take on a representational role.

Symbolsystem, in dem sich diese vollziehen könnten („There is no computation without representation“, Fodor, 1981, S. 122).<sup>21</sup> Die komputationale Auffassung des Geistes impliziert daher die Annahme eines ‘mentalen Symbolsystems’, eines Symbolsystems ‘in unserem Gehirn’ und dieses Symbolsystem bestehe aus ‘mentalen Repräsentationen’.<sup>22</sup> Im Kontext wahrnehmungspsychologischer Theorien scheint dann im Wesentlichen an diesen Ansatz angeknüpft zu werden:

There must be symbols inside our heads for the things we see, symbols which themselves are unlike the things they represent. There must be a symbolic description in the brain of the outside world, a description cast in symbols which stand for the various aspects of the world of which sight makes us aware (Frisby, 1980, S. 8).

Bezüglich der ‘Form’, in der diese internen Repräsentationen vorlägen, lassen sich ganz unterschiedliche Ansichten finden. Manche Autoren beschreiben diese als Präpositionen (eine Ansicht, die an die *Language of Thought*-Hypothese von Fodor anschließt), andere als bildliche Repräsentationen, etwa die sogenannten *mental images* (z.B. Shepard, 1978). Einen zentralen Stellenwert hat das Konzept der internen Repräsentationen schließlich in der komputationalen Theorie der Wahrnehmung von David Marr:

From a philosophical point of view, the approach that I describe is an extension of what have sometimes been called representational theories of mind. ... Modern representational theories conceive of the mind as having access to systems of internal representations; mental states are characterized by asserting what the internal representations currently specify, and mental processes by how such internal representations are obtained and how they interact. (Marr, 1982, S. 6)

Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information. We have already seen that a process may be thought of as a mapping from one representation to another, and in the case of human vision, the initial representation is in no doubt – it consists of arrays of image intensity values as detected by the photoreceptors in the retina. (Marr, 1982, S. 31)

---

<sup>21</sup>Ebenso Palmer (1999, S.77): „We have claimed that the fundamental components of an information processing system are *representations* and *processes*“

<sup>22</sup>Dies mündet bei Fodor (1975) dann in die sogenannte *Language of Thought*-Hypothese; die Annahme einer *nicht im metaphorischen Sinne gemeinten* ‘Sprache des Geistes’ (das sogenannte *Mentalesisch*), die syntaktische *und* semantische Eigenschaften aufweisen soll. Mit der Literatur von Befürwortern und Gegnern des Repräsentationalismus ließen sich buchstäblich Bücherregale füllen. Da der Fokus der vorliegenden Arbeit auf den Kontext der Wahrnehmungspsychologie beschränkt ist, soll hier nicht weiter auf den allgemeinen Verlauf dieser Diskussion eingegangen werden, da sich diese meist nicht auf Probleme der Wahrnehmung bezog.

Wahrnehmung sei als Prozess der Informationsverarbeitung aufzufassen, wobei dieser aus formalen, *rein syntaktischen* Operationen der Transformation interner Repräsentationen bestehe. Der Ausgangspunkt dieses Prozesses sei das *retinal image*, welches die punktuellen Lichtintensitäten des auf die Retinae projizierten Lichtmusters repräsentiere und über mehrere Stufen der Informationsverarbeitung in eine explizite *symbolische Beschreibung* dessen „what is present in the world, and where it is“ (Marr, 1982, S. 3) transformiert werde. Die Termini „Repräsentation“ und „Beschreibung“ werden von Marr (1982) folgendermaßen eingeführt:

*A representation* is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this. And I shall call the result of using a representation to describe a given entity a *description* of the entity in that representation. (S. 20)

Das Dezimalsystem und das Binärsystem seien laut Marr etwa formale Systeme zur Repräsentation von Zahlen. Beiden Systemen liegt eine Spezifikation zugrunde, wie die Darstellung einer bestimmten Zahl in dem jeweiligen System konstruiert wird. Das Symbol „37“ wäre dann nach Marrs Terminologie eine Beschreibung der *Zahl* Siebenunddreißig im Dezimalsystem und das Symbol „100101“ eine Beschreibung derselben Zahl im Binärsystem.

Der Begriff der Repräsentation sensu Marr ist in gewisser Weise abstrakter als der semantische Repräsentationsbegriff. Er bezieht sich nicht auf zwei Entitäten  $X$  und  $Y$ , von denen eine die andere repräsentiert, sondern auf ein *formales System*, unter dessen Verwendung bestimmte Entitäten eines zu repräsentierenden Systems (etwa Zahlen) durch Elemente des repräsentierenden Systems (etwa das Binärsystem) beschrieben werden könnten. Damit ein formales System als eine Repräsentation zu kennzeichnen ist, müssten zudem Regeln vorliegen, die spezifizieren, welche formalen Beziehungen zwischen dem zu repräsentierenden System und dem Repräsentationssystem bestehen. Diese Regeln können offenbar – wie es Marrs eigene Beispiele des Dezimal- und Binärsystems nahelegen – auf reinen Konventionen beruhen, d.h. sie *werden gemacht* von Agenten, die etwas als Repräsentation oder Repräsentationssystem *benutzen* wollen. Repräsentationen sollen etwas ‘explizieren’ bzw. anhand von Repräsentationen soll etwas beschrieben werden, kurz: sie besitzen semantischen Gehalt. Das Symbol (die Ziffernfolge) „100101“ soll für die Zahl Siebenunddreißig stehen. Dies ist aber nur dann sinnvoll, wenn sie – ganz im Sinne des semantischen Repräsentationsbegriffs – als Repräsentationen benutzt werden, d.h. nur *für* jemanden etwas explizieren oder beschreiben können, der um die entsprechenden Regeln weiß und zudem weiß, dass etwas als durch diese Regeln spezifizierte Repräsentation beschrieben werden soll. Für die meisten Nichtinformatiker wird etwa der semantische Gehalt der Zeichenkette „100101“ gerade nicht die Zahl Siebenunddreißig sein, sondern die Zahl 100101. Eine Nachricht in

Morsecode (laut Marrs Charakterisierung müsste es sich beim Morsecode ebenfalls um eine Repräsentation handeln) beschreibt nur dann etwas, bzw. macht nur dann etwas explizit, wenn der Empfänger dieser Nachricht die Regeln der Verwendung des Morsecodes beherrscht. Davon zu sprechen, dass eine Nachricht in Morsecode *per se* etwas repräsentiere oder expliziere, ist ohne Bezug auf einen die Verwendungsregeln des Morsecodes beherrschenden Subjekts, *für den* die Nachricht in Morsecode etwas expliziert nicht sinnvoll. Marr ist demnach in seiner allgemeinen Beschreibung einer Repräsentation zwar abstrakter, aber nicht wesentlich vom semantischen Repräsentationsbegriff entfernt.

Im Kontext der komputationalen Auffassung der Wahrnehmung ist das Konzept der internen Repräsentationen, mit denen das Gehirn 'operiere', dann jedoch problematisch, da diese *physikalisch realisiert* sein sollen.<sup>23</sup> Im Fall des Digitalcomputers lassen sich die Symbolmanipulationen als physikalisch festgelegte Übergänge zwischen verschiedenen physikalischen Zuständen des Systems kennzeichnen (vgl. Newell, 1980). Ganz analog sollen sich im Fall des Menschen die 'komputationalen Operationen' als Übergänge zwischen verschiedenen physikalischen Zuständen des Gehirns (also verschiedenen neuronalen Aktivitätsmustern) kennzeichnen lassen. Im wahrnehmungspsychologischen Kontext soll das Ergebnis des komputationalen Prozesses aber semantischen Gehalt haben, es soll eine symbolische Beschreibung *von etwas*, nämlich den Gegenständen der Umwelt und deren räumlicher Anordnung sein. Diese Konzeption ist auf den üblichen, semantischen Begriff der Repräsentation angewiesen. Der Input dieses Prozesses soll das *retinal image* sein, eine 'interne Repräsentation', die physikalisch realisiert sei, und zwar in dem durch die Absorption des auf die Retinae projizierten Lichts ausgelösten Aktivitätsmuster bestimmter Neurone. Dieses Rezeptoraktivitätsmuster ist wohl das, was im Informationsverarbeitungsparadigma als Information angesehen wird, die zunächst im *retinal image* repräsentiert und dann im Rahmen weiterer *rein syntaktischer* Operationen zu einer expliziten symbolischen Beschreibung der Gegenstände der Umwelt und deren räumlicher Anordnung verarbeitet werde. Es werde keine Information im Rahmen dieses Prozesses generiert, sondern die Information liege bereits vor und werde in Verbindung mit verschiedenen, in das visuelle System implementierten Annahmen über die Umwelt gleichsam 'nutzbar' gemacht<sup>24</sup>

Es scheint, dass eine Auseinandersetzung mit dem Konzept der internen Repräsentationen nicht substantiell zu einer Erhellung des Informationskonzeptes (und umgekehrt) in der aktuellen Wahrnehmungspsychologie beitragen kann. Beide Konzepte sind untrennbar miteinander und mit ähnlichen konzeptuellen Schwierigkeiten verbunden.

<sup>23</sup>Siehe etwa Palmer (1999, S. 77), der den Begriff der internen Repräsentation einführt als „a physical entity that carries information about something“.

<sup>24</sup>„The role of the processing is not to create information, but to extract it, integrate it, make it explicit and usable“ (Ullmann, 1980, S. 367).

Möglicherweise lässt sich jedoch anhand der Vorstellung der physikalisch realisierten internen Repräsentationen die Frage nach der Bedeutung des Begriffs der Information im *cue*-Ansatz zuspitzen. Die entscheidende Frage ist, ob sich interne *physikalische Zustände* eines als rein physikalisch aufgefassten Systems (Gehirn oder Computer) sinnvoll als *repräsentational im semantischen Sinn* beschreiben lassen, auf den die Vertreter des *cue*-Ansatzes angewiesen sind. Oder, um das Informationskonzept wieder aufzugreifen: Lassen sich interne physikalische Zustände eines Systems *per se* sinnvoll als Informationen mit semantischem Gehalt, als Informationen *über Tiefe* beschreiben?

Es fällt in der Tat schwer, diesen Beschreibungen eine sinnvolle Bedeutung abzugewinnen. *Wir* können Computer *benutzen* um etwas zu repräsentieren, die strukturellen physikalischen Zustände des Computers oder Übergänge zwischen diesen allein sind nicht hinreichend für eine Repräsentation. Physikalische Zustände eines Computers lassen sich dann als repräsentational bezeichnen, wenn sie als Repräsentationen *benutzt* werden, wenn es geteilte Regeln der Verwendung dieser Zustände *als* Repräsentationen gibt. Der semantische Gehalt einer Repräsentation ist den physikalischen Eigenschaften dieser nicht intrinsisch, er ist durch den Gebrauch von etwas *als* Repräsentation bestimmt. Unser alltagssprachlicher Umgang mit der Funktionsweise von Computern scheint uns in diesem Fall in die Irre zu leiten, etwa wenn wir davon sprechen, dass *der Computer* ‘etwas berechne’ oder ‘Informationen verarbeite’ oder diese ‘speichere’. In einem strikten Sinn berechnet ein Computer (ebenso wenig wie ein Rechenschieber) *nichts* und er verarbeitet auch keine Informationen im semantischen Sinne oder speichert diese. *Wir benutzen Computer*, um Berechnungen anzustellen oder Informationen zu speichern, etwa indem *wir* Programme schreiben, deren Input und Output *für uns* repräsentationalen Charakter haben bzw. Informationen darstellen. Wir können jedoch unser Gehirn nicht in dem Sinne benutzen, um etwas zu repräsentieren, wie wir einen Computer oder unsere Hände benutzen.

Physikalische Zustände des Gehirns können nichts im semantischen Sinne repräsentieren, ohne die Voraussetzung machen zu müssen, dass es gleichsam einen von uns unabhängigen ‘Benutzer des Gehirns’ (einen Homunkulus) gibt, *für den* die verschiedenen physikalischen Zustände des Gehirns etwas repräsentieren, bzw. der diese verschiedenen physikalischen Zustände als Repräsentationen verwendet. Marr (1982, S. 6) spricht zwar von „the mind as having access to systems of internal representations“, diese Formulierung ist jedoch nicht weniger nebulös als die Beschreibung, dass der Geist das ‘Netzhautbild sehe’ oder die von Descartes geäußerte Vorstellung, dass der Seele in der Zirbeldrüse etwas zum Wahrnehmen ‘dargeboten’ werde. Im Fall der von Marr angeführten Repräsentationen (etwa der in Abbildung 2.3 auf Seite 134 dargestellte 2.5D-*sketch*) oder der im *cue*-Ansatz propagierten ‘Tiefenkarte’ ist dieses Problem auf den ersten Blick nicht deutlich ersichtlich, da diese häufig bildlich dargestellt werden

und wir *als Betrachter* dieser Darstellungen diese als Repräsentationen ansehen können, so wie wir die Höhenlinien in einer Wanderkarte als Repräsentationen der Höhe der repräsentierten Landschaft ansehen können. Die ‘internen Repräsentationen’ sehen wir jedoch nicht, wie wir Abbildung 2.3 sehen können, sie sollen doch im Rahmen des komputationalen Ansatzes der Wahrnehmung unsere Wahrnehmung erst verständlich machen. Eine Liste von Zahlen, die Tiefe repräsentieren sollen, kann diese Funktion nur erfüllen, wenn diese *als* Tiefe repräsentierende Zahlen interpretiert werden. Eine ‘Karte’ kann ebenso nur die Funktion erfüllen, eine Karte *von etwas* zu sein, etwas zu repräsentieren, wenn sie als eine solche interpretiert wird.

Ein zusätzliches Problem des Konzeptes interner Repräsentationen im Kontext visueller Wahrnehmung ist der angesprochene Umstand, dass, um etwas als Repräsentation für etwas davon Verschiedenes verwenden zu können, sowohl ein Wissen darüber vorliegen muss, dass etwas als Repräsentation verwendet wird, als auch ein davon unabhängiges Wissen darüber, *was* repräsentiert werden soll. Ich kann die Zeichenkette „100101“ weder als Repräsentation für die Zahl Siebenunddreißig, noch als Repräsentation für die Zahl 100101 verwenden, wenn ich nicht weiß, was Zahlen sind. In diesem Fall beschreiben diese Zeichenketten auch nichts, sie haben keinerlei Bedeutung. Im wahrnehmungspsychologischen Kontext hat dies zur Folge, dass die Annahme mentaler Repräsentationen, die eine symbolische Beschreibung der Gegenstände der Umwelt und deren räumlicher Anordnung darstellen sollen, voraussetzt, dass ein ‘generelles Wissen’ über die zu repräsentierende Umwelt und deren Gegenstände bereits vorliegen muss. Im Fall der Konstruktion künstlicher Sehsysteme wird dieses ‘Wissen’ vom Konstrukteur oder Programmierer implementiert, es ist aber *seiner Wahrnehmung entnommen*.

Verschiedene Vertreter des komputationalen Ansatzes der Wahrnehmung (etwa Ullman, 1991) haben versucht, die Annahme interner Repräsentationen über den Verweis auf empirisch feststellbare Regelmäßigkeiten zu rechtfertigen. So lässt sich etwa empirisch feststellen, dass unterschiedliche retinale Lichtmuster mit unterschiedlichen neuronalen Aktivitätsmustern einhergehen und relevante Korrelationen zwischen dem retinalen Lichtmuster und dem neuronalen Aktivitätsmuster bestehen. Da das retinale Lichtmuster, so die Argumentation, als (wenngleich nicht in eindeutiger Weise) durch die Gegenstände der Umwelt und deren räumlicher Anordnung bestimmt aufgefasst werden könne, ‘repräsentiere’ das neuronale Aktivitätsmuster eben diese Gegenstände der Umwelt und deren räumliche Anordnung, so wie etwa Rauch Feuer ‘repräsentiere’.<sup>25</sup> Ein solches Verständnis des Begriffs der Repräsentation scheint jedoch nicht mit den Beschreibungen des Outputs des komputationalen Prozesses als *symbolischer Beschreibung* in Einklang zu bringen sein und zudem den Begriff der Repräsentation zu

---

<sup>25</sup>Dretske (1983) vertritt im Rahmen seines Konzeptes der sogenannten *natural signs* eine ähnliche Auffassung der Repräsentation.

trivialisieren. Jeder beliebige physikalische Zustand eines hinreichend offenen Systems würde somit etwas repräsentieren oder Information über etwas sein, solange er mit einem anderen physikalischen Zustand eine regelmäßige Beziehung aufweist. Die Anzahl der weißen Blutkörperchen im Körper eines Menschen würden somit etwa eine Entzündung repräsentieren und der Schatten eines Gegenstandes den Stand der Sonne. Aber der Schatten eines Gegenstandes ist keine explizite symbolische Beschreibung *für* den Stand der Sonne, ebenso wenig wie die Anzahl der weißen Blutkörperchen eine explizite symbolische Beschreibung für eine Entzündung darstellt. Der Schatten ist im besten Falle ein *Anzeichen* für den Stand der Sonne, so wie auch Rauch eher ein Anzeichen von Feuer, aber keine symbolische Beschreibung *für* Feuer ist. Gleichzeitig ist es auch hier nur sinnvoll von einem Anzeichen zu sprechen, wenn jemand den Schatten als Schatten eines Gegenstandes *wahrnimmt* und sowohl über ein Wissen über den Gegenstand, die Sonne und die (in diesem Fall nicht triviale) Beziehung zwischen Schattenwurf eines Gegenstandes und Orientierung der Erde im Bezug zur Sonne weiß.

Wir scheinen hier einen zentralen Punkt des Konzepts der Tiefencues zu berühren, der sowohl in den im letzten Kapitel dargestellten Ansätzen (insbesondere bei den Theorien von Alhazen, Descartes, Reid, Müller, Lotze und Helmholtz), als auch in der aktuellen Formulierung, die mit den Konzepten der Information und der internen Repräsentationen operiert, immer wieder auftaucht. Die aktuelle Formulierung der Funktionsweise der Tiefencues ist darauf angewiesen, physikalischen Entitäten wie dem Gehirn oder dem visuellen System gleichsam *mentale Eigenschaften* zuschreiben zu müssen, damit diese die Funktionen erfüllen können, die ihnen im Rahmen des *cue*-Ansatzes zugewiesen werden. Das Gehirn oder das visuelle System soll in der Lage sein, Informationen zu extrahieren, ‘auszuwerten’ oder zu ‘interpretieren’ und symbolische Beschreibungen über die Gegenstände der Umwelt und deren räumliche Relationen zu ‘konstruieren’, welche unser Sehen letztlich verständlich machen sollen. Bei den früheren, im letzten Kapitel dargestellten Theorien lassen sich im Ansatz ähnliche Formulierungen finden, wenngleich in weniger technischem Vokabular. Es werden ‘Netzhautbilder verglichen’, die ‘Größe des Bildes eines Objektes’ auf der Retina mit der bekannten Größe des Objekts ‘verglichen’, Entfernungen ‘berechnet’, ‘Tiefe konstruiert’ usw. Eine sich durch sämtliche Ansätze durchziehende Beschreibung ist die der ‘Interpretation’, ein deutlich auf mentale Eigenschaften bezogener Begriff. An den entsprechenden Stellen der Darstellung dieser Theorien wurde jeweils darauf hingewiesen, dass diese Konzeption implizit voraussetzt, dass eine personenähnliche Instanz, ein Homunkulus, diese Tätigkeiten durchführt, der, um diese Tätigkeiten durchführen zu können, über solche mentalen Eigenschaften verfügen muss. Insbesondere muss dieser ein *Wahrnehmender* sein, der die retinalen Lichtmuster *als Bilder* von Objekten sieht,

um diese Tätigkeiten durchführen zu können.<sup>26</sup> Die Beschreibungen der Tiefencues weisen eine deutliche Nähe zum sogenannten *Homunkulus-Trugschluss* (Kenny, 1981) bzw. dem sogenannten *mereologischen Trugschluss* (Bennett & Hacker, 2001; Smit & Hacker, 2014) auf, der im folgenden Abschnitt, die Überlegungen der vorangegangenen Abschnitte aus allgemeinerer Perspektive zusammenfassend, dargestellt werden soll.

### 3.3. Tiefencues und das Problem homunkularer Redeweisen

Wir wollen zunächst das Problem des Homunkulus-Trugschlusses oder homunkularer Redeweisen aus allgemeinerer Perspektive behandeln und anschließend erörtern, in welcher Beziehung dieses Problem zum Konzept der Tiefencues steht. Was ist mit einem Homunkulus-Trugschluss gemeint? Dieser Terminus wurde von Kenny (1981) geprägt, der diesen folgendermaßen einführt:

I shall call the reckless application of human-being predicates to insufficient human-like objects the ‚homunculus fallacy‘, since its most naïve form is tantamount to the postulation of a little man within a man to explain experience and behavior. (S. 155)

Da insbesondere die Wendung „insufficient human-like objects“ sehr weite Charakterisierungen dieses Trugschlusses erlaubt, wird in der Philosophie des Geistes häufig eine andere, engere Charakterisierung gewählt, die sich folgendermaßen umschreiben lässt (vgl. Keil, 2003): Ein Homunkulus-Trugschluss liegt vor, wenn psychologische Prädikate, die allein *Personen* mit Sinn zuzuschreiben sind, auf *Teile* eines Lebewesens oder subpersonale Vorgänge (etwa Gehirnaktivitäten) angewendet werden.<sup>27</sup>

Überaus modern anmutende, sprachkritische Bemerkungen bezüglich dieser ungeRechtfertigten Anwendung psychologischer Prädikate auf Teile von Lebewesen lassen sich bereits in Aristoteles’ „*De Anima*“ finden:

So ist die Aussage, daß die Seele sich erzürne, ähnlich der, wie wenn man sagte, die Seele webe ein Tuch oder baue ein Haus; denn es ist vielleicht bes-

---

<sup>26</sup>Ebenso wurde darauf hingewiesen, dass selbst unter dieser Voraussetzung der Zusammenhang zwischen dem Wahrnehmen dieses Homunkulus und *unserem* Sehen nicht verständlich ist. Ähnliches gilt nun aber auch für die aktuellen Formulierungen der Tiefencues. Selbst wenn man voraussetzen würde, dass das Gehirn gleichsam mentale Eigenschaften hätte, ist nicht ersichtlich, wie diese Beschreibungen zu einem Verständnis unseres Sehens beitragen sollen.

<sup>27</sup>Smit und Hacker (2014, S. 1077) sprechen im gleichen Sinn von einem mereologischen Trugschluss: „If someone commits the mereological fallacy, then he ascribes psychological predicates to parts of an animal that apply only to the (behaving) animal as a whole“. Der von Stanisław Léśniewski geprägte Terminus „Mereologie“ bezeichnet ein Teilgebiet der Ontologie bzw. der angewandten Logik, das sich mit dem Verhältnis zwischen Teil und Ganzem befasst.

ser, nicht zu sagen, die Seele habe Mitleid, oder lerne, oder denke, sondern der Mensch mittels der Seele. (zitiert nach Keil, 2003, S. 1)

Wittgenstein (1953/1999) diskutiert dieses Problem unter anderem in seinen „Philosophischen Untersuchungen“, in denen es in §281 heißt:

Es kommt darauf hinaus: man könne nur vom lebenden Menschen und was ihm ähnlich ist (sich ähnlich benimmt) sagen, es habe Empfindungen; es sähe; sei blind; höre; sei taub; sei bei Bewußtsein, oder bewußtlos. (S. 368f.)

In der deutschsprachigen Literatur wird auch häufig von einem ‘Homunkulus-Fehlschluss’ gesprochen, was in diesem Zusammenhang durchaus diskutabel ist, da es sich ja nicht um die Anwendung eines logisch ungültigen Schlußschemas handelt. Das Problem bei einem Homunkulus-Trugschluss besteht eher darin, dass jemand, der solche Prädikate, die allein auf Personen sinnvoll anwendbar sind, auf Teile eines Lebewesens anwendet, eine personenähnliche Instanz (einen Homunkulus) innerhalb der Person annehmen *müsste*, um dieses Vorgehen auf sinnvolle Weise rechtfertigen zu können. Wir wollen im Folgenden weder von einem Trug-, noch einem Fehlschluss, sondern von *homunkularen Redeweisen* oder von *Homunkulismus* sprechen.

Typische und eindeutige Beispiele homunkularer Redeweisen findet man vielerorts in neurowissenschaftlichen Veröffentlichungen, in denen unbekümmert und unbesonnen davon gesprochen wird, dass das Gehirn ‘entscheide’, etwas ‘wisse’, ‘verstehe’, ‘abwäge’, uns ‘täusche’ etc. (siehe etwa Roth, 2001, Singer, 2002; zahlreiche weitere Beispiele aus den Neurowissenschaften lassen sich in Bennett & Hacker, 2001 finden). Roth (2001, S. 442) äußert sich etwa im Zusammenhang mit der Frage nach der Entscheidungsfreiheit einer Person hinsichtlich ihrer Handlungen folgendermaßen: „Dieser Willensakt tritt in der Tat erst auf, nachdem das Gehirn bereits entschieden hat, welche Bewegung es ausführen wird“. Die Entscheidung, eine bestimmte Handlung auszuführen, soll in dieser Argumentationslinie dadurch erklärt werden, dass unser Gehirn (und *nicht* wir als Personen) sich dafür ‘entscheide’, diese Handlung auszuführen. Es handelt sich hier um eine Aussage, die offensichtlich unsinnig ist; das Gehirn ist nicht das logisch angemessene Subjekt für das Prädikat „entscheiden“. Diese Aussage ist weder wahr noch falsch; sie hat, ebenso wie die Aussage, dass die Kreiszahl  $\pi$  grün ist oder ein Dreieck etwas entscheidet, schlicht keinen Sinn.

Das Problem des Homunkulismus ist kein empirisches Problem, sondern ein *konzeptuelles*, das auf einer unsinnigen Verwendung von Wörtern unserer Sprache beruht. Ob es sinnvoll ist, davon zu sprechen, dass unser Gehirn ‘entscheidet’, erfordert eine konzeptuelle Klärung, keine empirische Untersuchung. Es kann nicht auf empirischem Weg ‘entdeckt’ oder festgestellt werden, ob das Gehirn ‘entscheidet’, der Mensch eine Seele hat oder  $\pi$  grün ist, da nicht geklärt ist, was unter derartigen Aussagen zu verstehen ist. Solange nicht vom neurowissenschaftlich argumentierenden Forscher geklärt

wird, was er unter solchen Aussagen wie „das Gehirn entscheidet“ verstehen will und welche konzeptuelle und eventuell auch empirische Verbindung er zu entsprechenden Aussagen *über Personen* knüpft, sind diese Aussagen ganz ohne Sinn und damit weder wahr noch falsch. Eine konzeptuelle Klärung hat unter anderem zum Gegenstand, den sinnadäquaten Verwendungsweisen der Wörter unserer Sprache nachzugehen, um feststellen zu können, ob in bestimmten Aussagen Begriffe ihrem angestammten Kontext entfremdet und in nicht gerechtfertigter Weise auf andere Zusammenhänge angewendet und so „the bounds of sense“ (Smit & Hacker, 2014, S. 1080) übertreten werden.

Worin besteht also das Problem, wenn davon gesprochen wird, dass das Gehirn ‘entscheide’, etwas ‘wisse’, ‘verstehe’ oder uns ‘täusche’? Die durch diese Begriffe bezeichneten Leistungen lassen sich als *psychologische Attribute*, die Prädikate selbst als *intentionale* oder *psychologische Prädikate* kennzeichnen. Weitere Beispiele für psychologische Prädikate sind etwa „Schmerzen haben“, „wünschen“ oder „glauben“. Wir haben in der Regel ein intuitives und unmittelbares Verständnis davon, was es heißt, wenn *wir* Schmerzen haben oder uns etwas wünschen und wir sprechen anderen Personen solche psychologischen Attribute in der Regel zu, wenn es in ihnen *möglich* ist, ein bestimmtes Verhalten zu zeigen oder auf bestimmte Weise zu reagieren. Wir schreiben jemandem das Haben von Schmerzen zu, wenn er etwa äußert, dass er Schmerzen hat, aufschreit und/oder einen bestimmten Gesichtsausdruck aufweist. Wir schreiben einer Person eine bestimmte Überzeugung zu, wenn sie sich äußert, dass sie glaube oder meine, dass dies oder jenes der Fall sei. Wir sprechen nicht davon, dass jemand gerade Schmerz empfindet, eine Überzeugung hat oder schlussfolgert, wenn er keinerlei zugehöriges Verhalten zeigt. Um mit Wittgenstein (1953/1999, S. 455) zu sprechen: „Ein ‚innerer Vorgang‘ bedarf äußerer Kriterien“. Wenn wir jemandem erklären wollen, was es bedeutet, wenn man sagt, dass jemand Schmerzen empfindet oder etwas entscheidet, so geschieht dies über die Beschreibung entsprechenden Verhaltens oder das Zeigen auf entsprechendes Verhalten. Es handelt sich in diesen Fällen jedoch nicht um induktive Evidenz (Bennett & Hacker, 2001), wir schreiben den Personen nicht *aufgrund* des von ihnen gezeigten Verhaltens solche psychologischen Prädikate zu. Dies würde voraussetzen, dass zwei Phänomene (etwa das Empfinden von Schmerz und entsprechendes ‘Schmerzverhalten’ oder das Fällen einer Entscheidung und das entsprechende Verhalten) *empirisch überprüfbar* miteinander korrelieren. Es ist jedoch keine empirische feststellbare Tatsache, dass das Fällen von Entscheidungen, das Haben von Schmerz, das Interpretieren von etwas etc. mit entsprechendem Verhalten korreliert. Der Grund, weshalb wir bestimmte psychologische Prädikate Personen zuschreiben, wenn sie bestimmtes Verhalten zeigen, ist ein logisch-grammatikalischer. Die Bedeutung eines psychologischen Prädikates ist auf *konzeptueller Ebene* mit dem entsprechenden Verhalten verbunden. Dies bedeutet nicht, dass das Haben von Schmerzen oder das Wünschen von etwas mit dem ent-

sprechenden Verhalten *identifiziert* wird. Die äußeren Kriterien sind konstitutiv für die Bedeutung des zugehörigen psychologischen Prädikates. Zeigt jemand das für das Fällen von Entscheidungen oder das Haben von Schmerz ‘angemessene’ Verhalten, so ist dies auf logisch-grammatikalischer Ebene eine gute Evidenz dafür, ihm das Fällen einer Entscheidung bzw. das Haben von Schmerzen zuzuschreiben. Es kann natürlich vorkommen, dass wir jemandem, der das einem psychologischen Prädikat zugehörige Verhalten zeigt, in dieser Situation dieses Prädikat *fälschlicherweise* zuschreiben; so kann etwa ein Schauspieler auf der Bühne das dem Empfinden von Schmerz angemessene Verhalten zeigen, ohne Schmerzen zu haben oder jemand kann behaupten, er glaube dies oder jenes, aber lügen. Dass wir jemandem ein bestimmtes psychologisches Prädikat fälschlicherweise zugeschrieben haben, lässt sich jedoch wiederum nur mit dem gezeigten Verhalten dieser Person rechtfertigen, etwa indem die Person zugibt, dass sie gelogen bzw. den Schmerz nur gespielt hat. Wenn in einer bestimmten Situation die Verhaltenskriterien dafür erfüllt sind, dass jemand Schmerzen empfindet, etwas entscheidet, interpretiert oder schlussfolgert und keine gegenteilige behaviorale Evidenz vorliegt, so ist es aus konzeptuellen Gründen gerechtfertigt, dieser Person zuzuschreiben, dass sie gerade Schmerzen empfindet, etwas entscheidet, interpretiert oder schlussfolgert.<sup>28</sup>

Im Fall des Gehirns ist es nun allerdings vollkommen unklar, wann wir davon sprechen wollen, dass dieses eine Entscheidung trifft oder etwas interpretiert. Wie soll sich feststellen lassen, ob das Gehirn gerade eine Entscheidung trifft oder ob es gerade keine Entscheidung trifft? Welches Kriterium soll herangezogen werden, um dem Gehirn das Treffen einer Entscheidung zuzuschreiben? Es gibt hier keine auch nur vergleichbaren ‘behavioralen Kriterien’, wie sie bei Personen vorliegen. Der Ausdruck, dass das Gehirn etwas entscheide ist daher ebenso ohne Sinn wie der Ausdruck, dass ein Dreieck etwas entscheide – er ist ein Beispiel *konzeptueller Verwirrung* (Bennett & Hacker, 2001).

Es ließe sich nun einwenden, dass es ja gerade in den Wissenschaften aus theoretischen Gründen fruchtbar sein könnte, bestimmte, in unserem alltäglichen Sprachgebrauch vorkommende Begriffe mit gänzlich anderen Inhalten zu verbinden (der Begriff der Kraft in der klassischen Physik mag als Beispiel dienen). Im Fall des Gehirns könnte man die Sprechweise „etwas entscheiden“ mit gänzlich neuen, rein physiologischen Inhalten verbinden. Wenn dem so wäre, dann ließe sich lediglich einwenden, dass diese Verwendungsweisen gerade in diesem Forschungsbereich zu einiger terminologischer Verwirrung Anlass geben könnten. Die diese Wendungen gebrauchenden Autoren lassen aber nicht erkennen, dass sie dieser Sprechweise in Bezug auf das Gehirn eine *andere*

---

<sup>28</sup>„It *makes sense* to ascribe a psychological attribute to another being truly *or* falsely, only if it is *possible* for that being to display such behaviour as would *count* as good evidence for the ascription of the psychological attribute, i.e. the appropriate forms of behaviour must be in the creature’s behavioural repertoire“ (Hacker, o. J., S. 6).

Bedeutung als in Bezug auf Personen geben wollen. Im Gegenteil, die Ausführungen deuten eher darauf hin, dass diese Sprechweise hier dieselbe Bedeutung haben soll.

Zudem werden homunkulare Redeweisen im neurobiologischen Sprachumfeld in der Regel in einem gewissen *Erklärungszusammenhang* verwendet; sie sollen dazu dienen, bestimmte, als erklärungsbedürftig angesehene Phänomene, die als wesentlich *intentionale Phänomene* (wie Glauben, Wissen und Entscheiden) zu kennzeichnen sind, zu erklären. Die Entscheidung, eine bestimmte Handlung auszuführen, soll beispielsweise dadurch erklärt werden, dass unser Gehirn eine Art Entscheidung getroffen habe und wir diese lediglich ausführten. Das als erklärungsbedürftig angesehene Konzept wird in dieser Argumentationsfigur auf subpersonaler Ebene nochmals verwendet und als Erklärung ausgegeben. Der naheliegende Einwand gegen dieses Vorgehen ist, dass es in keiner Weise erklärungskräftig ist, da es auf einen unendlichen Regreß hinausläuft. Es wird nichts erklärt, sondern das zu erklärende Phänomen oder die zu erklärende geistige Leistung wird lediglich auf eine andere Instanz ‘verschoben’, es hat den Nachteil „of merely postponing the central problem before us by positing unanalysed man-analogues as systematic elements in that which we are trying to analyse, namely man“ (Dennett, 1969/1996, S. 87). Eine geistige Leistung einer Person wird als ein bestimmter ‘Prozess’ *innerhalb* dieser Person analysiert oder zu erklären versucht. Die in der Beschreibung dieses Prozesses verwendeten Prädikate können jedoch mit Sinn nur Personen zugeschrieben werden, also müsste, um diese Zuschreibung zu rechtfertigen, etwas ‘hinreichend Personenähnliches’ *innerhalb der Person* angenommen werden, man landet wieder bei einem Homunkulus-Trugschluss. Nun ist diese Annahme weder plausibel, noch würde, selbst wenn sie plausibel wäre, dadurch etwas erklärt werden, denn nun müsste die geistige Leistung, die bezüglich einer Person als erklärungsbedürftig angesehen wurde, bezüglich der hypostasierten Homunkulus-Instanz verständlich gemacht werden.

Homunkulare Redeweisen tauchen in der Regel in Bezug auf ein spezifisches *philosophisches Problem* auf, nämlich, grob gesprochen, im Kontext der Frage, wie der menschliche Geist ‘funktioniert’. Den auf homunkulare Redeweisen zurückgreifenden Erklärungsansätzen scheint insbesondere in den Neurowissenschaften die Motivation zugrunde zu liegen, geistige oder intentionale Phänomene zu naturalisieren, indem sie im Rahmen physikalisch beschreibbarer Prozesse (bestimmten Gehirnaktivitäten) zu analysieren und zu erklären versucht werden. Intentionalität ist kein Bestandteil der physischen Welt, sie lässt sich nicht in ein physikalisches Weltbild integrieren, daher, so scheint es, wird versucht, diese auf physikalisch beschreibbare ‘Bestandteile’ des Menschen zurückzuführen, die nicht selbst intentionale Eigenschaften haben sollen. Wir wollen uns nun, an diese allgemeinen Überlegungen anschließend, wieder dem im Zentrum dieser Arbeit stehenden Thema zuwenden, dem Konzept der Tiefencues.

### Homunkulare Redeweisen im Kontext des Konzepts der Tiefencues

Insbesondere in der Kognitionspsychologie scheint man jedoch nicht auf intentionales Vokabular vollständig verzichten zu wollen, es soll ja gerade versucht werden, die *geistigen Leistungen des Menschen* zu analysieren, aber innerhalb eines als physikalisch zu kennzeichnenden Weltbildes. In wahrnehmungspsychologischen Theorien, insbesondere in modernen, dem komputationalen Ansatz verpflichteten Theorien, lassen sich in der Regel homunkulare Redeweisen nicht so eindeutig identifizieren wie im Fall des ‘entscheidenden Gehirns’. Es wird üblicherweise nicht versucht, das Sehen einer Person dadurch verständlich zu machen, dass die Seele, das Gehirn oder das visuelle System schlicht ‘sieht’. Im Rahmen der Darstellung der historischen Entwicklung des Konzepts der Tiefencues im vorangehenden Kapitel wurde an entsprechenden Stellen deutlich zu machen gesucht, dass es sich um homunkulare Redeweisen handelt, indem darauf hingewiesen wurde, dass viele der Beschreibungen Wahrnehmung in gewisser Weise bereits voraussetzen. Insbesondere an Stellen, an denen sich die Beschreibungen auf bestimmte Eigenschaften des retinalen Lichtmusters beziehen, wurde darauf hingewiesen, dass diese Beschreibungen voraussetzen, dass das retinale Lichtmuster von einer hypostasierten Instanz *gesehen* werden muss. In Helmholtz’ Beschreibung der binokularen Disparität etwa wird davon gesprochen, dass es über einen ‘Vergleich’ der beiden ‘Netzhautbilder’ zur Wahrnehmung von Tiefe komme. Die Beschreibungen des retinalen Lichtmusters als ‘Netzhautbild’, das anschließend ‘interpretiert’, ‘ausgewertet’ oder ‘verglichen’ wird, sind als homunkular zu kennzeichnen. Das retinale Lichtmuster als ‘Abbild der äußeren Objekte’ anzusehen, deren ‘Größe’ in diesem Bild mit der bekannten Größe der Objekte ‘verglichen’ werde, setzt einen Homunkulus voraus, der diese Tätigkeiten durchführen muss.

Auch in aktuellen wahrnehmungstheoretischen Ansätzen lassen sich gelegentlich offensichtlich homunkulare Redeweisen ausmachen, so etwa bei Gregory (1998, 2002), der das visuelle System wie einen Naturwissenschaftler beschreibt. Es stelle Hypothesen auf und ‘bewerte’ diese im Licht der ‘Evidenz’ („The visual brain has the same kind of problem for accepting or rejecting evidence“, Gregory, 1998, S. 11). Offensichtlich homunkular ist auch folgende Beschreibung von Hoffman (2000):

Sehen ist nicht nur ein Vorgang passiver Wahrnehmung, sondern ein intelligenter Prozeß aktiver Konstruktion. Sie sehen immer nur das, was ihre visuelle Intelligenz konstruiert. Wie Wissenschaftler mit hoher Intelligenz nützliche Theorien konstruieren, die sich auf experimentelle Ergebnisse stützen, so konstruiert Ihr visuelles System mit hoher Intelligenz nützliche visuelle Welten, die sich auf die Bilder in Ihre Augen stützen. (S. 10f.)

Cutting und Vishton (1995, S. 70) sprechen im Zusammenhang mit der Wahrnehmung von Tiefe dem visuellen System gar die Fähigkeit zu, die mit Hilfe der Tiefencues ‘abgeleitete’ dreidimensionale Struktur einer Szene zu ‘verstehen’: „Why does the human visual system make use of such a variety of sources of information – often called *cues* – in understanding and in deriving the structure in depth of a complex natural scene?“.

In aktuellen, dem komputationalen Ansatz verpflichteten Wahrnehmungstheorien (dies beinhaltet auch die aktuelle Formulierung des Konzepts der Tiefencues) lassen sich solche offensichtlich homunkularen Redeweisen üblicherweise nicht finden. Es wird zwar in der Regel (wie in Abschnitt 3.1 dargestellt wurde) davon gesprochen, dass das visuelle System etwas ‘interpretiere’ oder ‘schlussfolgere’, diese Wendungen werden jedoch im Rahmen abstrakter Konzepte wie dem der Information und der Informationsverarbeitung, der ‘internen Repräsentation’ und der Berechnung formuliert. Diese Konzepte sind nicht offen intentional, in den vorangehenden Abschnitten wurde jedoch dafür argumentiert, dass diese auf einer *versteckten* Annahme einer personalen oder personenähnlichen Instanz beruhen, da sie mit einem semantischen Gehalt verbunden sind. Das visuelle System operiere laut dem komputationalen Ansatz mit Symbolen, die bestimmte Eigenschaften der Objekte der Außenwelt repräsentierten. Es wurde darauf hingewiesen, dass diese Beschreibungen darauf hindeuten, dass diese mentalen oder internen Repräsentationen syntaktische *und* semantische Eigenschaften aufweisen sollen.<sup>29</sup> Der Informationsverarbeitungsansatz behauptet durch den Rückgriff auf die generelle Annahme, dass das Gehirn als Computer aufzufassen sei, dass *physische Strukturen* wie das Gehirn oder der Computer mit Symbolen operieren, die für etwas außerhalb ihrer selbst stehen, die einen semantischen Gehalt haben sollen. Bestimmte, durch die Absorption von Lichtquanten ausgelöste neuronale Aktivitätsmuster werden als physikalische Realisierungen von Symbolen, die Übergänge zwischen verschiedenen Hirnzuständen als syntaktische Transformationen dieser Symbole angesehen. In diesem Sinne wird von ‘Interpretation’ oder von ‘Schlussfolgerungen’ im komputationalen Ansatz der Wahrnehmung gesprochen. Wir wollen uns dieser Auffassung nun näher zuwenden, indem wir untersuchen wollen, ob diese generellen Annahmen, dass das Gehirn als ein Digitalcomputer und geistige Prozesse wie der der Wahrnehmung als rein komputational aufzufassen sind, als homunkulare Redeweisen aufzufassen sind.

Unsere Wahrnehmungen haben einen mentalen Gehalt, sie haben Bedeutung. Mit dem Gedankenexperiment des „Chinesischen Zimmers“ hatte Searle (1980, 1982) im Zusammenhang mit der Funktionsweise von Computern darauf hingewiesen, dass Syntax nicht hinreichend für Semantik ist. Die Fähigkeit, Symbole regelgerecht zu manipulieren, garantiert nicht, dass diese Symbole für das symbolmanipulierende System

---

<sup>29</sup>Wie bereits erwähnt postulieren verschiedene Autoren (etwa Fodor, 1975, 1987; Barlow, 2002) sogar explizit eine ‘Sprache des Gehirns’ mit syntaktischen und semantischen Eigenschaften.

irgendeine Bedeutung haben. Die Funktionsweise von Computern lässt sich als ein Manipulieren von Zeichenketten nach bestimmten Regeln kennzeichnen, diese Zeichenketten haben aber für einen Computer keinerlei Bedeutung, sie sind *für* den Computer keine Repräsentationen von etwas; bloße Zeichenmanipulation ‘generiert’ keine Semantik. Computer operieren nicht in dem Sinne mit Repräsentationen, auf den der komputationale Ansatz der Wahrnehmung angewiesen ist. Nicht dem Computer kann sinnvoll das Operieren mit Repräsentationen zugeschrieben werden, sondern allein dem Konstrukteur/Programmierer oder dem Benutzer des Computers. *Für diese* haben die Zeichenketten des Computers eine Bedeutung, sie *verwenden* die Zeichenketten als Repräsentationen im vollen Sinn. Diese Zeichenketten können für beliebige Dinge stehen, mit unterschiedlicher Bedeutung versehen werden, je nachdem, für was der Programmierer sie verwenden möchte. Die Zeichenmanipulationen beim Ablaufen eines Programms haben keinen intrinsischen semantischen Gehalt, ein semantischer Gehalt kann diesen nur durch Personen *verliehen* werden. Der semantische Gehalt der Wahrnehmung kann demnach nicht durch ein rein syntaktisches Manipulieren von Symbolen verständlich gemacht werden, ohne einen Homunkulus voraussetzen zu müssen, *für den* diese Symbolmanipulationen Bedeutung haben.

Man kann dieses Argument sogar verschärfen und die Frage aufwerfen, ob die Gleichsetzung der Funktionsweise des Gehirns bzw. des visuellen Systems mit der Funktionsweise eines Computers *selbst ohne Verweis auf Semantik* nicht zwangsläufig Homunkulismus zur Folge habe, da sie voraussetzt, dass physische Vorgänge, wie sie im Gehirn oder im Computer stattfinden, *intrinsischerweise* syntaktische Eigenschaften hätten. Ob etwas ein komputationales, ein regelgeleitetes, symbolmanipulierendes System ist, ist jedoch nicht durch die physischen Eigenschaften dieses Systems bestimmt. Kein physischer Prozess ist intrinsischerweise eine regelgeleitete Symbolmanipulation. Es lässt sich auch nicht durch Rekurs auf die physischen Eigenschaften eines Gegenstandes bestimmen, ob dieser ein Zeichen, ein Symbol ist. „Syntax and symbols are not defined in terms of physics ... syntax is not intrinsic to physics“ (Searle, 1990, S. 35). Die physikalischen Zustände eines Systems können als syntaktisch, als regelgeleitet *interpretiert* werden, sie haben aber intrinsisch keine syntaktischen Eigenschaften, sie werden ihnen *zugewiesen*. Prinzipiell können alle möglichen, hinreichend offenen physikalischen Prozesse, an denen sich etwas unterscheiden lässt, als syntaktisch interpretiert werden. Die physikalischen Zustände eines Computer sind daher nicht *per se* als syntaktisch zu kennzeichnen, es bedarf stets eines *interpretierenden Beobachters*, der diese physikalischen Zustände *als* syntaktisch interpretiert:

There is no way you could discover that something is intrinsically a digital computer because the characterization of it as a digital computer is always relative to an observer who assigns a syntactical interpretation to the pu-

rely physical features of the system. The characterization of a process as a computational is a characterization from outside; and the identification of the process as computational does not identify an intrinsic feature of the physics, its is essentially an observer relative characterization. (Searle, 1990, S. 28)

Nun besteht allerdings ein fundamentaler Unterschied zwischen dem, was wir üblicherweise als Computer bezeichnen, und dem visuellen System bzw. dem Gehirn. Im Fall des Computers sind *wir*, als mit geistigen Fähigkeiten ausgestattete Programmierer oder Benutzer, diejenigen, die die physikalischen Zustände als syntaktisch und semantisch interpretieren und als solche verwenden:

In the case of the computer, an outside agent encodes some information in a form that can be processed by the circuitry of the computer. That is, he or she provides a syntactical realization of the information that the computer can implement in, for example, different voltage levels. The computer then goes through a series of electrical stages that the outside agent can interpret both syntactically and semantically, even though, of course, the hardware has no intrinsic syntax or semantics. Finally, an output is produced in the form of physical phenomena which an observer can interpret as symbols with a syntax and a semantics. (Searle, 1990, S. 34)

Aber wer soll diese Rolle im Fall des Gehirns einnehmen? Wenn es stets eines solchen Beobachters der physikalischen Eigenschaften oder Zustände eines Objektes bedarf, um diese als syntaktisch und semantisch zu interpretieren, scheint das Homunkulus-Problem unausweichlich, es ist der Auffassung des Gehirns als eines informationsverarbeitenden Systems endemisch. Diese Auffassung setzt voraus, dass es einen mit ‘geistigen Fähigkeiten’ ausgestatteten Beobachter gibt, der die physikalischen Zustände des Gehirns, die durch die Absorption von Licht auf den Retinae ausgelösten elektrischen Impulse der Neuronen, als syntaktisch und semantisch interpretiert, etwa *als* ‘Information über Tiefe’. Wir selbst können uns nicht als Benutzer unseres Gehirns oder visuellen Systems ansehen, in der Weise, in der wir einen Computer, eine Zange oder unsere Hände benutzen.<sup>30</sup> Dem Gehirn (gleiches gilt für einen Computer) kann ohne die Annahme eines Homunkulus nicht *per se* zugeschrieben werden, dass es etwas ‘berechne’ oder mit Repräsentationen operiere oder auch nur bloße Symbolmanipulationen vornehme. Berechnungen durchführen zu können oder etwas zu repräsentieren, sind *keine* physikalischen Eigenschaften, diese Beschreibungen sind notwendig auf ein intentionales Subjekt, eine mit geistigen Fähigkeiten ausgestattete Person angewiesen, die

<sup>30</sup> „The brain is not an organ with which we can do anything, though we cannot do anything without our brain“ (Smit & Hacker, 2014, S. 1082).

*etwas* als einen komputationalen Prozess, als eine Berechnung oder eine Repräsentation interpretiert oder verwendet. Das Gehirn aufgrund seiner physikalischen Eigenschaften als ein informationsverarbeitendes System anzusehen ist weder falsch noch richtig, es handelt sich um eine *beobachterrelative Zuschreibung*.

Die aktuellen, dem Informationsverarbeitungsparadigma zuzuordnenden Beschreibungen des Konzepts der Tiefencues sind daher als homunkulare Redeweisen zu kennzeichnen. Ein Extrahieren und Verarbeiten von (u.a.) ‘Informationen über Tiefe’ aus dem *retinal image* durch das visuelle System, das als Ergebnis die ‘Konstruktion’ einer symbolischen Beschreibung der Umwelt durch das visuelle System haben soll, setzt die Annahme einer mit geistigen Fähigkeiten ausgestatteten personenähnlichen Instanz voraus. Im Vergleich zu früheren Beschreibungen der Tiefencues, die darauf hindeuten, dass ein Homunkulus vorausgesetzt werden muss, der das retinale Lichtmuster als ein Bild *sieht* (um etwa bestimmte ‘Vergleiche’ vorzunehmen), muss in aktuellen, dem Informationsverarbeitungsparadigma verpflichteten Formulierungen ein eher ‘lesender’ bzw. physikalische Zustände *als etwas interpretierender* Homunkulus vorausgesetzt werden. Das visuelle System selbst als mit einem Symbolsystem operierend zu beschreiben, ist nicht sinnvoll, ohne einen solchen Homunkulus anzunehmen. Das Wahrnehmungssystem als physiologisch bestimmbarer Teil eines Menschen ist nicht das logisch angemessene Subjekt solcher Beschreibungen. Die als zentrales Problem der Tiefenwahrnehmung angesehene Frage, wie das Gehirn aus den zweidimensionalen Lichtmustern eine dreidimensionale symbolische Beschreibung der Welt ‘berechne’, ist, man kann Searle (1990, S. 28) hier nur zustimmen, ebenso sinnvoll wie die Frage: „How do nails compute the distance they are to travel in the board from the impact of the hammer and the density of the board?“. Dass die zweite Frage absurd ist, wird wohl auf wenig Widerrede stoßen, die erste Frage hingegen füllt wahrnehmungspsychologische Standardwerke. Zudem ist – wie bereits mehrfach erwähnt – die Kennzeichnung des Ergebnisses des Wahrnehmungsprozesses im komputationalen Paradigma als ‘symbolische Beschreibung der Umwelt’ merkwürdig, denn das Ergebnis des Wahrnehmungsprozesses ist *keine* symbolische Beschreibung, sondern eben eine *Wahrnehmung*. Mit anderen Worten: Selbst wenn man vom Problem der homunkularen Redeweisen absehen würde, wäre nicht klar, wie die ‘Konstruktion’ einer symbolischen Beschreibung der Umwelt durch das visuelle System zu unserem *Sehen* von Objekten führen sollte.

Nun könnte man als Vertreter solcher homunkularer Redeweisen entgegnen, dass es sich bei der vorgebrachten Kritik gleichsam um sprachliche Haarspalterei handle und Redeweisen dieser Art eher *metaphorisch* und *pragmatisch* zu verstehen seien, so wie man im Alltag etwa die Wendung gebrauche, dass einem eine innere Stimme etwas sage oder die Sinne sich an etwas erfreuten. Niemand, der solche homunkularen Redeweisen benutze, glaube tatsächlich, dass das visuelle System oder das Gehirn in

der üblichen Bedeutung dieses Ausdruckes etwas interpretiere, die Funktionsweise ließe sich aber so beschreiben, *als ob* es dies täte. Das Ziel der hier vorgebrachten Einwände gegen homunkulare Redeweisen soll es auch nicht sein, bestimmte Wendungen aus unserem Sprachgebrauch zu verbannen, da sie aus grammatikalisch-logischen Gründen nicht ‘erlaubt’ seien. Der relevante Punkt in diesem Fall ist nicht, welche Aussagen in Wahrnehmungstheorien enthalten sind, sondern welchem *Zweck* diese dienen sollen, insbesondere, welche theoretischen Erklärungsansprüche mit diesen Aussagen verbunden werden. Es spricht, wie bereits oben gesagt, prinzipiell nichts dagegen, bestehende Konventionen unserer Sprache zu erweitern oder zu verändern und es lassen sich ohne Zweifel Beispiele finden, in denen Metaphern in der Wissenschaft auf fruchtbare und erkenntnisbefördernde Weise benutzt wurden. Um jedoch psychologische Prädikate auf Subsysteme von Personen wie das Gehirn oder das visuelle System mit Sinn anwenden zu können, müssten neue oder andere Regeln der Verwendung für diese Ausdrücke zumindest angedeutet werden, um ihnen eine neue oder andere Bedeutung zu geben. Im wahrnehmungstheoretischen Kontext scheint es jedoch so zu sein, dass die in homunkularen Redeweisen vorkommenden psychologischen Prädikate entweder dieselbe Bedeutung haben sollen, die sie bei der Anwendung auf Personen haben (etwa bei Gregory, 1998, 2002), oder, im Rahmen des Informationsverarbeitungsansatzes, Begriffe wie der der Interpretation rein *komputationale* Prozesse bezeichnen sollen. In beiden Fällen, so wurde versucht zu zeigen, handelt es sich um homunkulare Redeweisen. Im ersten Fall soll – falls diese Begriffe lediglich als Metaphern dienen sollen – das ‘als ob’, der metaphorische Teil dieser Beschreibungen, die gesamte Erklärungsleistung erbringen. Wenn visuelle Wahrnehmung durch das Bilden von Hypothesen und das Bewerten vorliegender Evidenz durch das visuelle System verständlich gemacht werden soll, diese Begriffe jedoch nicht wörtlich zu nehmen seien, ist nicht deutlich, inwiefern es sich dabei um einen fruchtbaren Ansatz handeln soll.<sup>31</sup> Im zweiten Fall bekommen die Begriffe eine klarere Bedeutung, nämlich unter Bezugnahme auf symbolische Repräsentationen, mit denen das visuelle System operiere, und deren Transformationen. Die Brauchbarkeit dieses Ansatzes lässt sich aufgrund der zahlreichen Erträge zugehöriger empirischer Untersuchungen nicht bestreiten, der Ansatz selbst weist jedoch den logischen Mangel des Homunkulismus auf.

Gleichzeitig lassen sich in der Philosophie des Geistes auch eher offensive Einwände gegen den Vorwurf des Homunkulismus in kognitions- oder neurowissenschaftlichen Theorien finden.<sup>32</sup> Dennett (1981, 1991) etwa wendet sich gegen den Einwand, dass

<sup>31</sup>Keil (2003, S. 10) führt zu diesem Punkt folgende Anekdote aus dem 19. Jahrhundert zur Funktionsweise des Rundfunks an: „Stell dir einen ganz langen Dackel vor, der von Boston nach New York reicht. Wenn man ihn in Boston am Schwanz zieht, dann bellt es in New York. Und jetzt denk dir den Dackel weg: So funktioniert Rundfunk!“

<sup>32</sup>Eine äußerst lesenswerte Auseinandersetzung um die vorgebrachten Argumente zur *mereological fallacy* in den Neurowissenschaften lässt sich in Bennett, Dennett, Hacker und Searle (2007) finden.

das Postulieren eines Homunkulus zur Erklärung einer bestimmten kognitiven, intentionalen Leistung zwangsläufig zu einem unendlichen Regreß führe. Er behauptet, dass die von dem postulierten Homunkulus auszuführende Leistung schrittweise in immer 'anspruchlosere' Tätigkeiten aufgegliedert werden könnte, die von einem „team or committee of *relatively* ignorant, narrow-minded, blind homunculi“ ausgeführt werden könnten, „to produce the intelligent behavior of the whole“ (Dennett, 1981, S. 123). Auf der untersten Stufe dieser 'Arbeitsteilung' komme so der Regreß zu einem Ende. Die von den „stupid homunculi“ ausgeführten Subprozesse seien rein mechanisch, „no part ... has to understand what it is doing and why“ (Dennett, 1991, S. 91). Der Homunkulus könne so auf der untersten Stufe durch eine Maschine ersetzt werden:

When the level is reached where the homunculi are no more than adders and subtractors, by the time they need only the intelligence to pick the larger of two numbers when directed to, they have been reduced to functionaries who can be replaced by a machine. (Dennett, 1981, S. 80)

Dieses Gegenargument kann meines Erachtens letztlich nicht überzeugen, da sie das aufgeworfene Problem nicht lösen kann. Dennetts Vorstellung läuft darauf hinaus, dass durch das Verhalten des gesamten Systems der „stupid homunculi“ Semantik gleichsam graduell 'generiert' werde. Aber *welche?* Es gibt potentiell unendlich viele Möglichkeiten, die für die Homunkuli bedeutungslosen Symbolmanipulationen mit einem semantischen Gehalt zu *versehen*, dieser ergibt sich nicht aus der Symbolmanipulation selbst. Die vorgenommenen Symbolmanipulationen könnten für alles Mögliche stehen, etwa das Lösen einer Differentialgleichung oder die Züge in einem Schachspiel. Semantik, wir kehren hier wieder auf das Argument des chinesischen Zimmers zurück, ist der Syntax nicht intrinsisch. Weiterhin müssten auch hier zunächst die physikalischen Zustände des Systems als regelgeleitet, als syntaktisch *interpretiert* werden, was Dennett in seiner Beschreibung schon voraussetzt. Interessanterweise charakterisiert Dennett die 'dummen Homunkuli' auf der untersten Stufe zudem immer noch als intentional, als Empfänger von Anweisungen, so dass es sich hier nicht um eine rein mechanistische Beschreibung handelt, sondern um eine weitere homunkulare Redeweise. Die Frage ist nicht, 'wie viel' Intentionalität oder Intelligenz man einem Homunkulus zuschreiben muss, sondern ob intentionale und repräsentationale Phänomene, so wie Dennett behauptet, naturalisierbar sind. Intentional charakterisierte Homunkuli bleiben solche, unabhängig von der Komplexität der Aufgabe, die sie zu erfüllen haben.

Es lässt sich allerdings die allgemeine Frage aufwerfen, ob solche mit dem Problem homunkularer Redeweisen verbundenen theoretischen Erörterungen für die Kognitions- oder Neurowissenschaften überhaupt von Relevanz sind. So lassen sich durchaus Stimmen finden, die solche Analysen im Endeffekt als unergiebig, ja geradezu hinderlich für einen Erkenntnisfortschritt auf diesen Gebieten ansehen:

It is therefore fortunate that neurobiologists are not philosophers, for they might otherwise find themselves immersed, like the philosophers, in an endless and ultimately fruitless discussion on the meaning of words such as ‚unconscious‘ or ‚inference‘ or ‚knowledge‘ and ‚information‘ instead of trying to unravel important facts about the brain. They would, in brief, end up contributing as meagerly to an understanding of the brain and the mind as philosophers have. (Zeki, 1993, S. 7)

Sollte sich der an Wahrnehmung interessierte Forscher auf ein Zusammentragen von ‚Fakten‘ beschränken, in der Hoffnung, durch die zunehmende Menge an empirischen Ergebnissen zu einem zunehmenden Verständnis von visueller Wahrnehmung zu gelangen? In diesem Kapitel wurde u.a. dafür argumentiert, dass jede empirische Aussage eines gewissen konzeptuellen Rahmens, gewisser theoretischer Vorstellungen bedarf, die gleichsam das generelle Schema der Beschreibung der erhaltenen Ergebnisse bestimmen. Nur vor dem Hintergrund bestimmter theoretischer Vorstellungen lassen sich ‚Fakten‘ feststellen und ‚enträtseln‘. Es lässt sich nicht empirisch feststellen, ob das visuelle System etwas ‚berechnet‘, mit ‚internen Repräsentationen‘ operiert oder Informationen ‚interpretiert‘, es lässt sich aber erörtern, ob die diesen Aussagen zugrundeliegenden theoretischen Vorstellungen sinnvoll sind oder nicht. Die geäußerte Kritik, das soll betont werden, bezieht sich nicht auf die *empirischen Erträge* der Wahrnehmungsforschung, sondern auf den konzeptuellen Rahmen, auf begriffliche Voraussetzungen und Implikationen, auf dessen Grundlage empirische Befunde beschrieben und interpretiert werden. Nicht sinnvolle konzeptuelle Rahmen können möglicherweise für die in Frage stehende empirische Wissenschaft selbst schädlich sein, da sie nicht sinnvolle Fragen und entsprechende Forschungsprogramme zur Folge haben könnten. Ein gutes Beispiel hierfür kann man in der durch Keplers und Descartes‘ Entdeckung des invertierten ‚Netzhautbildes‘ aufgeworfenen Frage sehen, auf welche Weise dieses invertierte ‚Bildchen der äußeren Gegenstände‘ im Wahrnehmungsprozess wieder ‚umgedreht‘ wird, damit wir die Welt nicht als auf dem Kopf stehend wahrnehmen. Dass diese Frage als nicht sinnvoll und nicht fruchtbar erkannt wurde, ist nicht auf zunehmende empirische Ergebnisse zurückzuführen, sondern auf theoretische Erörterungen.

## 4 | Empirische Untersuchung

„Das Denken ist ein rätselhafter Vorgang, von dessen vollem Verständnis wir noch weit entfernt sind.‘ Und nun stellt man Experimente an. Offenbar, ohne sich bewußt zu sein, *worin* das Räselhafte des Denkens für uns liegt. Die experimentelle Methode tut *etwas*; daß sie das Problem nicht löst, schiebt man darauf, daß sie noch in ihren Anfängen liegt. Es ist, als wollte man durch chemische Experimente feststellen, was Materie, und was Geist ist.“

---

*Wittgenstein (1984, S.197)*

In diesem Kapitel soll exemplarisch anhand einer eigenen empirischen Untersuchung zur Bildwahrnehmung verdeutlicht werden, auf welche Weise das *cue*-Konzept, unter Berücksichtigung der in den vorangegangenen Abschnitten angestellten Überlegungen, zur Formulierung von Hypothesen und der Interpretation empirischer Befunde herangezogen werden kann. Zunächst soll zum besseren Verständnis näher auf die spezifischen Eigenheiten der Wahrnehmung linearperspektivischer oder realistischer Bilder eingegangen werden, anschließend sollen die experimentelle Untersuchung und die erhaltenen Befunde dargestellt und im Rahmen der Ausrichtung dieser Arbeit diskutiert werden.

### 4.1. Exkurs: Bildwahrnehmung

Bilder sind in unserer immer stärker visuell orientierten Kultur beinahe allgegenwärtig und spielen eine zentrale Rolle in so gut wie allen Bereichen des menschlichen Lebens; man denke etwa an Fernsehen, Kino oder Fotografien und Abbildungen in Zeitschriften und Illustrierten. Dennoch spielt innerhalb des Feldes der Wahrnehmungspsychologie die theoretische Behandlung der Wahrnehmung von Bildern nur eine äußerst marginale Rolle; in den üblichen Standardlehrbüchern zur visuellen Wahrnehmung taucht dieses Thema in der Regel nicht auf. Dies ist bemerkenswert, da bei experimentellen Untersuchungen zur visuellen Wahrnehmung nahezu ausschließlich *Bilder* als Reize verwendet werden, ohne hinreichend theoretisch geklärt zu haben, ob sich bei der Wahrnehmung von Bildern möglicherweise andere Phänomene zeigen, als es bei der Wahrnehmung

natürlicher Szenen der Fall ist. Ein möglicher Grund für eine derart 'stiefmütterliche' Behandlung dieses Themas könnte die implizite Überzeugung vieler Wahrnehmungsforscher sein, dass sich die Wahrnehmung von Bildern in der Tat nicht wesentlich von der Wahrnehmung natürlicher Szenen unterscheidet und somit jede Wahrnehmungstheorie diesen Bereich gleichsam automatisch mit einschließt, ohne dass dieser eine separate Behandlung erfahren müsste. Ob diese Überzeugung gerechtfertigt ist, beziehungsweise ob das derzeit übliche theoretische Schema zur Erklärung visueller Wahrnehmung in der Tat geeignet ist, die mit der Betrachtung von Bildern einhergehenden Phänomene verständlich zu machen, ist zumindest fraglich.

Das derzeit übliche theoretische Schema des Wahrnehmungsprozesses wurde im Rahmen des letzten Kapitels bereits deutlich zu machen versucht. Licht werde von einem physikalischen Objekt oder einer Szene (dem sogenannten *distalen Reiz*) reflektiert und ein Teil dieses reflektierten Lichts, den Gesetzen der Strahlenoptik folgend, auf die beiden Retinae des Betrachters projiziert. Dieses auf die Retinae projizierte Lichtmuster (der sogenannte *proximale Reiz*) löse nun bestimmte elektrochemische Reaktionen in den Netzhautrezeptoren aus, welche über die mit den Rezeptoren verbundenen Nervenbahnen in bestimmte Areale des Gehirns weitergeleitet und schließlich dort als 'Informationen' über den distalen Reiz 'verarbeitet' würden. Im Rahmen dieses Prozesses spielt das im Fokus der vorliegenden Arbeit stehende Konzept der Tiefencues eine wichtige Rolle. Wie im letzten Kapitel dargestellt, komme es zu einer 'Interpretation' der verschiedenen Tiefencues als 'Informationen über Tiefe' sowie zu einer Integration dieser Informationen durch das visuelle System. Das Ergebnis dieser Informationsverarbeitung soll dann das Haben einer ganzheitlichen Wahrnehmung einer spezifischen Szene oder eines Objektes (das 'Perzept') sein. Üblicherweise wird der kritische Übergang von elektrochemischen Vorgängen im Hirn zum Haben einer Wahrnehmung seitens des Betrachters von einer ausführlichen theoretischen Behandlung ausgeklammert, indem etwa davon gesprochen wird, dass das visuelle System das Perzept schlicht 'generiere'. In der Regel enthalten die derzeit üblichen theoretischen Ansätze auch eine mehr oder minder explizite Überzeugung, dass eine gewisse Korrespondenz zwischen dem physikalischen Objekt oder der Szene und dem zugehörigen Wahrnehmungseindruck bestehe.

Bestimmte, mit der Wahrnehmung von Bildern verbundene Phänomene, stellen den so orientierten Wahrnehmungsforscher allerdings vor besondere theoretische Herausforderungen. Eines der bei der Betrachtung von Bildern sicher auffälligsten Phänomene, insbesondere bei Betrachtung linearperspektivischer Bilder oder Fotografien, auf die wir uns in der Behandlung beschränken wollen, ist, dass ein mehr oder minder prägnanter *Tiefeneindruck* mit der Betrachtung verbunden ist, obwohl es sich bei einem Bild physikalisch um eine *Ebene* handelt, deren Oberfläche in bestimmter Weise bearbeitet

wurde. Die Tiefencues scheinen in diesem Fall, folgt man dem üblichen theoretischen Ansatz, gleichsam ‘irreführende Informationen’ zu ‘liefern’. Bei der Betrachtung realistischer Bilder entsteht der Eindruck, als ob es gleichsam ‘in dem Bild weitergehe’ oder man ‘in das Bild hineinsehen’ könne, obgleich es sich um eine Fläche, um ein physikalisches Objekt in fester Entfernung handelt. Betrachtet man etwa Abbildung 4.1 auf der nächsten Seite, eine 1942 hergestellte Fotografie des *Grand Teton Nationalparks* (Wyoming, USA) von Ansel Adams, so kann man sich des Eindrucks nicht erwehren, dass die dargestellten Berge weiter entfernt sind als der Wald, sich der Fluss gleichsam immer weiter in die Tiefe schlängelt usw. Eine lange Zeit vorherrschende und auch heute noch anzutreffende Auffassung dieses Phänomens ist, dass die Betrachtung eines realistischen Bildes im Wesentlichen so aufzufassen sei, als würde der Betrachter die im Bild dargestellte Szene gleichsam ‘in der Zeit eingefroren durch ein offenes Fenster’ betrachten. Diese, auch als ‘Alberti-Fenster’ bekannte Vorstellung geht auf Leon Battista Alberti, einen italienischen Kunsttheoretiker der Frührenaissance, zurück.

Die linearperspektivische Darstellung eines Objektes oder einer Szene soll zum Ziel haben, unter gewissen Restriktionen wie einäugiger Betrachtung von einem bestimmten Punkt aus, idealiter *dasselbe* Lichtmuster auf die Retina des Betrachters des Bildes zu projizieren, das auch bei einäugiger Betrachtung der dargestellten Szene projiziert worden wäre.<sup>1</sup> Laut dieser theoretischen Konzeption unterscheidet sich somit – zumindest im Prinzip – die Wahrnehmung eines Bildes, insbesondere des damit verbundenen Tiefeneindrucks (der im Folgenden auch als *Bildraum* bezeichnet werden soll) *nicht* von der Wahrnehmung einer natürlichen Szene. Der Bildraum ließe sich gemäß dieser Konzeption als ‘optische Täuschung’ kennzeichnen, ganz analog etwa den Wahrnehmungseindrücken, die mit dem berühmten Ames-Raum verbunden sind. Genauere Analysen des mit der Bildbetrachtung einhergehenden Wahrnehmungseindrucks weisen jedoch stark daraufhin, dass diese Fenstermetapher unangemessen ist. So ist der mit einem realistischen Bild verbundene Wahrnehmungseindruck – mit Ausnahme sogenannter *trompe l’œil*-Bilder, von denen noch die Rede sein wird – von einer ‘optischen Täuschung’ im üblichen Sinn doch deutlich verschieden. Im Gegensatz zu einer ‘Täuschung’ weist der mit einem Bild verbundene Wahrnehmungseindruck einen bestimmten *dualen Charakter* auf, womit hier gemeint ist, dass der Betrachter das Bild als eine *Fläche* im physikalischen Raum erlebt *und gleichzeitig* als etwas, das mit einem bestimmten Tiefeneindruck, dem Bildraum, verbunden ist.

Ein weiteres mit der Betrachtung realistischer Bilder oder Fotografien verbundenes Phänomen ist, dass sich bei Bewegung des Betrachters vor dem Bild der zugehörige Wahrnehmungseindruck in der Regel und in gewissen Grenzen nicht auffallend zu ver-

---

<sup>1</sup>In Abschnitt 4.1.1 werden wir näher auf die geometrischen Grundlagen der linearperspektivischen Darstellung eingehen.



Abbildung 4.1. Ansel Adams, „The Tetons and the Snake River“ (1942). Quelle: *The National Archives Catalog*.

ändern scheint. Die grundsätzliche Idee der linearperspektivischen Darstellung einer Szene, nämlich durch diese idealiter dasselbe Lichtmuster auf die Retina zu projizieren wie die dargestellte Szene selbst, schreibt jedoch vor, dass das Bild aus einem ganz bestimmten Punkt – dem sogenannten *Augpunkt* – heraus betrachtet werden müsse. Betrachtet man das Bild außerhalb des Augpunktes, so stellen sich geometrisch beschreibbare systematische ‘Verzerrungen’ des auf die Retina des Betrachters projizierten Lichtmusters ein, die gemäß der Idee des Alberti-Fensters sich auch in einem unterschiedlichen Wahrnehmungseindruck niederschlagen sollten. Die wahrgenommenen Veränderungen bei Betrachtung eines Bildes außerhalb des Augpunktes sind allerdings vergleichsweise moderat; die wahrgenommene Tiefenstruktur scheint relativ ‘robust’ gegenüber der Bewegung vor dem Bild zu sein.

Diese beiden bildspezifischen Phänomene, der duale Charakter der Bildwahrnehmung und die relative Robustheit des Wahrnehmungseindrucks bei Positionswechsel des Betrachters, sollen im Folgenden ausführlicher dargestellt werden, da sie wesentlich für Idee und Diskussion der experimentellen Untersuchung sind. Zunächst soll jedoch in angemessener Kürze die Konstruktion linearperspektivischer Bilder erläutert werden, da diese sowohl für das Verständnis der experimentellen Untersuchung notwendig ist,

als auch eine Grundlage für die Diskussion der beiden zu behandelnden bildspezifischen Phänomene darstellt.

### 4.1.1. Grundlagen der zentralprojektiven Darstellung

Bei der Zentralprojektion handelt es sich um eine geometrische Methode, dreidimensionale Gebilde nach bestimmten Prinzipien auf eine Ebene, die sogenannte *Bildebene*, abzubilden. Am prominentesten wird die Zentralprojektion ‘als Handwerk’ in der bildenden Kunst bei der Erstellung linearperspektivischer Bilder genutzt. Ein linearperspektivisches Bild ist ein Bild, das nach der geometrischen Idee der Zentrakprojektion gefertigt wurde. Erklärtes Ziel einer solchen linearperspektivischen Darstellung ist es, wie bereits erwähnt, eine bestimmte dreidimensionale Szene derart auf eine Fläche abzubilden oder zu projizieren, so dass beim Betrachter dieser Fläche, zumindest im Prinzip und unter Voraussetzung bestimmter Randbedingungen wie monokularer Betrachtung von einer bestimmten Position aus, das gleiche Lichtmuster auf die Retina projiziert wird, wie bei Betrachtung der Originalszene. Die Originalszene und die linearperspektivische Darstellung dieser müssten daher idealiter einen *identischen Effekt* auf das visuelle System haben. Niederée und Heyer (2003) sprechen in diesem Zusammenhang auch davon, dass es sich bei Originalszene und der linearperspektivischen Darstellung dieser um *äquivalente Stimuli* handle bzw. das Ziel der linearperspektivischen Darstellung darin bestehe, einen zu einer Szene äquivalenten Stimulus herzustellen.<sup>2</sup>

Die Entwicklung einer *mathematischen Theorie* der Zentralprojektion lässt sich zurückverfolgen bis in die italienische Frührenaissance, wobei Florenz als Zentrum dieser Entwicklung auszumachen ist. Üblicherweise wird die 1435 erschienene Schrift „*De pictura*“ von Leon Battista Alberti als erste systematische Auseinandersetzung mit den geometrischen Prinzipien der Zentralprojektion angesehen (Andersen, 2007).<sup>3</sup> Diese mathematische Theorie der Zentralprojektion soll nun in jenen Aspekten, welche zum Verständnis der empirischen Untersuchung sowie der Diskussion der oben erwähnten bildspezifischen Phänomene erforderlich sind, dargestellt werden. Es sollen zunächst zwei eher anschauliche Charakterisierungen der Idee der Zentralprojektion gegeben werden, denen sich dann eine für den mathematisch interessierten Leser gedachte, formale Darstellung anschließt.

---

<sup>2</sup>In der Regel ist es allerdings nicht möglich, mit den Mitteln der linearperspektivischen Darstellung *dasselbe* Lichtmuster auf die Retina zu projizieren, wie es im Fall der Originalszene geschehen würde. So lassen sich etwa bestimmte Farbsättigungen oder Kontraste nicht bildlich reproduzieren.

<sup>3</sup>Die Frage nach dem ‘Ursprung’ der perspektivischen Darstellung wird allerdings seit langem diskutiert, wobei die unterschiedlichen Positionen sich weitestgehend dahin unterscheiden, was deren Vertreter unter ‘Ursprung’ und ‘perspektivischer Darstellung’ verstehen. Siehe zu dieser Diskussion etwa Panofsky (1927/1980), Kubovy (1988), Kemp (1990), Brener (2004), Andersen (2007) und Damisch (2010).

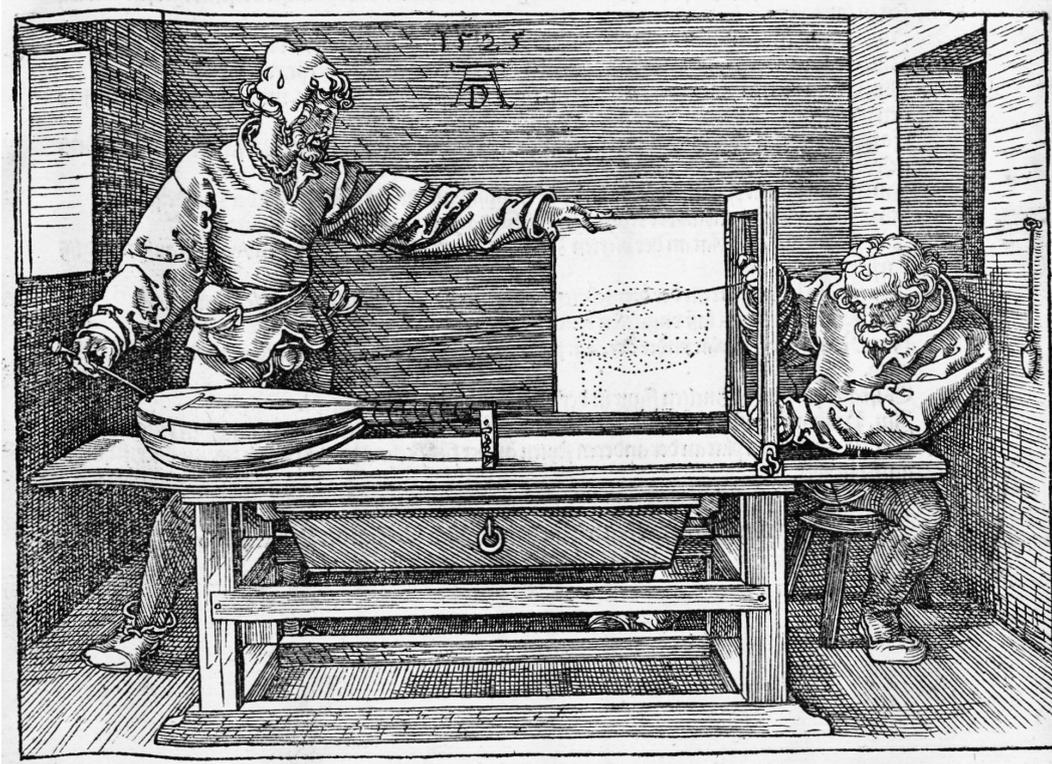


Abbildung 4.2. Abbildung 1 aus Seite 181 von Albrecht Dürers 1525 erschienener Schrift „Unerweysung der Messung, mit dem Zirckel und Richtscheyt, in Linien, Ebenen unnd gantzen corporen“. Erläuterungen im Text. Quelle: *Wikimedia Commons*

Das Prinzip der Zentralprojektion lässt sich in besonders anschaulicher Weise durch einen Stich Albrecht Dürers verdeutlichen (vgl. Abbildung 4.2). Dieses Bild lässt sich gleichsam als Anleitung zur Erstellung einer linearperspektivisch korrekten Darstellung eines dreidimensionalen Objektes (hier einer Laute) auffassen und bietet gleichzeitig eine anschauliche und klare Darstellung des Projektionsbegriffes. Eine mit einem Gewicht versehene Schnur ist anhand einer Öse an einem festen Punkt installiert. Zwischen diesem festen Punkt und der abzubildenden Laute befindet sich die Bildebene, welche senkrecht zu jener Ebene ist, auf der die Laute liegt. Die perspektivisch korrekte Darstellung der Laute wird nun (näherungsweise) erzeugt, indem das nicht mit dem Gewicht versehene Ende dieser Schnur von der Person links im Bild an hinreichend viele Punkte der Laute angelegt wird und die Schnittpunkte der Schnur mit der Bildebene auf dieser von der Person rechts im Bild markiert werden. Elementar dabei ist, dass über das Gewicht die Schnur so unter Spannung gehalten wird, dass sie die kürzeste Verbindung zwischen dem fest installiertem Punkt und dem abzubildenden Punkt der Laute darstellt. Wir werden im Rahmen der eher formaleren Analyse der Zentralprojektion sehen, dass der feste Punkt, an dem die Schnur installiert ist, den sogenannten *Augpunkt* repräsentiert und die Schnur selbst die sogenannten *Projektionsstrahlen*, mit

deren Hilfe die einzelnen Punkte der Oberfläche der Laute auf die Bildebene projiziert werden.

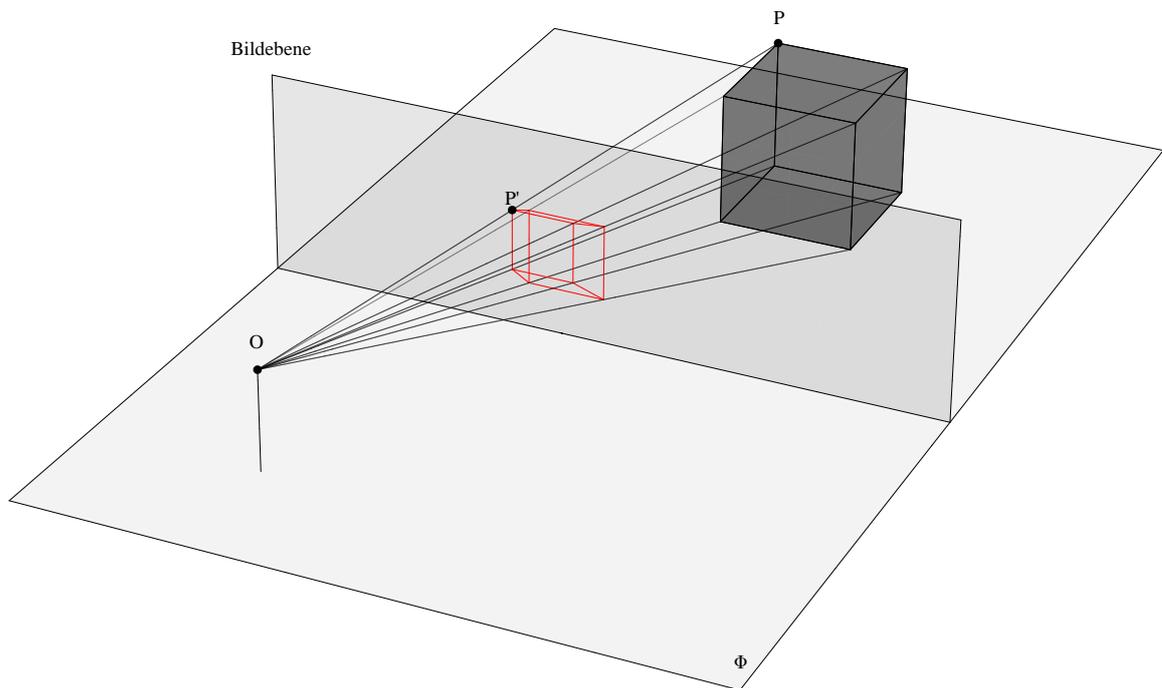
Ein wenig abstrakter liesse sich die Zentralprojektion auch so beschreiben (siehe Abbildung 4.3 auf der nächsten Seite): Gegeben sei ein beliebiger geometrischer Körper (etwa ein Würfel), der mittels Zentralprojektion auf eine Bildebene abgebildet werden soll. Weiterhin sei  $O$  ein ausgezeichneter Punkt, der Augpunkt, der sich im Vergleich mit dem abzubildenden Körper auf der anderen Seite der Bildebene befindet. Verbindet man nun hinreichend viele Punkte auf der Oberfläche des betrachteten Körpers durch Strahlen mit dem Augpunkt, so nennt man diese Strahlen auch *Projektionsstrahlen* und die Menge aller Projektionsstrahlen auch die *Sehpyramide* oder den *Sehkegel*. Die Zentralprojektion des geometrischen Körpers lässt sich nun charakterisieren als *Schnitt* der Bildebene mit dem Sehkegel. Die Punkte, in denen die Projektionsstrahlen die Bildebene durchstossen, nennt man auch die Bildpunkte und die Menge aller Bildpunkte auch das *Zentralbild* des Körpers auf der Bildebene.

Um einige wesentliche Eigenschaften der Zentralprojektion herausarbeiten zu können, soll nun eine formalere Darstellung erfolgen. Es werde im Folgenden ein dreidimensionaler Raum  $\mathcal{V}$  sowie ein ausgezeichneter Punkt  $O$  in  $\mathcal{V}$  (der Augpunkt) betrachtet. Weiterhin seien  $\Xi$  und  $\Phi$  zwei Ebenen in  $\mathcal{V}$ . Die Ebene  $\Xi$  soll im Folgenden als *Bildebene* und  $\Phi$  als *Grundebene* bezeichnet werden, wobei der Augpunkt  $O$  nicht Element der Bildebene sein soll. Der Schnitt von  $\Xi$  und  $\Phi$  wird gelegentlich auch als *Grundlinie* bezeichnet. Zum Zwecke der Anschaulichkeit soll zusätzlich angenommen werden, dass es sich bei  $\Xi$  um eine zu  $\Phi$  orthogonale Ebene handelt. Für jeden Punkt  $P$  in  $\mathcal{V}$ , der bezüglich  $O$  nicht auf derselben Seite von der Bildebene  $\Xi$  liegt, werde nun diejenige Gerade betrachtet, die durch  $O$  und  $P$  gegeben ist. Der Schnittpunkt  $P'$  dieser Geraden mit der Bildebene heiße Projektion von  $P$  auf  $\Xi$  bezüglich  $O$  und soll im Folgenden auch kurz als *Zentralbild* des Punktes  $P$  auf der Bildebene bezeichnet werden. Das eben Gesagte lässt sich auch als zentraler Satz der Zentralprojektion kennzeichnen, da sich sämtliche der im Folgenden darzustellenden Eigenschaften der Zentralprojektion aus diesem ableiten lassen.<sup>4</sup>

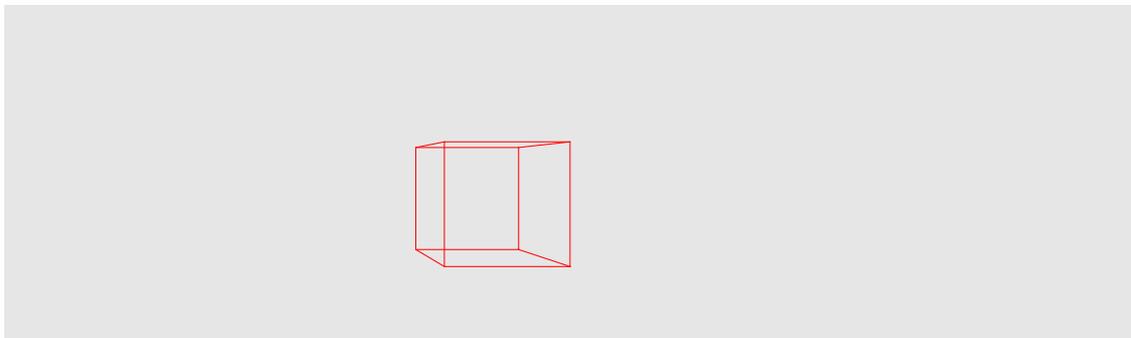
Betrachtet man eine nicht zur Bildebene parallele Gerade  $g$  in  $\mathcal{V}$ , so nennt man den Schnittpunkt von  $g$  mit der Bildebene auch den *Spurpunkt* von  $g$ . Betrachtet man nun jene Gerade, die parallel zu  $g$  und durch den Augpunkt  $O$  verläuft, so nennt man den Schnittpunkt dieser Geraden mit der Bildebene auch den *Fluchtpunkt* der Geraden  $g$ . Der Fluchtpunkt von  $g$  ist somit auch der Fluchtpunkt *aller* zu  $g$  parallelen Geraden. Aufgrund der oben vorgenommenen Einschränkung der Betrachtung auf nur

---

<sup>4</sup>Die vorgenommene Einschränkung auf Punkte in  $\mathcal{V}$ , die bezüglich  $O$  nicht auf derselben Seite der Bildebene liegen, ist hier der Anschaulichkeit der Darstellung geschuldet. Die Zentralprojektion ließe sich auch allgemeiner einführen, allerdings wäre dies mit einem wesentlich größeren Begriffsaufwand verbunden, der für unsere Zwecke nicht gerechtfertigt erscheint.



(a)



(b)

Abbildung 4.3. Anschauliche Darstellung des Prinzips der Zentralprojektion als Schnitt durch den Sehkegel. Jeder Punkt des auf die Bildebene zu projizierenden Würfels wird über einen Projektionsstrahl mit dem Augpunkt verbunden (a). Die Schnittpunkte jedes Projektionsstrahls mit der Bildebene bilden die Bildpunkte des Würfels auf der Bildebene und die Menge der Bildpunkte das Zentralbild des Würfels auf der Bildebene (b).

jene Punkte in  $\mathcal{V}$ , die bezüglich  $O$  nicht auf derselben Seite der Bildebene liegen, soll im Folgenden  $g$  stets als *Halbgerade* aufgefasst werden, d.h. als die Menge all jener Punkte in  $g$ , die bezüglich des Augpunktes nicht auf derselben Seite der Bildebene liegen. Die Projektion dieser Punkte auf die Bildebene (bezüglich  $O$ ) soll im Folgenden auch kurz als *Zentralbild von  $g$*  bezeichnet und durch  $g'$  abgekürzt werden. Spurpunkt und Fluchtpunkt von  $g$  lassen sich als ausgezeichnete Punkte des Zentralbildes dieser

Geraden in der Bildebene auffassen, da sich das Zentralbild  $g'$  als Strecke zwischen Spurpunkt und Fluchtpunkt ausschließlich des Fluchtpunktes kennzeichnen lässt, wie durch Abbildung 4.4 verdeutlicht werden soll.

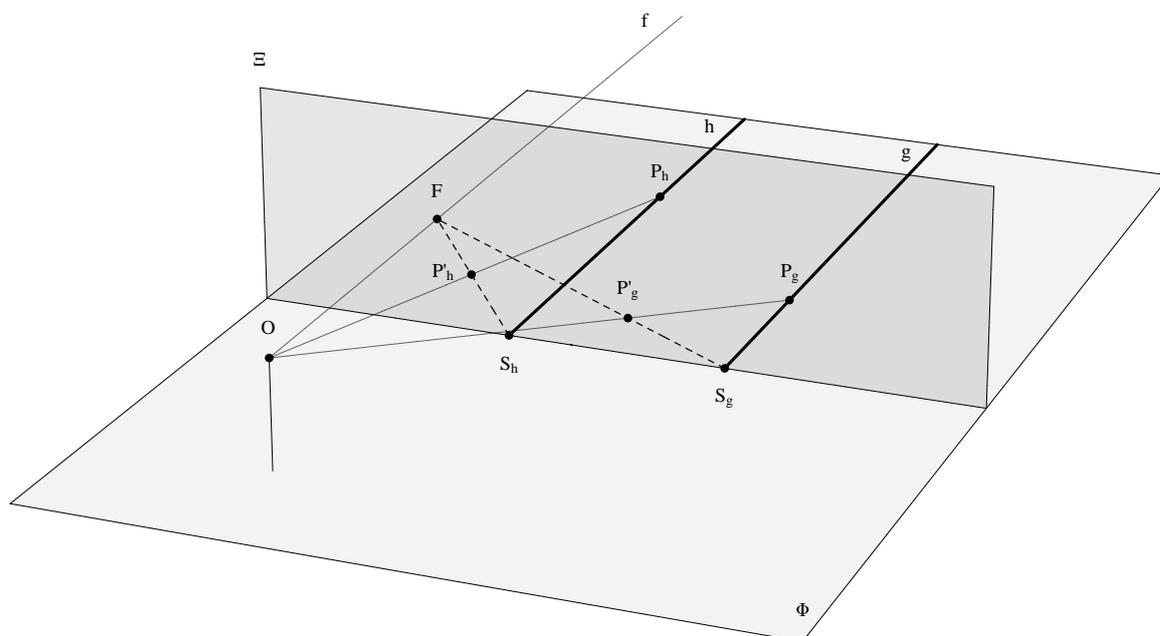


Abbildung 4.4. Geometrische Darstellung des Prinzips der Zentralprojektion. Siehe Text für Erläuterungen.

In Abbildung 4.4 seien  $g$  und  $h$  zwei parallele Geraden in der Grundebene  $\Phi$ , die jedoch nicht zur Bildebene  $\Xi$  parallel seien.  $S_g$  bezeichne den Spurpunkt von  $g$  und  $S_h$  den Spurpunkt von  $h$ . Weiterhin seien  $P_g$  und  $P_h$  Punkte auf  $g$  bzw.  $h$ , die entsprechenden Zentralbilder in der Bildebene seien mit  $P'_g$  bzw.  $P'_h$  bezeichnet. Für die abgeschlossene Strecke zwischen zwei Punkten  $A$  und  $B$  soll im Folgenden auch kurz  $[A, B]$  geschrieben werden, die halboffene Strecke zwischen  $A$  und  $B$  ausschließlich des Punktes  $A$  bzw. ausschließlich des Punktes  $B$  soll durch  $(A, B)$  bzw.  $[A, B)$  gekennzeichnet werden. Die Zentralbilder  $P'_g$  und  $P'_h$  der Punkte  $P_g$  bzw.  $P_h$  sind nun gerade die Schnittpunkte der Strecke  $[P_g, O]$  bzw.  $[P_h, O]$  mit der Bildebene. Die durch  $O$  und parallel zu  $g$  und  $h$  verlaufende Gerade ist mit  $f$  bezeichnet und der Schnittpunkt  $F$  dieser Geraden mit der Bildebene als der gemeinsame Fluchtpunkt von  $g$  und  $h$  eingeführt worden. Man kann sich nun anschaulich klar machen, dass bei Veränderung der Lage eines Punktes  $P_g$  auf der Geraden  $g$  die Veränderung des zugehörigen Zentralbildes  $P'_g$  auf die halboffene Strecke  $(S_g, F)$  in der Bildebene beschränkt ist. Nähert sich  $P_g$  der Bildebene, so nähert sich  $P'_g$  immer stärker dem Spurpunkt von  $g$  an. Bei einer verschwindenden Distanz zwischen Bildebene und  $P_g$  sind die Punkte  $P_g$  und  $P'_g$  identisch und identisch

dem Spurpunkt von  $g$ . Je weiter sich der Punkt  $P_g$  von der Bildebene entfernt, so nähert sich  $P'_g$  immer stärker dem Fluchtpunkt  $F$  an. Wächst die Distanz zwischen  $P_g$  und der Bildebene über alle Grenzen, so ergibt sich als Grenzwert der Bildpunkte  $P'_g$  der Fluchtpunkt  $F$ . Analoge Ausführungen gelten für den Punkt  $P_h$  und die Gerade  $h$ , sowie für alle weiteren, zu  $g$  parallelen Geraden. Alle zu  $g$  parallelen Halbgeraden laufen somit auf denselben Fluchtpunkt  $F$  zu. Betrachtet man *sämtliche* Geraden, welche parallel zur Grundebene verlaufen, so nennt man die Menge aller zugehörigen Fluchtpunkte auch den *Horizont*. Man beachte, dass die 'Höhe' des Horizontes in der Bildebene, d.h. die Lotdistanz zwischen Horizont und Grundebene von der Lotdistanz des Augpunktes zur Grundebene abhängig ist.

Eine nicht zuletzt für die Wahrnehmungsforschung überaus wichtige Eigenschaft der Zentralprojektion, die sich daraus ergibt, dass es sich anschaulich gesprochen um eine Abbildung von einem dreidimensionalen Raum in eine Ebene handelt, ist, dass es zu *einem beliebigen Zentralbild unendlich viele* mögliche dreidimensionale Körper oder Szenen gibt, deren Zentralprojektion gerade dieses Zentralbild ist. Dies ist im Kontext von Theorien visueller Wahrnehmung auch als '*inverse projection-Problem*' bekannt. Abbildung 4.5 auf der nächsten Seite, eine Erweiterung einer Darstellung aus der berühmten Abhandlung von Taylor (1719) über die linearperspektivische Darstellung, verdeutlicht dies auf anschauliche Weise. Dargestellt ist hier ein Zentralbild auf einer Bildebene, der Punkt  $O$  bezeichnet auch hier wieder den Augpunkt. Man beachte, dass dieses Zentralbild nun sowohl das Zentralbild des Quaders mit den Punkten  $A, B, C$  und  $D$  als auch des etwas weiter von der Bildebene entfernten Körpers mit den Punkten  $V, W, X$  und  $Y$  ist, da es zu jedem Punkt des einen Körpers einen 'korrespondierenden Punkt' des anderen Körpers gibt, der auf *demselben* Projektionsstrahl liegt und somit auf denselben Punkt in der Bildebene abgebildet wird (die Punkte  $B$  und  $V$  etwa bilden ein solches Paar korrespondierender Punkte, deren Zentralbild identisch ist). Man macht sich nun leicht klar, dass es zu jedem beliebigen Punkt  $P$  auf einem beliebigen Projektionsstrahl unendlich viele andere Punkte auf demselben Projektionsstrahl gibt, deren Zentralbild identisch ist dem Zentralbild von  $P$ . Oder, mit anderen Worten, zu einem beliebigen Zentralbild  $P'$  gibt es unendlich viele mögliche geometrische Objekte, deren Zentralbild  $P'$  ist.

Zusammenfassend lassen sich die folgenden Eigenschaften der Zentralprojektion festhalten:

1. Die Zentralbilder paralleler Geraden, die nicht parallel zur Bildebene sind, laufen in der Bildebene auf einen gemeinsamen Punkt zu, dem zu diesen Geraden zugehörigen Fluchtpunkt. Die Zentralbilder paralleler Geraden, die parallel zur Bildebene sind, sind auch in der Bildebene parallel.

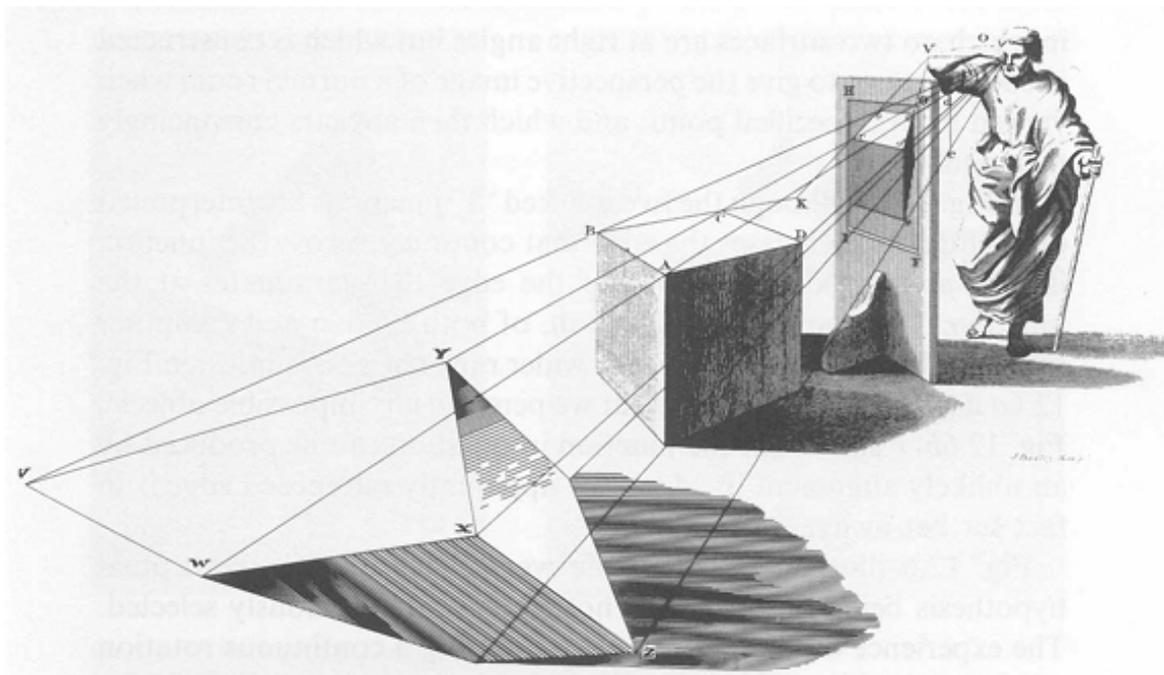


Abbildung 4.5. Anschauliche Darstellung des *inverse projection*-Problems. Abbildung aus Braddick und Atkinson (1982). Erläuterungen im Text.

2. Der Horizont ist gerade die Menge aller Fluchtpunkte jener Geraden, die parallel zur Grundebene verlaufen. Der Horizont liegt stets in der Höhe des Augpunktes, unterschiedlich hohe Augpunkte erzeugen somit Horizonte mit unterschiedlicher Lotdistanz zur Grundlinie.
3. Jeder Punkt auf der Bildebene ist ein Fluchtpunkt einer Geraden.
4. Zu einem Zentralbild  $P'$  gibt es unendlich viele mögliche dreidimensionale geometrische Objekte, deren Zentralbild gerade  $P'$  ist.

Anzumerken ist noch, dass bei der Fotografie mit einem Normalobjektiv durch die Konstruktion des optischen Systems der Kamera und den zugehörigen optischen Gesetzen die Fotografie automatisch eine linearperspektivische Darstellung der fotografierten Szene darstellt.

Die eingangs erwähnte Idee der Erzeugung äquivalenter Stimuli durch linearperspektivische Darstellung kann nun besser verständlich werden. Die Gesetze der Strahlenoptik besagen, dass bei Betrachtung eines beliebigen Objektes oder einer Szene das durch das Objekt oder die Szene reflektierte Licht in 'geraden Strahlen' auf die Retinae des Betrachters projiziert werden. Wird nun eine linearperspektivische Darstellung einer Szene mit einem Auge von der Position des Augpunktes aus betrachtet, so wird – zumindest im Prinzip – durch das Zentralbild das gleiche Lichtmuster auf die Retina des Auges projiziert, wie durch die Szene selbst, insofern diese einäugig von der Position des Augpunktes aus betrachtet würde. Bei Betrachtung einer linearperspektivischen

Darstellung *außerhalb* des Augpunktes kann man sich hingegen durch die obigen Ausführungen zur Zentralprojektion leicht klar machen, dass das durch das Bild projizierte Lichtmuster nicht das gleiche wie das durch die Originalszene projizierte ist, sondern es stellen sich in Abhängigkeit von der Position des Betrachters relativ zum Augpunkt des Bildes gleichsam ‘systematische Verzerrungen’ des projizierten Lichtmusters ein. Würde man nun in einer zugegeben relativ simplen Auffassung visueller Wahrnehmung davon ausgehen, dass das retinale Lichtmuster den damit einhergehenden Wahrnehmungseindruck auf eindeutige Weise bestimmt oder Wahrnehmung gar darin besteht, dass der Beobachter gleichsam das auf die Retina projizierte Lichtmuster *als Bild sieht*, so müsste man erwarten, dass sich diese geometrischen ‘Verzerrungen’ auch im mit dem Zentralbild einhergehenden *Wahrnehmungseindruck* niederschlagen sollten. Etwas allgemeiner gesprochen sollten nach dieser Auffassung *stets* systematische ‘Verzerrungen’ (im Vergleich zur dargestellten Szene) *wahrgenommen* werden, wenn ein zentralperspektivisch korrektes Bild nicht aus dem Augpunkt heraus betrachtet wird.<sup>5</sup>

Die alltägliche Erfahrung mit realistischen Bildern oder Fotografien macht jedoch deutlich, dass die wahrgenommenen ‘Verzerrungen’ – zumindest in gewissen Grenzen und mit bestimmten Ausnahmen – nicht so stark ausfallen, wie man aufgrund der geometrischen Überlegungen erwarten könnte.<sup>6</sup> Bewegt man sich beispielsweise vor einem zentralperspektivisch korrekt ausgeführten Bild, so verändert sich der zugehörige Wahrnehmungseindruck in der Regel in nicht deutlich auffallender Weise, obwohl sich das auf die Retina projizierte Lichtmuster bei jedem Schritt deutlich verändert; eine Erfahrung, von der man sich durch den Gang einer Ausstellung mit perspektivisch gemalten Bildern oder Fotografien leicht überzeugen kann. Dieses Phänomen des sich nicht auffällig verändernden Wahrnehmungseindrucks bei Betrachtung perspektivisch korrekter Bilder *außerhalb* des Augpunktes soll im Folgenden, in Anlehnung an Kubovy (1988), auch als *wahrgenommene Robustheit der Perspektive* oder kurz *Perspektivenrobustheit* bezeichnet werden.

---

<sup>5</sup>Bereits Leonardo da Vinci wies darauf hin, dass perspektivische Darstellungen als ‘verzerrt’ wahrgenommen würden, falls die Betrachtung der Darstellung nicht vom Augpunkt aus erfolge. So heißt es in der Übersetzung seiner auf 1487-1490 datierten auch als „*Codex Ashburnham*“ bekannten Ausführungen über die perspektivische Darstellung:

A picture of objects in perspective will look more lifelike when seen from the point from which the objects were drawn. If you want to represent an object near you which is to have the effect of nature, it is impossible that your perspective should not look wrong, with every false relation and disagreement of proportion that can be imagined in a wretched work, unless the spectator, when he looks at it, has his eye at the very distance and height and direction where the eye or the point of sight was placed in doing the perspective. (Richter, 1883, S. 271)

Man beachte, dass auch in dieser Passage die Idee der Erzeugung äquivalenter Stimuli über die zentralperspektivische Darstellung deutlich wird („which is to have the effect of nature“).

<sup>6</sup>Ausnahmen hiervon bilden sogenannte *trompe l’œil*-Bilder, auf die in Abschnitt 4.1.3 näher eingegangen wird.

### 4.1.2. Perspektivenrobustheit

‘Perspektivenrobustheit’ ist sicher kein unmissverständlicher Terminus, daher sollen noch kurz einige Erläuterungen folgen. Erstens ist festzuhalten, dass sich dieser Terminus natürlich nicht auf das Bild selbst bezieht, sondern auf den mit der Betrachtung eines linearperspektivischen Bildes verbundenen Wahrnehmungseindruck. Zweitens soll mit ‘robust’ hier gemeint sein, dass die bei Betrachtung eines linearperspektivischen Bildes außerhalb des Augpunktes auftretenden Verzerrungen des Wahrnehmungseindrucks üblicherweise nicht so stark ausfallen, wie man sie aufgrund der geometrischen Veränderungen des auf die Retina(e) projizierten Lichtmusters erwarten könnte. Insbesondere soll der Terminus ‘Perspektivenrobustheit’ im Folgenden nicht als zweiwertig (Robustheit liegt vor oder liegt nicht vor) verstanden werden, sondern ein Wahrnehmungsphänomen bezeichnen, welches gleichsam in ‘unterschiedlich starker Ausprägung’ vorliegen kann.

Die wahrgenommene Robustheit der Perspektive ist eines der am stärksten untersuchten Felder innerhalb der Bildwahrnehmungsforschung, die ersten systematischen Beobachtungen lassen sich bis mindestens in die zweite Hälfte des 19. Jahrhunderts zurückverfolgen (de La Gournerie, 1859). Es liegt mittlerweile eine vergleichsweise beträchtliche Anzahl an empirischen Studien vor (etwa Perkins, 1973; Hagen, 1976; Rosinski & Farber, 1980; Rosinski, Mulholland & Degelmann, 1980; Wallach & Marshall, 1986; Goldstein, 1987; Vishwanath, Girshick & Banks, 2005), denen sich folgendes Befundmuster entnehmen lässt:

1. Die Perspektivenrobustheit scheint im Vergleich umso schwächer ausgeprägt, je weiter sich der Beobachter vom Augpunkt eines Bildes entfernt. Insbesondere bei ‘relativ starken’ Abweichungen in horizontaler Richtung zur Bildebene deuten die Ergebnisse darauf hin, dass die im Bild dargestellten Objekte von den Probanden als zunehmend ‘verzerrt’ (im Vergleich zur Betrachtung aus dem Augpunkt heraus) wahrgenommen werden.
2. Die Ausprägung der wahrgenommenen Robustheit der Perspektive scheint davon abzuhängen, ob der Betrachter das Bild monokular oder binokular betrachtet. Bei binokularer Betrachtung deuten die Ergebnisse auf eine insgesamt stärkere Perspektivenrobustheit hin als bei monokularer Betrachtung.
3. Die Ausprägung der Perspektivenrobustheit scheint abhängig davon zu sein, in welchem Maße – etwas abstrakt gesprochen – Attribute des Bildträgers durch experimentelle Manipulation hervorgehoben wurden, die den *Objektcharakter* des Bildträgers betonen. Dies wurde meist dadurch realisiert, dass der Rahmen des Bildes für die Probanden sichtbar war( ein Attribut, welches den Charakter des

Bildes als Objekt im physikalischen Raum betont) oder die Probanden das Bild durch ein Reduktionsrohr betrachteten (und dadurch der Rahmen des Bildes nicht sichtbar war). Die Ergebnisse deuten darauf hin, dass im Fall eines sichtbaren Rahmens die Perspektivenrobustheit stärker ausgeprägt ist als bei Betrachtung durch ein Reduktionsrohr.

Insbesondere die letzten beiden Punkte sind im Zusammenhang mit der Ausrichtung dieser Arbeit und dem noch eingehender zu behandelnden Konzept der Dualität der Bildwahrnehmung von Relevanz, da sie mit dem Konzept der Tiefencues verbunden sind.<sup>7</sup>

Der Befund, dass die wahrgenommene Robustheit der Perspektive bei monokularer Betrachtung und unter Hervorhebung von Attributen, die den Objektcharakter des Bildes betonen (etwa ein für den Betrachter sichtbarer Rahmen), am stärksten ausgeprägt zu sein scheint, lässt sich in einen Zusammenhang mit dem auch als *monocular depth enhancement* (Niederée & Heyer, 2003) bekannten Phänomen bringen. Damit wird das Phänomen bezeichnet, dass der mit Betrachtung einer zentralperspektivischen Darstellung einer Szene einhergehende Tiefeneindruck merklich ‘verstärkt’ werden kann, wenn das Bild erstens monokular und zweitens durch ein Reduktionsrohr betrachtet wird (Koenderink, van Dorn & Kappers, 1995, siehe auch Schlosberg, 1941). Die Prägnanz des mit einem realistischen Bild verbundenen Tiefeneindrucks nimmt mit zunehmenden, dem Betrachter verfügbaren potentiellen ‘Informationen’ ab. Bei der Bildwahrnehmung lässt sich daher eine im Vergleich zur ‘alltäglichen Wahrnehmung’ *umgekehrte* Beziehung zwischen den dem Betrachter vorliegenden potentiellen ‘Informationen’ und der Prägnanz des Tiefeneindrucks feststellen, da bei der Betrachtung von natürlichen Szenen die Prägnanz des Tiefeneindrucks bei binokularer Betrachtung in der Regel am stärksten ausgeprägt ist (Koenderink et al., 1995). Man beachte, dass im Rahmen des Konzepts der Tiefencues bei binokularer Betrachtung dem visuellen System sowohl die binokularen als auch die monokularen Tiefencues als potentielle ‘Information über Tiefe’ zur Verfügung stehen, im Fall der monokularen Betrachtung nur die monokularen Tiefencues. In dem dem Konzept der Dualität der Bildwahrnehmung gewidmeten Abschnitt wird dieser Punkt wieder aufgegriffen werden. An dieser Stelle soll sich zunächst der Frage zugewendet werden, wie die Perspektivenrobustheit üblicherweise erklärt wird und ob diese Erklärungsansätze das Phänomen verständlich machen können.

---

<sup>7</sup>Eine nicht zu vernachlässigende Variable in diesem Zusammenhang stellt auch die *Aufmerksamkeit* des Betrachters hinsichtlich möglicher Veränderungen seines Wahrnehmungseindrucks dar, die allerdings schwer im experimentellen Kontext zu operationalisieren ist. So ist etwa vorstellbar, dass einem diesbezüglich geübten Beobachter mögliche, mit Betrachtung eines Bildes außerhalb des Augpunktes einhergehende Veränderungen des Wahrnehmungseindrucks stärker auffallen als einem diesbezüglich ungeübten Beobachter.

Zur Erklärung des Phänomens der Perspektivenrobustheit wird in der bestehenden Literatur üblicherweise auf einen der folgenden Ansätze zurückgegriffen (vgl. Rogers, 1995; Vishwanath et al., 2005):

- i. Bilder würden im alltäglichen Leben normalerweise aus Positionen betrachtet, die nicht ‘allzu weit’ vom Augpunkt des Bildes entfernt seien. Die damit verbundenen ‘Verzerrungen’ auf Ebene des retinalen Lichtmusters seien daher so gering, dass sie vom Betrachter gleichsam ‘vernachlässigt’ würden (etwa Gombrich, 1960; Busey, Brady & Cutting, 1990).<sup>8</sup>
- ii. Perspektivenrobustheit sei gegeben, da der Betrachter bestimmte ‘Vorannahmen’ oder Erwartungen über die im Bild dargestellten Objekte habe (etwa Symmetrie oder Orthogonalität), die zur Folge hätten, dass die geometrischen ‘Verzerrungen’ bei Betrachtung des Bildes außerhalb des Augpunktes ‘vernachlässigt’ oder auf nicht näher ausgeführte Weise ‘ausgeglichen’ würden (etwa Perkins, 1973).
- iii. Es komme zu wahrgenommener Robustheit der Perspektive, da das visuelle System über einen ‘bildspezifischen Wahrnehmungsmechanismus’ in der Lage sei, den Augpunkt des Bildes zu ermitteln (oder zumindest hinreichend genau zu schätzen) und anhand dieser ‘Information’ die mit Betrachtung außerhalb des Augpunktes einhergehenden ‘Verzerrungen’ zu ‘korrigieren’ oder zu ‘kompensieren’ (etwa Pirenne, 1970; Hagen, 1976; Rosinski & Farber, 1980; Kubovy, 1988).

Im Gegensatz zu den ersten beiden Ansätzen hat der letzte Erklärungsansatz, der, nebenbei bemerkt, ein prägnantes Beispiel einer homunkularen Redeweise darstellt und allgemein auch als *Kompensationshypothese* (Rogers, 1995; Vishwanath et al., 2005) bekannt ist, in der Literatur eine starke Verbreitung und Akzeptanz erfahren. Über die genaue Funktionsweise dieses ‘bildspezifischen Mechanismus’ scheint Uneinigkeit zu bestehen, in der Mehrzahl der Fälle wird er als ein ‘unbewusster Prozess’ angesehen (Pirenne, 1970; Kubovy, 1988), der auf wiederholten Erfahrungen mit Bildern basiere (Hagen, 1976). Bezüglich der Frage, welche ‘Informationen’ vom visuellen System genutzt werden, um gemäß dieser Hypothese den Augpunkt eines Bildes zu ‘ermitteln’, lassen sich wiederum (mindestens) zwei verschiedene theoretische Ansätze unterscheiden. Mit dem sogenannten *pictorial-compensation*-Ansatz (etwa Kubovy, 1988) ist die Annahme verbunden, dass das visuelle System geometrische Informationen aus dem Zentralbild selbst ‘nutze’, um den Augpunkt zu ‘ermitteln’. Rein analytisch ist

---

<sup>8</sup>Dieser Erklärungsansatz ist jedoch unplausibel, da sich genügend Beispiele von Bildern finden lassen, die sich nur deutlich außerhalb des Augpunktes betrachten lassen, mit deren Betrachtung jedoch keine auffälligen Verzerrungen des Wahrnehmungseindrucks verbunden sind. Ein exemplarisches Beispiel stellen Teile des Freskenzyklus „*Storie della Vera Croce*“ von Piero della Francesca dar. Einige dieser sich in der Kirche *San Francesco* in Arezzo befindenden Fresken hängen in einer Höhe von etwa zwölf Metern und weisen einen Augpunkt auf, der sich etwa zehn Meter über dem Betrachter befindet.

es etwa möglich, anhand von drei verschiedenen Fluchtpunkten jeweils orthogonal zur Grundlinie verlaufender paralleler Geraden den Augpunkt eines Bildes zu bestimmen (Sedgwick, 1980). Würde man nun zusätzlich annehmen, dass das visuelle System über ‘Informationen’ bezüglich der Position des Betrachters relativ zur Bildebene verfügt, wäre es in diesem konzeptuellen Rahmen gesprochen zumindest logisch nicht ausgeschlossen, dass durch ‘geometrische Analysen’ in einem erstem Schritt der Augpunkt ‘ermittelt’ wird und anschließend – ebenfalls durch ein geometrisches Verfahren – die ‘Verzerrungen’ gleichsam ‘herausgerechnet’ werden. Vertreter des sogenannten *surface-compensation*-Ansatzes (etwa Perkins, 1973; Rosinski & Farber, 1980; Wallach & Marshall, 1986) gehen hingegen davon aus, dass das visuelle System keine geometrischen Informationen aus dem Zentralbild selbst, sondern lediglich ‘Informationen’ über die Position des Betrachters relativ zum Mittelpunkt des Bildes nutze, um, in Verbindung mit der Annahme, dass der Augpunkt eines Bildes sich *stets* auf dem durch den Bildmittelpunkt verlaufenden Normalenvektor befinde, den Augpunkt zu ‘schätzen’ und anhand dieser Schätzung die ‘Verzerrungen’ zu ‘kompensieren’.

In einer von Vishwanath et al. (2005) durchgeführten Untersuchung, die die empirische Tauglichkeit der unterschiedlichen Vorhersagen der beiden Varianten der Kompensationshypothese zum Gegenstand hatte, zeigte sich, dass keine dieser Vorhersagen mit dem Verhalten der Probanden kompatibel war. Die Autoren schlagen auf Basis ihrer Befundlage daraufhin ein Modell vor, das sie als *local-slant*-Hypothese bezeichnen. Laut der *local-slant* Hypothese komme es im Rahmen der wahrgenommenen Robustheit der Perspektive nicht zu einer ‘Schätzung’ oder ‘Ermittlung’ des Augpunktes durch das visuelle System, sondern es werde bei Betrachtung eines *beliebigen* Punktes auf der Bildebene der eingeschlossene Winkel zwischen Sichtlinie des Betrachters und des diesem Punkt entspringenden Normalenvektors ‘geschätzt’ und anhand dieser Schätzung die ‘Verzerrung’ lokal ‘herausgerechnet’.

Eine enorme theoretische Herausforderung für sämtliche Varianten der Kompensationshypothese stellen sogenannte *anamorphotische Bilder* dar (Niederée & Heyer, 2003), etwa das berühmte, 1553 vollendete Gemälde „Die Gesandten“ von Hans Holbein dem Jüngeren (siehe Abbildung 4.6 auf der nächsten Seite). Dieses Bild stellt zwei französische Gesandte am Hofe Heinrichs des VIII. von England dar, die vor einem Regal mit verschiedenen astronomischen und mathematischen Messinstrumenten sowie anderen Gegenständen wie einem Gesangsbuch und einer Laute posieren. Im Vordergrund des Bildes lässt sich außerdem ein verzerrt wirkendes, ‘eiförmiges Objekt’ ausmachen, das bei frontaler Betrachtung zunächst nicht zu identifizieren ist. Betrachtet man das Gemälde allerdings von der Horizontalen in einem Winkel von ca. 30° vom rechten Bildrand aus, lässt sich das dargestellte Objekt recht eindeutig als Totenschädel aus-



Abbildung 4.6. „Die Gesandten“ (1533) von Hans Holbein dem Jüngeren. Öl auf Holz, 206 × 209 cm. Quelle: *Wikimedia Commons*.

machen.<sup>9</sup> Die geometrische Grundlage der Konstruktion dieses Bildes ist, dass alle Elemente des Bildes mit Ausnahme der Darstellung des Totenkopfes von einem Augpunkt  $O_1$  konstruiert wurden, der sich in etwa auf dem durch den Mittelpunkt des Bildes verlaufenden Normalenvektor befindet, der Totenkopf hingegen von einem deutlich anderen Augpunkt  $O_2$  aus (so dass der eingeschlossene Winkel zwischen durch den Bildmittelpunkt verlaufenden Normalenvektor und der Strecke zwischen Augpunkt und Bildmittelpunkt in etwa  $30^\circ$  beträgt). Jede Form einer Kompensationshypothese im obigen Sinne kann den leicht selbst beobachtbaren Umstand, dass man bei frontaler Betrachtung des Bildes ein verzerrtes Objekt und keinen Totenkopf wahrnimmt, in Nähe von  $O_2$  hingegen die Darstellung des Totenkopfes als Totenkopf sieht, nicht erklären.

<sup>9</sup>Zur vielfältigen Symbolik des Gemäldes, insbesondere des Totenschädels, siehe etwa North (2002).

Zieht man den *pictorial-compensation*-Ansatz zur Erklärung heran, besteht zunächst die Frage, *welcher* der beiden Augpunkte  $O_1$  und  $O_2$  denn vom visuellen System ‘ermittelt’ werde. Unabhängig von der Beantwortung dieser Frage sollte jedoch nie der Fall eintreten, dass man bei frontaler Betrachtung den Totenkopf nicht sieht; bei Betrachtung aus  $O_2$  heraus hingegen schon. In beiden Situationen sollte man laut des *pictorial-compensation*-Ansatzes das Gleiche sehen, entweder den Totenkopf oder das eiförmige Objekt. Im Rahmen des *surface-compensation*-Ansatzes sollte man den Totenkopf hingegen *nie* sehen können; aus allen möglichen Positionen der Betrachtung sollte ein ‘eiförmiges Objekt’ wahrgenommen werden. Aus analogen Gründen kann auch die von Vishwanath et al. (2005) vorgeschlagene *local-slant*-Hypothese die Wahrnehmung des Totenkopfes bei Betrachtung aus  $O_2$  heraus nicht erklären, so dass das Phänomen der Perspektivenrobustheit (selbst innerhalb des *cue*-Ansatzes) noch einiger theoretischer Klärung bedarf.

### 4.1.3. Die Dualität der Bildwahrnehmung

Wir wollen uns nun dem bereits angesprochenen Konzept des *dualen Charakters* der Wahrnehmung von Bildern eingehender zuwenden, da auf dieses bei der Gestaltung und der Diskussion der empirischen Untersuchung zurückgegriffen wird. Mit diesem Begriff sollte das Phänomen umschrieben werden, dass bei der Betrachtung von realistischen Bildern der zugehörige Wahrnehmungseindruck in der Regel *sowohl* die Wahrnehmung des Bildes als einer Fläche im physikalischem Raum, *als auch* die Wahrnehmung des zugehörigen Bildraumes beinhaltet, wobei diese beiden Aspekte des Wahrnehmungseindrucks auf subtile Weise ineinander verwoben sind.<sup>10</sup> In Anlehnung an Niederée und Heyer (2003) sollen diese beiden Aspekte des mit der Bildbetrachtung einhergehenden Wahrnehmungseindrucks im Folgenden auch kurz als *planarer beziehungsweise räumlicher Aspekt* desselben bezeichnet werden.

Der Begriff der Dualität bezieht sich dabei allein auf den mit der Bildbetrachtung einhergehenden *Wahrnehmungseindruck*, das Perzept, und nicht auf das Bild als physikalischer Gegenstand. Darüberhinaus soll der duale Charakter der Bildwahrnehmung nicht so verstanden werden, dass im Wahrnehmungseindruck entweder allein der planare Aspekt *oder* allein der räumliche Aspekt oder ein beständiger Wechsel zwischen diesen aufzufinden ist, sondern in der Regel beide Aspekte *gleichzeitig*, in *einem als*

<sup>10</sup>Das Konzept einer Dualität der Bildwahrnehmung lässt sich zurückverfolgen bis zu den subtilen phänomenologischen Analysen der Wahrnehmung von Bildern durch Edmund Husserl in seinen 1904/05 in Göttingen gehaltenen Vorlesungen „Phantasie und Bildbewusstsein“ (Husserl, 1980/2006). Dieses Konzept wurde u.a. von Polanyi (1970) und Pirenne (1970) wieder aufgegriffen, differenzierte wahrnehmungspsychologische Analysen lassen sich bei Niederée und Heyer (2003) finden.

*einheitlich erlebten Perzept* vorhanden sind.<sup>11</sup> Es ist zwar möglich, die *Aufmerksamkeit* bei der Betrachtung eines Bildes stärker auf einen der beiden Aspekte zu richten (üblicherweise ist der räumliche Aspekt bei der Bildbetrachtung der salientere), im Regelfall werden jedoch stets beide Aspekte, gleichsam in unterschiedlicher ‘Gewichtung’, erlebt. Wäre dem nicht so, müsste unser alltäglich beobachtbarer Umgang mit realistischen Bildern ein vollkommen anderer sein. Im Falle eines nicht vorhandenen planaren Aspekts des Wahrnehmungseindrucks käme es stets zu ‘optischen Täuschungen’ bei der Bildbetrachtung. Der Betrachter würde etwa versuchen, Objekte im Bild zu ergreifen oder in die Darstellung einer Landschaft hineinzulaufen und unser üblicher Umgang mit Bildern als Objekten im physikalischen Raum (man denke etwa an das Sortieren von Fotos) wäre nicht ohne Weiteres möglich. Wäre hingegen nur der planare Aspekt vorhanden, würden Bilder nichts weiter sein als bloße Dekorationsobjekte oder bearbeitete Oberflächen; ein Erleben der endlos erscheinenden Landschaft, wie sie sich bei Betrachtung des in Abbildung 4.1 auf Seite 184 dargestellten Ausschnitts des *Grand Teton-Nationalparks* einstellt, wäre dann nicht möglich.

Ein weiterer Umstand, der ohne das Konzept des dualen Charakters der Bildwahrnehmung unverständlich wäre, ist, dass es uns in der Regel keine Probleme bei der Betrachtung eines Bildes bereitet, wenn der Bildraum und der physikalische Raum, in dem sich das Bild als Objekt befindet, *inkompatibel* sind, was häufig der Fall ist. Nimmt man etwa Abbildung 4.7 auf der nächsten Seite, eine mit einem deutlichen Tiefeneindruck einhergehende Fotografie des Marktplatzes von Halle und befestigt diese an der Decke eines Raumes, so entsteht eine Situation, die bei der Betrachtung natürlicher Szenen im physikalischen Raum mit großer Wahrscheinlichkeit nicht auftreten würde: Der Betrachter blickt im physikalischen Raum nach oben, der wahrgenommene Bildraum legt allerdings nahe, dass man von oben auf den Marktplatz *herabblickt*. Dennoch hat man als Betrachter in dieser Situation insgesamt einen *einheitlichen* Wahrnehmungseindruck, der einem in keiner Weise ‘falsch’ oder zumindest deutlich ungewöhnlich vorkommt. Aber er weist einen dualen Charakter auf. Der mit der Fotografie verbundene Tiefeneindruck und der Eindruck des Bildes als Gegenstand im physikalischen Raum sind in ihm gleichzeitig und miteinander verbunden präsent. Ohne die Annahme eines dualen Charakters der Bildwahrnehmung würde man gänzlich andere Phänomene in dieser Situation erwarten, etwa dass die Fotografie allein als Fläche ohne Tiefeneindruck wahrgenommen wird oder eine überaus ungewöhnliche ‘Täuschung’ auftreten sollte, nämlich, dass der Betrachter gleichsam fürchten müsste, beim Blick *nach oben* von der Marktkirche *herunterzufallen*.

---

<sup>11</sup>Mit der Dualität der Bildwahrnehmung ist mithin nicht das auch als *Bistabilität* bekannte Phänomen gemeint, welches sich etwa bei der Betrachtung des Necker-Würfels einstellt.



Abbildung 4.7. Fotografie des Marktplatzes von Halle (Saale) von der Marktkirche aus.

Es lassen sich allerdings auch Beispiele von Objekten finden, die gemeinhin als Bilder gelten, mit deren Betrachtung aber *kein* Perzept mit einem dualen Charakter in dem beschriebenen Sinn einhergeht. Einen solchen Fall stellt etwa das berühmte Deckenfresko „Die Apotheose des heiligen Ignatius“ von Andrea Pozzo (1642-1709) in der Kirche *Sant’Ignazio di Loyola* in Rom dar. Bei diesem handelt es sich um eine mit aller Kunst ausgeführte linearperspektivische Darstellung eines religiösen Motivs, wobei die Architektur der Kirche durch die Malerei gleichsam fortgesetzt wird (siehe dazu auch Pirenne, 1970). Im Boden der Kirche ist eine Marmorscheibe eingelassen, deren Position mit der des Augpunktes des Freskos identisch ist. Betrachtet man das Deckenfresko von diesem Punkt aus, so erscheint die gemalte Architektur (etwa Säulen) als real im strikten Sinne einer vollkommenen ‘Täuschung’ des Betrachters. Die gemalte Architektur wird nicht als von der tatsächlichen Architektur der Kirche verschieden erlebt. Im Kontext des Dualitätskonzeptes ist in diesem Fall allein der räumliche Aspekt präsent, d.h. das Bild wird nicht zur selben Zeit als ein von der Darstellung verschiedenes Objekt im physikalischen Raum erlebt. Fälle wie dieser werden auch als *trompe l’œil*-Bilder – oder besser: *trompe l’œil*-Effekt eines Bildes – bezeichnet. Bei Bildern mit einhergehendem *trompe l’œil*-Effekt (in diesem strikten Sinne einer ‘Täuschung’ des Betrachters) kommt es bei Betrachtung außerhalb des Augpunktes interessanterweise *nicht* zu einer wahrgenommenen Robustheit der Perspektive, wie sie im letzten Abschnitt angesprochen wurde. Bewegt man sich etwa im Fall des Pozzoschen Deckenfreskos aus dem

durch die Marmorscheibe gekennzeichneten Augpunkt hinreichend weit heraus, so werden nahezu unmittelbar und mit drastischer Deutlichkeit ‘Verzerrungen’ der gemalten Architektur wahrgenommen; die tatsächliche Architektur kann vom Betrachter ohne Weiteres von der gemalten unterschieden werden (für eine detailliertere Beschreibung des Wahrnehmungseindrucks bei Betrachtung des Deckenfreskos außerhalb des Augpunktes siehe Pirenne, 1970). Auf der anderen Seite des Spektrums hat insbesondere die bildende Kunst des 20. Jahrhunderts Bilder hervorgebracht, die explizit die Erzeugung eines wahrgenommenen Bildraumes beim Betrachter vermeiden sollen und so allein der planare Aspekt des zugehörigen Perzeptes vom Betrachter erlebt wird; ein berühmtes Beispiel ist etwa das 1915 für einen handfesten Skandal in der Kunstszene sorgende Bild „Schwarzes Quadrat auf weißem Grund“ von Kasimir Malewitsch. Niederée und Heyer (2003) schlagen daher vor, diese Extremfälle als Endpunkte eines Kontinuums der mit Bildern einhergehenden Wahrnehmungseindrücke aufzufassen. Der duale Charakter der Wahrnehmung von Bildern lasse sich dann zwischen diesen beiden Polen (*bearbeitete Oberfläche* und *optische Täuschung*) auffinden, wobei der planare und der räumliche Aspekt im Perzept in unterschiedlichem Maße präsent sein können, aber *stets* beide Aspekte vorhanden sind.

Lassen sich Möglichkeiten finden, das Konzept der Dualität der Bildwahrnehmung in das übliche allgemeine theoretische Schema der visuellen Wahrnehmung zu integrieren? Zunächst ist festzuhalten, dass das Dualitätskonzept prinzipiell mit der Idee der Erzeugung *äquivalenter Stimuli* durch linearperspektivische Darstellung einer Szene in einem Bild und somit mit der Metapher des Alberti-Fensters inkompatibel ist. Beim Betrachten einer linearperspektivischen Darstellung wird diese eben nicht für die dargestellte Szene selbst gehalten. Wie im letzten Kapitel dargestellt, ist der übliche Ansatz zur Erklärung der Wahrnehmung von Tiefe oder Räumlichkeit eng mit dem Konzept der Tiefencues verbunden. Die Wahrnehmung von Tiefe, auch die mit der Bildbetrachtung verbundene wahrgenommene Tiefe, basiere wesentlich auf den verschiedenen, in einer Sehsituation vorliegenden Tiefencues, die im Rahmen der *cue*-Integration vom visuellen System auf bestimmte Weise kombiniert würden, um eine möglichst reliable und spezifische ‘Repräsentation’ einer Szene beim Betrachter zu erzeugen. Im speziellen Fall der Betrachtung von Bildern liegen diesem Ansatz zufolge *konfligierende* Tiefencues oder ein ‘*cue*-Konflikt’ vor. In der Sprechweise dieses Ansatzes ‘deuten’ verschiedene bildspezifische monokulare Tiefencues, wie etwa dargestellte Verdeckung, dargestellter Schatten, in einem Fluchtpunkt konvergierende Linien etc., auf *Tiefe* hin, wohingegen andere Tiefencues, etwa binokulare Disparität oder fehlende Bewegungsparallaxe, auf die Betrachtung einer Fläche ‘hindeuten’. Nun ließe sich im Rahmen dieses Schemas zwar das in Abschnitt 4.1.2 angesprochene Phänomen des *monocular depth enhancement* verständlich machen, indem angenommen wird, dass bei monoku-

larer Betrachtung eines realistischen Bildes durch ein Reduktionsrohr wesentliche, auf die Betrachtung einer Fläche 'hindeutende' Tiefencues ausgeschaltet werden und so der Tiefeneindruck verstärkt wird; das Konzept eines *dualen Charakters* des Wahrnehmungseindrucks ist jedoch durch die Annahme *einer spezifischen* Repräsentation einer Szene als Ergebnis der *cue*-Integration von vornherein ausgeschlossen.

Wenn man das Konzept der Dualität der Bildwahrnehmung mit dem *cue*-Ansatz verbinden wollte (worauf das Dualitätskonzept nicht angewiesen ist), so könnte man laut Niederée und Heyer (2003) das übliche theoretische Schema erweitern. Ein möglicher Weg der Erweiterung wäre laut den Autoren etwa die Annahme, dass bei Vorliegen konfligierender *cues* dieser Konflikt vom visuellen System nicht in einer Weise 'aufgelöst' wird, die *eine* spezifische Repräsentation einer Szene seitens des Betrachters zur Folge hat, sondern über Prozesse der *cue*-Integration *zwei oder mehr* hinreichend kohärente Teilmengen an Tiefencues gebildet werden (die Autoren nennen dies auch 'cue-Segregation'), von denen jede einzeln *eine* spezifische 'Repräsentation' zur Folge haben würde, die aber gleichzeitig nebeneinander bestehen und im Sinne eines 'cross-talks' (Sedgwick, 2003) miteinander interagieren können. Eine solche Erweiterung müsse aber erstens zusätzliche Annahmen darüber enthalten, auf welche Weise diese gleichzeitig nebeneinander bestehenden Teilmengen an *cues* dennoch zu einem *einheitlichen* Wahrnehmungseindruck führen, in dem sowohl ein planerer wie auch ein räumlicher Aspekt präsent sind. Zweitens müsse die Möglichkeit zugelassen werden, dass zwischen diesen Teilmengen bestimmte 'Hierarchieverhältnisse' bestehen, um unserem Erleben Rechnung tragen zu können, dass wir bei Betrachtung eines Bildes das Objekt 'Bild' als dem physikalischen Raum zugehörig ansehen, den zugehörigen Bildraum jedoch nicht (man erinnere sich an das obige Beispiel der Fotografie des Marktplatzes von Halle). Zusammenfassend kann man dem Resümee der Autoren allerdings nur zustimmen, dass eine solche Erweiterung ohne weitere theoretische Klärung zunächst insgesamt mehr Fragen aufwirft, als sie zu beantworten in der Lage ist.

Es ließe sich auch die grundsätzliche Frage stellen, ob es sich bei dem dualen Charakter der Wahrnehmung von Bildern zwar durchaus um ein phänomenologisch fundiertes Phänomen handelt, das jedoch nur einen äußerst ungewöhnlichen Sonderfall der menschlichen visuellen Wahrnehmung darstellt und daher eher als Aberration vernachlässigt, anstatt *zwingend* in eine allgemeine Wahrnehmungstheorie zu integrieren gesucht werden sollte. Von der eingangs erwähnten Tatsache abgesehen, dass Bilder bzw. bewegte Bilder (Film und Fernsehen) einen wesentlichen Teil unseres alltäglichen Lebens ausmachen und daher eine Theorie visueller Wahrnehmung auch die damit verbundenen Phänomene verständlich machen sollte, lassen sich ähnliche Dualitätsphänomene auch außerhalb des Kontextes der Bildbetrachtung ausmachen (vgl. Polanyi, 1970). Beispielsweise lässt sich beim Verfolgen eines Theaterstücks ein vergleichba-

res Phänomen beobachten. Der etwa die Rolle des Hamlet spielende Schauspieler wird *gleichzeitig* wahrgenommen als ein Schauspieler auf einer Theaterbühne, der eben diese Rolle spielt und in gewisser Weise eben auch als die dargestellte Person Hamlet selbst, der in Helsingör den Tod seines Vaters zu rächen trachtet, wobei auch hier diese beiden Aspekte in einem *einheitlichen* Wahrnehmungseindruck auf komplexe Weise miteinander verbunden sind. Wäre nur einer dieser beiden Aspekte präsent, wäre unser Erleben eines aufgeführten Theaterstücks ein ganz anderes. Wäre allein der ‘Schauspieleraspekt’ präsent, so sollten Phänomene wie ein Mitleiden mit den dargestellten Personen (etwa beim berichteten Suizid Ophelias oder beim Tode Hamlets) oder überhaupt das Erleben von Spannung im Verlaufe eines Theaterstücks nicht in der Weise auftreten, wie sie es üblicherweise tun.<sup>12</sup> Käme es andererseits dazu, dass wir die Schauspieler nicht als Schauspieler auf einer Bühne wahrnehmen würden, sondern allein als die dargestellten Personen, so würde man erwarten, dass die Zuschauer versuchen würden, etwa Hamlet beim Duell gegen Laertes zu unterstützen oder Hamlet vor dem Trinken des vergifteten Bechers zu warnen. Ähnliches gilt in diesem Kontext auch für die Requisiten einer Aufführung: Ein auf der Bühne stehender Stuhl wird wahrgenommen als eben ein Stuhl, der auf der Bühne des Theaters steht, indem man sich selbst als Zuschauer befindet, und *gleichzeitig* als der Thron des Königs von Schloss Helsingör. Analoge Beobachtungen lassen sich vornehmen, wenn man etwa eine Skulptur oder sein eigenes Bild in einem Spiegel betrachtet (vgl. Niederée & Heyer, 2003).

Ein weiteres dualitätsartiges Phänomen, das außerhalb eines ‘künstlerischen Kontextes’ anzusiedeln ist, lässt sich beim Betrachten von Eisenbahnschienen beobachten. In diesem einheitlichen Wahrnehmungseindruck sind ebenfalls (mindestens) zwei unterschiedliche Aspekte auszumachen, welche miteinander verwoben scheinen. Einerseits werden die Eisenbahnschienen als gleichsam in einem unendlich fernen Punkt konvergierend wahrgenommen, andererseits werden sie, und zwar *gleichzeitig*, als parallel verlaufend erlebt. Auch hier lässt sich durch Verschiebung der Aufmerksamkeit der eine oder der andere Aspekt stärker in den Vordergrund rücken, dennoch scheinen stets beide Aspekte vorhanden zu sein, und zwar als nicht nebeneinander bestehend, sondern miteinander verwoben. Man könnte in diesem Fall zwar argumentieren, dass es sich hier möglicherweise um einen vom dualen Charakter der Bildwahrnehmung abweichenden Fall der Dualität handelt, da hier das Wissen (oder die Erfahrung) eine Rolle zu spielen scheint, dass die Bahnschwellen stets dieselbe Länge haben und die Schienen daher parallel verlaufen *müssen*. Dieser Punkt scheint indes nicht zentral für das Phänomen

---

<sup>12</sup>Ein weiteres Indiz für einen dualen Charakter der Wahrnehmung lässt sich in diesem Zusammenhang beobachten, wenn man als Zuschauer etwa aufgrund des Todes von Ophelia Trauer empfindet. Diese Trauer ist Trauer, aber von anderem Charakter als die Trauer um einen ‘tatsächlichen Menschen’. Die ineinander verwobenen Wahrnehmungseindrücke eines durch einen Schauspieler dargestellten Todes und des Todes der dargestellten Person selbst könnten eine solche Trauer, die man auch als ‘*als-ob-Trauer*’ bezeichnen könnte, hervorrufen.

selbst zu sein und es lassen sich mit einiger Überlegung zahlreiche andere Beispiele für als dual zu kennzeichnende Wahrnehmungseindrücke finden. Zudem scheint die dem Dualitätskonzept zugrunde liegende Idee nicht darauf beschränkt zu sein, dass in einem als einheitlich erlebten Wahrnehmungseindruck genau *zwei* unterschiedliche Aspekte desselben gleichzeitig präsent sind, sondern die diesem Konzept zugrundeliegende Idee lässt sich natürlich auf *mehr* als zwei gleichzeitig präsente Aspekte eines als einheitlich erlebten Wahrnehmungseindrucks erweitern.<sup>13</sup>

Zusammenfassend lassen sich hinreichend viele Hinweise auf einen dualen Charakter der Wahrnehmung finden, um zu argumentieren, dass es sich hierbei eben *nicht* um einen von der alltäglichen Wahrnehmung abweichenden Sonderfall handelt, sondern Dualität möglicherweise ein allgemeines und auch zentrales Charakteristikum des menschlichen Sehens darstellt und daher auch durch eine Theorie visueller Wahrnehmung verständlich gemacht werden sollte. Dies ist bisher allerdings erst in unzureichendem Maße geschehen, wobei dieser Umstand möglicherweise darauf zurückzuführen ist, dass innerhalb der Wahrnehmungsforschung der Bildwahrnehmung und insbesondere explizit phänomenologischer Analysen des mit der Betrachtung von Bildern einhergehenden Wahrnehmungseindrucks bisher eine deutlich untergeordnete Rolle zukam.

## 4.2. Einleitende Bemerkungen zur empirischen Untersuchung

Um in die eingehendere Darstellung des empirischen Teils dieser Arbeit übergehen zu können, bedarf es noch einiger einleitender Bemerkungen. Die empirische Untersuchung ist thematisch mit den im vorangegangenen Exkurs zur Bildwahrnehmung dargestellten Konzepten der Perspektivenrobustheit und der Dualität der Bildwahrnehmung (und damit auch mit dem Konzept der Tiefencues) verbunden. Insbesondere im Rahmen der Perspektivenrobustheit spielen geometrische, mit der Zentralprojektion verbundene Überlegungen eine wichtige Rolle. Da für das Verständnis des grundsätzlichen Vorgehens der Untersuchung solche Überlegungen notwendig sind, soll im Folgenden der für die vorliegende Untersuchung relevante Spezialfall der Zentralprojektion einer Kugel eingehender behandelt werden. Im Anschluss soll in Kürze eine Arbeit von Vishwanath et al. (2005) vorgestellt werden, an die die vorliegende Untersuchung in wesentlichen Punkten angelehnt ist.

---

<sup>13</sup>So spricht etwa Mausfeld (2003, 2010, 2011) in diesem Zusammenhang von *conjoint representations* oder *multiperspectivity*. Ob es sich hierbei in der Tat lediglich um eine Erweiterung des Dualitätskonzeptes von Niederée und Heyer (2003) auf zwei oder mehr gleichzeitig in einem einheitlichen Wahrnehmungseindruck präsente unterschiedliche Aspekte handelt, sei dahingestellt.

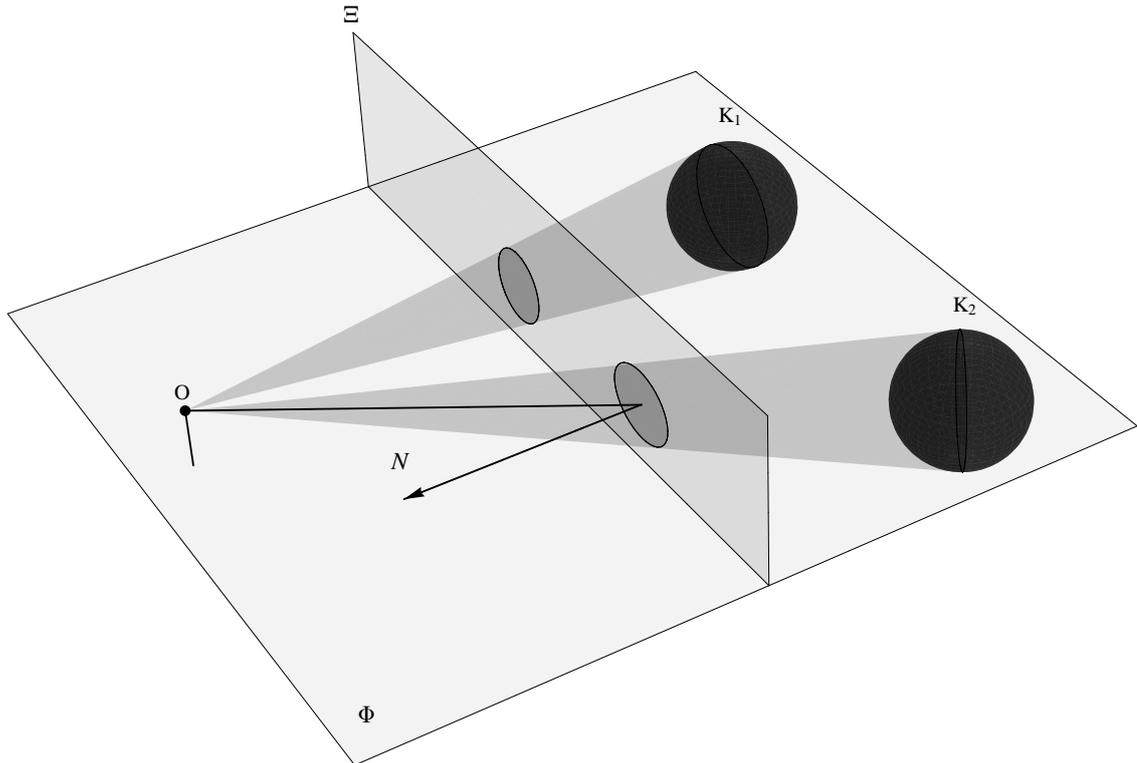
## Zentralprojektion einer Kugel

Wie in dem den Grundlagen der Zentralprojektion gewidmeten Abschnitt 4.1.1 werde wieder in einem Raum  $\mathcal{V}$  eine Bildebene  $\Xi$ , eine Grundebene  $\Phi$  sowie ein Augpunkt  $O$  mit Lotabstand  $H$  von  $\Phi$  betrachtet. Darüber hinaus enthalte  $\mathcal{V}$  zwei Kugeln  $K_1, K_2$  mit identischem Durchmesser  $A = 2H$ , deren Zentren sich erstens beide in Höhe des Augpunktes befinden und zweitens den gleichen Abstand  $d$  von der Bildebene aufweisen. Das Zentrum von  $K_1$  werde mit  $Z_{K_1}$ , das von  $K_2$  mit  $Z_{K_2}$  bezeichnet, die Zentralbilder dieser Zentren mit  $Z'_{K_1}$  und  $Z'_{K_2}$ . Der eingeschlossene Winkel zwischen der Strecke  $[O, Z'_{K_1}]$  und dem in  $Z'_{K_1}$  entspringenden Normalenvektor zur Bildebene betrage  $0^\circ$ , der eingeschlossene Winkel zwischen  $[O, Z'_{K_2}]$  und dem in  $Z'_{K_2}$  entspringenden Normalenvektor zur Bildebene sei größer als  $0^\circ$ . Abbildung 4.8 auf der nächsten Seite zeigt diese Situation auf schematische Weise.

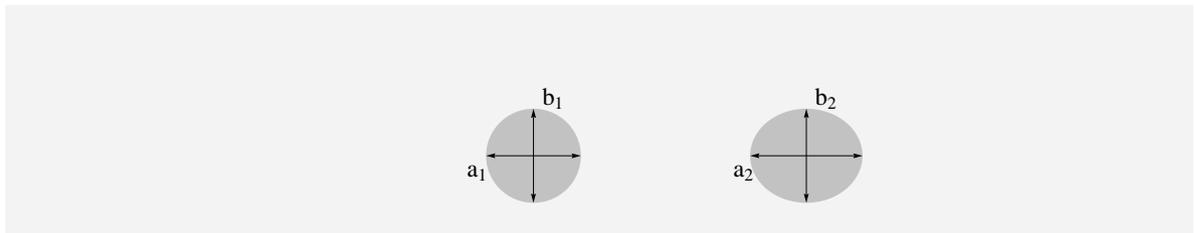
Die Zentralbilder  $K'_1, K'_2$  der beiden Kugeln  $K_1, K_2$ , d.h. die Schnitte der entsprechenden Sehkegel mit der Bildebene, lassen sich allgemein als Ellipsen in  $\Xi$  kennzeichnen. Im Folgenden soll die Länge der horizontalen Hauptachse einer beliebigen Ellipse mit  $a$  und die Länge der vertikalen Hauptachse mit  $b$  bezeichnet werden. Zudem soll aus Darstellungsgründen die horizontale Hauptachse einer beliebigen Ellipse auch als *Hauptachse* und die vertikale Hauptachse als *Nebenachse* bezeichnet werden.

Eine für die vorliegende Untersuchung zentrale Frage ist nun, wie sich allgemein die Längen von Haupt- und Nebenachse der beiden Zentralbilder  $K'_1, K'_2$  in der obigen Situation bestimmen lassen. Anhand von Abbildung 4.8 kann man sich anschaulich leicht klarmachen, dass für diese spezielle Situation erstens Haupt- und Nebenachse von  $K'_1$  gleich lang sein müssen ( $a_1 = b_1$ ), es sich somit bei  $K'_1$  um einen Kreis handelt, und zweitens die beiden Nebenachsen der Zentralbilder  $K'_1$  und  $K'_2$  eine identische Länge aufweisen sollten ( $b_1 = b_2$ ). Abbildung 4.8 (b) lässt sich darüberhinaus entnehmen, dass offenbar aufgrund der verschiedenen großen eingeschlossenen Winkel die Länge der Hauptachsen  $a_1$  und  $a_2$  von  $K'_1$  bzw.  $K'_2$  unterschiedlich ist. Diese Beziehung soll nun anhand von Abbildung 4.9 auf Seite 207 genauer untersucht werden, die die in Abbildung 4.8 (a) dargestellte Szene in der Aufsicht zeigt.

Für diese spezielle Situation lässt sich zeigen (vgl. Gårding, 1992; Vishwanath et al., 2005; Wokes & Palmer, 2008), dass für die Länge  $a$  der Hauptachse und die Länge  $b$  der Nebenachse des Zentralbildes  $K'$  einer Kugel  $K$  mit Durchmesser  $A$  folgende



(a)



(b)

Abbildung 4.8. Schematische Darstellung der Zentralprojektion zweier Kugeln  $K_1$  und  $K_2$  auf die Bildebene  $\Xi$ . Abbildung (a) zeigt die im Text beschriebene Situation, (b) die Zentralbilder  $K'_1$  und  $K'_2$  der beiden Kugeln in der Bildebene mit den entsprechenden Hauptachsenlängen  $a_i$  und Nebenachsenlängen  $b_i$  für das Zentralbild  $K'_i$  der Kugel  $K_i$ . Weitere Erläuterungen im Text.

Beziehungen gelten:<sup>14</sup>

$$a \approx A \cdot \frac{q}{q+d} \cdot \cos(\gamma_O)^{-1} \quad (4.1)$$

$$b = A \cdot \frac{q}{q+d} \quad , \quad (4.2)$$

<sup>14</sup>Es sei nochmals darauf hingewiesen, dass die Beziehungen (4.1) und (4.2) nur für den geschilderten Spezialfall gelten. Insbesondere gilt (4.2) nur dann, wenn sich das Zentrum der Kugel in Höhe des Augpunktes befindet.

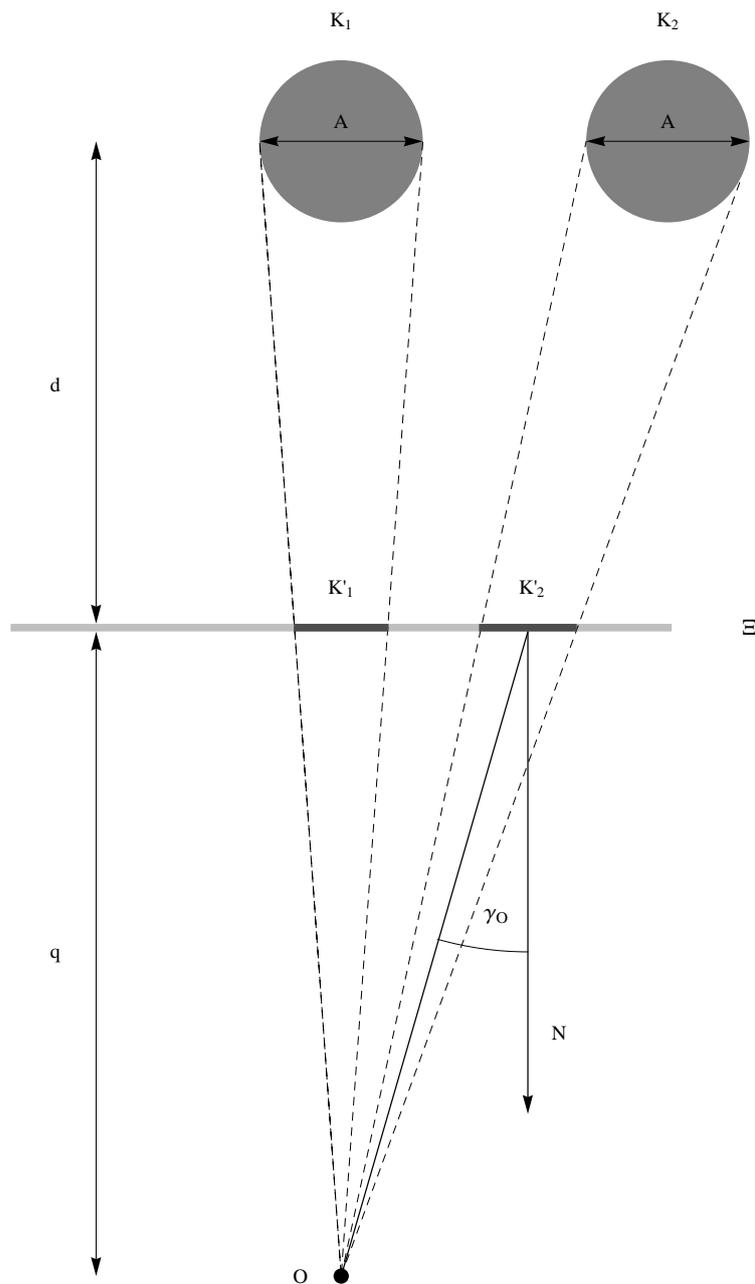


Abbildung 4.9. Aufsicht der in Abbildung 4.8 (a) dargestellten Szene. Erläuterungen im Text.

wobei  $q$  den Abstand des Augpunktes  $O$  von der Bildebene  $\Xi$ ,  $d$  den Abstand des Zentrums  $Z_K$  der Kugel  $K$  von  $\Xi$  und  $\gamma_O$  den eingeschlossenen Winkel zwischen der durch  $O$  und dem Zentrum des Zentralbildes  $Z'_K$  verlaufenden Gerade und dem in  $Z'_K$  entspringenden Normalenvektor zu  $\Xi$  bezeichne. In Abbildung 4.9 etwa sind für die Kugeln  $K_1, K_2$  mit identischem Durchmesser  $A$  die Größen  $q$  und  $d$  identisch, der Winkel  $\gamma_O$  bezüglich  $K_1$  ist Null, der Winkel  $\gamma_O$  bezüglich  $K_2$  ist hingegen größer Null.

Wir wollen im Folgenden die in Gleichung (4.1) ausgedrückte Approximation als *Gleichheit* behandeln, da die Abweichungen von der Gleichheit für unsere Zwecke als

praktisch vernachlässigbar angesehen werden können. Folgende Eigenschaften ergeben sich aus (4.1) und (4.2):

1. Es sei zunächst  $\gamma_O = 0$ . In diesem Fall sind die Gleichungen (4.1) und (4.2) identisch, so dass  $a = b$  gilt. Betrachtet man nun lediglich den Faktor  $q/(q+d)$ , so gilt:

- i. Für alle  $d > 0$  gilt:

$$\frac{q}{q+d} \rightarrow 1, \quad \text{falls } q \rightarrow \infty \quad \text{und}$$

$$\frac{q}{q+d} \rightarrow 0, \quad \text{falls } q \rightarrow 0$$

- ii. Für alle  $q > 0$  gilt:

$$\frac{q}{q+d} \rightarrow 0, \quad \text{falls } d \rightarrow \infty \quad \text{und}$$

$$\frac{q}{q+d} \rightarrow 1, \quad \text{falls } d \rightarrow 0$$

Aus i. und ii. ergibt sich, dass, bei festem Abstand  $d$  zwischen Zentrum der Kugel  $Z_K$  und der Bildebene, die Längen  $a$  und  $b$  des Zentralbildes  $K'$  sich mit wachsender Distanz  $q$  des Augpunktes von der Bildebene dem Durchmesser  $A$  der Kugel  $K$ , und mit kleiner werdendem  $q$  der Null annähern. Bei festem Abstand  $q$  zwischen Augpunkt und Bildebene verschwinden die Längen  $a$  und  $b$  mit wachsender Distanz  $d$  zwischen  $Z_K$  und der Bildebene und nähern sich  $A$ , falls  $d$  gegen Null geht.

2. Für alle Winkel  $0^\circ < \gamma_O < 90^\circ$  gilt:

$$\frac{1}{\cos(\gamma_O)} \rightarrow \infty, \quad \text{falls } \gamma_O \rightarrow 90^\circ \quad \text{und}$$

$$\frac{1}{\cos(\gamma_O)} \rightarrow 1, \quad \text{falls } \gamma_O \rightarrow 0^\circ$$

Daher gilt für alle  $q, d > 0$ , dass die Länge  $a$  der Hauptachse des Zentralbildes von  $K$  über alle Grenzen wächst, falls  $\gamma_O$  gegen  $90^\circ$  läuft. Umgekehrt geht  $a$  gegen  $A \cdot \frac{q}{q+d}$ , falls  $\gamma_O$  gegen Null geht. Insbesondere gilt also für das Verhältnis von Haupt- und Nebenachse des Zentralbildes von  $K$  bei beliebigem  $d$  und  $q$  größer

Null:

$$\frac{a}{b} = 1, \quad \text{falls } \gamma_O = 0, \quad (4.3)$$

$$\frac{a}{b} > 1, \quad \text{falls } \gamma_O > 0. \quad (4.4)$$

Die in Abbildung 4.9 auf Seite 207 dargestellte Situation soll nun dahingehend erweitert werden, dass zusätzlich ein monokularer Betrachter  $B$  eingeführt wird, dessen Auge denselben Lotabstand  $H$  von der Grundebene  $\Phi$  aufweist wie der Augpunkt  $O$ . Es soll im Folgenden jedoch nur noch eine Kugel  $K$  und deren Zentralbild  $K'$  betrachtet werden. Der Abstand zwischen der Bildebene  $\Xi$  und dem Nodalpunkt des Auges von  $B$  werde mit  $p$  und der eingeschlossene Winkel zwischen der durch den Nodalpunkt des Auges von  $B$  und dem Zentrum  $Z'_K$  des Zentralbildes von  $K$  verlaufenden Gerade und dem in  $Z'_K$  entspringenden Normalenvektor mit  $\gamma_B$  bezeichnet.<sup>15</sup> Der Winkel  $\gamma_B$  soll im Folgenden auch der *Blickwinkel* genannt werden. Betrachtet dieser monokulare Betrachter nun das Zentralbild  $K'$  der Kugel  $K$  mit den Achsenlängen  $a$  und  $b$ , so kommt es zu einer Projektion dieses Zentralbildes auf die Netzhaut von  $B$ . Der visuelle Winkel der Projektion der Hauptachse  $a$  des Zentralbildes  $K'$  auf die Retina soll im Folgenden mit  $\alpha$  und der der Projektion der Nebenachse  $b$  auf die Retina mit  $\beta$  bezeichnet werden. Abbildung 4.10 auf der nächsten Seite zeigt schematisch diese Situation in der Aufsicht.

Es lässt sich zeigen, dass für den geschilderten Fall folgende Beziehungen gelten (vgl. Vishwanath et al., 2005):

$$\alpha \approx a \cdot \frac{f}{p} \cdot \cos(\gamma_B) \quad (4.5)$$

$$\beta \approx b \cdot \frac{f}{p}, \quad (4.6)$$

wobei  $f$  die Brennweite des sogenannten *Normalauges* bezeichne, das ein schematisches Modell des durchschnittlichen menschlichen Auges in der physiologischen Optik darstellt. Das menschliche Auge wird in diesem Modell als eine einfache Linse mit der

<sup>15</sup>Der Begriff des Nodal- oder Knotenpunktes stammt aus der physikalischen Optik und bezeichnet zwei ausgezeichnete Punkte auf der optischen Achse eines optischen Systems. Da dieser Begriff für das Verständnis dieses Abschnittes nicht wesentlich ist, soll an dieser Stelle auf eine Definition verzichtet werden. Die beiden Nodalpunkte des menschlichen Auges liegen sehr nahe beieinander (etwa 0.2mm) nahezu unmittelbar hinter der Augenlinse. Aufgrund dieser geringen Distanz voneinander werden die beiden Nodalpunkte des Auges aus Gründen der Einfachheit häufig auch mit einem einzigen Nodalpunkt identifiziert, der etwa 7mm hinter dem Hornhautscheitel und 16 bis 17mm vor der Netzhaut liegt. Für ausführlichere Darstellungen siehe etwa Graham (1965) oder Hecht (2009).

Brennweite  $f$  aufgefasst, wir wollen im Folgenden auch hier aus praktischen Gründen die Approximation als Gleichheit behandeln. Den Gleichungen (4.5) und (4.6) lässt sich entnehmen, dass sich erstens bei festem  $a, b$  und konstantem Blickwinkel  $\gamma_B$  die Größen  $\alpha$  und  $\beta$  umgekehrt proportional zum Abstand  $p$  des Nodalpunktes zur Bildebene verhalten. Zweitens ist die Größe  $\alpha$  wesentlich vom Blickwinkel  $\gamma_B$  abhängig. Erhöht sich bei konstantem  $a$  und  $p$  der Winkel  $\gamma_B$  ( $0 \leq \gamma_B < 90$ ), so verringert sich  $\alpha$ , wobei  $\alpha \rightarrow 0$ , falls  $\gamma_B \rightarrow 90$ .

Eine besondere und für die Theorie der Zentralprojektion wichtige Situation stellt nun der Fall dar, in dem der Nodalpunkt des (monokularen) Beobachters  $B$  sich an derselben Position wie der Augpunkt  $O$  befindet, also  $p = q$  und  $\gamma_B = \gamma_O$  gilt. In diesem Fall ist die Projektion des Zentralbildes  $K'$  der Kugel  $K$  auf die Retina von  $B$  (idealiter) gleich der Projektion der Kugel  $K$  auf die Retina von  $B$ . Anschaulich gesprochen wird in dieser Situation durch das Zentralbild  $K'$  (idealiter) *dasselbe* retinale Lichtmuster erzeugt wie durch  $K$  selbst. Insbesondere ergeben sich aus dem bisher Gesagten für die hier betrachtete Situation die folgenden, für die empirische Untersuchung relevanten Beziehungen: Unabhängig von den Größen  $q, d$  und  $f, p$  gilt:

$$\frac{\alpha}{\beta} = \frac{a}{b} \cdot \cos(\gamma_B) = \frac{A \cdot \cos(\gamma_O)^{-1} \cdot \cos(\gamma_B)}{A}. \quad (4.7)$$

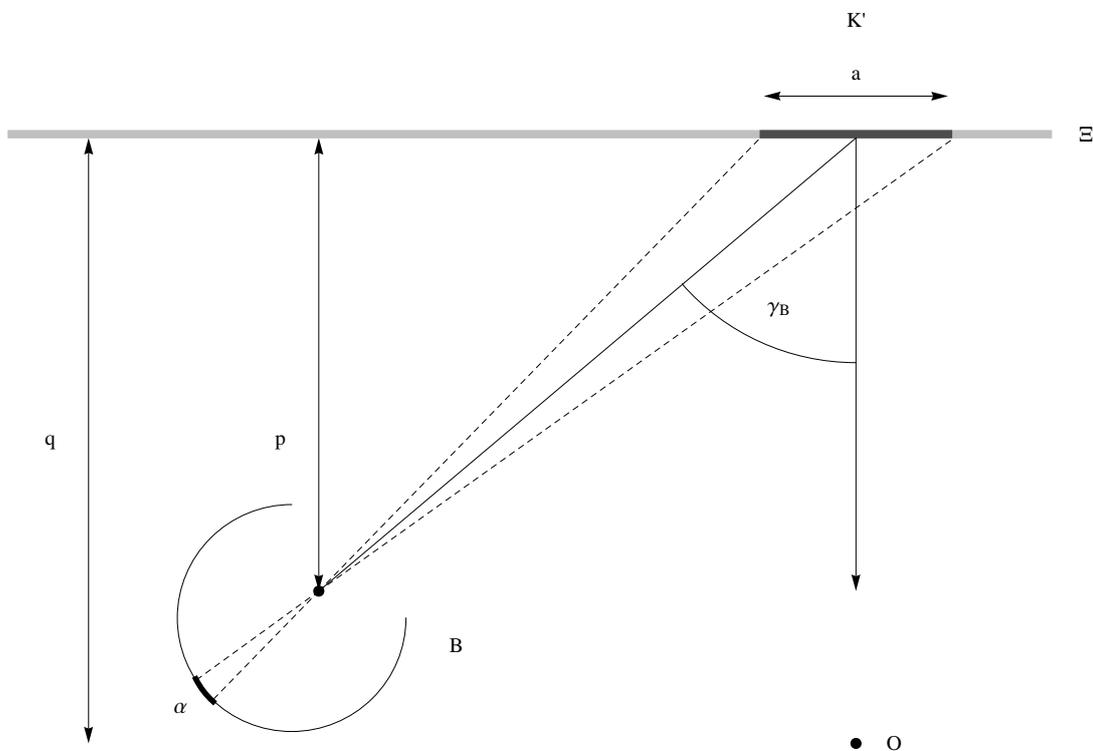


Abbildung 4.10. Erweiterung der in Abbildung 4.8 (a) dargestellten Situation um einen (monokularen) Betrachter  $B$ . Weitere Erläuterungen im Text.

Dadurch ergibt sich:

$$\frac{\alpha}{\beta} = \frac{A}{A} = 1, \text{ falls } \gamma_B = \gamma_O \quad \text{und} \quad \frac{\alpha}{\beta} \neq 1, \text{ falls } \gamma_B \neq \gamma_O, \quad (4.8)$$

sowie

$$\frac{\alpha}{\beta} = \frac{a}{b}, \text{ falls } \gamma_B = 0 \quad \text{und} \quad \frac{\alpha}{\beta} < \frac{a}{b}, \text{ falls } \gamma_B > 0. \quad (4.9)$$

Anschaulich gesprochen besagt Gleichung (4.8), dass das Zentralbild  $K'$  (mit beliebigem Verhältnis  $a/b$ ) der Kugel  $K$  nur in dem Fall als ‘Kreis’ ( $\alpha/\beta = 1$ ) auf die Retina eines monokularen Beobachters projiziert wird, falls  $\gamma_B = \gamma_O$  gilt. Bei monokularer Betrachtung des Zentralbildes  $K'$  außerhalb des Augpunktes ( $\gamma_O \neq \gamma_B$ ) ergeben sich somit systematische ‘Verzerrungen’ des retinalen Lichtmusters im Vergleich zur Projektion von  $K$  auf die Retina des Betrachters. Die vorliegende Untersuchung wird unter anderem das bereits erläuterte Phänomen der *Perspektivenrobustheit* zum Gegenstand haben, d.h. in diesem Zusammenhang die Frage, inwieweit sich diese *geometrischen* Veränderungen des auf die Retina projizierten Lichtmusters auch in einer Veränderung des Wahrnehmungseindrucks niederschlagen.

## Die Untersuchung von Vishwanath et al. (2005)

Vishwanath et al. (2005) versuchten aufgrund der gerade dargestellten Beziehungen, auf die wir hier zurückgreifen wollen, die Robustheit der Perspektive unter verschiedenen Sehbedingungen quantitativ folgendermaßen zu erfassen: Sie präsentierten ihren Probanden eine computergenerierte zentralperspektivische Darstellung einer Szene, wobei der Augpunkt dieser Darstellung sich auf dem durch den Mittelpunkt des Bildes verlaufenden Normalenvektor befand. Zentrales Element dieses Bildes war eine Ellipse (deren Zentrum mit dem Mittelpunkt des Bildes identisch war) mit fester Nebenachsenlänge  $b$ , deren Hauptachsenlänge  $a$  jedoch von den Probanden verändert werden konnte. Die Probanden betrachteten die zentralperspektivische Darstellung stets aus der gleichen Position, nämlich so, dass (grob gesprochen) ihre Sichtlinie durch den Mittelpunkt des Bildes, und damit auch durch das Zentrum der Ellipse verlief. Aufgabe der Probanden war es, die Länge  $a$  der Hauptachse der Ellipse so zu verändern, dass sie als Darstellung einer *Kugel* erscheint. Variiert wurde dabei erstens der Blickwinkel  $\gamma_B$  (vgl. Abbildung 4.10 auf der vorherigen Seite), unter dem die Probanden das Bild betrachteten. Dies wurde durch *Drehung* des die Darstellung präsentierenden Monitors im oder gegen den Uhrzeigersinn um eine gedachte vertikale Achse realisiert, so dass der eingeschlossene Winkel zwischen Blickrichtung auf die Mitte des Bildes und dem durch die Mitte des Bildes verlaufenden Normalenvektor gerade identisch dem

zu realisierenden Winkel  $\gamma_B$  war. Anschaulich gesprochen wurden hier unterschiedliche Blickwinkel nicht durch Bewegung des Bildbetrachters, sondern durch Drehung der Bildebene realisiert. Zweitens wurde die Sehbedingung, unter der den Probanden das Bild präsentiert wurde, variiert. Die Probanden betrachteten das Bild binokular, monokular oder monokular durch ein Reduktionsrohr (im Folgenden als *Aperturbedingung* bezeichnet). Die Probanden mussten die ihnen gestellte Aufgabe unter jeder möglichen Winkelbedingung und jeder möglichen Sehbedingung bewältigen.

Aufgrund der im vorangehenden Abschnitt dargestellten geometrischen Beziehungen lässt sich für jede Winkelbedingung vorhersagen, wie groß das Verhältnis  $a/b$  von Haupt- und Nebenachse – für das wir im Folgenden auch kurz  $\tau$  schreiben wollen – von den Probanden eingestellt werden müsste, wenn es zu keinerlei Perspektivenrobustheit kommen sollte, d.h. wenn die Probanden gleichsam die Projektion des Zentralbildes auf ihrer Retina ‘sehen’ würden. Dieser Auffassung gemäß, die wir im Folgenden auch als *geometrische Vorhersage* bezeichnen wollen, sollte eine Ellipse unter sämtlichen Winkelbedingungen genau dann als eine Kugel wahrgenommen werden, wenn  $\tau$  so eingestellt wird, dass das Verhältnis der Projektionen  $\alpha/\beta$  gerade Eins ist.

Wie in Gleichung (4.7) bereits dargestellt wurde, gilt für beliebige Winkel  $\gamma_B$  ( $0 \leq \gamma_B < 90$ ), dass das Verhältnis der Projektionen  $\alpha/\beta$  in folgender Beziehung zum Längenverhältnis von Haupt- und Nebenachse  $\tau$  der dargestellten Ellipse steht:

$$\frac{\alpha}{\beta} = \tau \cdot \cos(\gamma_B). \quad (4.10)$$

Bei nicht vorhandener Perspektivenrobustheit würde man daher laut der skizzierten geometrischen Vorhersage (wenn also das Urteil der Probanden *allein* von den Größen  $\alpha$  und  $\beta$  abhängig wäre) für Blickwinkel  $\gamma_B$  ( $0 \leq \gamma_B < 90$ ) erwarten, dass von den Probanden eine Ellipse mit einem Hauptachsen-Nebenachsenlängenverhältnis von

$$\tau = \frac{1}{\cos(\gamma_B)} \quad (4.11)$$

als Kugel wahrgenommen werden sollte, da allein in diesem Fall  $\alpha/\beta = 1$  gilt. Bei Vorliegen gleichsam ‘vollständiger Perspektivenrobustheit’ würde man hingegen erwarten, dass unter *sämtlichen* Winkelbedingungen  $\gamma_B$  ( $0 \leq \gamma_B < 90$ ) ein Verhältnis von  $\tau = 1$  von den Probanden als Kugel wahrgenommen wird. Abbildung 4.11 auf der nächsten Seite zeigt die mit diesen beiden theoretischen Annahmen verbundenen zu erwartenden Einstellungen der Probanden in Abhängigkeit des Winkels  $\gamma_B$ .

Abbildung 4.12 auf Seite 214 zeigt eine *idealisierte* Darstellung der empirischen Befunde von Vishwanath et al. (2005) im Vergleich zu den bei vollständiger Robustheit

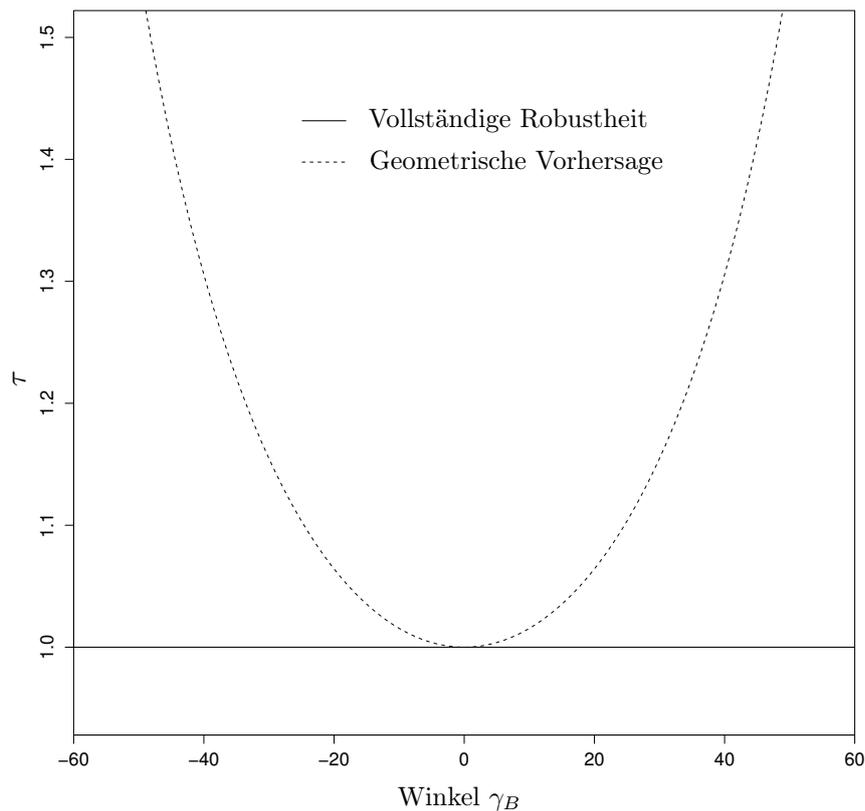


Abbildung 4.11. Darstellung der zu erwartenden  $\tau$ -Einstellungen in Abhängigkeit des Winkels  $\gamma_B$  bei geometrischer Vorhersage und Vorliegen vollständiger Perspektivenrobustheit. Die Unterscheidung zwischen positiven und negativen Winkeln  $\gamma_B$  soll kennzeichnen, ob der Monitor im oder gegen den Uhrzeigersinn um den entsprechenden Winkel gedreht wird.

und geometrischer Vorhersage zu erwartenden Einstellungen. Aus Darstellungsgründen wird im Folgenden auf eine gesonderte Kennzeichnung der bei vollständiger Robustheit zu erwartenden Einstellungen verzichtet. Die bei vollständiger Robustheit zu erwartenden Einstellungen sind im Folgenden als nicht separat gekennzeichnete, parallel zur  $x$ -Achse verlaufende und die  $y$ -Achse im Punkt  $\tau = 1$  schneidende Gerade eingezeichnet.

Wie sich Abbildung 4.12 entnehmen lässt, war das anhand der  $\tau$ -Einstellungen von den Probanden gezeigte Ausmaß der Perspektivenrobustheit sowohl von der Winkelbedingung als auch von der Sehbedingung abhängig. Unter allen drei Sehbedingungen nahmen die Probanden mit zunehmendem absoluten Winkel  $\gamma_B$  Einstellungen vor, die in der Tendenz zunehmend größer als das bei vollständiger Perspektivenrobustheit zu erwartende Verhältnis von  $\tau = 1$  waren. Zudem zeigte sich, dass eine wesentliche Abhängigkeit zwischen dem Ausmaß der Perspektivenrobustheit und der Sehbedingung besteht: In jeder Winkelbedingung ungleich Null war die Perspektivenrobustheit in jener Sehbedingung, in der die Probanden den Reiz monokular durch ein Reduktionsrohr betrachteten (die Aperturbedingung), am geringsten und in der binokularen Sehbedingung am stärksten ausgeprägt, wobei sich das mit der monokularen Sehbedingung ver-

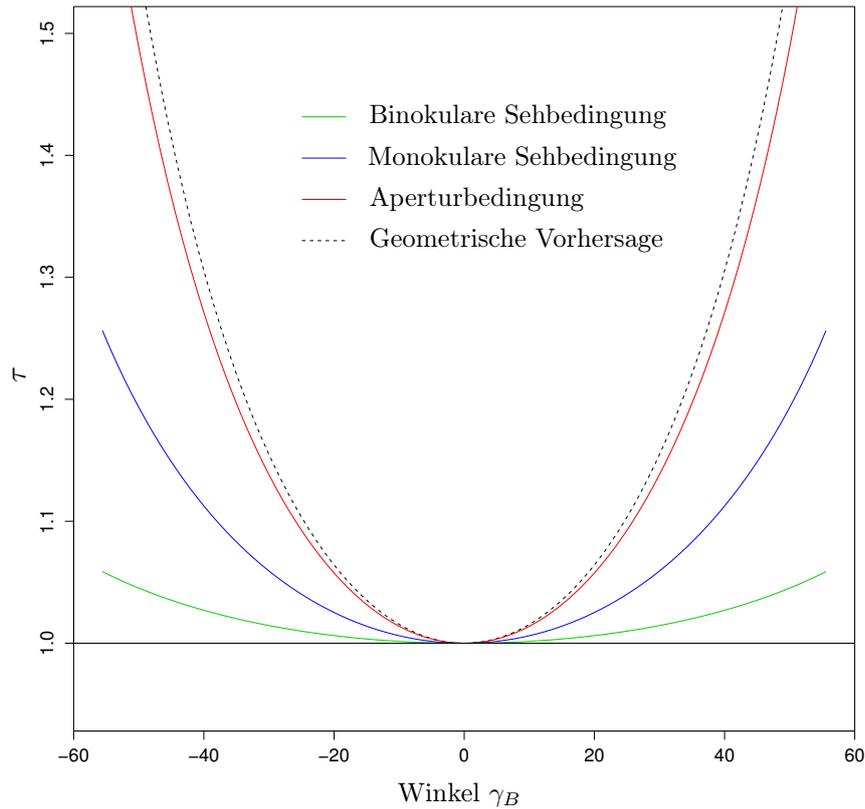


Abbildung 4.12. Idealisierte Darstellung der Ergebnisse von Vishwanath et al. (2005). Abgetragen ist für verschiedene Winkel  $\gamma_B$  und die verschiedenen Sehbedingungen das Verhältnis  $\tau$  jener Ellipse, die von den Probanden als Kugel wahrgenommen bzw. als eine solche beurteilt wird.

bundene Ausmaß an Perspektivenrobustheit in etwa zwischen diesen beiden einordnen lässt. Vishwanath et al. (2005) interpretieren ihre Ergebnisse im Fall der binokularen Sehbedingung als Vorliegen vollständiger Perspektivenrobustheit, obwohl insbesondere bei absoluten Winkeln  $\gamma_B > 30$  zum Teil recht deutliche Abweichungen von den bei vollständiger Robustheit zu erwartenden Einstellungen vorlagen.<sup>16</sup> Keine wesentliche Rolle spielte hingegen der Umstand, ob die Probanden den kritischen Reiz auf einem Monitor präsentiert bekamen, der im oder gegen den Uhrzeigersinn um  $\gamma_B$  gedreht wurde.

Aufgrund der gut bestätigten Befundlage zum Phänomen des *monocular depth enhancement* lässt sich argumentieren, dass das unterschiedliche Ausmaß der Perspektivenrobustheit unter den verschiedenen Sehbedingungen (nicht aber den verschiedenen Winkelbedingungen) mit unterschiedlich *prägnanten Tiefeneindrücken* seitens der Probanden einhergeht. In der binokularen Sehbedingung sollte dieser Argumentation zufolge der mit dem präsentierten Bild verbundene Tiefeneindruck verglichen mit den anderen Sehbedingungen am wenigsten prägnant, in der Aperturbedingung hingegen

<sup>16</sup>„When observers had rich slant information (binocular viewing with frame visible), the aspect ratio settings were invariant over viewing angle, particularly when viewing angle was less than  $|45^\circ|$ ; that is, observers based their judgements solely on a retinal image shape that was first adjusted for incorrect viewing position“ (Vishwanath et al., 2005, S. 1404).

am prägnantesten sein.<sup>17</sup> Die empirischen Befunde scheinen nun darauf hinzudeuten, dass bei einem sehr prägnanten Tiefeneindruck das Ausmaß der Perspektivenrobustheit am schwächsten ausgeprägt ist, mit abnehmender Prägnanz des Tiefeneindrucks allerdings zunimmt. Folgt man dieser Argumentation, könnte man das Ausmaß der Perspektivenrobustheit auch als ein Indiz für die Prägnanz des mit der Bildbetrachtung einhergehenden Tiefeneindrucks auffassen.

### 4.3. Ziele und inhaltliche Hypothesen

Die vorliegende Untersuchung verfolgt zwei Ziele. Zum Ersten soll der Versuch unternommen werden, den im letzten Abschnitt dargestellten Teil der von Vishwanath et al. (2005) berichteten empirischen Befunde zur Perspektivenrobustheit zu replizieren. Der Hintergrund dieser Überlegung ist, dass die Autoren in dem für die vorliegende Untersuchung relevanten Telexperiment lediglich drei Probanden untersucht hatten, die zum Teil differente Ergebnisse produzierten. Darüberhinaus lässt sich in Vishwanath et al. (2005) keine Auskunft darüber finden, wie viele Durchgänge pro Bedingung von den Probanden absolviert werden mussten, d.h. wie viele Datenpunkte pro Proband und Bedingung letztendlich vorlagen, was eine differenzierte Einordnung der von den Autoren berichteten Ergebnisse erschwert. Der Versuch der Replikation des von den Autoren berichteten Befundmusters soll daher dazu dienen, aufgrund einer größeren Datengrundlage möglichst differenzierte Aussagen zur Perspektivenrobustheit in diesem Kontext zu treffen. Aus diesem Grund soll sich bei der vorliegenden Untersuchung an die im letzten Abschnitt dargestellte generelle Vorgehensweise der Untersuchung von Vishwanath et al. (2005) gehalten werden.

Angelehnt an die Ausführungen zum dualen Charakter der Bildwahrnehmung und zur Perspektivenrobustheit soll zum Zweiten überprüft werden, inwieweit sich eine experimentelle Manipulation auf das Ausmaß der Perspektivenrobustheit auswirkt, welche dazu dienen soll, den wahrgenommenen Charakter der ‘Oberflächenhaftigkeit’ der den Probanden präsentierten Bilder zu erhöhen. Dies soll durch eine graphische Bearbeitung erreicht werden, welche die computergenerierten zentralperspektivischen Darstellungen so aussehen lässt, als ob sie gleichsam mit Pinselstrichen auf eine Leinwand gemalt worden wären. Im Kontext des Dualitätskonzeptes sollte eine solche Manipulation den mit den Bildern einhergehenden räumlichen Aspekt des Wahrnehmungseindrucks verringern (und damit den planeren Aspekt erhöhen), da im Rahmen des

---

<sup>17</sup>Man erinnere sich, dass das Phänomen des *monocular depth enhancement* üblicherweise über das Konzept der Tiefencues zu erklären gesucht wird. Bei monokularer Betrachtung eines zentralperspektivischen Bildes durch ein Reduktionsrohr würden jene Tiefencues ‘ausgeschaltet’ (etwa Konvergenz), die darauf ‘hindeuten’ würden, dass es sich bei dem Betrachteten, dem Bildträger, um eine *Fläche* handelt, was zu einer verstärkten Prägnanz des mit der Betrachtung des Bildes einhergehenden Tiefeneindrucks führe.

Konzepts der Tiefencues und möglicher *cue*-Integrations- und *cue*-Segregationsprozesse eine Hervorhebung des Oberflächencharakters eines Bildes gleichsam als ‘*cue gegen Tiefe*’ (oder ‘Flächencue’) fungieren sollte, der darauf ‘hindeutet’, dass es sich um ein Bild, d.h. um eine bearbeitete Fläche im physikalischen Raum handelt. Aufgrund der im Rahmen des Konzepts der Tiefencues behaupteten engen Verflechtung der bei der Bildbetrachtung vorliegenden Menge an ‘Informationen’, die darauf ‘hindeuten’, dass es sich bei dem Betrachteten um ein Bild *als bearbeitete Fläche* handelt, und dem Ausmaß an zugehöriger Perspektivenrobustheit, würde man erwarten, dass diese Manipulation zu einem im Vergleich generell niedrigeren Ausmaß der Perspektivenrobustheit führt. Bereits Pirenne (1970) und Polanyi (1970) hatten im Rahmen eines dem vorgestellten Dualitätskonzept ähnlichen Ansatzes die Vermutung geäußert, dass der wahrgenommene ‘Leinwandcharakter’ eines Bildes – im Sinne des Ausmaßes, in dem dem Betrachter das Bild als bearbeitete *Oberfläche* präsent ist – einerseits mit der Prägnanz des Tiefeneindrucks und andererseits mit dem Ausmaß der Perspektivenrobustheit bei Betrachtung außerhalb des Augpunktes zusammenhänge:

A painting’s liability to be distorted by angular vision is thus linked to its deceptiveness, and a subsidiary awareness of the canvas is supposed both to protect a painting against such distortion and to prevent it from being deceptive. (Polanyi, 1970, S. 226)

Nach Wissen des Autors ist dieser Vermutung bisher jedoch nicht weiter nachgegangen worden. Die in der vorliegenden Untersuchung vorgenommene Manipulation des ‘Leinwandcharakters’ zentralperspektivischer Bilder ist im Bereich der Forschung zur Bildwahrnehmung daher als neuartig zu charakterisieren, da in bisherigen Studien der mit der Betrachtung eines Bildes verbundene Tiefeneindruck allein über die Sehbedingung (binokular, monokular, monokular mit Reduktionsrohr) zu variieren gesucht wurde.

Etwas technischer ausgedrückt ergeben sich für die vorliegende Untersuchung die folgenden unabhängigen Variablen:

1. *Sehbedingung*:

Die Probanden sehen die zentralperspektivische Darstellung einer Szene entweder mit beiden Augen (*binokulare Sehbedingung*), mit einem Auge (*monokulare Sehbedingung*) oder mit einem Auge durch ein Reduktionsrohr (*Aperturbedingung*). Diese Variable lässt sich im Rahmen des *cue*-Ansatzes gleichsam als Standard der experimentellen Manipulation der dem visuellen System verfügbaren ‘Informationen’ kennzeichnen. Unter der Aperturbedingung, in der u.a. der Rahmen des Bildes für den monokularen Betrachter nicht sichtbar ist, seien jene ‘Informationen’, die auf ein Bild als ebenes physikalisches Objekt im Raum ‘hindeuten’

würden, am geringsten ausgeprägt, in der binokularen Sehbedingung hingegen seien diese am stärksten ausgeprägt. Der binokulare Tiefencue der Konvergenz würde unter der binokularen Bedingung etwa darauf ‘hindeuten’, dass es sich bei dem betrachteten Bild um eine bearbeitete Fläche handelt. Im Rahmen des Dualitätskonzepts gesprochen, soll diese Variable dazu dienen, die ‘Gewichtung’ zwischen dem planaren und dem räumlichen Aspekt des zugehörigen Wahrnehmungseindrucks zu variieren.

2. *Winkelbedingung:*

Den Probanden wird ein Bild aus stets gleicher Distanz, aber aus verschiedenen Blickwinkeln  $\gamma_B$ , d.h. unterschiedlich stark in horizontaler Richtung vom Augpunkt des Bildes abweichenden Positionen dargeboten. Dies wird dadurch realisiert, dass der Monitor, auf dem die Bilder dargeboten werden, um den entsprechenden Winkel um eine gedachte vertikale Achse im oder gegen den Uhrzeigersinn gedreht wird. Insgesamt werden sieben verschiedene Blickwinkel  $\gamma_B$  realisiert, wobei  $\gamma_B \in \{-45, -30, -15, 0, 15, 30, 45\}$ .

3. *Reizart:*

Die präsentierten Bilder weisen entweder einen durch graphische Manipulation hervorgehobenen ‘Oberflächencharakter’ auf (*Leinwandbedingung*), oder nicht (*Standardbedingung*). Die Variation der Reizart soll, wie die Variable *Sehbedingung*, dazu dienen, das Verhältnis zwischen planarem und räumlichem Aspekt des zugehörigen Wahrnehmungseindrucks zu manipulieren. In der Sprechweise des *cue*-Ansatzes sollte der hervorgehobene ‘Oberflächencharakter’ vom visuellen System als zusätzliche ‘Information’ darüber genutzt werden, dass es sich bei dem Betrachteten um eine bearbeitete Oberfläche handelt.

Insgesamt ergeben sich bei vollständiger Kreuzung demzufolge 42 ( $3 \times 7 \times 2$ ) verschiedene Bedingungen, welche von *sämtlichen* Probanden durchlaufen werden sollen. Wie später anhand des genauen Vorgehens deutlich werden wird, handelt es sich um eine sehr aufwendige Untersuchung; eine Versuchsperson benötigte etwa 25 Stunden, um sämtliche Bedingungen zu durchlaufen. Die Entscheidung für ein *within-subjects*-Design ist darin begründet, dass bei Untersuchungen zur visuellen Wahrnehmung häufig nicht geringe individuelle Unterschiede aufzufinden sind, welche die hypostasierten experimentellen Effekte bei Mittelung der Daten über sämtliche Versuchspersonen nicht deutlich hervortreten lassen könnten, obgleich jeder einzelne Proband hypothesenkonformes Verhalten gezeigt hat.

Die zu erhebende abhängige Variable stellt auch hier das von den Probanden (über ein noch näher zu erläuterndes Verfahren) eingestellte Verhältnis  $\tau$  zwischen Länge der Haupt- und Nebenachse jener Ellipse dar, die ihnen unter der entsprechenden Bedin-

gung als eine Kugel erscheint. Die Abweichung dieser Größe von dem bei vollständiger Perspektivenrobustheit zu erwartenden Verhältnis  $\tau = 1$  soll dann als Indikator für das Ausmaß der Perspektivenrobustheit unter der entsprechenden Bedingung interpretiert werden, wobei die Größe der Abweichung als in umgekehrt proportionaler Beziehung zum Ausmaß der Perspektivenrobustheit stehend aufgefasst werden soll. Darüberhinaus soll das Ausmaß der Perspektivenrobustheit aufgrund der dargelegten Überlegungen auch als Prägnanz oder ‘Stärke’ des mit der Bildbetrachtung unter entsprechender Seh- und Reizbedingung verbundenen Tiefeneindrucks interpretiert werden, wobei auch hier eine umgekehrt proportionale Beziehung postuliert wird.

Aufgrund der vorangestellten theoretischen Überlegungen wird folgendes Befundmuster erwartet:

1. Unter *sämtlichen* Seh- und Reizbedingungen sollte das Ausmaß der Perspektivenrobustheit bei (absolut) zunehmender Winkelgröße  $\gamma_B$  *abnehmen*. Dies sollte sich bei zunehmender Größe von  $|\gamma_B|$  in zunehmend stärker von Eins in positiver Richtung abweichenden  $\tau$ -Einstellungen durch die Probanden niederschlagen.
2. Über alle Winkelbedingungen hinweg betrachtet sollte das Ausmaß der Perspektivenrobustheit unter der binokularen Sehbedingung insgesamt stärker ausgeprägt sein als unter der monokularen Sehbedingung und der Aperturbedingung. Unter der Aperturbedingung sollte die Perspektivenrobustheit am schwächsten ausgeprägt sein. Auch diese Beziehungen sollte sich in entsprechend unterschiedlich stark von Eins in positiver Richtung abweichenden  $\tau$ -Einstellungen äußern.
3. Über sämtliche Winkelbedingungen hinweg betrachtet sollte unter *jeder* Sehbedingung die Perspektivenrobustheit in der Leinwandbedingung insgesamt niedriger ausgeprägt sein als in der Standardbedingung, was sich ebenfalls in entsprechenden  $\tau$ -Einstellungen seitens der Probanden niederschlagen sollte.

Zusammenfassend wird die in Abbildung 4.13 auf der nächsten Seite idealisiert dargestellte Struktur der Daten erwartet.

## 4.4. Methode

### Allgemeine Charakterisierung des Experimentes

Das rechnergestützte Experiment ist, nicht zuletzt wegen des Ziels einer Replikation der berichteten Befunde, hinsichtlich des Vorgehens an die von Vishwanath et al. (2005) durchgeführte Untersuchung angelehnt. Den Versuchspersonen wurde unter sämtlichen experimentellen Bedingungen eine zentralperspektivisch korrekte Darstellung einer Szene dargeboten, die als zentrales und farblich deutlich hervorgehobenes Element eine

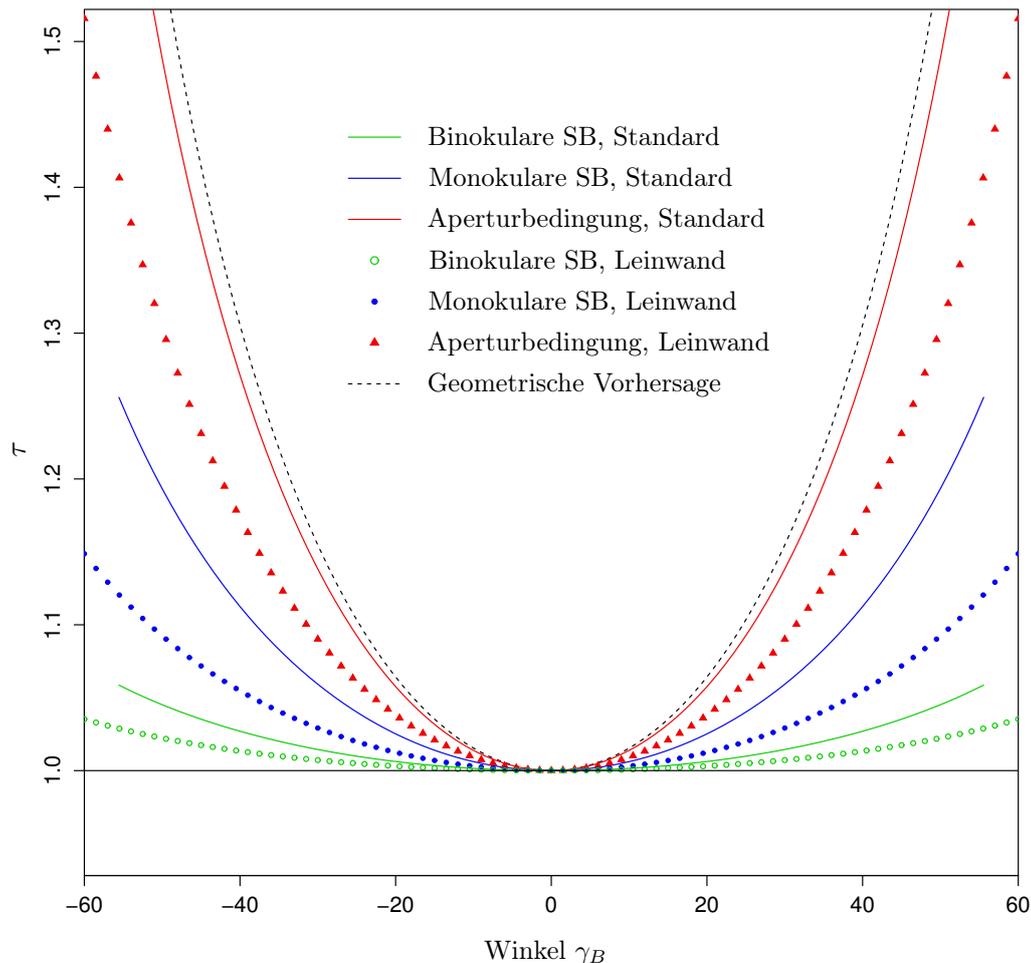


Abbildung 4.13. Idealisierte Darstellung der erwarteten Datenstruktur. Abgetragen ist für verschiedene Winkel  $\gamma_B$ , die verschiedenen Sehbedingungen (binokular, monokular und Aperturbedingung) sowie die beiden Reizbedingungen (Standard und Leinwand) das Verhältnis  $\tau$  jener Ellipse, die von den Probanden als Kugel wahrgenommen bzw. als eine solche beurteilt wird.

Ellipse enthält. Aufgabe der Probanden war es, über ein adaptives psychophysikalisches Verfahren (vgl. Treutwein, 1995; Leek, 2001) die Länge der Hauptachse dieser Ellipse so zu verändern, dass sie ihnen unter der jeweiligen Bedingung als eine Kugel erscheint. Erhoben wurde jeweils, bei welchem Hauptachsen-Nebenachsenverhältnis  $\tau$  die dargestellte Ellipse den Probanden als eine Kugel erscheint. Die Versuchspersonen sahen diese Szene binokular, monokular oder monokular mit Reduktionsrohr (UV *Sehbedingung*), unter verschiedenen Blickwinkeln (UV *Winkelbedingung*) und mit hervorgehobenem oder nicht hervorgehobenem ‘Leinwandcharakter’ (UV *Reizart*). Sämtliche Probanden durchliefen alle der 42 möglichen experimentellen Bedingungen. Für jeden Probanden wurden unter jeder Bedingung 10 Datenpunkte ( $\tau$ -Einstellungen) erhoben, die Ergebnisse wurden für jeden Probanden zunächst getrennt ausgewertet.

## Geräte und Aufbau

Sämtliche Reize wurden auf einem 21-Zoll Iiyama *Vision Master*<sup>™</sup> Pro 510 CRT-Monitor mit einer Bildwiederholungsrate von 100 Hertz dargeboten. Der Monitor wurde von einer NVIDIA<sup>®</sup> GeForce<sup>®</sup> 8800 GTS 512 Grafikkarte mit einer räumlichen Auflösung von 1280×960 Pixeln und einer Farbauffösung von 32 Bit pro Phosphorkanal angesteuert. Als Experimentalrechner wurde ein mit einem AMD Athlon<sup>™</sup> 64X2 Dual Core 4600+ Prozessor und 2 GB RAM ausgestatteter PC verwendet. Die Kalibrierung des Monitors wurde mithilfe eines *Cambridge Research Systems ColorCAL* Kolorimeters vorgenommen und in regelmäßigen Abständen überprüft (vgl. Brainard, 1989).<sup>18</sup> Der Monitor wurde 45 Minuten vor Beginn einer jeden Experimentalsitzung eingeschaltet, um eine gleichmäßige Strahlung der Kathodenstrahlröhre zu gewährleisten. Es wurde sichergestellt, dass der Bildschirm bei einem RGB-Palettenwert von [0, 0, 0] kein Eigenleuchten, d.h. eine Luminanz von 0 aufweist. Die CIE ( $x, y$ )-Chromatizitätskoordinaten des ‘Outputs’ der drei Phosphorkanäle des Monitors bei jeweils maximaler Intensität sind in Tabelle 4.1 mit zugehöriger Luminanz  $L$  in Candela pro Quadratmeter ( $cd/m^2$ ) angegeben.

Tabelle 4.1.

*Die  $xyL$ -Koordinaten des ‘Outputs’ der drei Phosphorkanäle bei maximaler Intensität*

Phosphorkanal	Palettenwert	$x$	$y$	$L$
R	[255, 0, 0]	.62	.34	22.93
G	[0, 255, 0]	.28	.61	66.9
B	[0, 0, 255]	.15	.07	8.05

Um die Aufmerksamkeit und die Blickrichtung der Probanden auf die präsentierten Reize zu konzentrieren, befand sich zwischen Monitor und Versuchsperson ein fest installierter,  $54 \times 45 \times 120$  cm großer Reduktionstunnel, der zudem innen geschwärzt war, um störende Reflexionen innerhalb des Tunnels zu vermeiden. Aus dieser Entfernung weist der Bildschirm eine Ausdehnung von  $19.1^\circ \times 14.5^\circ$  Sehwinkel auf. Auf der dem Monitor gegenüberliegenden Seite des Reduktionstunnels befanden sich zwei in diesen eingelassene, horizontal angeordnete und verschiebbare ‘Sehlöcher’ mit einem Durchmesser von jeweils circa 4 cm. Der linke Teil von Abbildung 4.14 auf der nächsten Seite zeigt eine schematische Darstellung der in dieser Untersuchung verwendeten Apparatur.

<sup>18</sup>An dieser Stelle sei Dr. Franz Faul vom *Visual Perception Lab* der Christian-Albrechts-Universität Kiel gedankt, der mir freundlicherweise ein Programm zur Gammakorrektur zur Verfügung gestellt hat.

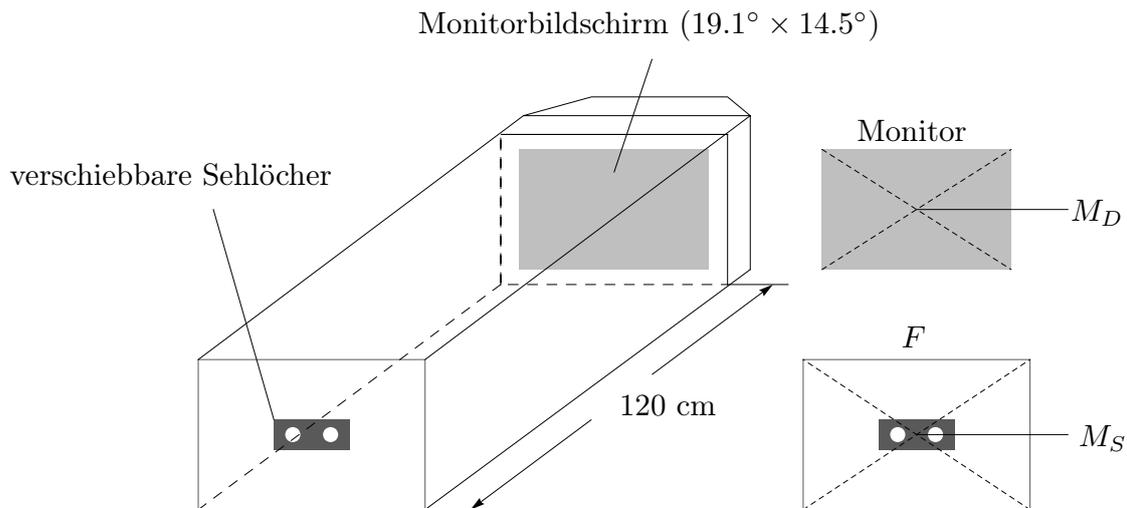


Abbildung 4.14. Schematische Darstellung der in der Untersuchung verwendeten Apparatur. Siehe Text für weitere Erläuterungen.

Für die folgende Beschreibung sollen aus Darstellungsgründen nun zusätzlich einige Abkürzungen eingeführt werden. Der Schnittpunkt der beiden Diagonalen des Monitordisplays soll mit  $M_D$  gekennzeichnet und auch als *Monitormittelpunkt* bezeichnet werden. Die frontale Ebene des Reduktionstunnels, auf der die Sehlöcher angebracht sind, soll mit  $F$  und der Schnittpunkt der Diagonalen von  $F$  mit  $M_S$  bezeichnet werden (vgl. rechter Teil der Abbildung 4.14). Die gesamte Apparatur war unter sämtlichen Experimentalbedingungen stets so aufgebaut, dass die Normale  $\mathfrak{N}_S$  von  $F$  zum Punkt  $M_S$  auch den Monitormittelpunkt  $M_D$  schneidet.

Je nach Sehbedingung blickten die Probanden durch beide in den Reduktionstunnel eingelassenen Sehlöcher (binokulare Sehbedingung), oder durch lediglich ein Sehloch (monokulare Sehbedingung und Aperturbedingung). Im Fall der binokularen Sehbedingung wurden die beiden Sehlöcher horizontal so ausgerichtet, dass jedes Sehloch vom Punkt  $M_S$  gerade die Hälfte des durchschnittlichen Augenabstandes ( $65/2$  mm) entfernt war. Unter der monokularen Sehbedingung und der Aperturbedingung wurde das eine Sehloch, durch das die Probanden blicken sollten, so verschoben, dass dessen Mittelpunkt im Punkt  $M_S$  lag. Unter der Aperturbedingung wurde das so ausgerichtete Sehloch zusätzlich mit einem 10 cm langen Reduktionsrohr von 3.5 cm Durchmesser versehen, um sicherzustellen, dass allein die auf dem Bildschirm dargestellte Szene für die Probanden sichtbar ist und nicht zusätzliche Elemente wie etwa der Rahmen des Monitors. Der Monitor stand auf einer mit Hilfslinien versehenen Fläche und wurde zur Realisierung der verschiedenen Winkelbedingungen über diese Hilfslinien um den entsprechenden Winkel im oder gegen den Uhrzeigersinn um eine gedachte, feste vertikale Achse rotiert, die durch den Punkt  $M_D$  verläuft. Bei jeder Rotation wurde sorgsam darauf geachtet, dass die Normale  $\mathfrak{N}_S$  stets durch den Punkt  $M_D$  verläuft. Der Blickwinkel

$\gamma_B$  ist nun definiert durch den von  $\mathfrak{N}_S$  und der Normalen  $\mathfrak{N}_D$  des Monitordisplays zum Punkt  $M_D$  eingeschlossenen Winkel. Abbildung 4.15 zeigt auf schematische Weise den Aufbau der Apparatur bei einer Winkelbedingung  $\gamma_B > 0$ .

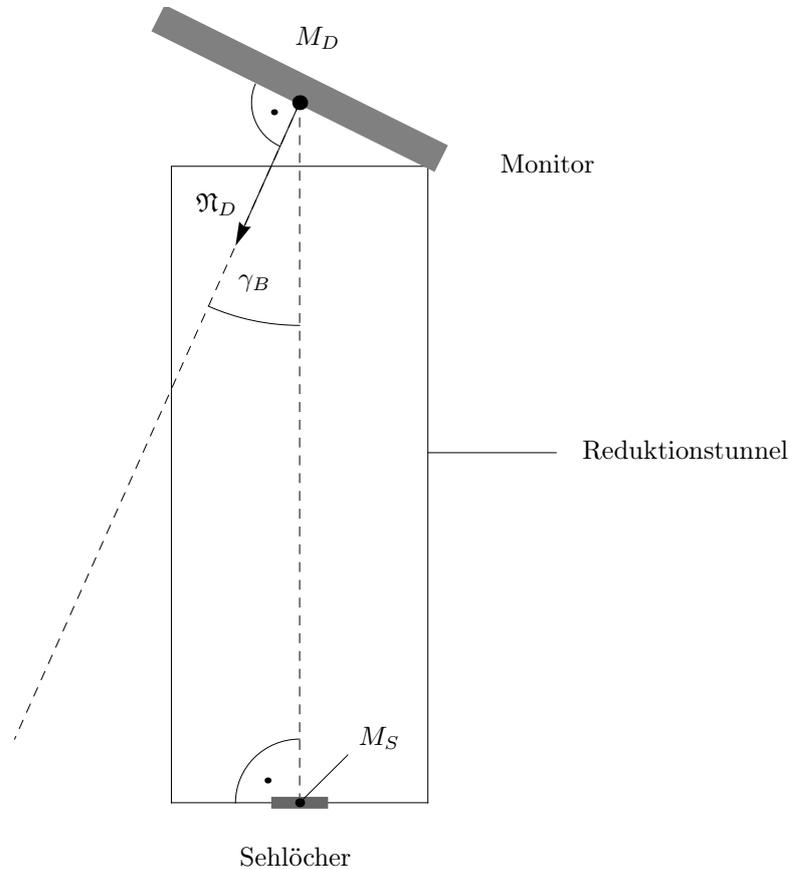


Abbildung 4.15. Schematische Darstellung (Aufsicht) der in den Experimenten verwendeten Apparatur bei Winkelbedingung  $\gamma_B > 0$ .

Die Apparatur befand sich in einem Raum, der komplett abgedunkelt werden konnte, möglicherweise störende Lichtquellen wie LED-Anzeigen an Rechner, Maus oder Tastatur wurden während des Experimentes abgedeckt. Als einzige Lichtquelle wurde eine schwach leuchtende, auf die Decke des Raums gerichtete Tischlampe genutzt, um sicherzustellen, dass in der binokularen und der monokularen Sehbedingung der Rahmen des Monitors für die Probanden sichtbar ist.

## Reizmaterial

Die den Probanden dargebotenen Bilder wurden per *ray tracing* (vgl. Shirley & Morley, 2003) unter Verwendung des Programms POV-Ray<sup>TM</sup> erstellt und wiesen in einer Auflösung von  $1280 \times 960$  Pixeln dieselbe räumliche Ausdehnung wie die gesamte Bildschirmfläche auf, also eine Größe von  $19.1^\circ \times 14.5^\circ$  Sehwinkel.<sup>19</sup> Die präsentierten Bilder

<sup>19</sup>Bei POV-Ray handelt es sich um freie Software des GNU-Projekts, die aktuelle Version lässt sich unter <http://www.povray.org> finden.

waren zentralperspektivisch korrekte Darstellungen einer Szene mit verschiedenen Objekten wie etwa Würfeln verschiedener Größe und Ausrichtung. Dabei wurde darauf geachtet, dass sämtliche dieser relevanten geometrischen Objekte auch unter der Aperturbedingung sichtbar waren, d.h. anschaulich gesprochen sich nicht ‘allzu nah’ an den äußeren Rändern der Darstellung befinden. Der Augpunkt wurde so gewählt, dass er auf der Normalen  $\mathfrak{N}_D$  liegt. Die Bilder sollten einen möglichst starken Tiefeneindruck erzeugen, dies wurde über verschiedene gestalterische Elemente, nämlich unter Verwendung vieler *pictorial depth cues* wie Schattenwurf, Verdeckung, *aerial perspective*, Texturgradient und kontrastreiche Darstellung der Ebene, auf der sich die Objekte befinden, zu erreichen versucht. Zudem wurde darauf geachtet, dass sich mehrere verschiedene Fluchtpunkte in der dargestellten Szene befinden, da dies laut Vishwanath et al. (2005) eine Bedingung für Perspektivenrobustheit sei.

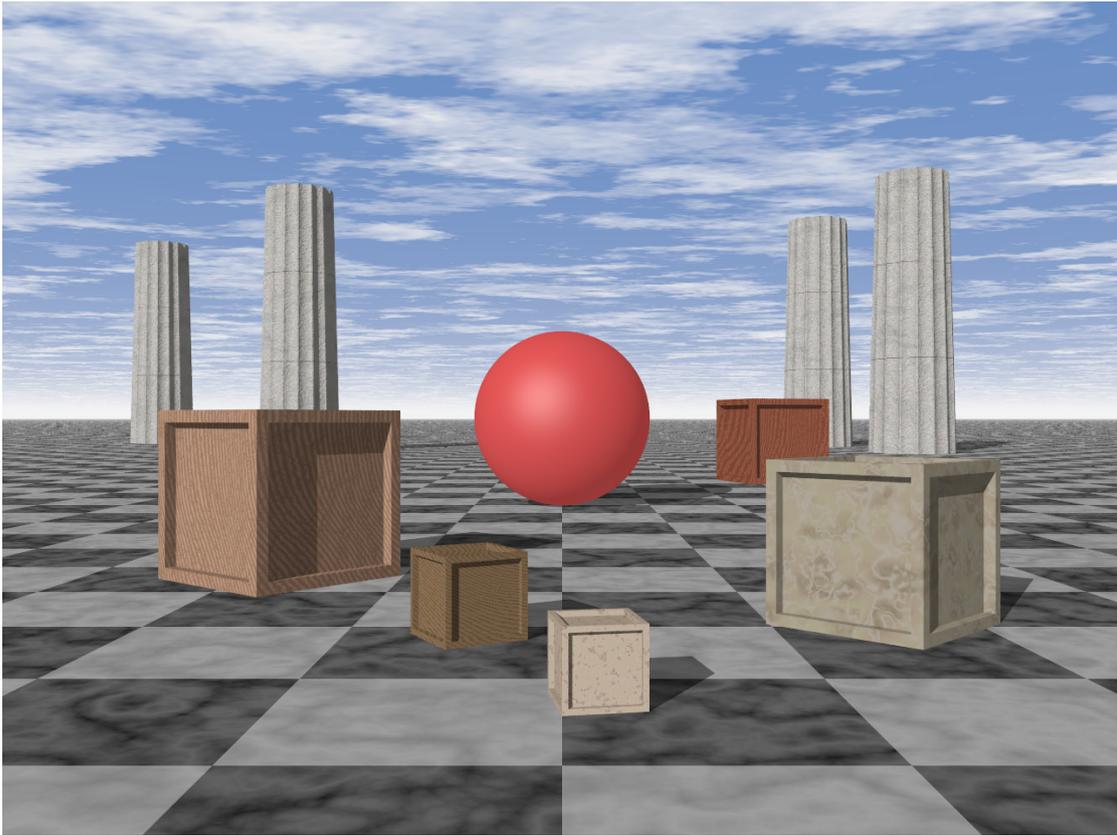
Zentrales Element der präsentierten Bilder war eine farblich vom Rest der Szene deutlich abgehobene Ellipse mit von Bild zu Bild unterschiedlicher Hauptachsenlänge  $a$ , aber konstanter Länge der Nebenachse  $b$ . Dieses zentrale Element der Bilder soll im Folgenden auch als *kritischer Reiz* bezeichnet werden. Der Mittelpunkt dieser Ellipse lag bei jedem erstellten Bild im Mittelpunkt des Monitordisplays  $M_D$  und somit auf einer Linie mit dem Augpunkt. Die  $(x, y, L)$ -Farbkoordinaten des kritischen Reizes betrugen in jedem Bild ( $x = .56, y = .34, L = 21.6$ ). Insgesamt wurden 121 solcher Bilder erstellt, die sich jeweils allein in der Länge der Hauptachse der Ellipse unterschieden. Das Spektrum des Verhältnisses von Länge der Hauptachse zu Länge der Nebenachse der Ellipsen reichte von  $\tau = 0.9$  (Bild 1) bis  $\tau = 1.5$  (Bild 121). Die Veränderung des Achsenverhältnisses von Bild  $i$  zu Bild  $i + 1$  ( $1 \leq i < 121$ ) betrug 0.005. Bei einem Hauptachsen-Nebenachsenverhältnis von Eins wies der kritische Reiz (in diesem speziellen Fall ein Kreis) eine Größe von etwa  $3^\circ$  Sehwinkel auf.

Insgesamt war es das übergeordnete Ziel, Bilder zu erstellen, die einerseits nicht zu komplex sind, um mögliche Ablenkungen von der Ellipse als kritischem Reiz zu verhindern, andererseits sollten sie dennoch ‘hinreichend komplex’ sein, um als realistische Darstellung einer Szene mit möglichst starkem Tiefeneindruck zu erscheinen. Abbildung 4.16 auf der nächsten Seite zeigt exemplarisch zwei dieser erstellten Bilder mit unterschiedlichem Hauptachsen-Nebenachsenverhältnis des kritischen Reizes. Diese Bilder wurden *nur* unter der Standardbedingung dargeboten.

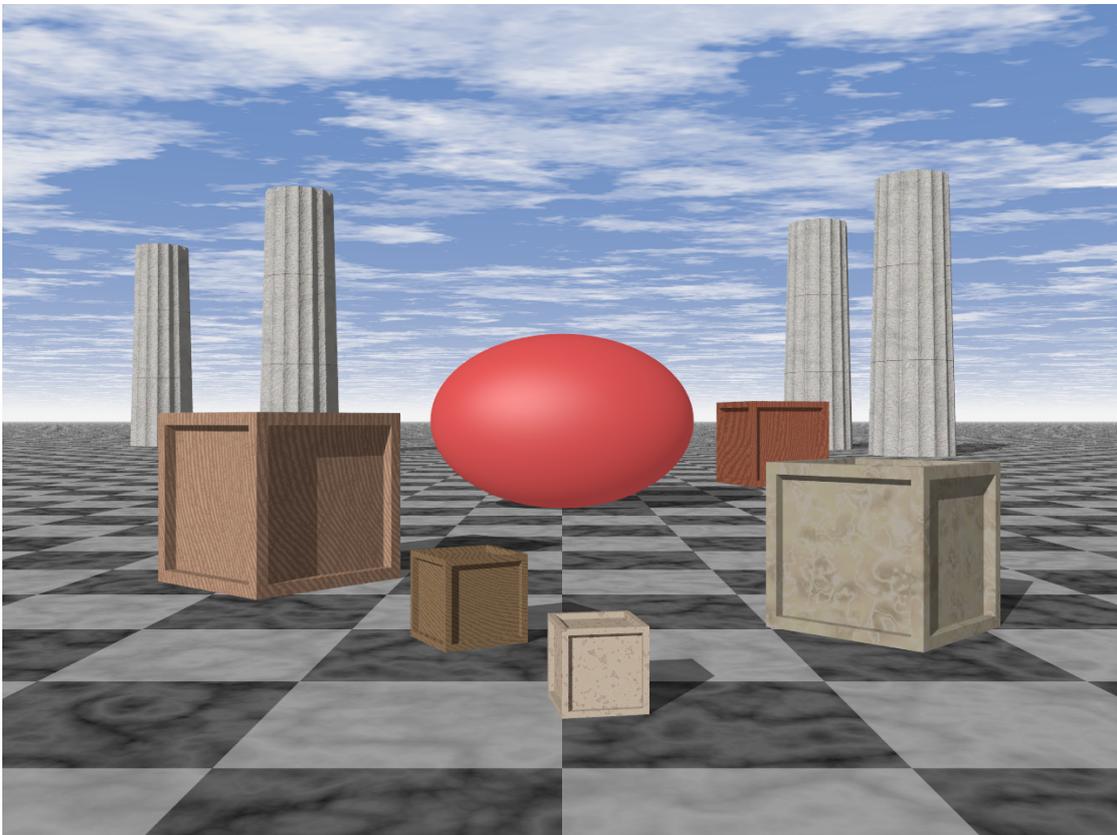
Für die Leinwandbedingung wurde jedes der 121 für die Standardbedingung erstellten Bilder mittels eines Filters der ebenfalls frei erhältlichen *GNU Image Manipulation Software* (GIMP) so verändert, dass die Bilder wirkten, als ob sie auf eine Leinwand gemalt worden wären.<sup>20</sup> Auf die Geometrie der dargestellten der Szene hatte diese

---

<sup>20</sup>Die *GNU Image Manipulation Software* lässt sich auf <http://www.gimp.org> finden. Der verwendete Filter ist der sogenannte ‘Leinwandfilter’, der mit den Einstellungen ‘Oben-rechts’ und ‘Tiefe



(a)



(b)

Abbildung 4.16. Zwei Bilder aus der Standardbedingung. Das Verhältnis von Haupt- zu Nebenachse der Ellipse beträgt  $\tau = 1$  (a) bzw.  $\tau = 1.5$  (b).

Veränderung keinerlei Auswirkung. Abbildung 4.17 auf der nächsten Seite zeigt exemplarisch zwei Bilder aus der Leinwandbedingung mit unterschiedlichem Hauptachsen-Nebenachsenverhältnis des kritischen Reizes.

## Durchführung

### Allgemeiner Ablauf

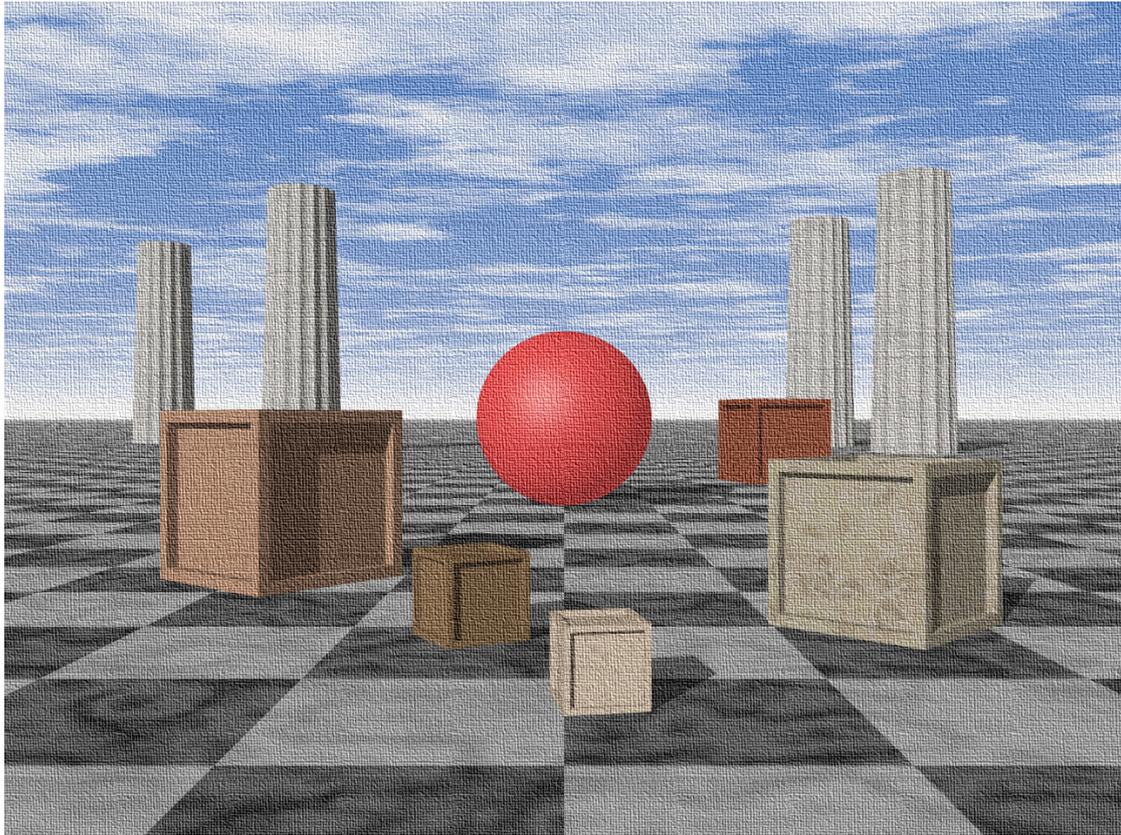
Jede Versuchsperson absolvierte für jede der 42 experimentellen Bedingungen eine Sitzung, die selbst wiederum aus 10 Durchgängen bestand. Die Reihenfolge der zu absolvierenden Bedingungen wurde dabei für jede Versuchsperson per Zufall ermittelt. Den Probanden wurde zu Beginn der ersten Sitzung die an sie gestellte Aufgabe erklärt, Ziel und inhaltliche Hypothesen des Experimentes wurden nicht mitgeteilt, insbesondere wurde nicht erwähnt, dass sie die Bilder unter den verschiedenen Bedingungen aus verschiedenen Blickwinkeln betrachten werden. Vor Beginn jeder Sitzung wurden die Versuchspersonen für fünf Minuten an die Lichtbedingungen im schwach beleuchteten Labor adaptiert, anschließend konnte der erste Durchgang per Tastendruck des Probanden gestartet werden. Dem Probanden wurde daraufhin für fünf Sekunden eines der 121 erstellten Bilder der der Sitzung zugehörigen Reizart unter der der Sitzung zugehörigen Seh- und Winkelbedingung präsentiert. So wurde beispielsweise in der der Bedingungskombination „Standardreiz, binokulare Sehbedingung,  $\gamma_B = 15^\circ$ “ zugehörigen Sitzung eines der 121 für die Standardreizbedingung erstellten Bilder unter einem Blickwinkel  $\gamma_B$  von 15 Grad präsentiert, welches der Proband binokular betrachten konnte. Welches der 121 Bilder der entsprechenden Reizbedingung dem Probanden zu Beginn des Durchgangs präsentiert wurde, wurde von dem verwendeten und noch näher zu erläuternden adaptiven Verfahren der *double random staircases* (vgl. Cornsweet, 1962; Falmagne, 1985) gesteuert.

Die Versuchspersonen hatten nach der Darbietung des Bildes die Aufgabe, innerhalb von fünf Sekunden per Tastendruck zu entscheiden, ob die jeweils dargestellte Ellipse (der kritische Reiz) ihnen „zu schmal“ oder „zu breit“ für eine Kugel erscheint.<sup>21</sup> Voruntersuchungen des Autors wiesen daraufhin, dass für diese Entscheidung ein Zeitraum von fünf Sekunden ausreichend lang ist, keine der Versuchspersonen des durchgeführten Experimentes empfand diesen retrospektiv als zu kurz. In Abhängigkeit von der Antwort des Probanden wurde nach der abgegebenen Entscheidung und über das adaptive Verfahren gesteuert, ein weiteres der 121 Bilder für fünf Sekunden präsentiert und die Versuchsperson musste nach Darbietung innerhalb von fünf Sekunden erneut die oben beschriebene Entscheidung vornehmen. Diese Prozedur wurde so lange wiederholt, bis

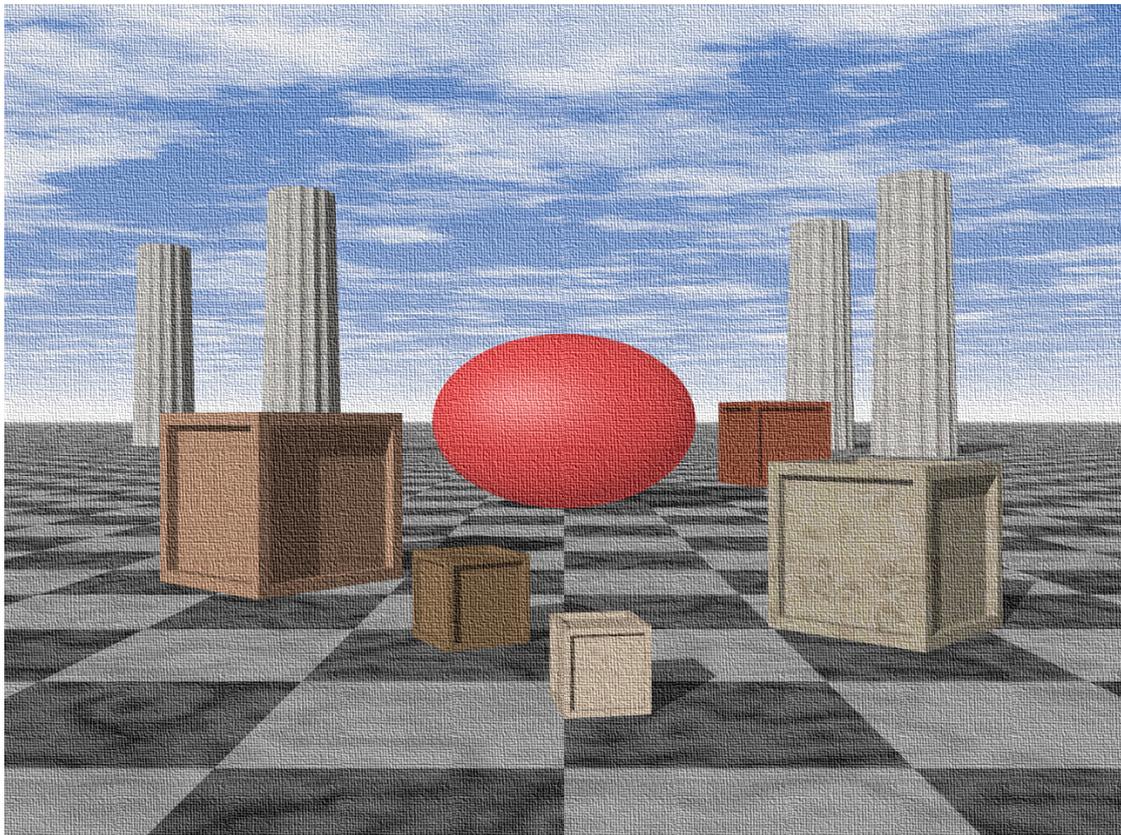
---

= 5“ auf die Reize der Standardbedingung angewendet wurde. Für eine exakte Beschreibung des Algorithmus dieses Filters siehe den auf der obigen Webseite einsehbaren Quellcode von GIMP.

<sup>21</sup>Es handelt sich demnach um eine sogenannte *two-alternative forced choice*-Aufgabe (2AFC).



(a)



(b)

Abbildung 4.17. Zwei Bilder aus der Leinwandbedingung. Das Verhältnis von Haupt- zu Nebenachse der Ellipse beträgt auch hier  $\tau = 1$  (a) bzw.  $\tau = 1.5$  (b).

das adaptive Verfahren ein vorher definiertes Kriterium erreicht hatte. Das Erreichen dieses Kriteriums markierte das Ende eines Durchgangs einer Sitzung. Anschaulich und vereinfacht gesprochen bestimmte innerhalb einer Sitzung die Antwort des Probanden auf ein Bild das anschließend präsentierte Bild. Beurteilte der Proband den kritischen Reiz etwa als „zu schmal“, so wurde ihm anschließend ein Bild mit einem kritischen Reiz dargeboten, der ein im Vergleich zum vorher dargebotenen kritischen Reiz größeres Hauptachsen-Nebenachsenverhältnis aufwies, beurteilte er den kritischen Reiz als „zu breit“, so wurde ihm anschließend ein kritischer Reiz dargeboten, der ein geringeres Hauptachsen-Nebenachsenverhältnis aufwies. Das ‘Durchschreiten’ eines derartigen *staircases* liefert für den Probanden ein Datum, nämlich das gesuchte  $\tau$  jener Ellipse, die dem Probanden unter der entsprechenden Bedingung als Kugel erscheint. Nach jedem Durchgang einer Sitzung wurde eine zweiminütige Pause eingelegt. Für jede der 42 Sitzungen wurden 10 Durchgänge durchgeführt, so dass am Ende der Untersuchung für jede Versuchsperson und jede Bedingung zehn Datenpunkte ( $\tau$ -Werte) vorlagen. Eine Sitzung mit 10 Durchgängen dauerte im Schnitt etwa 35 Minuten, das gesamte Experiment mit 42 Sitzungen nahm für eine Versuchsperson insgesamt ca. 25 Stunden in Anspruch.

Die Entscheidung für ein Staircase-Verfahren (welches den zeitlichen Aufwand deutlich erhöht) und gegen eine Methode, die es den Versuchspersonen etwa erlauben würde, das Verhältnis  $\tau$  der dargestellten Ellipse per Tastendruck und ohne Zeitvorgabe selbst so zu verändern, dass diese ihnen als Kugel erscheint, ist darin begründet, dass erstens mögliche Adaptationseffekte verhindert werden sollten und zweitens den Probanden nicht die Möglichkeit gegeben werden sollte, eventuell Elemente des Bildes als mögliche Hilfsmittel zur Bestimmung des Hauptachsen-Nebenachsenverhältnisses der dargestellten Ellipse zu nutzen. Zudem ist die mit dem *staircase*-Verfahren verbundene Urteilsaufgabe einfacher und verlässlicher von den Probanden zu erledigen, als die abstraktere Aufgabe, die Ellipse so einzustellen, dass sie ‘als Kugel erscheint’.

### **Anschauliche Erläuterung des verwendeten *staircase*-Verfahrens**

Es soll nun kurz anhand eines auf die vorliegende Untersuchung bezogenen Beispiels das verwendete Verfahren der *double random staircases* anschaulich erläutert werden. Die Idee dieser Methode lässt sich folgendermaßen umschreiben (siehe auch Cornsweet, 1962): Gesucht werde in unserem Fall jenes Verhältnis  $\tau$  von Hauptachse zu Nebenachse, so dass die zugehörige Ellipse der Versuchsperson unter der entsprechenden Bedingung als eine Kugel erscheint. Eine Schätzung dieses Wertes soll nun vorgenommen werden, indem man sich über ein Staircase-Verfahren dem gesuchten Wert gleichsam ‘schrittweise’ annähert, wobei bei dieser Vorgehensweise entscheidend ist, dass die Reizdarbietung in jedem Darbietungsschritt innerhalb eines Durchgangs in

systematischer Weise von den vorangegangenen Urteilen der Versuchsperson abhängt. Es soll an dieser Stelle auf eine formale Art der Darstellung verzichtet und zudem nicht ausführlich auf die technischen Details dieses Verfahrens eingegangen werden, da dies u.a. einen zusätzlichen notationellen Aufwand mit sich bringen würde, der hier nicht zu rechtfertigen ist.

Zunächst ist es erforderlich, dass der Versuchsleiter einen ‘hinreichend großen Bereich’ festlegt, in dem sich das gesuchte  $\tau$  mit ‘hoher Wahrscheinlichkeit’ befinden wird. Aufgrund der geometrischen Überlegungen aus Abschnitt 4.2 sowie den von Vishwanath et al. (2005) berichteten Ergebnissen wurde das Intervall  $[0.9, 1.5]$  gewählt. In einem nächsten Schritt wurde dieses Intervall diskretisiert, indem festgelegt wurde, um welche Größe sich der Parameter  $\tau$  des kritischen Reizes von Bild zu Bild verändern soll. Diese Größe wurde, wie bereits dargestellt, mit 0.005 festgesetzt, so dass insgesamt 121 Bilder mit den zugehörigen Parametern  $\tau_1, \dots, \tau_{121}$  erstellt wurden, wobei  $\tau_1 = 0.9$  und  $\tau_{121} = 1.5$ . Die beiden äußeren Grenzen des so festgelegten Bereiches bilden nun die ‘Startpunkte’ für zwei *unabhängige* Staircases  $A$  und  $B$ , zwischen denen im weiteren Verlauf eines Durchgangs zufällig alterniert wird. Wir wollen annehmen, dass der Startpunkt für Staircase  $A$  das zu  $\tau_1$  und dementsprechend der Startpunkt für Staircase  $B$  das zu  $\tau_{121}$  zugehörige Bild ist. Über diese beiden Staircases wird nun die Reizdarbietung innerhalb eines Durchgangs gesteuert. Das weitere Vorgehen soll nun anhand Abbildung 4.18 auf der nächsten Seite veranschaulicht werden, die einen typischen Verlauf eines *double random-staircase*-Verfahrens innerhalb eines Durchgangs unter einer festen Bedingung darstellt. Die durch die Quadrate gegebene Kurve ist die des Staircases  $A$  und die durch die Kreise dargestellte Kurve ist die des Staircases  $B$ . Ausgefüllte Symbole stehen für das von den Probanden nach der durch die Staircases gesteuerten Reizdarbietung abgegebene Urteil ‘zu schmal’ und nicht ausgefüllte Symbole für das Urteil ‘zu breit’. Am unteren Bildrand lässt sich ablesen, bei welchem Schritt sich der Proband im Staircase  $A$  und am oberen Bildrand, bei welchem Schritt sich der Proband im Staircase  $B$  befindet.

Wir wollen nun den Verlauf von Staircase  $A$  betrachten, der in diesem Beispiel für die erste Reizdarbietung innerhalb des Durchgangs zufällig gewählt wurde. In Schritt 1 dieses Staircases wird das mit dem Startpunkt von  $A$  zugehörige Bild präsentiert ( $\tau = .9$ ). Der Proband gibt daraufhin das Urteil ab, dass ihm die dargestellte Ellipse als ‘zu schmal’ für eine Kugel erscheint (ausgefülltes Symbol). Im nächsten Schritt von  $A$  wird ein Bild mit einem größeren  $\tau$  präsentiert. Die Größe dieser Veränderung des Wertes  $\tau$  von einem Schritt zum nächsten wird im Rahmen dieses Verfahrens auch als *Schrittbreite* (*step size*) bezeichnet. Diese Schrittbreite bleibt über den Verlauf eines Staircases nicht gleich, sondern *verringert* sich immer dann um einen konstanten Faktor, wenn der Proband ein *anderes* Urteil als im vorangehenden Schritt abgibt. Auch in

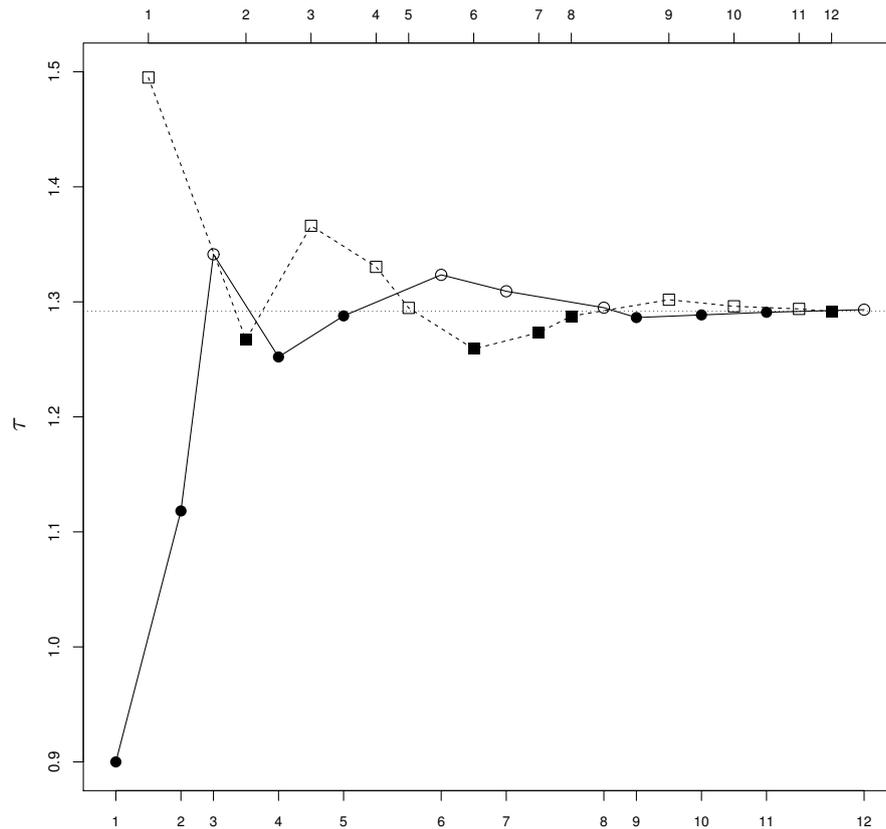


Abbildung 4.18. Typischer Verlauf eines *double random-staircase*-Verfahrens für unser Beispiel. Für weitergehende Erläuterungen siehe Text.

Schritt 2 von  $A$  beurteilt der Proband die Ellipse als ‘zu schmal’, woraufhin in Schritt 3 ein Bild mit einem noch größeren  $\tau$  präsentiert wird. Die Schrittbreite von Schritt 2 zu Schritt 3 ist aufgrund des gleich bleibenden Urteils identisch der Schrittbreite von Schritt 1 zu Schritt 2. In Schritt 3 gibt der Proband nun das Urteil ‘zu breit’ ab (nicht ausgefülltes Symbol). In Schritt 4 wird daraufhin ein Bild mit einem geringeren  $\tau$  als in Schritt 3 präsentiert, wobei die Schrittbreite von Schritt 3 zu Schritt 4 aufgrund des wechselnden Urteils des Probanden geringer ausfällt als die Schrittbreite von Schritt 1 zu 2 bzw. von Schritt 2 zu 3.

An Abbildung 4.18 lässt sich nun gut ablesen, wie im weiteren Verlauf von  $A$  durch dieses Vorgehen die Variabilität der Werte  $\tau$  stetig abnimmt und diese gegen Ende nur noch innerhalb eines ziemlich kleinen Intervalls variieren. Analoge Ausführungen gelten auch für Staircase  $B$ . Über die zufällige Alternierung zwischen  $A$  und  $B$  wird es so möglich, sich dem gesuchten Hauptachsen-Nebenachsenverhältnis jener Ellipse, die den Probanden als Kugel erscheint, sukzessive und in immer kleineren Schritten von den Startpunkten der beiden Staircases aus anzunähern. Zur Schätzung des gesuchten Wertes  $\tau$  wurde nun folgendes Abbruchkriterium für das Verfahren festgesetzt: Der Prozess bricht ab, wenn die Differenz zwischen dem durch Staircase  $A$  gewählten  $\tau$  und dem durch Staircase  $B$  gewählten  $\tau$  in sechs *aufeinanderfolgenden* Reizpräsentationen

kleiner oder gleich .015 beträgt. Die Schätzung des gesuchten  $\tau$  wird dann als Mittel dieser sechs letzten  $\tau$ -Werte bestimmt. Für unser Beispiel ist diese Schätzung ebenfalls in Abbildung 4.18 abgetragen, es ist gleichsam jener Wert, auf den die beiden Staircases, durch die Urteile der Probanden bestimmt, von ihren jeweiligen Startpunkten aus ‘zulaufen’. Diese Schätzung wurde dann interpretiert als jener gesuchte Wert  $\tau$ , so dass die zugehörige Ellipse der Versuchsperson unter der entsprechenden Bedingung als Kugel erscheint. Um sicherzustellen, dass das Verfahren nicht ‘endlos weiterläuft’, falls das festgelegte Kriterium nicht erreicht werden sollte, wurde zusätzlich festgelegt, dass das Verfahren nach 45 Bildpräsentationen automatisch abbricht und keine Schätzung des gesuchten  $\tau$  erfolgt. Das geforderte Abbruchkriterium wurde jedoch von allen Versuchspersonen in allen Durchgängen aller Bedingungen erreicht, wobei es im Schnitt nach etwa 25 Reizdarstellungen erreicht wurde.

## Teilnehmer

Es nahmen insgesamt zehn (4♂, 6♀) Versuchspersonen an der Untersuchung teil, die sich allesamt im Bachelor- oder Masterstudiengang Psychologie der Martin-Luther-Universität Halle-Wittenberg befanden. Der Altersdurchschnitt der Probanden betrug zum Zeitpunkt der Erhebung 23.4 Jahre mit einer Standardabweichung von ca. 3.2 Jahren. Alle Versuchspersonen konnten aufgrund der Ergebnisse des „Ishihara-Tests“ (Ishihara, 1998) als normal farbsichtig eingestuft werden. Probanden mit Fehlsichtigkeit trugen Korrekturgläser oder -linsen, um diese zu kompensieren. Die Probanden wurde für ihre Teilnahme an dem zeitlich sehr aufwendigen Experiment entlohnt (25 Versuchspersonenstunden, eine sowohl im Bachelor- als auch Master-Studiengang zu erbringende Studienleistung), niemand brach während der Datenerhebung die Teilnahme ab.

## 4.5. Ergebnisse

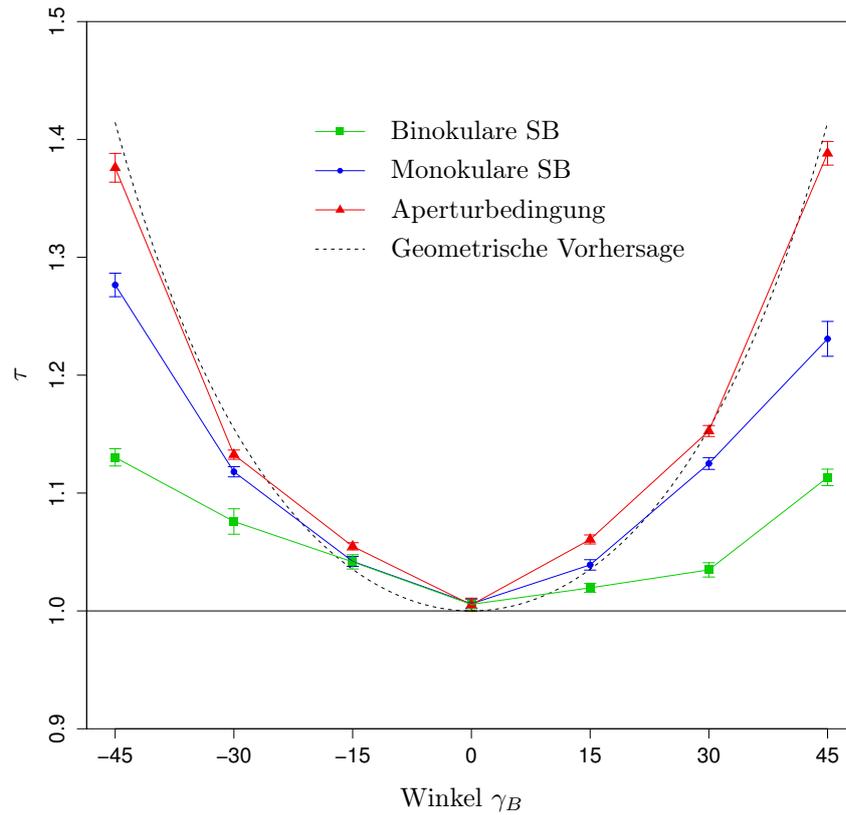
Durch das geschilderte Vorgehen lagen für jede der 10 Versuchspersonen unter jeder der 42 experimentellen Bedingungen 10 über das Staircase-Verfahren ermittelte Werte  $\tau$  vor, deren Mittelwert und Standardabweichung unter jeder Bedingung separat für jeden Probanden berechnet wurden. Zur zusätzlichen statistischen Absicherung der hier berichteten Befunde wurde für jede Versuchsperson separat eine inferenzstatistische Auswertung mittels einer mehrfaktoriellen Varianzanalyse mit den Faktoren *Winkelbedingung*, *Sehbedingung* und *Reizart* vorgenommen. Die statistischen Voraussetzungen dieses Verfahrens können in diesem Fall wegen der systematischen Verletzung der Voraussetzung der Varianzhomogenität als *nicht* erfüllt angesehen werden. Aufgrund der bei allen Versuchspersonen feststellbaren gleichen und markanten Effektmuster sol-

len die deskriptiven und inferenzstatistischen Ergebnisse der Varianzanalyse dennoch zur Unterstreichung des Gesamtbildes der Resultate herangezogen werden. Um den Lesefluss nicht zu unterbrechen, sind die ausführlichen Ergebnistabellen dieser zehn mehrfaktoriellen Varianzanalysen im Anhang A auf den Seiten 262 bis 265 zu finden.

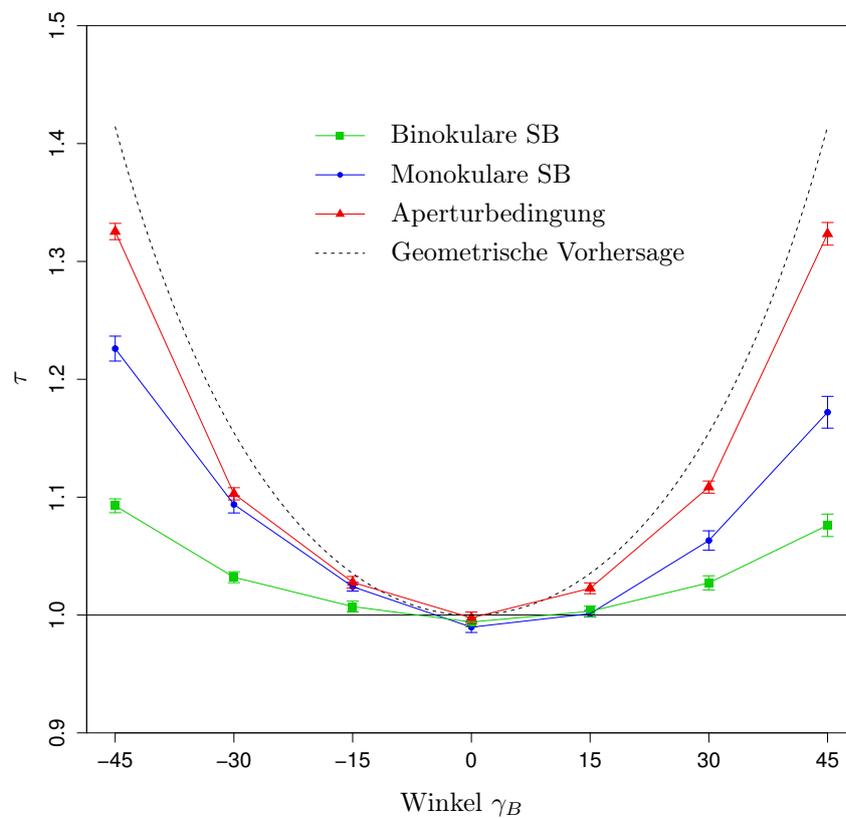
Die folgenden Abbildungen 4.19 bis 4.21 (S. 232 - 234) zeigen exemplarisch für drei der Probanden die über die jeweils 10 Durchgänge innerhalb einer Sitzung gemittelten Werte von  $\tau$  unter den verschiedenen Seh- und Winkelbedingungen, getrennt nach Standard- und Leinwandbedingung. Die Daten aller anderen Probanden lassen sich in Anhang A finden, die im Übrigen den hier gezeigten sehr ähnlich sind. Zusätzlich ist für sämtliche abgetragenen mittleren Einstellungen der zugehörige Standardfehler als halbe Länge der eingezeichneten Fehlerbalken dargestellt.

Bei sämtlichen Probanden ließ sich zunächst feststellen, dass unter einer beliebigen Seh- und Reizbedingung tendentiell die mittleren  $\tau$ -Einstellungen mit zunehmendem absolutem Winkel  $|\gamma_B|$  ebenfalls zunehmen, was hier als abnehmende Perspektivenrobustheit mit zunehmender Abweichung vom Augpunkt des Bildes interpretiert werden kann. Dabei zeigt sich deskriptiv kein merklicher Einfluss der ‘Richtung der Abweichung’ vom Augpunkt des Bildes (des Vorzeichens des Winkels  $\gamma_B$ ), die Mittelwertsverläufe unter jeder Seh- und Reizbedingung können als annähernd symmetrisch um die Einstellung bei  $\gamma_B = 0$  angesehen werden. Mit anderen Worten zeigt sich keine unterschiedliche Auswirkung der Rotation des Monitordisplays im oder gegen den Uhrzeigersinn. Eine vollständige oder ‘perfekte’ Perspektivenrobustheit im Sinne einer über *sämtliche* realisierten Winkelbedingungen hinweg konstanten Einstellung von  $\tau = 1$  zeigte sich unter keiner Seh- und Reizbedingung. Im Rahmen der inferenzstatistischen Auswertung ließ sich für *jede* Versuchsperson ein auf Niveau  $\alpha = .001$  signifikanter und überaus deutlicher Haupteffekt des Faktors *Winkelbedingung* feststellen (bei Proband SS etwa  $F_{6,378} = 982.5, \eta_p^2 = .94$ ).

Weiterhin zeigt sich ein deutlicher Einfluss der Sehbedingung auf die mittleren  $\tau$ -Einstellungen. Über sämtliche Winkelbedingungen hinweg betrachtet sind die mittleren  $\tau$ -Einstellungen unter der binokularen Sehbedingung am geringsten, in der Aperturbedingung am höchsten. Auffällig ist dabei die Systematik der verschiedenen Mittelwertsverläufe. Sowohl unter der Leinwandbedingung als auch unter der Standardbedingung sind unter jeder Winkelbedingung ( $\gamma_B = 0$  ausgeschlossen) die mittleren  $\tau$ -Einstellungen unter der Aperturbedingung tendenteill größer als unter der monokularen Sehbedingung und die mittleren  $\tau$ -Einstellungen unter der monokularen Sehbedingung wiederum tendenteill größer als unter der binokularen Sehbedingung. Die Größe dieser Unterschiede scheint zudem mit zunehmender Abweichung von  $\gamma_B = 0$  zunehmend markanter zu werden. Im Rahmen der inferenzstatistischen Auswertung zeigte sich bei *jeder* Versuchsperson ein auf Niveau  $\alpha = .001$  signifikanter und deutlicher Haupteffekt

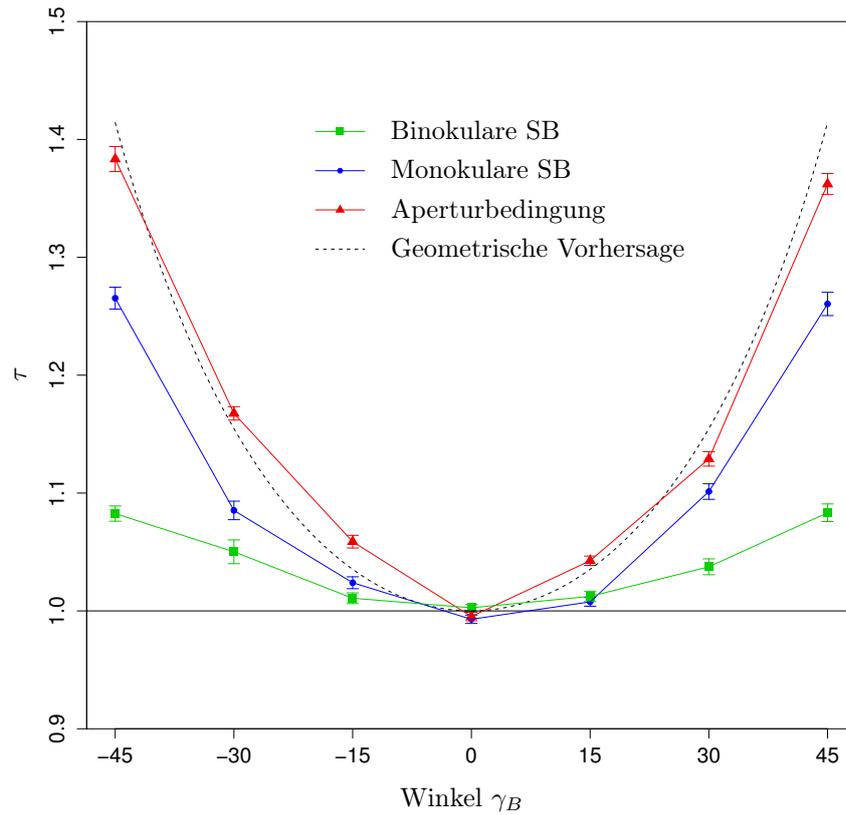


(a)

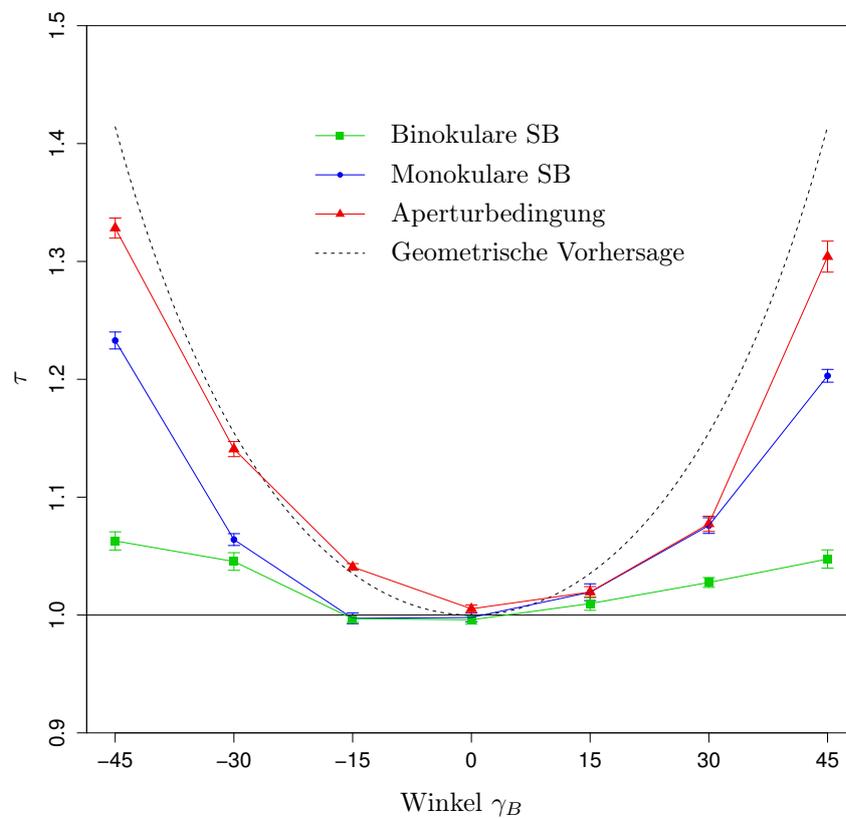


(b)

Abbildung 4.19. Mittelwerte der  $\tau$ -Einstellungen für Versuchsperson SS unter den verschiedenen Seh- und Winkelbedingungen in der Standardbedingung (a) und der Leinwandbedingung (b). Die halbe Länge der Fehlerbalken repräsentiert den jeweils zugehörigen Standardfehler.

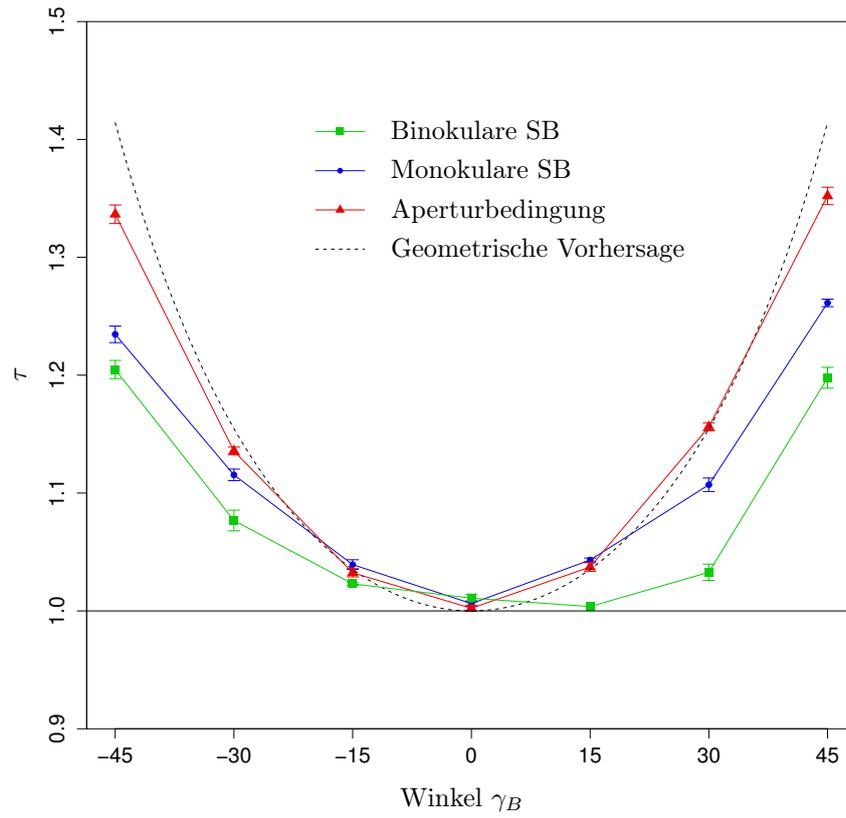


(a)

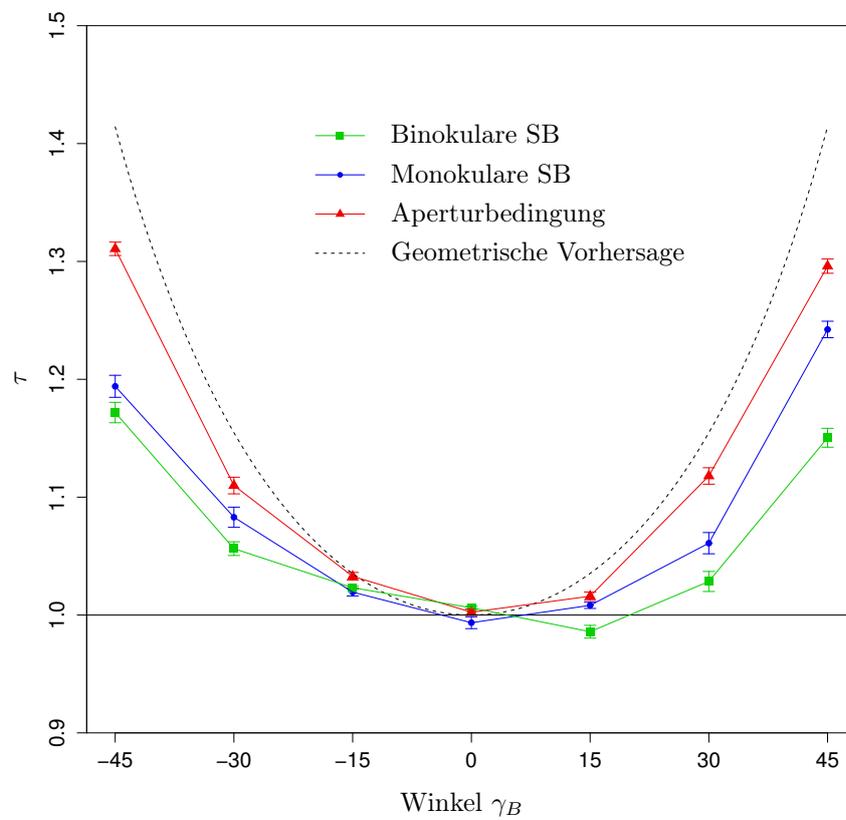


(b)

Abbildung 4.20. Mittelwerte der  $\tau$ -Einstellungen für Versuchsperson CL unter den verschiedenen Seh- und Winkelbedingungen in der Standardbedingung (a) und der Leinwandbedingung (b). Die halbe Länge der Fehlerbalken repräsentiert den jeweils zugehörigen Standardfehler.



(a)



(b)

Abbildung 4.21. Mittelwerte der  $\tau$ -Einstellungen für Versuchsperson FK unter den verschiedenen Seh- und Winkelbedingungen in der Standardbedingung (a) und der Leinwandbedingung (b). Die halbe Länge der Fehlerbalken repräsentiert den jeweils zugehörigen Standardfehler.

des Faktors *Sehbedingung* (bei Proband SS:  $F_{2,378} = 668.79, \eta_p^2 = .78$ ), sowie ein auf dem gleichen Niveau signifikanter und deutlicher Interaktionseffekt zwischen den Faktoren *Sehbedingung* und *Winkelbedingung* (bei Proband SS:  $F_{12,378} = 107.7, \eta_p^2 = .77$ ). Fasst man die  $\tau$ -Einstellungen als Indikatoren für das Ausmaß der Perspektivenrobustheit auf, so ist über alle Winkelbedingungen hinweg die Perspektivenrobustheit unter der binokularen Bedingung am stärksten und unter der Aperturbedingung am geringsten ausgeprägt, wobei die Größe der Unterschiedlichkeit des Ausmaßes der Perspektivenrobustheit unter den verschiedenen Sehbedingungen mit zunehmender Abweichung von  $\gamma_B = 0$  zunehmend markanter wird.

Wie lässt sich der Einfluss der Reizart auf die  $\tau$ -Einstellungen charakterisieren? Die Abbildungen 4.22 bis 4.24 (S. 236 - 238) zeigen exemplarisch für die betrachteten Probanden SS, CL und FK unter jeder Sehbedingung die mittleren  $\tau$ -Einstellungen unter der Leinwandbedingung und unter der Standardbedingung getrennt, für die restlichen Probanden siehe wieder Anhang A.

Vergleicht man die mittleren Einstellungen von  $\tau$  unter der Leinwand- mit denen unter der Standardbedingung, so wird deutlich, dass unter jeder Sehbedingung die von den Probanden vorgenommenen mittleren Einstellungen von  $\tau$  unter der Standardbedingung tendentiell höher ausfallen als unter der Leinwandbedingung, wobei auch hier diese Unterschiede umso markanter auszufallen scheinen, umso stärker der Blickwinkel  $\gamma_B$  vom Augpunkt des Bildes abweicht. Dieser Einfluss der Reizart auf die mittleren  $\tau$ -Einstellungen ist zwar als nicht so markant wie der der Sehbedingung zu charakterisieren, aber als Muster durchgängig bei sämtlichen Probanden beobachtbar. Im Rahmen der inferenzstatistischen Auswertung zeigte sich bei *jedem* Probanden ein auf Niveau  $\alpha = .001$  signifikanter Haupteffekt des Faktors *Reizart*, der jedoch weniger deutlich ist als die Haupteffekte der Winkelbedingung und der Sehbedingung (bei Proband SS:  $F_{1,378} = 225.9, \eta_p^2 = .37$ ). Zudem zeigte sich ein auf dem gleichen Niveau signifikanter Interaktionseffekt der Faktoren *Reizart* und *Winkelbedingung*, der in der Regel jedoch vergleichsweise gering ausfiel (bei Proband SS:  $F_{6,378} = 5.11, \eta_p^2 = .08$ ). Interpretiert man auch hier die  $\tau$ -Einstellungen als Ausdruck des Ausmaßes an Perspektivenrobustheit, so zeigt sich unter allen Sehbedingungen tendentiell ein stärkeres Maß an Perspektivenrobustheit unter der Leinwandbedingung als unter der Standardbedingung.

Trotz der zwischen den Probanden teilweise differierenden *absoluten Größe* der mittleren  $\tau$ -Einstellungen (siehe etwa die Daten von Proband CB in Anhang A) wurde aufgrund der deutlich ähnlichen Anordnung der Verläufe dieser Einstellungen entschieden, für jede mögliche Bedingungskonstellation den Mittelwert der mittleren  $\tau$ -Einstellungen aller Probanden zu ermitteln. Abbildung 4.25 auf Seite 240 zeigt die über alle Versuchspersonen gemittelten mittleren  $\tau$ -Einstellungen getrennt nach Standard-

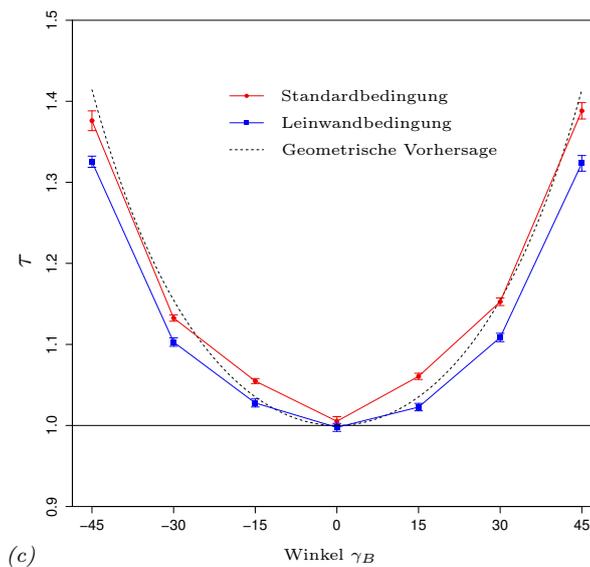
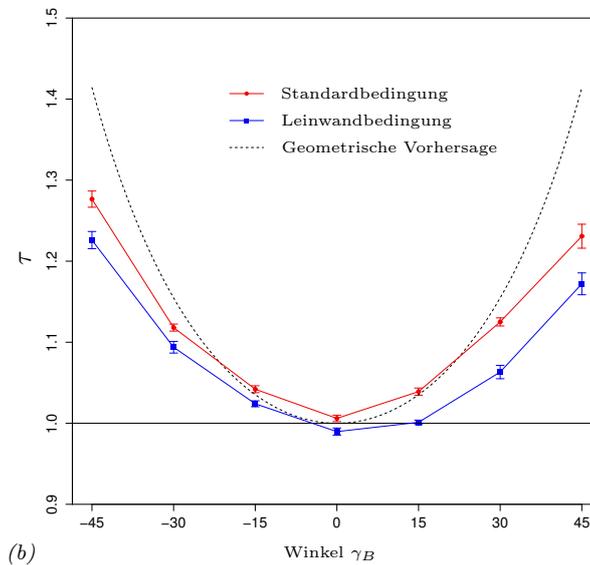
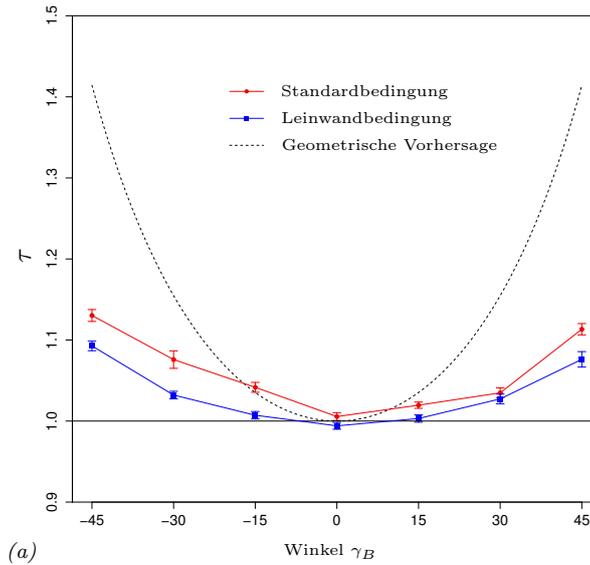


Abbildung 4.22. Mittlere  $\tau$ -Einstellungen des Probanden SS jeweils unter der Leinwand- und der Standardbedingung für die drei Sehbedingungen binokular (a), monokular (b) und Aperturbedingung (c). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.

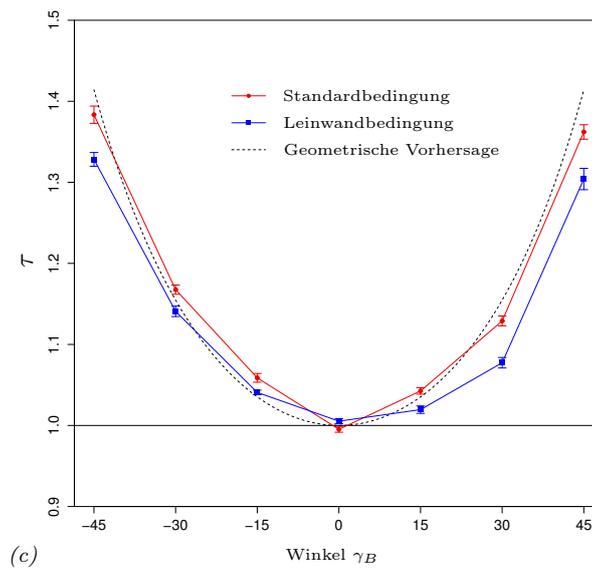
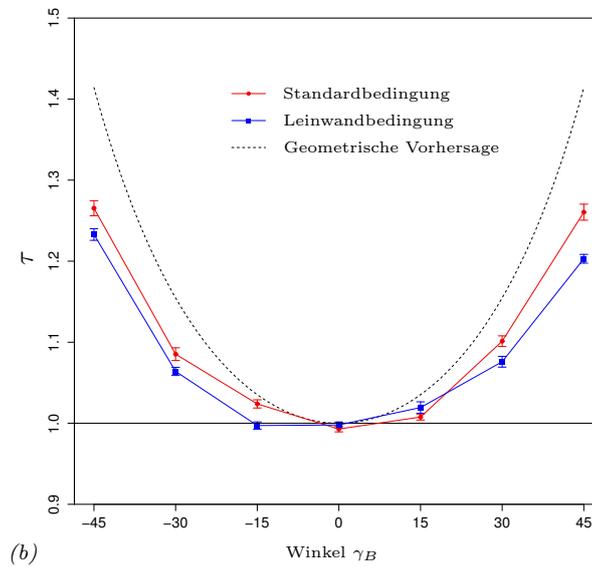
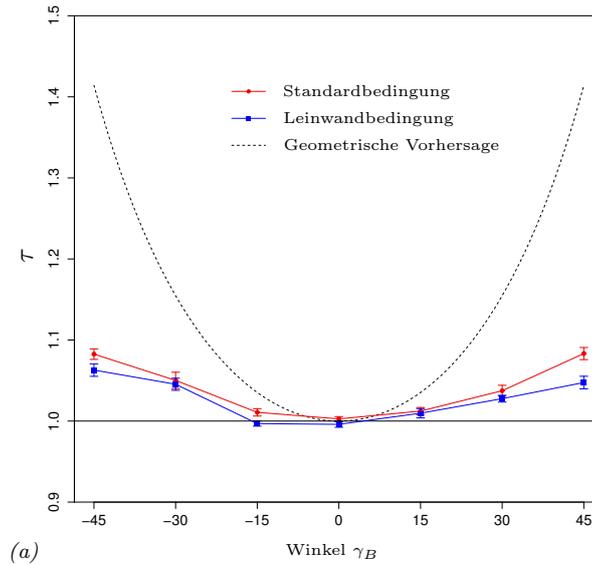


Abbildung 4.23. Mittlere  $\tau$ -Einstellungen des Probanden CL jeweils unter der Leinwand- und der Standardbedingung für die drei Sehbedingungen binokular (a), monokular (b) und Aperturbedingung (c). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.

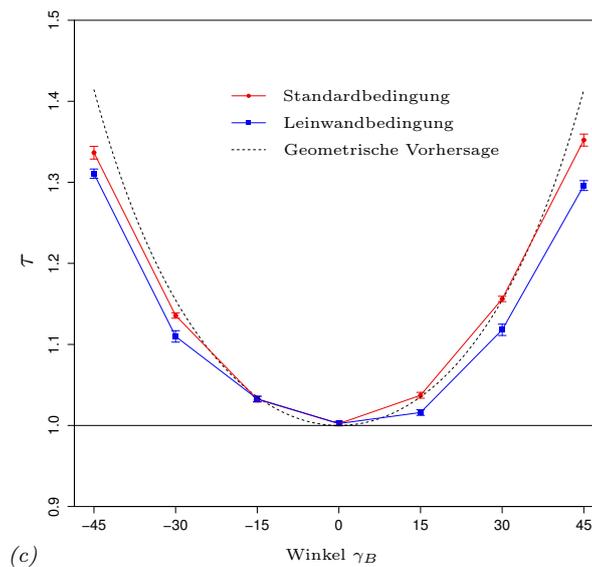
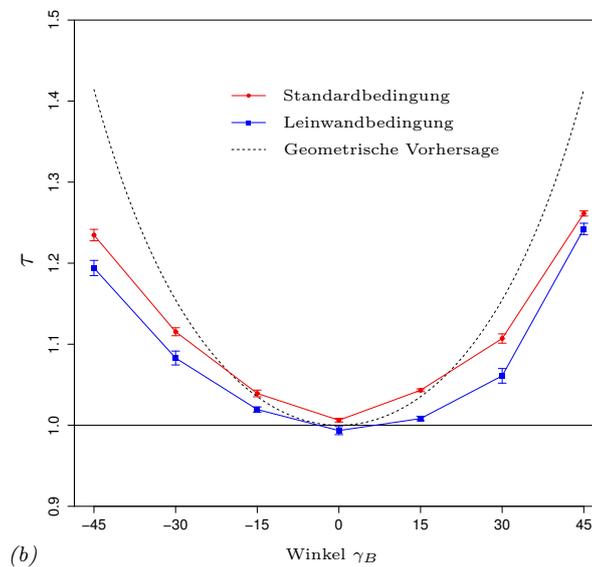
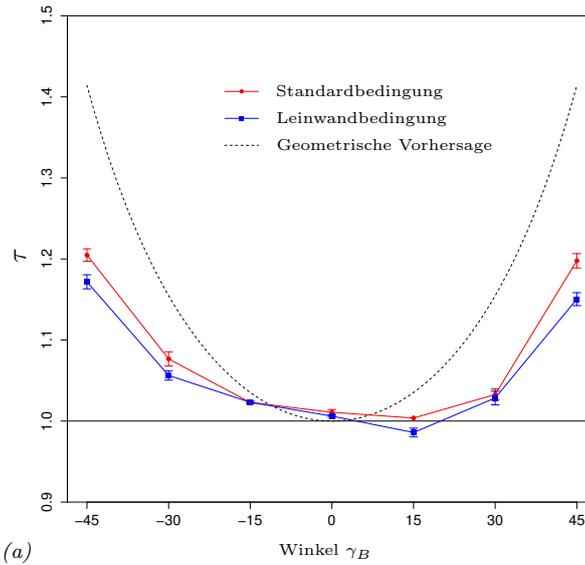


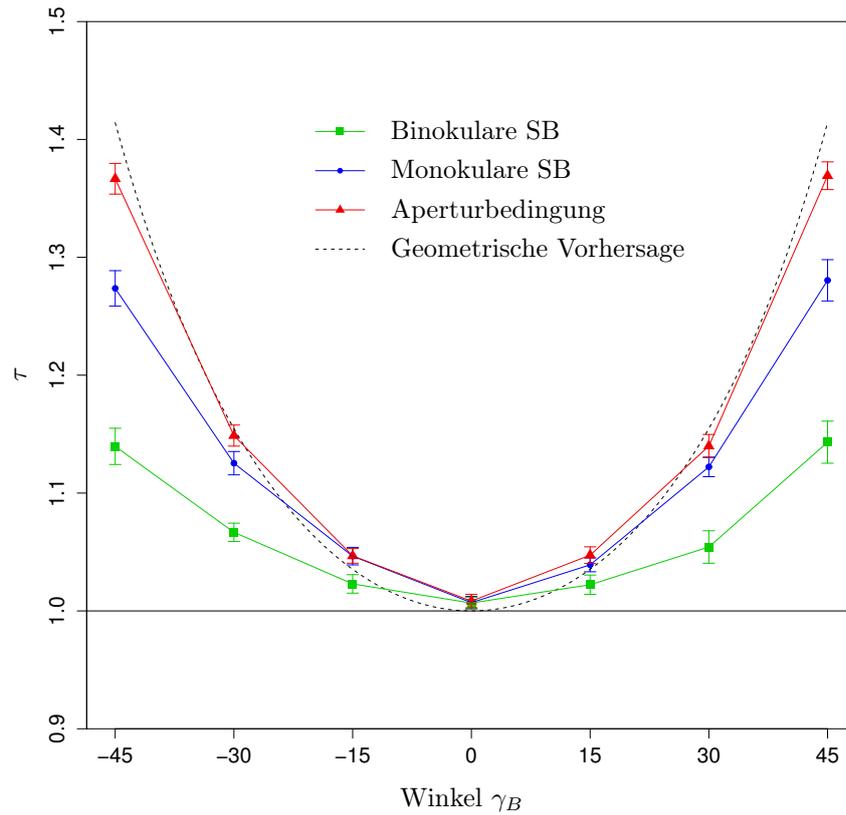
Abbildung 4.24. Mittlere  $\tau$ -Einstellungen des Probanden FK jeweils unter der Leinwand- und der Standardbedingung für die drei Sehbedingungen binokular (a), monokular (b) und Aperturbedingung (c). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.

und Leinwandbedingung, Abbildung 4.26 auf Seite 241 die über alle Probanden gemittelten mittleren  $\tau$ -Einstellungen getrennt für die verschiedenen Sehbedingungen.

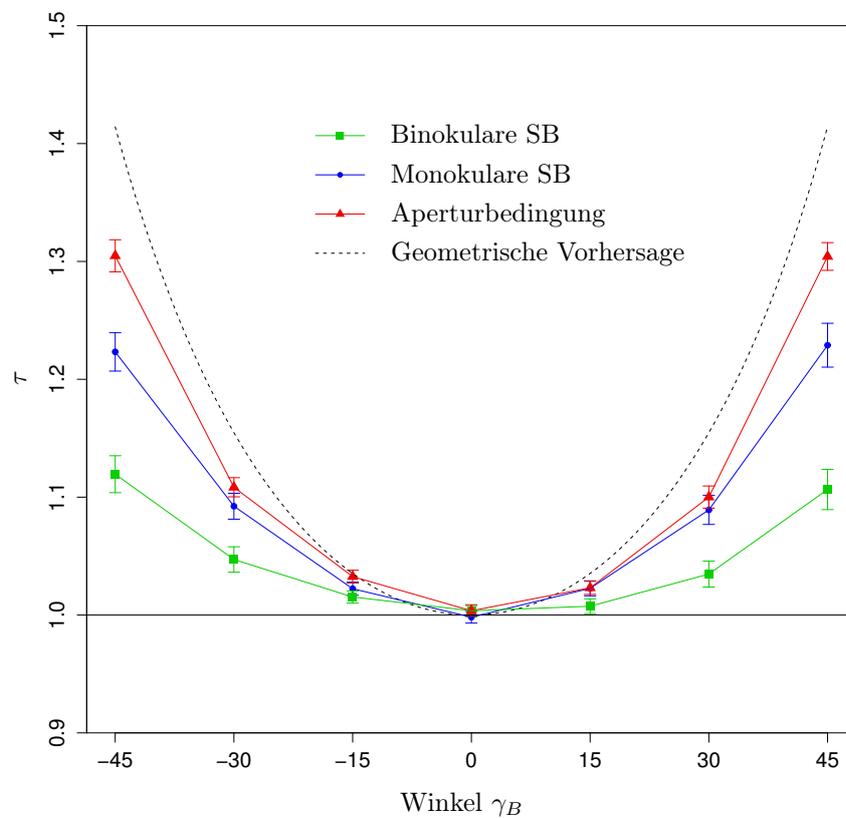
Anhand der aggregierten Daten tritt das bereits im Bezug auf die einzelnen Versuchspersonen Gesagte ebenfalls deutlich hervor, wir können daher an dieser Stelle auf eine Wiederholung verzichten. Um die Befunde bezüglich der Mittelwerte der mittleren Einstellungen von  $\tau$  abschließend graphisch prägnant zusammenzufassen, soll im Folgenden eine leicht andere Art der Darstellung erfolgen. Wir wollen zu diesem Zweck nun nicht mehr die mittleren  $\tau$ -Einstellungen betrachten, sondern deren Abweichung von Eins, d.h. von der bei 'vollständiger' Perspektivenrobustheit zu erwartenden Einstellung. Diese Abweichungen werden im Folgenden als  $\Delta$  bezeichnet. Weiterhin sollen die unter  $\gamma_B = 0$  erfolgten Einstellungen nicht weiter betrachtet werden, da diese in der Regel sehr nahe um  $\tau = 1$  lagen. Aufgrund der unter allen Seh- und Reizbedingungen beobachtbaren annähernden Symmetrie der Mittelwertsverläufe von  $\tau$  um  $\gamma_B = 0$  soll zudem lediglich der Betrag des Winkels  $\gamma_B$  betrachtet werden, d.h. es soll nicht mehr unterschieden werden, ob der Monitor im oder gegen den Uhrzeigersinn um den Winkel  $\gamma_B$  gedreht wurde. Abbildung 4.27 auf Seite 242 zeigt die über sämtliche Probanden gemittelten  $\Delta$ -Werte getrennt nach allen möglichen Kombinationen von Winkelbedingung  $|\gamma_B|$ , Sehbedingung und Reizbedingung in Form eines Balkendiagramms.

Zusammenfassend lässt sich in der vorliegenden Untersuchung folgendes Befundmuster beobachten:

1. *Effekt der Winkelbedingung:* Das Ausmaß der Perspektivenrobustheit ist unter jeder Seh- und Reizbedingung wesentlich abhängig vom Winkel  $\gamma_B$ , d.h. der horizontalen Abweichung der Position des Betrachters vom Augpunkt des präsentierten Bildes. Mit zunehmendem absoluten Winkel  $|\gamma_B|$  nehmen auch die von den Probanden vorgenommenen  $\tau$ -Einstellungen zu, was hier als abnehmende Perspektivenrobustheit mit zunehmender horizontaler Abweichung vom Augpunkt interpretiert werden kann. Dabei lässt sich kein Effekt des Vorzeichens von  $\gamma_B$  feststellen, die Verläufe der  $\tau$ -Einstellungen lassen sich als symmetrisch um die Einstellung bei  $\gamma_B = 0$  charakterisieren.
2. *Effekt der Sehbedingung:* Unabhängig von der Reizart lässt sich über alle Winkelbedingungen  $\gamma_B \neq 0$  hinweg feststellen, dass das Ausmaß der Perspektivenrobustheit unter der binokularen Sehbedingung am höchsten und unter der Aperturbedingung am geringsten ausgeprägt ist, wobei sich das Ausmaß der Perspektivenrobustheit unter der monokularen Sehbedingung in etwa dazwischen einordnen lässt. Die Größe dieses Unterschieds im Ausmaß der Perspektivenrobustheit zwischen den verschiedenen Sehbedingungen ist nicht über alle Winkelbedingungen  $\gamma_B \neq 0$  hinweg konstant, sondern nimmt mit zunehmender Abweichung von  $\gamma_B = 0$  zu.



(a)



(b)

Abbildung 4.25. Über alle Probanden gemittelte mittlere  $\tau$ -Einstellungen unter den verschiedenen Seh- und Winkelbedingungen in der Standardbedingung (a) und der Leinwandbedingung (b). Die halbe Länge der Fehlerbalken repräsentiert den jeweils zugehörigen Standardfehler.

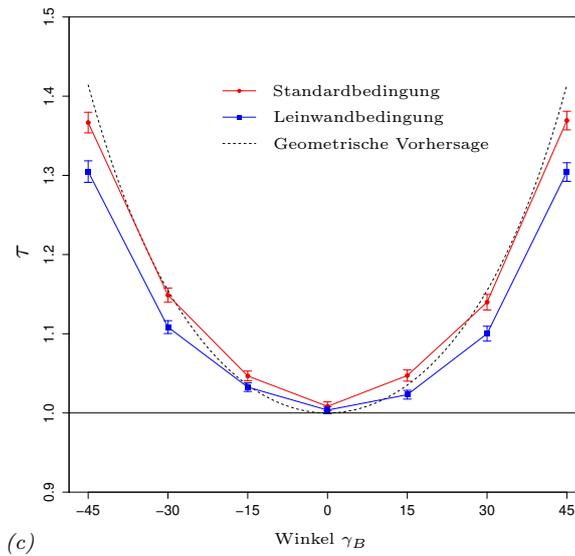
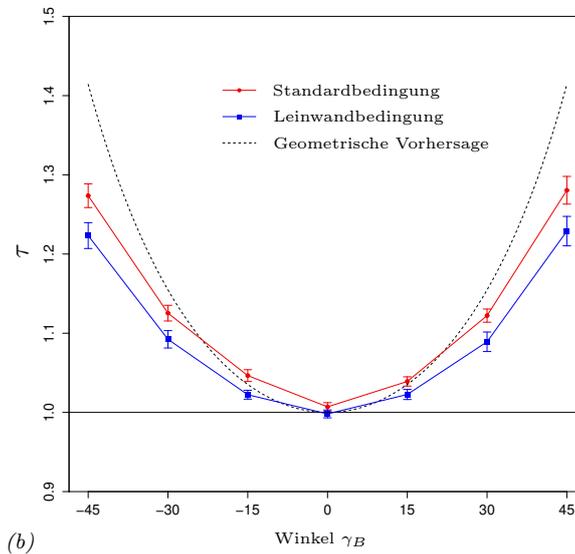
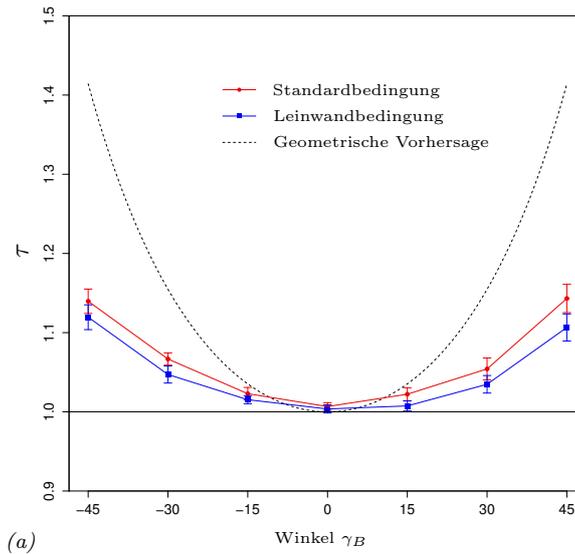


Abbildung 4.26. Über alle Probanden gemittelte mittlere  $\tau$ -Einstellungen unter der Leinwand- und der Standardbedingung getrennt dargestellt für die drei Sehbedingungen binokular (a), monokular (b) und Aperturbedingung (c). Die halbe Länge der Fehlerbalken repräsentiert den zugehörigen Standardfehler.

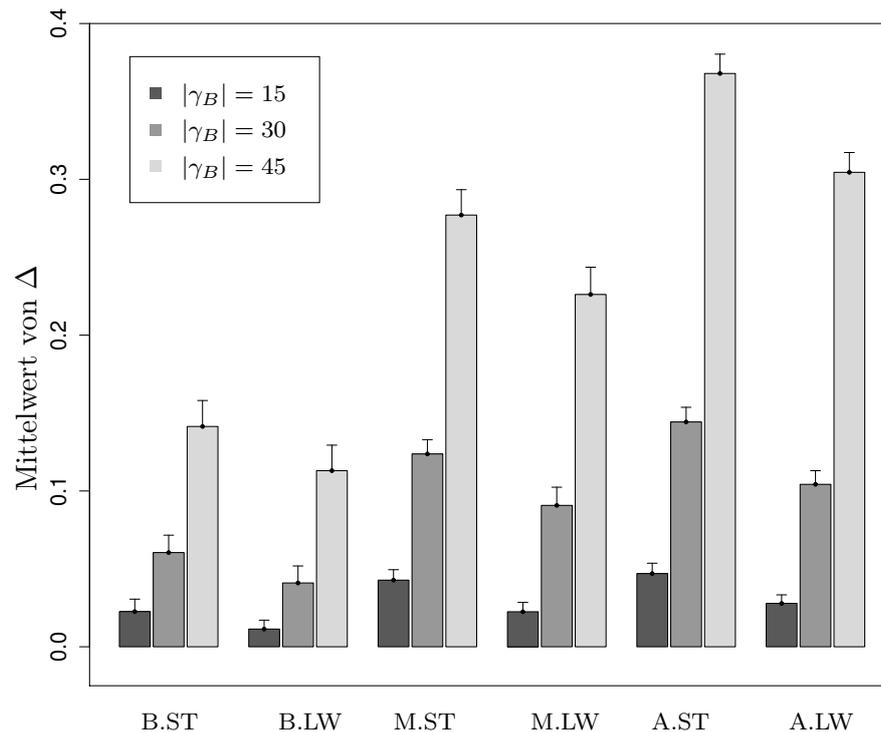


Abbildung 4.27. Die über sämtliche Probanden gemittelten Werte von  $\Delta$  unter allen möglichen Kombinationen der Ausprägungen der drei unabhängigen Variablen. Die erste Stelle der Bezeichnung unter den Balken gibt die Art der Sehbedingung an (B=binokular, M=monokular, A=Aperturbedingung), die Stelle nach dem Punkt in der Bezeichnung steht für die Art der Reizbedingung (ST=Standardbedingung, LW=Leinwandbedingung). Die Länge der Fehlerbalken repräsentiert den jeweiligen Standardfehler.

3. *Effekt der Reizart:* Unter jeder der drei Sehbedingungen lässt sich für alle Winkel  $\gamma_B \neq 0$  ein durchgängig höheres Ausmaß an Perspektivenrobustheit unter der Leinwandbedingung auffinden. Die Größe des Unterschieds zwischen Ausmaß der Perspektivenrobustheit unter der Standardbedingung und Ausmaß der Perspektivenrobustheit unter der Leinwandbedingung ist ebenfalls von der Winkelbedingung abhängig. Die Differenz zwischen den  $\tau$ -Einstellungen unter der Leinwand- und denen unter der Standardbedingung steigt unter jeder Sehbedingung mit zunehmendem absoluten Winkel  $|\gamma_B|$ .

## 4.6. Diskussion

In der vorgestellten Untersuchung sollte der Frage nachgegangen werden, inwieweit die mit der Betrachtung eines linearperspektivischen Bildes einhergehende wahrgenommene Robustheit der Perspektive von i.) der Abweichung vom Augpunkt des Bildes, ii.) der Sehbedingung (binokular, monokular, Aperturbedingung) und iii.) einer Manipulation des ‘Oberflächencharakters’ des Bildes abhängig ist. Dabei waren insbesondere die mit der Manipulation des Oberflächencharakters verbundenen Effekte von Interesse,

da eine solche Variation in bisherigen Untersuchungen zu diesem Bereich nicht vorgenommen wurde. Das Ausmaß der Perspektivenrobustheit sollte in dieser Untersuchung als in umgekehrt proportionaler Beziehung stehend zur Größe der von den Probanden vorgenommenen  $\tau$ -Einstellungen interpretiert werden. Hintergrund dieser Interpretation ist die (geometrische) Überlegung, dass Probanden bei ‘vollständiger Perspektivenrobustheit’ unter sämtlichen Blickwinkeln eine Ellipse mit einem Hauptachsen-Nebenachsenverhältnis von  $\tau \approx 1$  als zentralperspektivische Darstellung einer Kugel wahrnehmen sollten.

Insgesamt können die mit den experimentellen Manipulationen verbundenen *erwarteten* Effekte als konform mit den empirischen Ergebnissen angesehen werden. Wir wollen zunächst die mit der Winkelbedingung und der Sehbedingung verbundenen Effekte diskutieren. Das von Vishwanath et al. (2005) berichtete Ergebnismuster konnte im Wesentlichen repliziert werden, wobei in der vorliegenden Untersuchung in deutlich höherem Umfang Daten erhoben und in die Analyse einbezogen wurden. Es zeigte sich auch in der vorliegenden Untersuchung unter jeder Sehbedingung ein abnehmendes Ausmaß an Perspektivenrobustheit bei zunehmender Abweichung vom Augpunkt des Bildes. Über alle Winkel hinweg betrachtet waren zudem die verschiedenen Sehbedingungen mit verschieden stark ausgeprägter Perspektivenrobustheit verbunden. Man erinnere sich, dass sowohl die Hypothesen der Untersuchung von Vishwanath et al. (2005), als auch die der vorliegenden Untersuchung durch Rückgriff auf das Konzept der Tiefencues begründet worden sind.

Im Rahmen der Zielsetzung dieser Arbeit soll das vorliegende Experiment und die erhaltenen Ergebnisse wesentlich als ein Medium zur Verdeutlichung der in den vorangehenden Kapiteln herausgearbeiteten theoretischen Schwierigkeiten, die mit dem Konzept der Tiefencues verbunden sind, genutzt werden. Aus der Sicht des *cue*-Ansatzes würde man zur Erklärung der vorliegenden empirischen Befunde in dieser oder einer ähnlichen Weise wie folgt argumentieren: Die allgemein im Experiment festgestellte Abnahme der Perspektivenrobustheit bei zunehmender Abweichung vom Augpunkt des präsentierten Bildes ist auf die zunehmende systematische ‘Verzerrung’ des auf die Retina(e) projizierten Lichtmusters zurückzuführen (siehe die geometrischen Erörterungen in Abschnitt 4.2). Das visuelle System bezieht die Ausdehnung der Projektion der Ellipse auf die Retina(e) als ‘Input’ in die weiteren Schritte der Informationsverarbeitung ein. Wenn das visuelle System die Größen  $\alpha$  und  $\beta$  bzw. das Verhältnis  $\alpha/\beta$  generell als ‘Information’ über das im Bild dargestellte Objekt ‘nutzt’, dann sollten aufgrund der weiter oben diskutierten *geometrischen* Beziehungen zwischen  $\alpha/\beta$ ,  $\tau$  und  $\gamma_B$  bei zunehmender Abweichung vom Augpunkt des Bildes auch die  $\tau$ -Einstellungen der Probanden tendentiell zunehmen, damit die dargestellte Ellipse als eine Kugel wahrgenommen wird.

Wie lässt sich das in Abhängigkeit von der experimentellen Manipulation der Sehbedingung deutlich unterschiedliche Ausmaß der Perspektivenrobustheit im Rahmen des *cue*-Ansatzes erklären? Wir halten zunächst fest, dass die verschiedenen Sehbedingungen mit unterschiedlich starken ‘*cue*-Konflikten’ verknüpft sind. Die mit der zentralperspektivischen Darstellung verbundenen monokularen Tiefencues wie Texturgradient, Schattenwurf und Verdeckung sind dem visuellen System als ‘Informationen über Tiefe’ unter jeder Sehbedingung in gleicher Weise verfügbar. Diese haben, für sich genommen, einen mit der Bildbetrachtung verbundenen Tiefeneindruck zur Folge. Unter der binokularen Sehbedingung stehen dem visuellen System jedoch auch andere Tiefencues, wie Konvergenz und binokulare Disparität, zur Verfügung, die darauf ‘hindeuten’, dass es sich bei dem umrahmten Bild um eine bearbeitete Fläche, einen Gegenstand im physikalischen Raum handelt, das unter  $\gamma_B \neq 0$  als nicht frontoparallel zum Betrachter wahrgenommen wird. Diese zusätzlichen Hinweise oder Informationen – so die Vorstellung – werden nun vom visuellen System genutzt, um die systematischen Verzerrungen zu ‘kompensieren’ bzw. anhand des über den wahrgenommenen Monitorrahmen vom visuellen System ‘berechneten’ oder ‘geschätzten’ Blickwinkels ‘herauszurechnen’.<sup>22</sup> Dies führt zu einer im Vergleich zu den anderen Sehbedingungen generell stärker ausgeprägten Perspektivenrobustheit unter der binokularen Sehbedingung. Da in der monokularen Sehbedingung dem visuellen System die Tiefencues der Konvergenz und binokularen Disparität nicht zur Verfügung stehen, sind daher weniger ‘Informationen’ verfügbar, die durch das visuelle System genutzt werden könnten, um die durch den exzentrischen Betrachterstandpunkt bestimmte Verzerrung ‘herauszurechnen’. Dadurch kommt es zu einer im Vergleich zur binokularen Sehbedingung ‘geringeren Kompensation’ der Verzerrungen durch das visuelle System und weniger stark ausgeprägter Perspektivenrobustheit. Unter der Aperturbedingung kann zudem der sichtbare Rahmen nicht mehr als Information genutzt werden, daher fällt unter dieser Bedingung die ‘Kompensation’ am geringsten aus, was zu einer im Vergleich am geringsten ausgeprägten generellen Perspektivenrobustheit führt.

Der in der vorliegenden Untersuchung zusätzlich und zum ersten Mal aufgefundene Effekt der Reizart (Standardreiz/Leinwandreiz) ließe sich ebenfalls in dieser mit dem Konzept der Tiefencues verbundenen Argumentationslinie interpretieren. Über sämtliche Winkel  $\gamma_B \neq 0$  hinweg zeigte sich insgesamt, sowie unter jeder einzelnen Sehbedingung, eine stärker ausgeprägte Perspektivenrobustheit, wenn die linearperspektivischen Bilder einen hervorgehobenen Oberflächen- oder Leinwandcharakter aufwiesen. Diese hervorgehobene Flächenhaftigkeit könnte man nun ebenfalls als einen *cue* ansehen, der im Rahmen eines *cue*-Integrationsprozesses als zusätzliche Information vom visuellen

---

<sup>22</sup>Die mit einer solchen Kompensationshypothese verbundenen grundsätzlichen Schwierigkeiten wurden bereits in Abschnitt 4.1.2 diskutiert.

System genutzt werden kann. Ein hervorgehobener Leinwandcharakter könnte (wie die Konvergenz, die binokulare Disparität oder die Sichtbarkeit des Monitorrahmens) zusätzlich darauf ‘hindeuten’, dass es sich bei dem Betrachteten um eine in bestimmter Weise bearbeitete Fläche im physikalischen Raum handelt. Diese zusätzliche Information könnte so zu einer unter allen Sehbedingungen stärkeren ‘Kompensation’ der retinalen Verzerrungen durch das visuelle System und damit zu einer generell stärkeren Perspektivenrobustheit führen.

Die empirischen Befunde lassen sich auch in Beziehung setzen zu dem in Abschnitt 4.1.3 behandelten Konzept der Dualität der Bildwahrnehmung. Dieses Konzept selbst impliziert nicht die Vorstellung, dass das visuelle System auf *cues* operiere. Wenn man aber die Auffassung, dass die Bildbetrachtung mit einem dualen Charakter des Wahrnehmungseindrucks verbunden ist, hat, und nun versucht, den *cue*-Ansatz in diese zu integrieren, so liegt es nahe, die mit dieser Vorstellung verknüpften ‘Teilperzepte’ für den planaren und räumlichen Aspekt bei der Bildwahrnehmung mit unterschiedlichen *cues* zu verknüpfen, die man auch als *Tiefencues* und *Flächencues* bezeichnen könnte. Verfolgt man diesen Ansatz, so könnte man in Hinsicht auf die vorliegenden Ergebnisse argumentieren, dass die verschiedenen Kombinationen der Sehbedingung und der Reizart mit jeweils verschiedenen Prozessen der *cue*-Segregation einhergehen, die wiederum unterschiedliche ‘Gewichtungen’ des räumlichen und planaren Aspekts des mit der Bildbetrachtung verbundenen Wahrnehmungseindrucks zur Folge haben. Unter der binokularen Sehbedingung mit Leinwandreiz sollte dann aufgrund der verfügbaren, auf eine Fläche ‘hindeutenden’ Flächencues der planare Aspekt am prägnantesten, unter der Aperturbedingung mit Standardreiz hingegen am wenigsten prägnant sein. Bezüglich des räumlichen Aspekts sollte es sich genau umgekehrt verhalten. In dieser Vorstellung könnte der hervorgehobene Leinwandcharakter des präsentierten Bildes als ein solcher Flächencue angesehen werden, der die Prägnanz des räumlichen Aspekts unter allen Sehbedingungen abschwächt bzw. die Prägnanz des planaren Aspekts erhöht. In dieser Argumentation sollte somit die Manipulation der Sehbedingung und die Manipulation der Reizart unter jeder Sehbedingung mit verschieden prägnanten *Tiefeneindrücken* einhergehen.

Aus einer forschungsmethodischen Perspektive lassen sich bezüglich der vorliegenden Untersuchung noch drei Punkte diskutieren. Erstens liegt dieser Untersuchung die Idee zugrunde, dass das Phänomen der Perspektivenrobustheit über die von den Probanden vorgenommenen  $\tau$ -Einstellungen quantifiziert erfasst werden könne. Von einer eher phänomenologischen Haltung aus könnte man den Standpunkt vertreten, dass diese Vorgehensweise mögliche subtile Unterschiede der mit der Betrachtung linearperspektivischer Bilder außerhalb des Augpunkts verbundenen Wahrnehmungseindrücke von vornherein ausschließt, indem Perspektivenrobustheit als *ein* einheitliches Phä-

nomen behandelt wird, das lediglich in verschiedenen (quantitativen) Ausprägungen vorliegen kann. Möglicherweise führt jedoch die in dieser Untersuchung vorgenommene Manipulation der Reizart oder gar die Manipulation der Sehbedingung zu *qualitativ* verschiedenen Wahrnehmungseindrücken oder *Erscheinungsweisen der Bilder* in einem ähnlichen Sinn, in dem Katz (1930) von verschiedenen „Erscheinungsweisen der Farben“ spricht. Als extremes Beispiel solcher unterschiedlichen Erscheinungsweisen könnte man *trompe l'œil*-Bilder wie das in Abschnitt 4.1.3 angesprochene Pozzosche Deckenfresco heranziehen. Bewegt man sich hier (in möglichst kleinen Schritten) aus dem Augpunkt heraus, so scheinen im Vergleich zur Betrachtung eines ‘üblichen’ linearperspektivischen Bildes (d.h. eines Bildes, das eine Tiefenwahrnehmung hervorruft, die sich nicht konsistent in den Realraum des Betrachters einordnen lässt) die bewegungskorrelierten Veränderungen qualitativ von ganz eigener Art zu sein, so dass es unangebracht wäre, lediglich von stark oder schwach ausgeprägter Perspektivenrobustheit zu sprechen.

Der zweite Punkt betrifft das Vorgehen selbst. Den Probanden wurden linearperspektivisch korrekt konstruierte Bilder außerhalb des Augpunktes dargeboten. Über die von ihnen abgegebenen Urteile konnte die horizontale Ausdehnung des kritischen Reizes (der Ellipse) verändert werden, um so das Ausmaß der Perspektivenrobustheit zu erfassen. Die abgegebenen Urteile veränderten jedoch allein den kritischen Reiz, die restlichen Elemente des Bildes blieben *unverändert*. Bei sehr stark vom Augpunkt abweichenden Blickwinkeln sollte dann aber eine im Vergleich zur üblichen Bildbetrachtung phänomenal sehr ungewöhnliche Sehsituation auftreten, nämlich dass über die abgegebenen Urteile der kritische Reiz als unverzerrt (als eine Kugel) wahrgenommen werden sollte, die restlichen Elemente des Bildes (etwa die Würfel im Vordergrund) hingegen als *verzerrt*. Bei der üblichen Bildbetrachtung außerhalb des Augpunkts bezieht sich die Perspektivenrobustheit auf das *gesamte Bild*, in der vorliegenden Untersuchung auf ein Element innerhalb einer linearperspektivischen Darstellung einer Szene.

Drittens wäre es interessant, der Frage nachzugehen, welchen Einfluss unterschiedliche Formulierungen der Instruktion auf das Verhalten der Probanden hat. Die Formulierung der Instruktion stellt in wahrnehmungspsychologischen Urteilsaufgaben einen kritischen Faktor dar, unterschiedliche Formulierungen können unterschiedliche Urteile seitens der Probanden zur Folge haben (siehe etwa Gilinsky, 1955). In der vorliegenden Untersuchung sollten die Probanden bei jedem präsentierten Bild entscheiden, ob ihnen der kritische Reiz als ‘zu schmal’ oder ‘zu breit’ für eine *Kugel* erscheint. Welche ‘Kriterien’ die Probanden letztendlich herangezogen haben, um ein entsprechendes Urteil abzugeben, lässt sich den empirischen Befunden natürlich nicht entnehmen. Es ist jedoch nicht ausgeschlossen, dass die Probanden möglicherweise andere ‘Kriterien’ herangezogen hätten, wenn sie stattdessen entscheiden sollten, ob der kritische Reiz zu schmal oder zu breit für einen *Kreis* erscheint. Diese Formulierung der Instruktion

könnte möglicherweise zur Folge haben, dass die Probanden, etwas vage ausgedrückt, von vornherein die präsentierten Bilder ‘stärker’ als bearbeitete Oberflächen und die Elemente des Bildes ‘stärker’ als zweidimensionale geometrische Formen ‘ansehen’, als bei der hier verwendeten Instruktion. Ob sich dies auch in anderen Verläufen der  $\tau$ -Einstellungen als den hier berichteten ausdrücken würde, könnte Gegenstand einer weiteren Untersuchung sein. Generell ließen sich interessante Erweiterungen der vorliegenden Untersuchung vorstellen. So könnte man etwa eine zusätzliche Bedingung einführen, in der die ‘Anzahl’ der in der linearperspektivischen Darstellung ‘enthaltenen’ Tiefencues (Verdeckung, Schattenwurf etc.) systematisch variiert wird. Man könnte den Bildern gleichsam Stück für Stück die einzelnen Tiefencues ‘hinzufügen’ und beobachten, ob diese Manipulationen mit unterschiedlichen Verläufen der  $\tau$ -Einstellungen einhergehen. In einer weiteren Variation könnte man den kritischen Reiz nicht in der Mitte des Monitordisplays, sondern in unterschiedlicher Exzentrizität darbieten und die Probanden das Bild stets aus dem Augpunkt heraus betrachten lassen.

Wir wollen uns nun wieder auf das im Zentrum dieser Arbeit stehende Thema besinnen und abschließend die Frage aufwerfen, ob die oben präsentierten, auf dem *cue*-Konzept basierenden Interpretationen der empirischen Befunde der vorliegenden Untersuchung überzeugen können. Zumindest scheinen sie den Vorteil einer gewissen ‘Griffigkeit’ zu haben, da die präsentierten empirischen Befunde über ein Konzept (in Verbindung mit der Vorstellung eines ‘Kompensationsmechanismus’) verständlich gemacht werden sollen. Die empirischen Ergebnisse bestehen: Die in der vorliegenden Untersuchung vorgenommenen experimentellen Manipulationen führten in der Tat zu den erwarteten, systematisch und deutlich verschiedenen Verläufen der von den Probanden vorgenommenen  $\tau$ -Einstellungen. Die präsentierten Befunde können zudem neue und relevante empirische Erkenntnisse für den Forschungsbereich der Bildwahrnehmung liefern, da sie deutlich darauf hinweisen, dass eine in dieser Untersuchung zum ersten Mal vorgenommene Manipulation des Leinwandcharakters eines Bildes zu systematischen Veränderungen der Urteile der Probanden führt. Trotz des nicht bestreitbaren *heuristischen Wertes* des *cue*-Konzepts für die Formulierung neuer experimenteller Fragestellungen sind den unter Rückgriff auf dieses Konzept gegebenen Erklärungen wegen der in dem vorangehenden Kapitel diskutierten theoretischen (begriffslogischen) Schwierigkeiten doch enge Grenzen gesetzt. Sie können nur als explanatorische Notbehelfe angesehen werden. Mehr noch: Diese auf den ersten Blick einsichtig erscheinenden Erklärungen verschleiern die konzeptuellen Probleme und behindern dadurch meines Erachtens eine ernsthaftere Auseinandersetzung mit ihnen.

Wir wollen diese Probleme in direktem Bezug zu den hier angeführten Erklärungen der empirischen Befunde nochmals verdeutlichen. Zunächst scheint man sich in der vorgeschlagenen Interpretation des Effekts der Winkelbedingung nicht von der Analo-

gie des ‘Netzhautbildes’ lösen zu können. Davon zu sprechen, dass auf den Retina(e) Projektionen der dargestellten Ellipse mit beschreibbarer Ausdehnung dem visuellen System zur weitergehenden Analyse gleichsam ‘vorliegen’, ist irreführend, da eine solche Beschreibung voraussetzt, dass das retinale Lichtmuster *als Bild gesehen* wird, in dem sich bestimmte Elemente *als* Projektionen bestimmter *Objekte* ‘identifizieren’ lassen. Diese Beschreibung ist mit dem in Abschnitt 3.3 behandelten Problem des Homunkulismus behaftet. Im Kontext des komputationalen Ansatzes lässt sich dieses Problem leicht umformulieren; hier müssten die mit der Absorption der Lichtquanten durch die Photorezeptoren einhergehenden neuronalen Aktivitätsmuster vom visuellen System *als* Repräsentationen der betrachteten Szene bzw. der in dieser enthaltenen Objekte und deren räumlicher Anordnung ‘interpretiert’ und im Rahmen der weiteren Informationsverarbeitung als solche ‘verwendet’ werden. In Abschnitt 3.3 wurde aufgrund der Überlegungen zum Konzept der internen Repräsentationen argumentiert, dass diese Beschreibung des visuellen Systems ebenfalls als homunkulare Redeweise zu kennzeichnen ist. Die vorgeschlagenen Interpretationen zum Effekt der Sehbedingung und der Reizart sind in gleicher Weise mit diesem Problem behaftet, da hier über die postulierten *cue*-Integrationsprozesse auf den komputationalen Ansatz verwiesen wird. Selbst wenn man von diesen Problemen absehen würde, könnte dieser Ansatz den Zusammenhang zwischen den behaupteten Informationsverarbeitungsprozessen durch das visuelle System und unserer Wahrnehmung dennoch nicht verständlich machen; eine Schwierigkeit, auf die im Verlauf dieser Arbeit ebenfalls mehrfach hingewiesen wurde. Die in der Regel anzutreffenden Beschreibungen, etwa dass das visuelle System das Bild schlicht ‘interpretiere’ (Vishwanath et al., 2005), ein Perzept ‘generiere’ oder das Ergebnis des Informationsverarbeitungsprozesses eine symbolische Beschreibung der Dinge der Außenwelt sei, tragen hier nicht zur Erhellung bei.

Hinsichtlich des Vorschlags, den Oberflächencharakter des Bildes als einen *cue* anzusehen, lässt sich eine weitere Bemerkung machen, die sich in Übereinstimmung mit den in den vorangegangenen Kapiteln vorgenommenen Erörterungen bringen lässt. Die Beschreibung des Oberflächencharakters eines Bildes ist bereits der Wahrnehmung entnommen, ein Punkt, der insbesondere in der Behandlung der Verdeckung oder der linearperspektivischen Darstellung als Tiefencues angesprochen wurde. Der Oberflächencharakter eines Bildes ist (in dem dargestellten Sinn) kein Aspekt des physikalischen Gegenstandes bzw. des in dieser Untersuchung präsentierten Reizes, er ist – wie die Verdeckung – ein Aspekt *unserer Wahrnehmung*. In einer rein physikalischen Beschreibung des präsentierten Bildes gibt es weder sich teilweise verdeckende Objekte, noch einen hervorgehobenen Oberflächencharakter, dies sind Beschreibungen auf Ebene der Wahrnehmung des präsentierten Reizes. Sie sind auf das *Ergebnis* des Wahrnehmungsprozesses bezogen, werden aber gleichzeitig herangezogen, um das Zustandekommen

---

dieses Ergebnisses verständlich zu machen. Man scheint sich in diesem Erklärungsansatz gleichsam im Kreis zu bewegen. Vollkommen unsinnig scheint dieser Ansatz dennoch nicht zu sein, denn empirisch zeigt sich doch, dass bestimmte experimentelle Manipulationen mit bestimmten systematischen Veränderungen einhergehen. Um das Zitat am Anfang dieses Kapitels aufzugreifen: Die experimentelle Methode tut *etwas*. Man steht jedoch bezüglich der Interpretation dieser empirischen Ergebnisse offenbar vor immensen theoretischen Schwierigkeiten, die wiederum mit tiefen philosophischen Problemen in Verbindung zu stehen scheinen, auf die wir in dieser Arbeit an den entscheidenden Stellen immer wieder gestoßen sind.



## 5 | Fazit

„In the eye we have on the one hand light falling on this wonderful structure, and on the other hand we have the sensation of sight. We cannot compare these two things; they belong to opposite categories. The whole of metaphysics lies like a great gulf between them.“

---

*Maxwell (1857/1890, S. 227)*

Es sollen nun die in den einzelnen Kapiteln der vorliegenden Arbeit gewonnen Erkenntnisse vor dem Hintergrund der in der Einleitung aufgeworfenen Problemstellung auf prägnante Weise zusammengeführt und abschließend diskutiert werden. Das übergeordnete Ziel dieser vorwiegend theoretisch und konzeptuell orientierten Arbeit sollte es sein, das in aktuellen wahrnehmungstheoretischen Vorstellungen zur Tiefenwahrnehmung zentrale Konzept der Tiefencues einer kritischen Beleuchtung zu unterziehen. Die Notwendigkeit einer theoretische Analyse dieses Konzeptes wurde darin gesehen, dass einerseits das Konzept der Tiefencues in der gegenwärtigen Wahrnehmungsforschung eine außerordentlich zentrale Rolle spielt, andererseits aber dort eine eher unbefriedigende Enthaltbarkeit in Hinblick auf eine konzeptuell-logische Klärung des Status dieses Konzeptes zu konstatieren ist.

In dem umfangreichsten Teil dieser Arbeit wurde zunächst untersucht, ob sich in früheren theoretischen Ansätzen Antworten finden lassen. Es wurde die Entwicklung des Konzepts der Tiefencues und seiner Verwendungsweisen anhand ausgewählter historischer wahrnehmungstheoretischer Ansätze, in denen dem heutigen Konzept der Tiefencues funktional analoge oder gar äquivalente Konzepte eine zentrale Rolle einnehmen, dargestellt und im Rahmen jeder einzelnen Theorie erörtert, welche impliziten Annahmen diesem Konzept zugrundeliegen und mit welchen theoretischen Problemen diese Annahmen verbunden sind.<sup>1</sup> Zudem sollte geklärt werden, ob signifikante Veränderungen in den theoretischen Ansätzen auch signifikante Entwicklungen oder Veränderungen der mit diesem Konzept verbundenen Vorstellungen zur Folge hatten, oder nicht. Da am Ende eines jeden Unterkapitels des historischen Teils dieser Arbeit nach

---

<sup>1</sup>Wir wollen im Folgenden aus Gründen der Einfachheit die den Tiefencues funktional analogen Konzepten der historischen Ansätze ebenfalls als Tiefencues bezeichnen.

der Darstellung des jeweiligen historischen Ansatzes bereits eine Zusammenfassung und kritische Diskussion erfolgte, sollen hier weniger die innerhalb der einzelnen Theorien herausgearbeiteten Punkte im Vordergrund stehen, sondern jene, die sich im Verlauf der historischen Entwicklung des Konzepts der Tiefencues herauskristallisiert haben.

In sämtlichen dargestellten Ansätzen wird die Tiefenwahrnehmung als ein in besonderer Weise erklärungsbedürftiges Phänomen angesehen. Der Grund für diese Erklärungsbedürftigkeit wird spätestens seit Descartes' wahrnehmungstheoretischem Ansatz (bei Alhazen deutet sich dies allerdings schon an) darin gesehen, dass, in moderner Formulierung, das von den Objekten der Außenwelt reflektierte Licht in durch optische Gesetzmäßigkeiten beschreibbarer Weise auf die als zweidimensional zu kennzeichnenden Retinae der Augen projiziert wird und wir dennoch Objekte in räumlicher Ausdehnung und bestimmter Lage im Raum wahrnehmen. Ein bloßes 'Empfangen' des auf die Augen treffenden Lichts, des 'visuellen Stimulus', ist laut diesen Ansätzen für sich genommen nicht hinreichend, um die Reichhaltigkeit der zugehörigen Wahrnehmung erklären zu können. Wahrnehmung müsse daher als ein grundsätzlich *mehrstufiger Prozess* aufgefasst werden, in dessen Verlauf, ganz allgemein gesprochen, der gleichsam 'verarmte' visuelle Stimulus in einem dem Wahrnehmenden selbst nicht zugänglichen Prozess 'interpretiert' werde und so die Kluft zwischen diesem und dem Sehen räumlich ausgedehnter Gegenstände überbrückt werden kann. Im spezifischen Rahmen der Tiefenwahrnehmung findet man häufig Beschreibungen der Art, dass 'Räumlichkeit' oder 'die dritte Dimension' der Außenwelt aus den zweidimensionalen Projektionen 'erschlossen' oder 'konstruiert' werde müsse. Dies sei nur über bestimmte Hilfsmittel, den heute sogenannten Tiefencues, möglich.

Erste, dem heutigen Konzept der Tiefencues funktional ähnliche Vorstellungen konnten in der Theorie des Alhazen aufgezeigt werden. Die Vorstellung bestimmter, Tiefenwahrnehmung ermöglichender Hilfsmittel weist somit eine knapp tausendjährige Geschichte auf. Alhazen ging davon aus, dass Wahrnehmung, auch die Wahrnehmung von Tiefe, über einen Prozess des (syllogistischen) *Schlussfolgerns* geschehe, der automatisiert ablaufe und daher dem Wahrnehmenden nicht bewusst sei. Die Auffassung, dass Wahrnehmung erst durch gleichsam 'intellektuelle Fähigkeiten' wie Berechnungen vornehmen, Urteile fällen, Schlussfolgern oder Interpretieren möglich werde, zieht sich (mit gewisser Ausnahme des Ansatzes von Berkeley) wie ein roter Faden durch die weitere Entwicklung theoretischer Vorstellungen zur Tiefenwahrnehmung. Insbesondere in Helmholtz' Vorstellung von visueller Wahrnehmung als das Ergebnis eines *unbewussten Schlusses* tritt dieser Zug dann überaus deutlich in Erscheinung.

Im Verlauf der Entwicklungsgeschichte des Konzepts der Tiefencues zeigte sich erstens, dass die Anzahl der Tiefencues schrittweise zunahm, nicht selten durch neu gewonnene anatomische Erkenntnisse (etwa im Fall der Akkommodation) oder neue tech-

nische Entwicklungen (etwa des Stereoskops) bedingt. Dies führte zu einer im Hinblick auf ihre angenommenen Wirkungsweisen ziemlich heterogenen Gruppe von Elementen. Bereits bei Descartes lässt sich eine solche Auflistung der Tiefencues ausmachen, die ganz unterschiedlichen Beschreibungsebenen entstammen und deren jeweilige ‘Funktionsweise’ ganz unterschiedlicher Natur ist. Die Akkommodation als Tiefencue wird in rein physiologischen Begriffen beschrieben; die Fixation unterschiedlich weit entfernter Objekte führt zu unterschiedlicher Linsenkrümmung und diese könne als ‘Information’ über die Entfernung des gesehenen Objektes genutzt werden. Die Konvergenz als Tiefencue wird aufgrund *abstrakter geometrischer Überlegungen* als ein Tiefencue angesehen; die Entfernung eines beidäugig fixierten Objektes könne unter Ausnutzung geometrischer Gesetzmäßigkeiten gleichsam ‘berechnet’ werden. Der dritte von Descartes angeführte Tiefencue, die ‘erscheinende Größe’ eines Objektes und deren ‘Vergleich’ mit der ‘tatsächlichen Größe’ ist wiederum ganz anderer Natur, da in dieser Vorstellung die Annahme enthalten ist, dass das Objekt bereits *wahrgenommen* sein muss. Mit Ausnahme der Berkeleyschen Theorie, die in erster Linie eine Kritik der geometrischen Auffassung der Konvergenz zum Gegenstand hat und Tiefenwahrnehmung allein über den Tastsinn zu erklären sucht, lässt sich eine solche Liste bei allen nachfolgenden Theorien auffinden.

Zweitens ist festzustellen, dass die Tiefencues über die verschiedenen theoretischen Ansätze hinweg im Wesentlichen *tradiert* werden. Ein einmal eingeführter Tiefencue taucht auch in sämtlichen nachfolgenden Ansätzen auf, in der Regel ohne prägnante Veränderungen zu erfahren.<sup>2</sup> Ein gutes Beispiel stellt hier der zuerst explizit von Descartes eingeführte Tiefencue der ‘erscheinenden Größe’, oder, wie er später bezeichnet wird, der ‘retinalen Größe’ eines Objekts dar, der in ähnlicher Beschreibung bei Reid, Müller, Lotze und Helmholtz anzutreffen ist. Dies ist bemerkenswert, da sich außerhalb dieses Konzeptes doch deutliche Veränderungen in der allgemeinen Ausrichtung der theoretischen Ansätze zeigen, etwa die von Reid eingeführte Unterscheidung zwischen Empfindung und Wahrnehmung, die zunehmend physiologisch bestimmte theoretische Orientierung im 19. Jahrhundert, erste Kodierungsvorstellungen im Fall der *Lokalzeichen* Lotzes oder die innovative Zusammenführung verschiedener Elemente vorhergehender theoretischer Ansätze in der Zeichentheorie von Helmholtz.

Trotz der Zunahme der Anzahl der Tiefencues und trotz prägnanter Entwicklungen in der Theoriebildung hat das *cue*-Konzept selbst insgesamt wenig Veränderung und insbesondere auch keine zunehmende theoretische Erhellung erfahren. In sämtlichen dargestellten theoretischen Ansätzen ist dieses Konzept letztlich immer mit ähnlichen Schwierigkeiten in leicht unterschiedlicher Ausgestaltung verbunden. Als zentrales Pro-

---

<sup>2</sup>Eine diesbezügliche Ausnahme stellt die Konvergenz dar, die von Descartes zu Berkeley eine merkliche Neuinterpretation erfahren hat.

blem dieser historischen Ansätze wurde herausgestellt, dass der Status der Tiefencues innerhalb der Theorien weitgehend unklar ist. In den Beschreibungen ihrer angenommenen Funktionsweise wird deutlich, dass die Tiefencues bereits einen Wahrnehmungsakt voraussetzen, aber gleichzeitig zur Erklärung der Wahrnehmung herangezogen werden. Dies wurde insbesondere an Beschreibungen von Tiefencues wie der ‘retinalen Größe’ eines Objektes, der Verdeckung oder der binokularen Disparität verdeutlicht, welche nur dann die ihnen zugeschriebene Funktion erfüllen können, wenn das auf die Retinae projizierte Licht *als ein Bild* der äußeren Objekte *wahrgenommen* wird oder bestimmte Elemente des retinalen Lichtmusters als Projektionen bestimmter Eigenschaften der Objekte der Außenwelt, beziehungsweise Relationen zwischen diesen, identifiziert werden.<sup>3</sup> Schaut man sich die mit dem *cue*-Konzept verbundenen Vorstellungen an, so setzen sie – mehr oder weniger implizit – eine personenähnliche Instanz voraus, die wir als Homunkulus bezeichnet haben, die selbst wahrnimmt, und/oder über ‘intellektuelle Fähigkeiten’ verfügt, um die retinalen Lichtmuster als Bilder der Objekte der Außenwelt zu sehen oder zu identifizieren, diese zu ‘vergleichen’ oder bestimmte Berechnungen mit Elementen dieser Bilder anzustellen.

Die Tiefencues weisen in diesen historischen theoretischen Ansätzen einen gleichsam ‘janusartigen Charakter’ auf. Sie sollen einerseits der Ebene der ‘von außen’ auf den Wahrnehmenden einwirkenden physikalischen Welt zugehörig sein, andererseits, so wurde argumentiert, werden sie aber erst als durch Wahrnehmung konstituiert verständlich. In den Beschreibungen der Hilfsmittel ist die physikalische Beschreibungsebene mit einer phänomenalen Beschreibungsebene verwoben, obwohl der phänomenale Gehalt, das Haben einer Wahrnehmung, im Wesentlichen *durch* die physikalische Beschreibungsebene, durch das Einwirken von Licht auf die Augen, verständlich gemacht werden soll. Gleichzeitig scheinen einige der Autoren (etwa Reid, Lotze und Helmholtz) eine deutliche Zurückhaltung an den Tag zu legen, wenn es um die Frage geht, ob der Übergang von der Lichteinwirkung hin zum Perzept überhaupt geklärt werden kann.

In aktuellen Formulierungen des *cue*-Ansatzes lässt sich eine solche Zurückhaltung in der Regel nicht ausmachen. Im Zuge der kritischen Auseinandersetzung mit aktuellen Formulierungen der Tiefencues zeigte sich zunächst, dass sämtliche der in den historischen Ansätzen behandelten Tiefencues auch heute noch als solche angesehen werden und die entsprechenden Formulierungen an die früheren Verwendungsweisen anschließen, aber innerhalb eines anderen begrifflichen Rahmens, dem des Informationsverarbeitungsparadigmas. In diesem stellt sich das Konzept der Tiefencues nun deutlich abstrakter dar, indem es in unmittelbare Beziehung zum Informationskonzept und dem Konzept ‘interner Repräsentationen’ gesetzt wird. Über eine nähere Behand-

---

<sup>3</sup>Im Fall der Berkeleyschen Theorie, die Tiefenwahrnehmung nicht primär über das retinale Lichtmuster, sondern über den Tastsinn verständlich machen möchte, wurde argumentiert, dass dieses Vorgehen das Problem lediglich verschiebt.

lung dieser Konzepte und deren Anwendung im wahrnehmungstheoretischen Kontext zeigte sich, dass auch im Kontext der modernen Auffassung visueller Wahrnehmung als eines Informationsverarbeitungsprozesses ähnliche theoretische Probleme mit dem Konzept der Tiefencues verbunden sind wie in den früheren Ansätzen, diese aber nun in gleichsam ‘abstrakterem Gewand’ auftreten. In dieser Konzeption ist die implizite Annahme enthalten, dass physikalische Zustände des Gehirns als syntaktisch und semantisch *interpretiert* werden müssen, damit diese mit Sinn als Informationen über Tiefe oder Repräsentation der Objekte der Umwelt und deren räumlicher Anordnung aufgefasst werden können. Weder Syntax noch Semantik ist jedoch den physikalischen Zuständen eines Systems intrinsisch, so dass auch hier, um diese Auffassung rechtfertigen zu können, eine personenähnliche Instanz, ein Homunkulus vorausgesetzt werden müsste. Das moderne Konzept der Tiefencues weist logische Mängel auf, die in dieser Auffassung nicht zu beheben sind. Selbst wenn man diese (unplausible) Annahme eines Homunkulus explizit machen würde, würde sich erstens das Problem eines unendlichen Regresses stellen, da nun wiederum dessen Fähigkeit, Symbole zu verwenden und zu verstehen, erklärt werden müsste und zweitens müsste die Verbindung zwischen dem Interpretieren von Symbolen durch den Homunkulus und dem Sehen des Wahrnehmenden begreifbar gemacht werden.

Im empirischen Teil dieser Arbeit konnten diese theoretischen Überlegungen anhand eines selbst durchgeführten Experiments zum Forschungsbereich der Bildwahrnehmung, in dem explizit auf das Konzept der Tiefencues zurückgegriffen wird, exemplifiziert werden. Es sollte verdeutlicht werden, wie mit Hilfe der Tiefencues Hypothesen formuliert und zur Erklärung experimenteller Resultate eingesetzt werden. In dieser Untersuchung zum Phänomen der Perspektivenrobustheit bei Betrachtung linearperspektivischer Bilder wurde eine neuartige experimentelle Manipulation vorgenommen, die sich auf den wahrgenommenen Oberflächencharakter eines Bildes bezieht. Es zeigte sich, dass sich die mit dieser Manipulation einhergehenden empirischen Effekte zwar gut in Übereinstimmung bringen lassen mit den im Kontext des *cue*-Ansatzes zu erwartenden Effekten, die Erklärung dieser Befunde aber nur an der Oberfläche überzeugend ist.

Ein naheliegender Einwand gegen die in der vorliegenden Arbeit geäußerten Bedenken könnte nun wie folgt lauten: Wenn sich – wie viele empirische Studien doch zeigen – unter Verwendung dieses Konzeptes empirische Untersuchungen motivieren und auch experimentell stützen lassen, sind dann die hier ausgesprochenen Zweifel am Erklärungswert dieses Konzeptes überhaupt gerechtfertigt? Es soll an dieser Stelle noch einmal unterstrichen werden, dass die vorliegende Arbeit nicht das Ziel haben kann, den großen *heuristischen* Wert des *cue*-Konzeptes in wahrnehmungspsychologischen Untersuchungen zu leugnen, der sich in einer Fülle eindrucksvoller experimenteller Ergebnisse

niedergeschlagen hat. *Cue*-Heuristiken können, wie auch die vorliegende Untersuchung demonstriert, relevante Erkenntnisse zutage fördern. Die theoretischen Überlegungen bezüglich des *cue*-Konzeptes sollen vielmehr darauf aufmerksam machen, dass das Konzept nur oberflächlich zur 'Erklärung' solcher Befunde herangezogen werden kann. Die Tiefenwahrnehmung durch Tiefencues (als Informationen über Tiefe) 'erklären' zu wollen, ist, und das sollte diese Arbeit zeigen, mit gehörigen theoretischen Schwierigkeiten verbunden, die nicht einfach übergangen werden dürfen. Zieht man heuristische Konzepte dieser Art heran, so sollte klar sein und auch deutlich gemacht werden, dass die Forschung von einem wirklichen Verständnis der (Tiefen-)Wahrnehmung noch weit entfernt ist.

Ein wissenschaftspragmatischer Einwand könnte darin bestehen, dass, gerade bei noch weitgehend unverstandenen Forschungsgegenständen wie dem der visuellen Wahrnehmung, eine gewisse Unschärfe der Konzepte nicht nur nicht hinderlich sein müsse, sondern gerade durch ihre Unschärfe für die Erstellung von neuen Fragestellungen und Untersuchungen, durch welche die Unschärfe Schritt für Schritt eliminiert werden kann, von unschätzbarem heuristischen Wert sein können. Wie in dieser Arbeit herausgearbeitet werden sollte, ist das bei dem Konzept der Tiefencues nicht der Fall, es hat im Laufe seiner langen Geschichte keine wesentlichen Veränderungen erfahren, welche zu einem tieferen Verständnis der Tiefenwahrnehmung und einer in sich logisch konsistenten wahrnehmungspsychologischen Erklärung der Phänomene geführt hätte. Trotz der ungemainen Forschungsanstrengungen im Bereich der Tiefenwahrnehmung und der Unzahl an zugehörigen empirischen Ergebnissen aus der Psychologie, der Physiologie oder den Neurowissenschaften ist dieses Konzept immer noch mit ähnlichen theoretischen Schwierigkeiten verbunden, wie es vor knapp tausend Jahren der Fall war. Diese Schwierigkeiten sind meines Erachtens nicht auf eine ungenügende empirische Befundlage zurückführbar, vielmehr deuten sie fundamentale Probleme hin, die mit der theoretischen Erklärung visueller Wahrnehmung innerhalb eines naturwissenschaftlichen Rahmens zusammenhängen.

In aktuellen wahrnehmungspsychologischen Ansätzen wird visuelle Wahrnehmung wesentlich als ein Gegenstand einer naturwissenschaftlich beschreibbaren Welt aufgefasst, der somit prinzipiell auch in einem naturwissenschaftlichen Rahmen verständlich gemacht werden könne. Es wird daher versucht, einen kausalen oder physikalisch beschreibbaren Zusammenhang zwischen dem auf die Retinae projizierten Licht und dem Haben einer Wahrnehmung herzustellen. Das Verhalten von Licht, die Absorption der Lichtquanten durch die Photorezeptoren auf der Retina, die dadurch ausgelösten Aktionspotentiale und deren Weiterleitung über den Sehnerv in verschiedene Areale des Gehirns, all diese Vorgänge lassen sich mit naturwissenschaftlichen Mitteln in einen hinreichend verständlichen kausalen Zusammenhang bringen. Entsprechende empirische

Untersuchungen, etwa in den Neurowissenschaften, haben einen unbestreitbaren Wert, da sie diese *physikalischen Vorgänge* in zunehmender Genauigkeit begreifbar machen können. Es wäre daher prinzipiell denkbar, all den mit dem Konzept der Tiefencues verbundenen problematischen Beschreibungen der ‘Interpretation’ oder der ‘Auswertung’ eine konkrete Bedeutung zu geben, indem sie mit bestimmten, empirisch feststellbaren Gehirnvorgängen identifiziert werden. Die berechtigte Frage, wie man ohne auf Wahrnehmung, d.h. insbesondere ohne auf mentales Vokabular, zurückgreifen zu müssen, eine derartige Identifizierung herstellen kann, soll hier ganz außer Acht gelassen werden. Die mit der Vorstellung einer Identifizierbarkeit von neurophysiologischen und perzeptuellen Aktivitäten verbundenen Probleme zeigen, dass sich hier die mit dem *cue*-Konzept verquickten Probleme auf einer anderen Ebene wiederholen. Aber würde eine solche Beschreibung überhaupt zu einem tieferen Verständnis der Struktur und Dynamik unserer Wahrnehmung führen?

Im Kontext einer naturwissenschaftlichen Erklärung *unserer Wahrnehmung* (nicht unserer Gehirnaktivitäten) muss notwendigerweise eine Verbindung oder ein ‘Übergang’ zwischen den neuronalen Aktivitätsmustern im Gehirn und dem phänomenalen Gehalt einer Wahrnehmung hergestellt werden, man ist unmittelbar mit dem Leib-Seele-Problem (oder, wie es heute auch heißt, dem *mind-body*-Problem) konfrontiert. Eine rein physikalisch-physiologische Beschreibung der mit visueller Wahrnehmung *einhergehenden* Gehirnvorgänge verfehlt den Gegenstandsbereich einer Wahrnehmungspsychologie – zudem können diese Beschreibungen bereits von anderen Wissenschaften geleistet werden. Ein *eliminativer Materialismus*, wie er etwa von Churchland (1986) vertreten wird, ist für eine Wahrnehmungspsychologie vollkommen unbefriedigend. Es gilt, die Phänomene verständlich zu machen; neuronale Aktivität ist eben *kein* Sehen von etwas, es ist (mit einiger Sicherheit) eine *Voraussetzung* dafür, dass wir sehen. Gleichzeitig zeigt die intensive und anhaltende Diskussion in der Philosophie des Geistes, dass die mannigfaltigen (und mittlerweile kaum mehr zu überblickenden) Versuche, eine solche Verbindung zwischen Gehirnaktivität und mentalem Erleben herzustellen, letztlich keine durchweg befriedigenden Antworten geben können.<sup>4</sup>

Der phänomenale Gehalt der Wahrnehmung und die naturwissenschaftliche Beschreibung der physikalisch-physiologischen Vorgänge bei Eintritt des Lichts in die Augen sind kategorial verschieden voneinander, sie bilden unterschiedliche, inkommensurable Sprachebenen – oder unterschiedliche *Sprachspiele* im Sinne Wittgensteins. Die Wahrnehmungspsychologie ist aber aufgrund ihres heutigen Selbstverständnisses auf *beide* Beschreibungsebenen und eine Verbindung zwischen diesen angewiesen. Was begreif-

---

<sup>4</sup>Solche Versuche lassen sich etwa in einem *interaktionistischen Dualismus* (etwa Popper & Eccles, 1977/1987), verschiedenen Spielarten der *Identitätstheorie* (etwa Place, 1956; Smart, 1959) oder des *Eigenschaftsdualismus* ausmachen. Wir wollen diese Positionen hier nicht weiter behandeln; für einen Überblick über diese verschiedenen Ansätze siehe etwa Chalmers (2003).

bar gemacht werden soll, ist der phänomenale Gehalt unserer Wahrnehmung, das, was wir *sehen*, aber innerhalb eines physikalischen oder naturwissenschaftlichen Weltbildes, nämlich über die Einwirkung des physikalischen Reizes Licht auf unsere Augen. Prinzipiell ist diese Herangehensweise durchaus nachvollziehbar – wenn kein Licht auf unsere Augen trifft, sehen wir eben auch nichts. Die fundamentale Schwierigkeit in diesem Schema besteht darin, dass in irgendeiner Weise ein Übergang zwischen den beiden inkommensurablen Beschreibungsebenen hergestellt werden soll, was aufgrund dieser Inkommensurabilität in Probleme wie das der homunkularen Redeweisen führt. Würde man diese homunkularen Redeweisen in rein physikalisch-physiologische Termini ‘übersetzen’, ginge der phänomenale Gehalt, der zu erklären gesucht wird, verloren. Möchte man auf die phänomenale Beschreibungsebene nicht verzichten (und eine Wahrnehmungspsychologie *sollte* meines Erachtens darauf nicht verzichten), aber den phänomenalen Gehalt durch Termini aus der physikalischen Beschreibungsebene verständlich machen, steht man vor dem Problem einer sinnvollen, logisch tragfähigen Verbindung dieser beiden Ebenen.

Das Konzept der Tiefencues lässt sich meines Erachtens als ein solcher Versuch der Vermittlung zwischen beiden Ebenen ansehen, es soll gleichsam die Funktion eines Scharniers übernehmen, das beide sprachlichen Ebenen miteinander verbinden und ineinander überführbar machen soll. Die Beschreibungen der Tiefencues beziehen sich auf beide Ebenen gleichzeitig, sie sollen der Ebene der physikalisch beschreibbaren Vorgänge angehören, werden aber erst auf Ebene der Wahrnehmung verständlich. Das Konzept der Tiefencues kann aufgrund dieser ‘Scharnierfunktion’ daher überhaupt kein klares und verständliches Konzept sein, da ein Übergang zwischen physiologischer Aktivität und mentalem Erleben heute noch so unverständlich ist wie zur Zeit Alhazens. Fasst man Wahrnehmung als ein naturwissenschaftlich zu erklärendes Phänomen auf, ist man notwendigerweise auf solche Konzepte angewiesen.

Lässt sich dieses unlösbar erscheinende Problem vermeiden? Ein Versuch, das Problem der Wahrnehmung aus fundamental anderer Perspektive anzugehen, lässt sich in der von Edmund Husserl geprägten philosophischen Strömung der *Phänomenologie* ausmachen. Diese stellt im Kontext der Wahrnehmung den Versuch dar, den *phänomenalen Gehalt* der Wahrnehmung, seine Struktur und innere Logik, zunächst zum alleinigen Gegenstand der Betrachtung zu machen. In einer phänomenologischen Analyse der Wahrnehmung (etwa Gurwitsch, 1966; Husserl, 1973/1991, 2004, 1980/2006; Merleau-Ponty, 1966; Schmitz, 1978/2005) wird sich u.a. Existenzannahmen über die Dinge der Welt und möglichen Kausalzusammenhängen zwischen diesen als theoretische Vorannahmen über die Welt von vornherein *enthalten*. Sie sollen gleichsam ‘eingeklammert’ (nicht negiert) werden, um so – den wohl bekanntesten Ausspruch Husserls aufgreifend – „zu den Sachen selbst“, so wie sie sich als Phänomene in unserem Erleben

darstellen, und deren ‘inhärenter Logik’ zu gelangen.<sup>5</sup> Was an der phänomenologischen Haltung überzeugt, ist die Auffassung, dass Wahrnehmung aus prinzipiellen Gründen *nicht* als ein mit naturwissenschaftlicher Methode unabhängig vom Untersuchenden selbst zu untersuchender Gegenstand neben anderen zu verstehen ist, da sich der Gegenstandsbereich empirischer Untersuchungen (die ‘Außenwelt’) stets und notwendigerweise erst über die Wahrnehmung konstituiert. Wenn wir als Wahrnehmungstheoretiker und -forscher über Wahrnehmung reden, so können wir das notgedrungen nur aus der Position einer Person heraus, die selbst ständig wahrnimmt und nur deshalb über Wahrnehmung reden und empirische Untersuchungen anstellen kann, *weil* sie ständig wahrnimmt. All unsere Beschreibungen und Beobachtungen sind bedingt durch unsere Wahrnehmung, sie sind nur durch Wahrnehmung möglich. Wahrnehmung kann somit nicht als von der Wahrnehmung des Beschreibenden unabhängiger Gegenstand aufgefasst werden, es ist kein ‘Objekt in der Welt’, sondern durch Wahrnehmung sind uns überhaupt erst Objekte in der Welt gegeben, die wir empirisch untersuchen und an denen wir empirische ‘Gesetzmäßigkeiten’ feststellen können. Merleau-Ponty (1966) macht das in einer eindrucksvollen Passage deutlich:

Das objektive Denken ignoriert das Subjekt der Wahrnehmung. Die Welt betrachtet es als etwas fertig Gegebenes, als das Milieu aller nur möglichen Vorkommnisse, als deren eines die Wahrnehmung behandelt wird. So etwa betrachtet der empiristische Philosoph das Subjekt X im Begriffe, wahrzunehmen, und sucht, was da vor sich geht, zu beschreiben: *es gibt* da Empfindungen, die als Zustände oder Seinsweisen des Subjekts, wahrhaftige mentale Dinge sind. Das wahrnehmende Subjekt ist der Ort dieser Dinge und der Philosoph beschreibt die Empfindungen und ihr Substrat, so wie man die Fauna eines fernen Landes beschreibt – ohne irgend daran zu denken, daß er ja selber wahrnimmt, selbst wahrnehmendes Subjekt ist, und daß die Wahrnehmung, so wie er sie selbst erlebt, all das, was er von der Wahrnehmung überhaupt sagt, in Frage stellt. Denn von innen gesehen, verdankt sich die Wahrnehmung durchaus nicht alledem, was wir anderswoher schon wissen mögen von der Welt, von den Reizen, so wie die Physik sie beschreibt, und von den Sinnesorganen, so wie die Biologie sie beschreibt. Sie gibt sich nicht zunächst als ein Vorkommnis in der Welt, auf das z.B. die Kategorie der Kausalität anzuwenden wäre, sondern als in jedem Augenblicke von neuem die Welt erst schaffend oder rekonstituierend. ... Alles

---

<sup>5</sup>Es handelt sich hier wohlgerne um eine eher ‘schlagwortartige’ Charakterisierung der phänomenologischen *Haltung*. Wir können (und wollen) an dieser Stelle nicht näher auf die phänomenologische Methode eingehen, da hierfür u.a. ein vollkommen anderer begrifflicher Rahmen geschaffen werden müsste. Für eine gut lesbare Einführung in das Denken Husserls und die phänomenologische Methode siehe Husserl (1968/2003).

Wissen begründet sich erst in den Horizonten, die die Wahrnehmung uns eröffnet. Die Wahrnehmung selbst ist nicht zu beschreiben als eines unter jenen Fakten, die in der Welt vorkommen, da wir im Bilde der Welt nie jene leere Stelle zu unterdrücken vermögen, die wir selber sind. (S. 244)

Ob und inwieweit eine solche phänomenologische Haltung für eine Wahrnehmungspsychologie fruchtbar gemacht werden könnte, sind Fragen, denen man weiter nachgehen könnte; in der Geschichte der Wahrnehmungspsychologie lassen sich aber durchaus beeindruckende Beispiele phänomenologisch ausgerichteter Untersuchungen finden, etwa Gelb (1929), Katz (1930) oder die Arbeiten von Michotte (siehe Thinès, Costall & Butterworth, 1991). Den phänomenologischen Standpunkt von vornherein als ‘unwissenschaftlich’ oder ‘subjektiv’ abzutun ist aufgrund des besonderen, philosophisch überaus problematischen Charakters einer Untersuchung der Wahrnehmung durch stets Wahrnehmende nicht gerechtfertigt.

Welches Fazit lässt sich aus der vorliegenden Arbeit ziehen? Betrachtet man die aktuelle Literatur im Bereich der Tiefenwahrnehmung, so entsteht zumindest der Eindruck, als gäbe es im Wesentlichen nur noch Detailfragen zu klären, etwa welcher Tiefencue am ‘effizientesten’ zur Wahrnehmung von Tiefe ‘beiträgt’; Fragen, die durch weitergehende empirische Untersuchungen beantwortet werden könnten. Begriffe wie „Informationen über Tiefe“, „mentale Repräsentation von Tiefe“ oder Wendungen wie die des ‘Perzept generierenden Gehirns’ werden in diesem Zusammenhang verwendet, als ob klar und verständlich wäre, was damit gemeint ist. In der vorliegenden Arbeit wurde versucht, die mit dem in Theorien der Tiefenwahrnehmung zentralen Konzept der Tiefencues verbundenen theoretischen Probleme herauszuarbeiten, um zumindest deutlich zu machen, dass dieses Konzept nicht in naiver Weise als eine hinreichende ‘Erklärung’ für die Wahrnehmung von Tiefe herangezogen werden kann. Es kann aus Mangel an Alternativen als zweckmäßige Heuristik dienen und gleichzeitig darauf hinweisen, welche fundamentalen, möglicherweise prinzipiell nicht zu umgehenden Probleme mit dem Versuch einer Erklärung visueller Wahrnehmung in einem naturwissenschaftlichen Rahmen verbunden sind. Ein wirkliches Verständnis unseres Sehens ist aufgrund der mit diesem Konzept verbundenen Probleme nicht zu erlangen. Es wäre meines Erachtens daher angemessen, sich einer gewissen Bescheidenheit oder Zurückhaltung bezüglich des Erklärungswertes dieses Konzeptes zu bedienen – unser Sehen ist und bleibt vermutlich so rätselhaft wie faszinierend.

# A | Daten

Die Tabellen A.1 bis A.10 (S. 262 - 265) geben die Ergebnisse der für jede Versuchsperson durchgeführten dreifaktoriellen Varianzanalyse an, die Abbildungen A.1 bis A.10 (S. 266 - 275) zeigen für sämtliche Versuchspersonen die mittleren Einstellungen von  $\tau$  unter der Standard- und Leinwandbedingung sowie getrennt nach den verschiedenen Sehbedingungen.

Tabelle A.1.

*Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson CB*

Quelle	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$	$\eta_p^2$
<i>RA</i>	1	.067	.067	209.4	< .001	.012	.356
<i>SB</i>	2	.209	.105	325.7	< .001	.038	.632
<i>WB</i>	6	4.85	.808	2515	< .001	.873	.975
<i>RA</i> × <i>SB</i>	2	.004	.002	5.71	< .01	< 10 <sup>-3</sup>	.029
<i>RA</i> × <i>WB</i>	6	.024	.004	12.49	< .001	.004	.165
<i>SB</i> × <i>WB</i>	12	.253	.021	65.79	< .001	.046	.676
<i>RA</i> × <i>SB</i> × <i>WB</i>	12	.025	.002	6.58	< .001	.004	.173
Residuen	378	.121	.0003				

Tabelle A.2.

*Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson CL*

Quelle	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$	$\eta_p^2$
<i>RA</i>	1	.051	.051	106.42	< .001	.01	.22
<i>SB</i>	2	.905	.452	946.63	< .001	.177	.834
<i>WB</i>	6	3.08	.513	1074	< .001	.602	.945
<i>RA</i> × <i>SB</i>	2	.006	.003	6.32	< .01	.001	.032
<i>RA</i> × <i>WB</i>	6	.029	.005	10.4	< .001	.005	.142
<i>SB</i> × <i>WB</i>	12	.851	.071	148.5	< .001	.167	.825
<i>RA</i> × <i>SB</i> × <i>WB</i>	12	.009	.0007	1.533	.11	.001	.046
Residuen	378	.181	.0005				

Tabelle A.3.

*Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson CR*

Quelle	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$	$\eta_p^2$
<i>RA</i>	1	.136	.136	218.84	< .001	.036	.367
<i>SB</i>	2	.606	.303	488.95	< .001	.16	.721
<i>WB</i>	6	2.24	.373	602.67	< .001	.592	.905
<i>RA</i> × <i>SB</i>	2	.012	.006	9.31	< .001	.003	.047
<i>RA</i> × <i>WB</i>	6	.05	.008	13.53	< .001	.013	.177
<i>SB</i> × <i>WB</i>	12	.483	.04	64.93	< .001	.128	.673
<i>RA</i> × <i>SB</i> × <i>WB</i>	12	.024	.002	3.22	< .001	.006	.092
Residuen	378	.234	.0006				

Tabelle A.4.

Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson DS

Quelle	df	SS	MS	F	p	$\eta^2$	$\eta_p^2$
RA	1	.151	.151	374.19	< .001	.02	.497
SB	2	1.102	.551	1368.9	< .001	.148	.879
WB	6	5.14	.857	2130.1	< .001	.689	.971
RA × SB	2	.016	.008	18.86	< .001	.002	.091
RA × WB	6	.038	.006	15.64	< .001	.005	.199
SB × WB	12	.861	.072	178.24	< .001	.115	.85
RA × SB × WB	12	.004	.0003	.86	.59	< 10 <sup>-3</sup>	.027
Residuen	378	.152	.0004				

Tabelle A.5.

Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson FK

Quelle	df	SS	MS	F	p	$\eta^2$	$\eta_p^2$
RA	1	.062	.062	165.03	< .001	.013	.295
SB	2	.328	.164	437.22	< .001	.071	.702
WB	6	3.86	.644	1714.2	< .001	.826	.965
RA × SB	2	.002	.001	2.99	.051	< 10 <sup>-3</sup>	.015
RA × WB	6	.016	.003	6.95	< .001	.003	.1
SB × WB	12	.257	.021	56.96	< .001	.06	.645
RA × SB × WB	12	.01	.001	2.22	< .05	.002	.066
Residuen	378	.142	.0004				

Tabelle A.6.

Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson JD

Quelle	df	SS	MS	F	p	$\eta^2$	$\eta_p^2$
RA	1	.078	.078	151.77	< .001	.013	.286
SB	2	.498	.249	483.66	< .001	.084	.719
WB	6	4.46	.744	1444.1	< .001	.754	.958
RA × SB	2	.022	.011	21.77	< .001	.004	.103
RA × WB	6	.02	.003	6.35	< .001	.003	.092
SB × WB	12	.627	.052	101.47	< .001	.106	.763
RA × SB × WB	12	.019	.002	3.07	< .001	.003	.089
Residuen	378	.195	.0005				

Tabelle A.7.

*Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson JG*

Quelle	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$	$\eta_p^2$
<i>RA</i>	1	.045	.045	66.27	< .001	.011	.149
<i>SB</i>	2	.392	.196	289.68	< .001	.096	.605
<i>WB</i>	6	2.98	.496	733.01	< .001	.729	.921
<i>RA</i> × <i>SB</i>	2	.006	.003	4.44	< .05	.001	.023
<i>RA</i> × <i>WB</i>	6	.047	.008	11.46	< .001	.011	.154
<i>SB</i> × <i>WB</i>	12	.348	.029	42.87	< .001	.085	.576
<i>RA</i> × <i>SB</i> × <i>WB</i>	12	.008	.001	.992	.45	.002	.031
Residuen	378	.256	.0007				

Tabelle A.8.

*Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson JW*

Quelle	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$	$\eta_p^2$
<i>RA</i>	1	.096	.096	173.43	< .001	.028	.315
<i>SB</i>	2	.603	.301	546.75	< .001	.178	.743
<i>WB</i>	6	1.94	.323	586.1	< .001	.573	.903
<i>RA</i> × <i>SB</i>	2	.015	.008	13.58	< .001	.004	.067
<i>RA</i> × <i>WB</i>	6	.029	.005	9.04	< .001	.009	.125
<i>SB</i> × <i>WB</i>	12	.48	.04	72.58	< .001	.142	.697
<i>RA</i> × <i>SB</i> × <i>WB</i>	12	.014	.001	2.05	< .05	.004	.061
Residuen	378	.208	.0006				

Tabelle A.9.

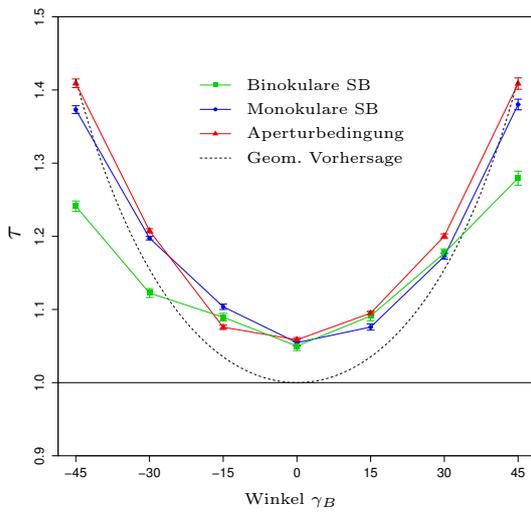
*Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson NS*

Quelle	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$	$\eta_p^2$
<i>RA</i>	1	.056	.056	108.81	< .001	.012	.224
<i>SB</i>	2	.403	.201	393.29	< .001	.085	.675
<i>WB</i>	6	3.65	.608	1188.2	< .001	.766	.95
<i>RA</i> × <i>SB</i>	2	.006	.003	5.96	< .01	.001	.031
<i>RA</i> × <i>WB</i>	6	.021	.004	6.89	< .001	.004	.099
<i>SB</i> × <i>WB</i>	12	.423	.035	68.87	< .001	.089	.686
<i>RA</i> × <i>SB</i> × <i>WB</i>	12	.011	.001	1.84	< .05	.002	.055
Residuen	378	.194	.0005				

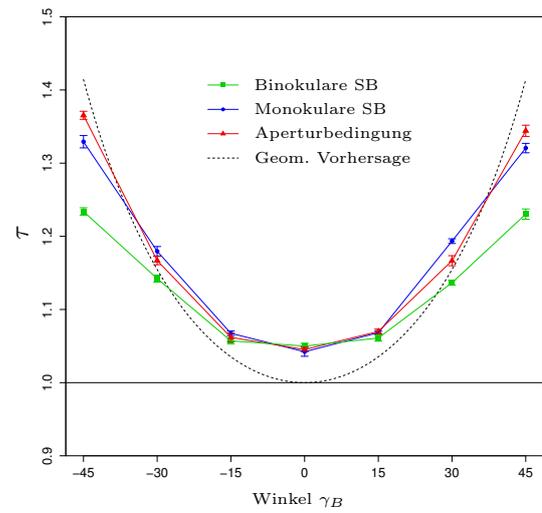
Tabelle A.10.

*Ergebnisse der dreifaktoriellen Varianzanalyse mit den Faktoren Reizart (RA), Sehbedingung (SB) und Winkelbedingung (WB) für Versuchsperson SS*

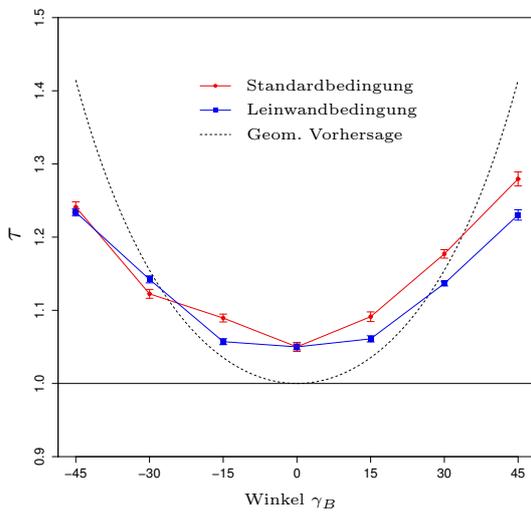
Quelle	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$	$\eta_p^2$
<i>RA</i>	1	.123	.123	225.97	< .001	.025	.374
<i>SB</i>	2	.726	.363	668.79	< .001	.146	.779
<i>WB</i>	6	3.19	.533	982.48	< .001	.642	.94
<i>RA</i> × <i>SB</i>	2	.003	.001	2.57	.078	.001	.013
<i>RA</i> × <i>WB</i>	6	.017	.003	5.11	< .001	.003	.075
<i>SB</i> × <i>WB</i>	12	.701	.058	107.73	< .001	.141	.774
<i>RA</i> × <i>SB</i> × <i>WB</i>	12	.011	.001	1.69	.067	.002	.051
Residuen	378	.205	.0005				



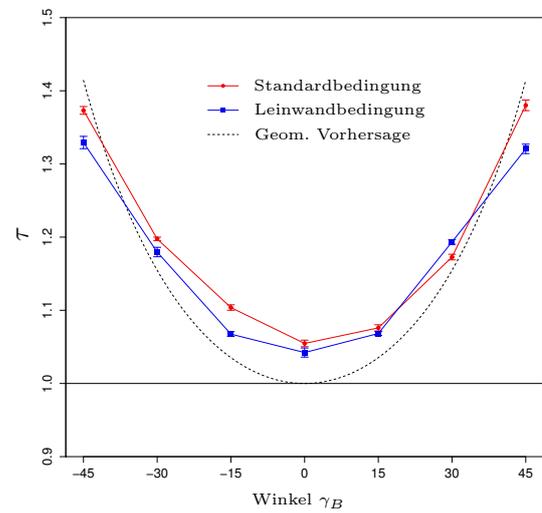
(a) Standardbedingung



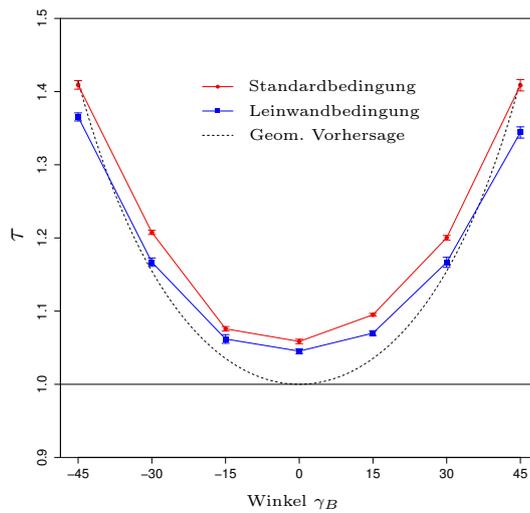
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

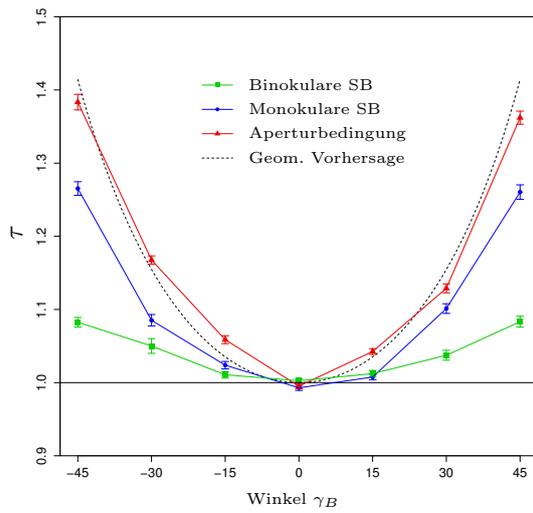


(d) Monokulare Sehbedingung

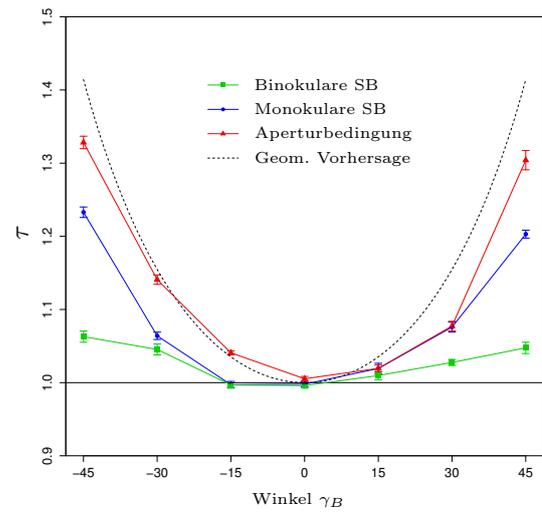


(e) Aperturbedingung

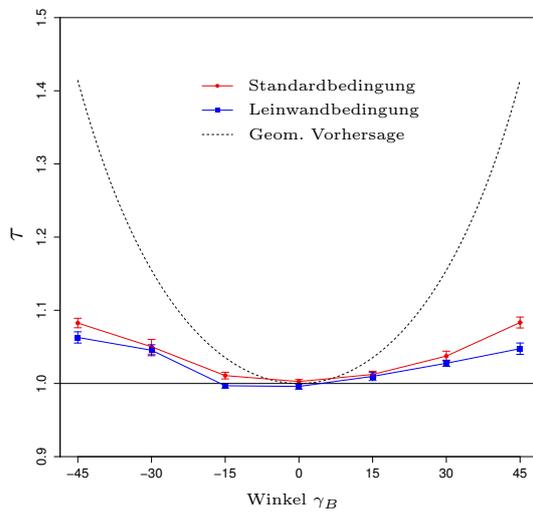
Abbildung A.1. Mittlere Einstellungen von  $\tau$  für Versuchsperson CB unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



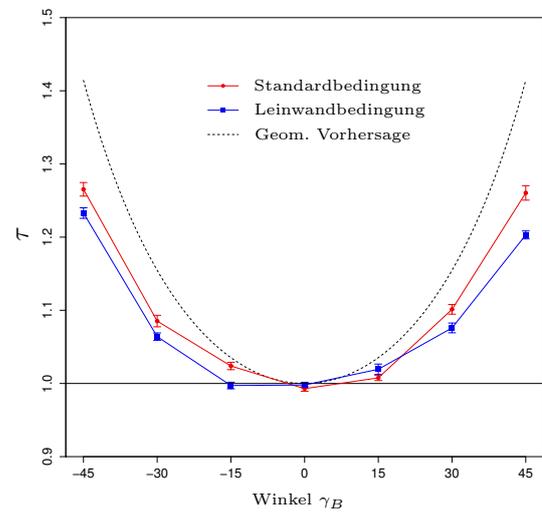
(a) Standardbedingung



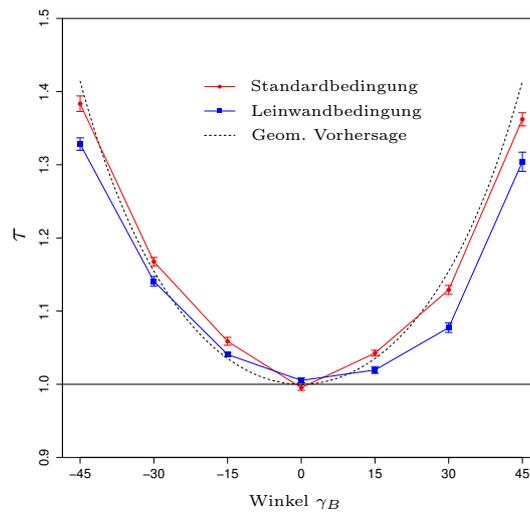
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

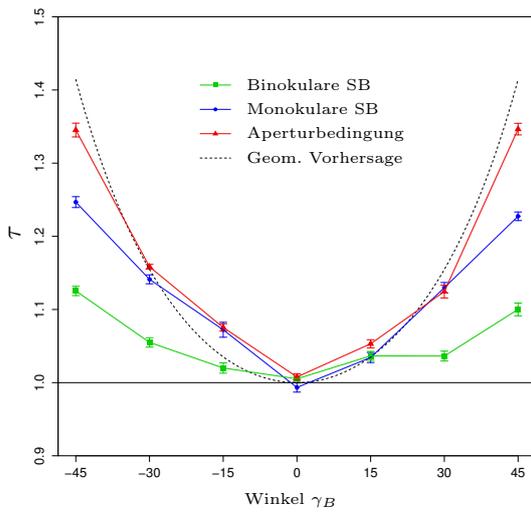


(d) Monokulare Sehbedingung

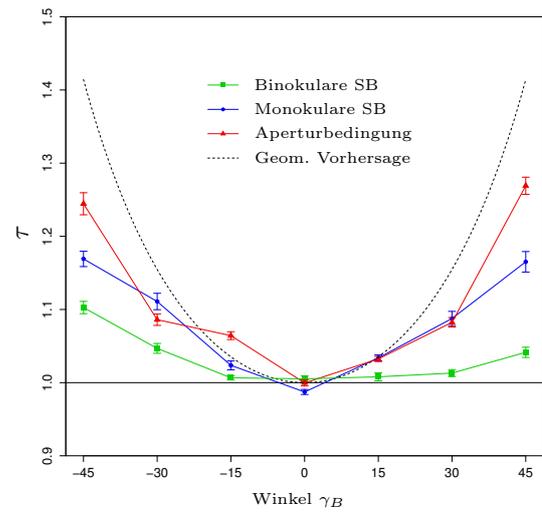


(e) Aperturbedingung

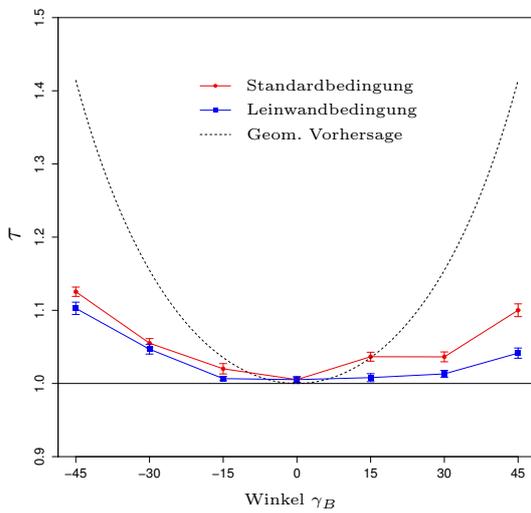
Abbildung A.2. Mittlere Einstellungen von  $\tau$  für Versuchsperson CL unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



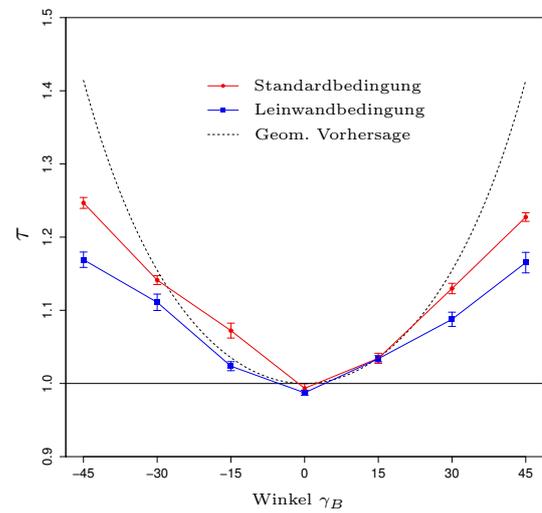
(a) Standardbedingung



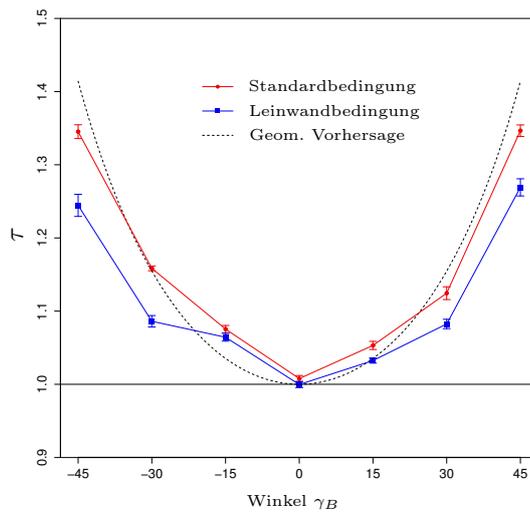
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

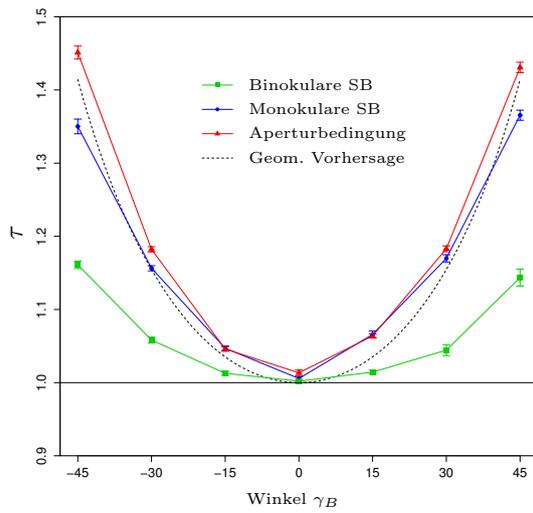


(d) Monokulare Sehbedingung

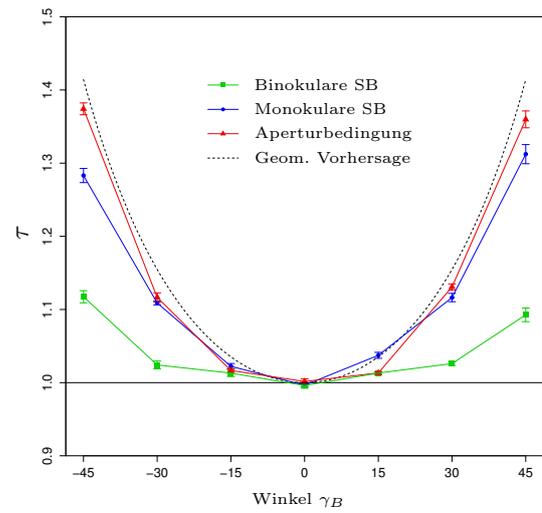


(e) Aperturbedingung

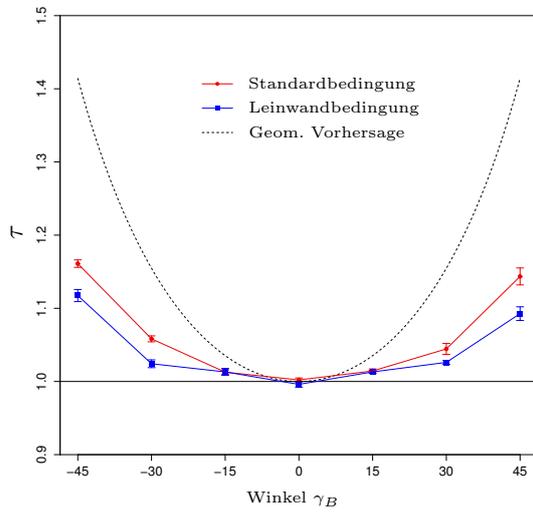
Abbildung A.3. Mittlere Einstellungen von  $\tau$  für Versuchsperson CR unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



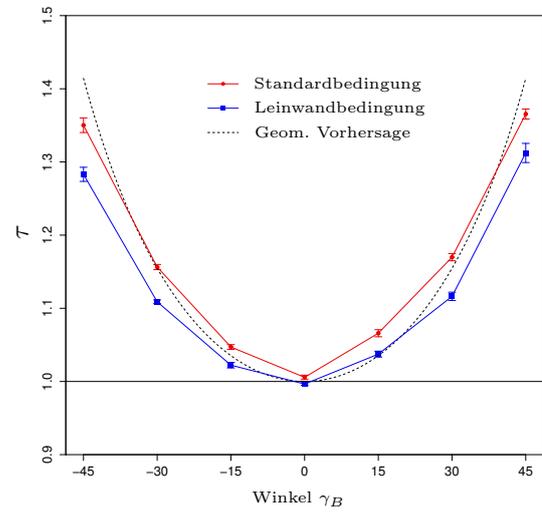
(a) Standardbedingung



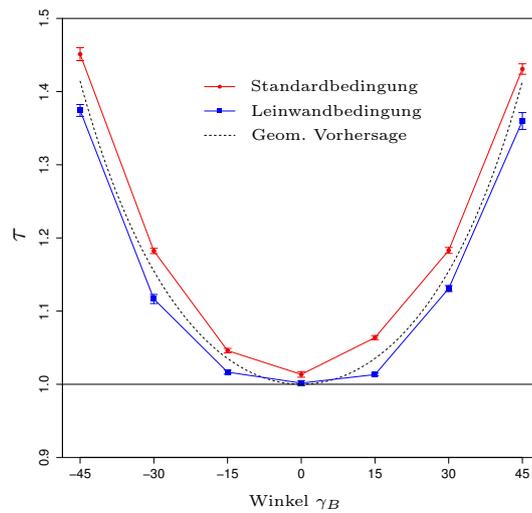
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

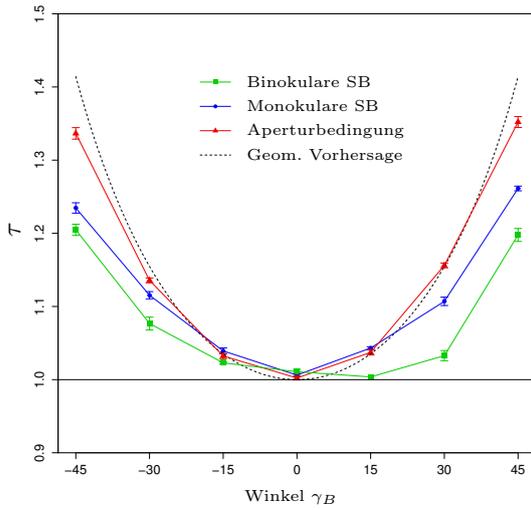


(d) Monokulare Sehbedingung

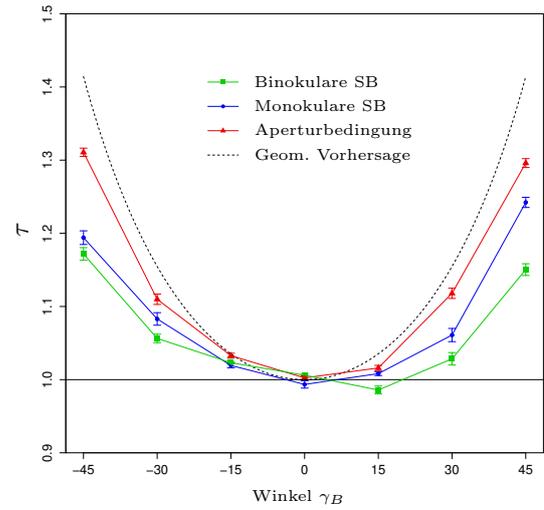


(e) Aperturbedingung

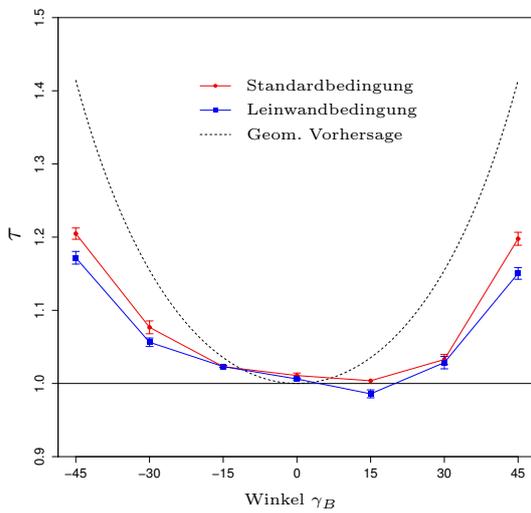
Abbildung A.4. Mittlere Einstellungen von  $\tau$  für Versuchsperson DS unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



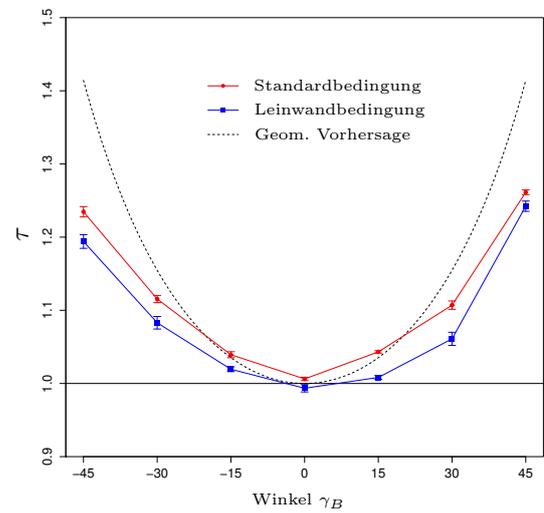
(a) Standardbedingung



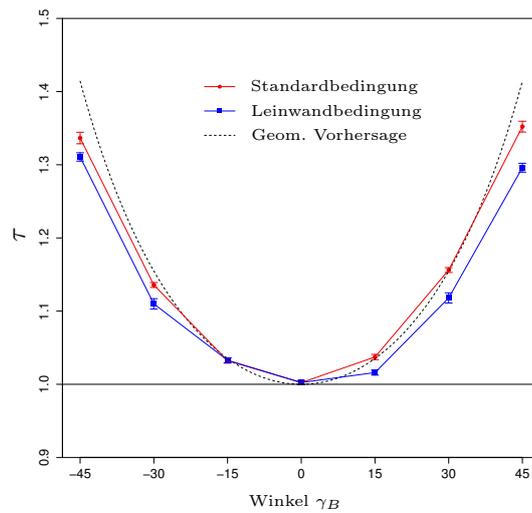
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

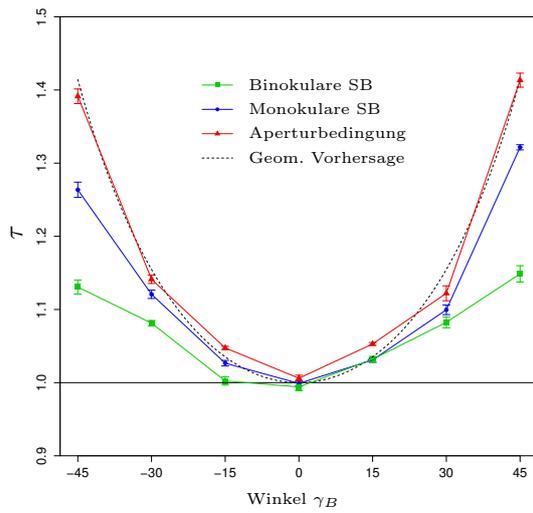


(d) Monokulare Sehbedingung

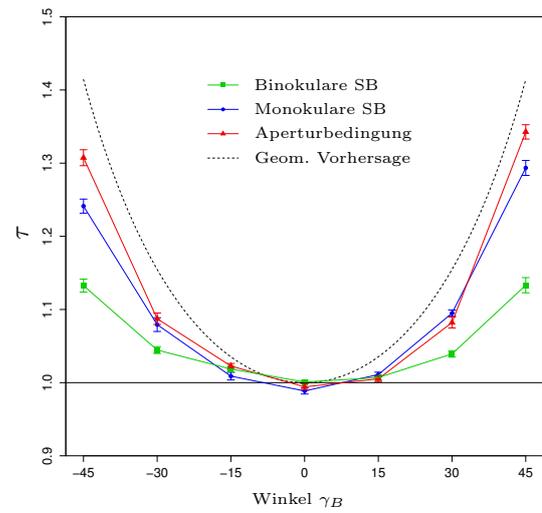


(e) Aperturbedingung

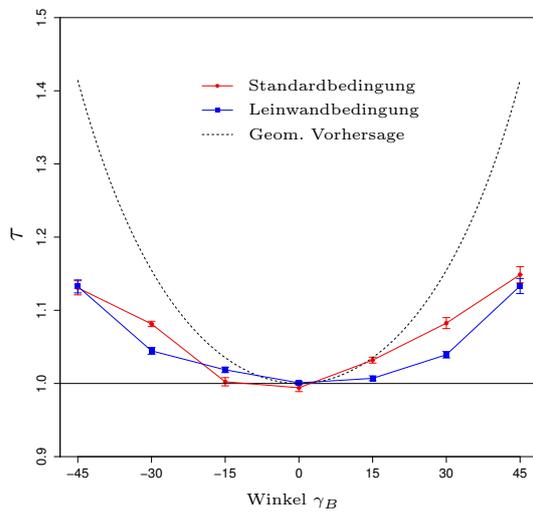
Abbildung A.5. Mittlere Einstellungen von  $\tau$  für Versuchsperson FK unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



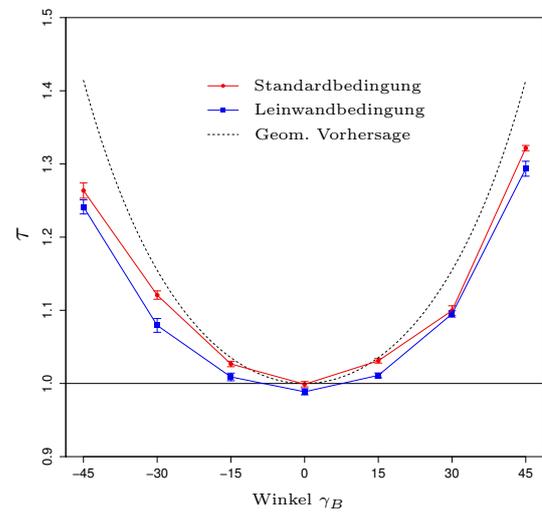
(a) Standardbedingung



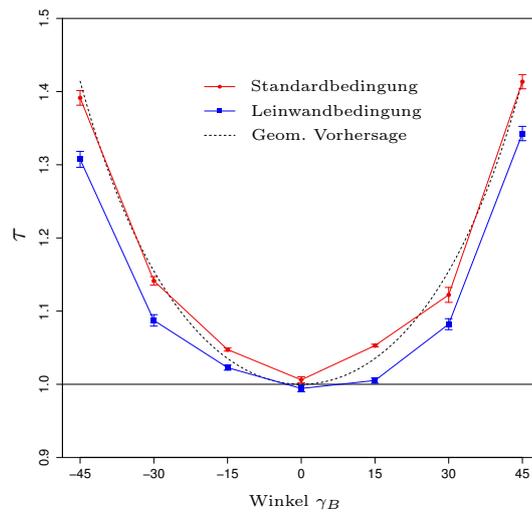
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

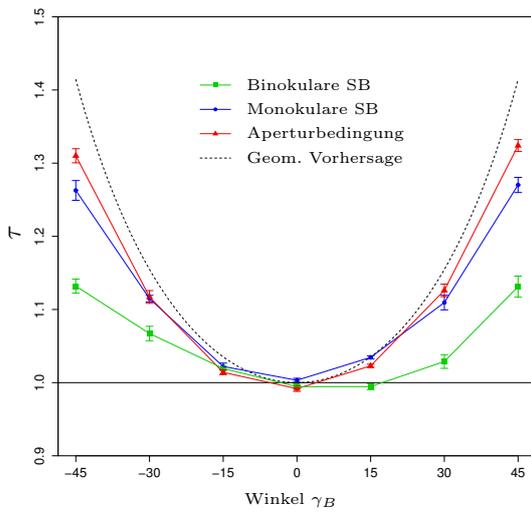


(d) Monokulare Sehbedingung

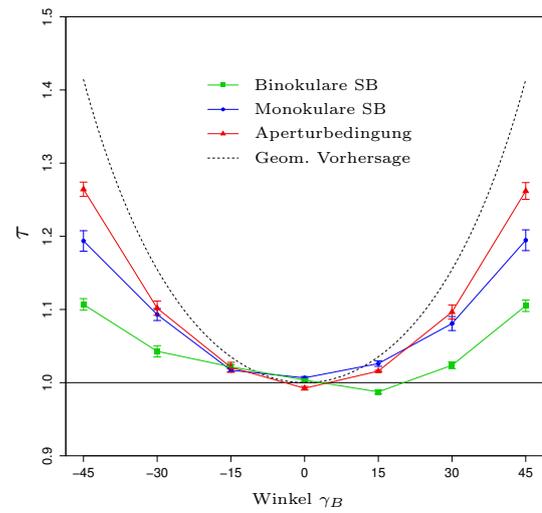


(e) Aperturbedingung

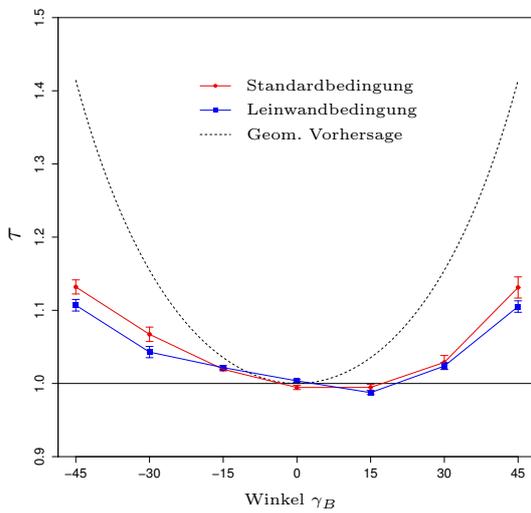
Abbildung A.6. Mittlere Einstellungen von  $\tau$  für Versuchsperson JD unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



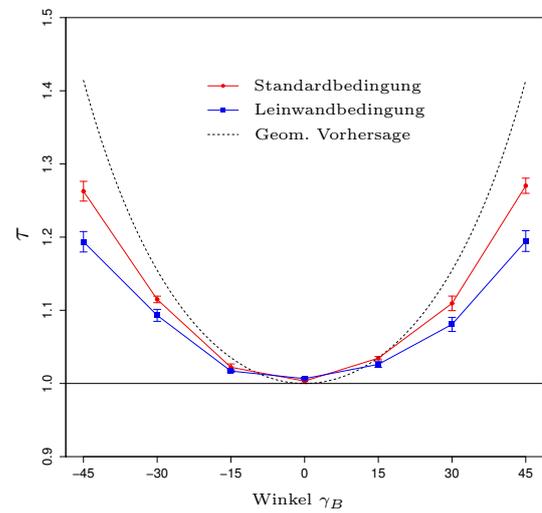
(a) Standardbedingung



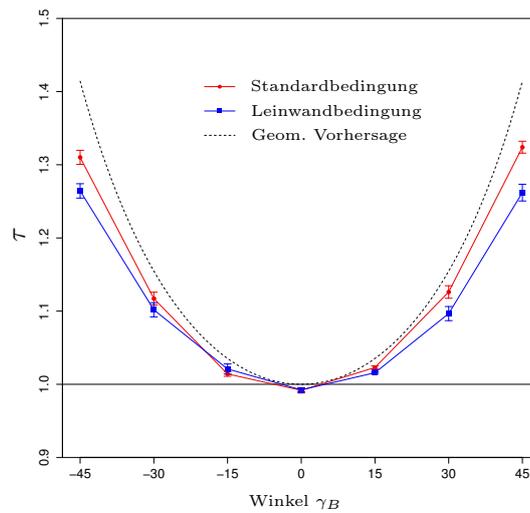
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

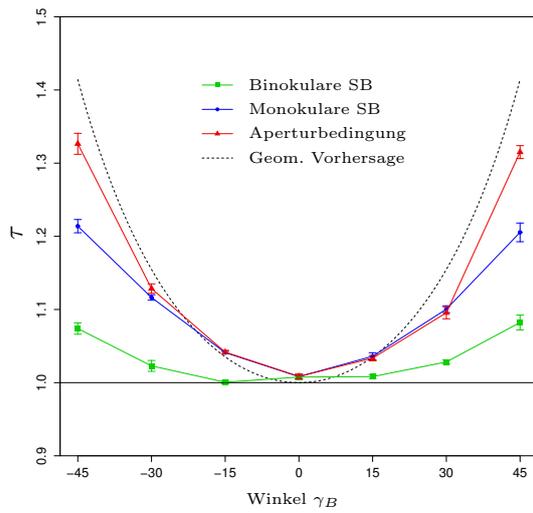


(d) Monokulare Sehbedingung

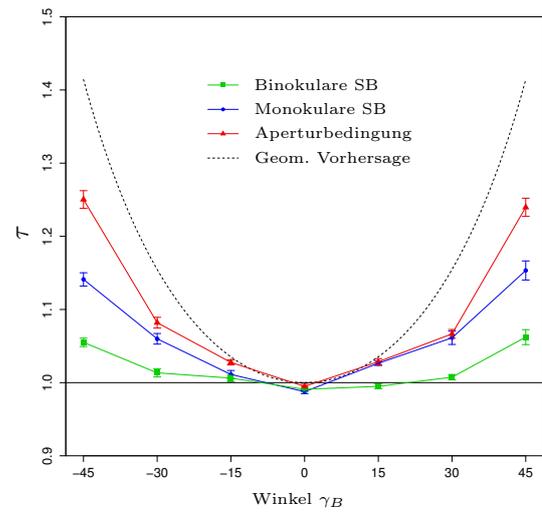


(e) Aperturbedingung

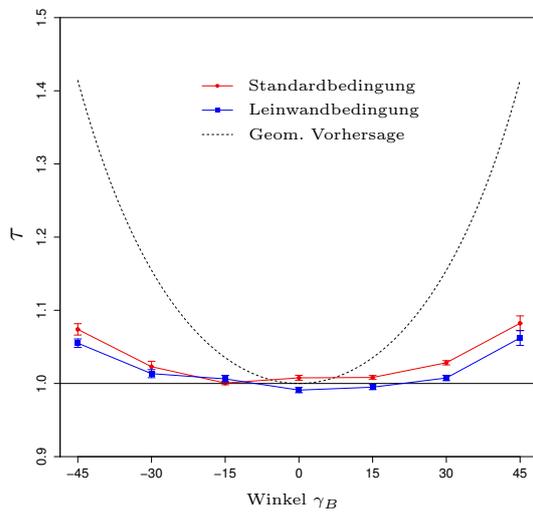
Abbildung A.7. Mittlere Einstellungen von  $\tau$  für Versuchsperson JG unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



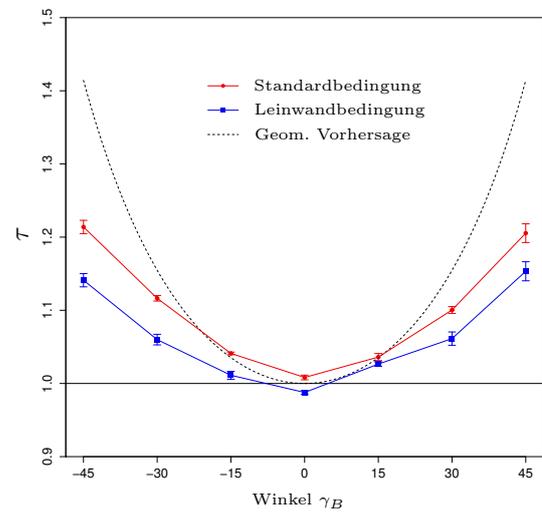
(a) Standardbedingung



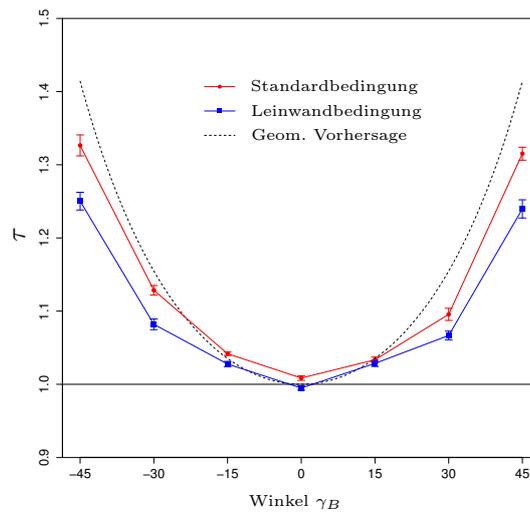
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

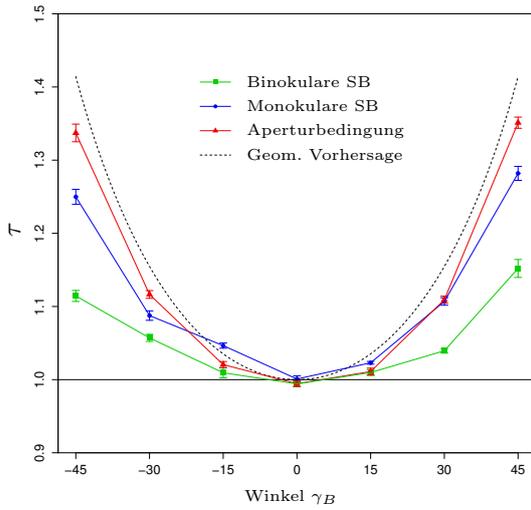


(d) Monokulare Sehbedingung

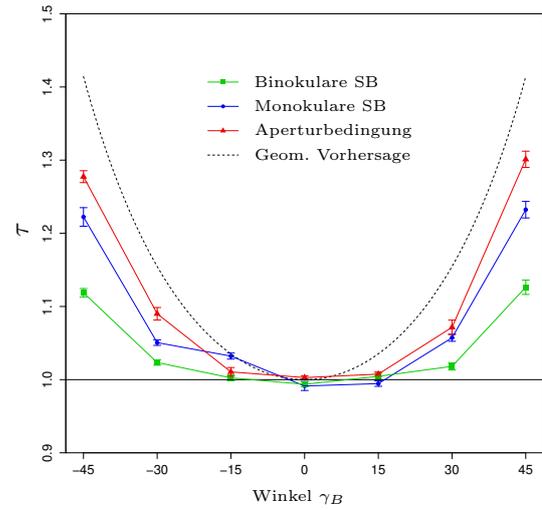


(e) Aperturbedingung

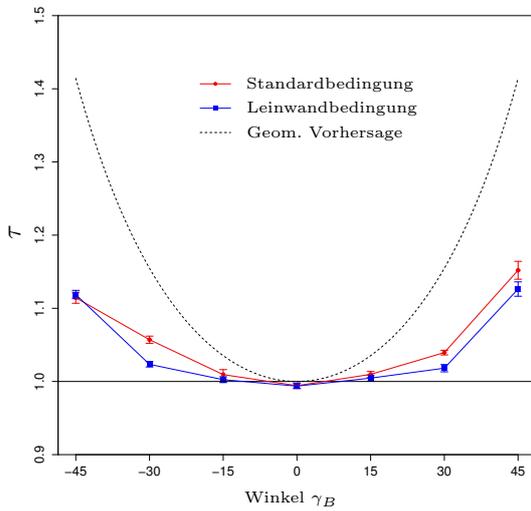
Abbildung A.8. Mittlere Einstellungen von  $\tau$  für Versuchsperson JW unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



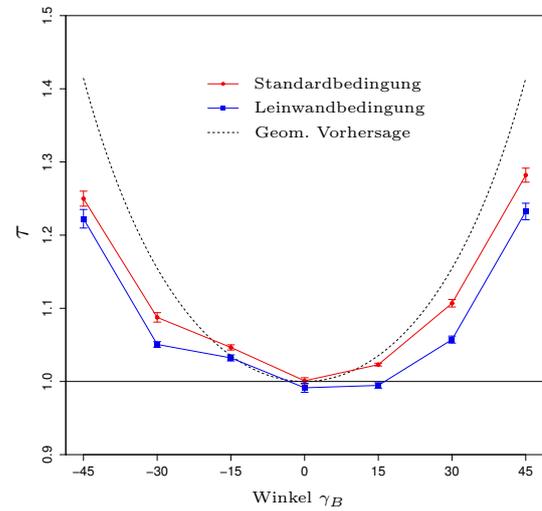
(a) Standardbedingung



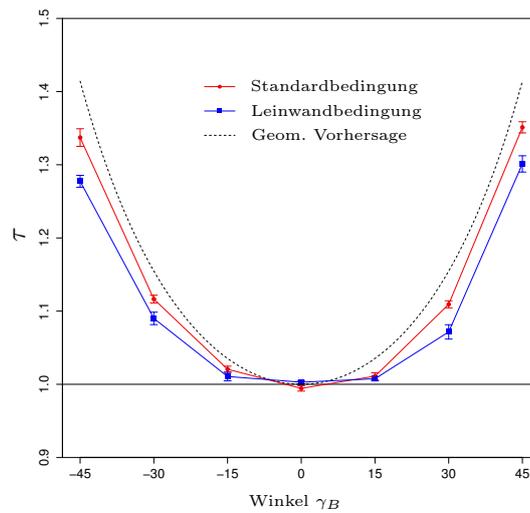
(b) Leinwandbedingung



(c) Binokulare Sehbedingung

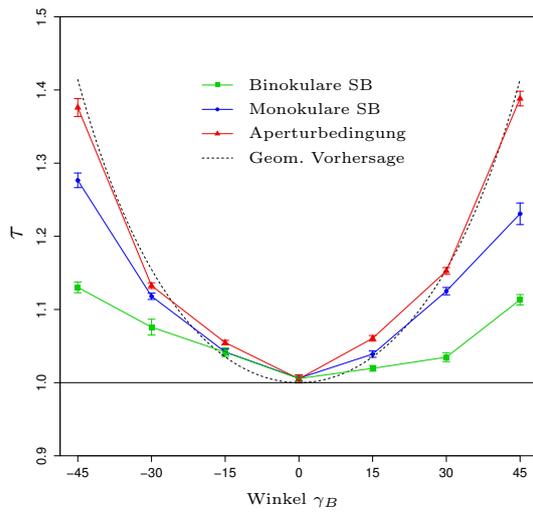


(d) Monokulare Sehbedingung

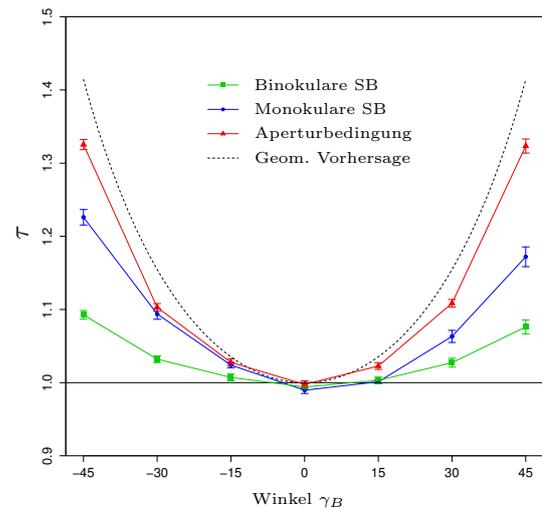


(e) Aperturbedingung

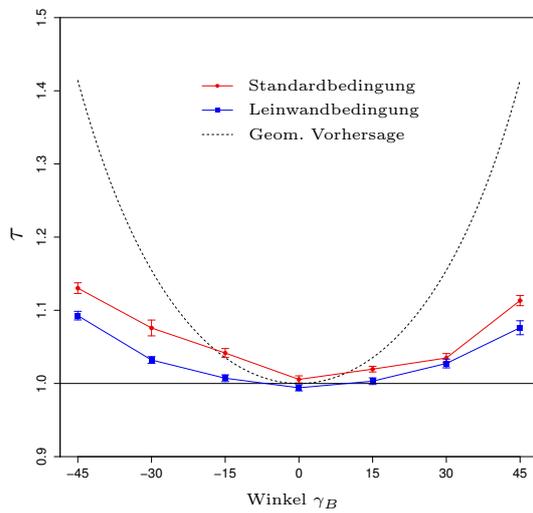
Abbildung A.9. Mittlere Einstellungen von  $\tau$  für Versuchsperson NS unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



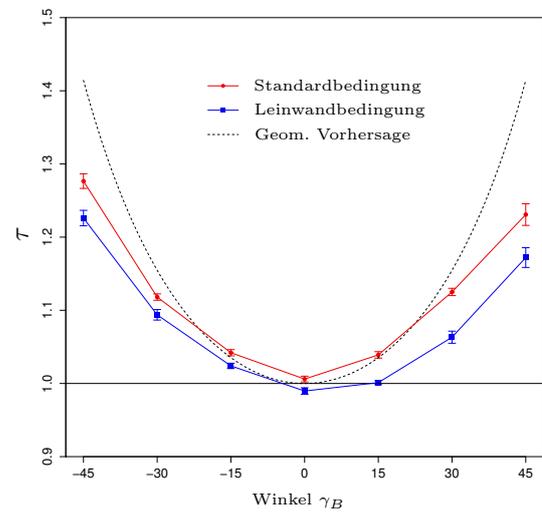
(a) Standardbedingung



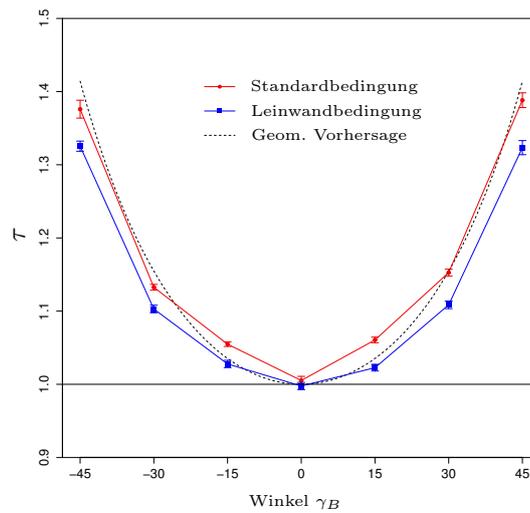
(b) Leinwandbedingung



(c) Binokulare Sehbedingung



(d) Monokulare Sehbedingung



(e) Aperturbedingung

Abbildung A.10. Mittlere Einstellungen von  $\tau$  für Versuchsperson SS unter der Standard- (a) und Leinwandbedingung (b) sowie getrennt nach den verschiedenen Sehbedingungen (c) - (e). Die halbe Länge der Fehlerbalken repräsentiert den Standardfehler.



# Literaturverzeichnis

- Alhazen. (2001). *Alhacen's theory of visual perception: A critical edition, with English translation and commentary, of the first three books of Alhacen's De aspectibus, the medieval Latin version of Ibn al-Haytham's Kitāb-al-Manāzīr* (Bd. 2; A. M. Smith, Hrsg.). Philadelphia, PA: American Philosophical Society.
- Allison, R. S. & Howard, I. P. (2000). Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision Research*, 40, 1869-1886.
- Andersen, K. (2007). *The geometry of an art. The history of the mathematical theory of perspective from Alberti to Monge*. New York, NY: Springer.
- Arditti, A. (1986). Binocular vision. In *Handbook of perception and human performance* (Bd. 1, S. 23.1-23.41). New York, NY: Wiley.
- Ash, R. B. (1990). *Information theory*. New York, NY: Dover.
- Atherton, M. (1990). *Berkeley's revolution in vision*. Ithaca, NY: Cornell University Press.
- Atherton, M. (2002). The origins of the sensation/perception distinction. In D. Heyer & R. Mausfeld (Hrsg.), *Perception and the physical world: Psychological and philosophical issues in perception* (S. 3-20). Chichester, UK: Wiley.
- Atkins, J. E., Fiser, J. & Jacobs, R. A. (2001). Experience-dependent visual cue integration based on consistencies between visual and haptic percepts. *Vision Research*, 41, 449-461.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61, 183-193.
- Augustinus, A. (1989). *Bekenntnisse* (K. Flasch & B. Mojsisch, Hrsg.). Stuttgart: Reclam.
- Baker, G. P. & Hacker, P. M. S. (1982). The grammar of psychology: Wittgenstein's Bemerkungen über die Philosophie der Psychologie. *Language & Communication*, 2 (3), 227-244.
- Barlow, H. (2002). Cognition as code-breaking. In D. Heyer & R. Mausfeld (Hrsg.), *Perception and the physical world* (S. 21-36). Chichester, UK: Wiley.
- Bauer, H. (1911). *Die Psychologie Alhazens*. Zugriff auf <http://www.archive.org>
- Bennett, M. R., Dennett, D. C., Hacker, P. M. S. & Searle, J. R. (Hrsg.). (2007).

- Neuroscience and philosophy*. New York, NY: Columbia University Press.
- Bennett, M. R. & Hacker, P. M. S. (2001). Perception and memory in neuroscience: A conceptual analysis. *Progress in Neurobiology*, 65, 499-543.
- Ben-Zeev, A. (1987). A critique of the inferential paradigm in perception. *Journal for the Theory of Social Behaviour*, 17, 243-263.
- Berkeley, G. (1987). *Versuch über eine neue Theorie des Sehens und Die Theorie des Sehens oder der visuellen Sprache verteidigt und erklärt* (W. Breidert, Hrsg.). Hamburg: Meiner.
- Bertamini, M., Martinovic, J. & Wuerger, S. M. (2008). Integration of ordinal and metric cues in depth processing. *Journal of Vision*, 8 (2).
- Birbaumer, N. & Schmidt, R. (2010). *Biologische Psychologie* (7. Aufl.). Berlin: Springer.
- Bogdan, R. J. (1994). *Grounds for cognition: How goal-guided behavior shapes the mind*. Hillsdale, NJ: Erlbaum.
- Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. New York, NY: Appleton-Century-Crofts.
- Braddick, O. J. & Atkinson, J. (1982). Higher functions in vision. In H. B. Barlow & J. D. Mollon (Hrsg.), *The senses* (S. 212-238). Cambridge, UK: Cambridge University Press.
- Brainard, D. H. (1989). Calibration of a computer controlled color monitor. *Color Research and Application*, 14, 23-34.
- Brener, M. E. (2004). *Vanishing points*. Jefferson, NC: McFarland.
- Bruce, V., Green, P. R. & Georgeson, M. A. (2003). *Visual perception* (4. Aufl.). New York, NY: Psychology Press.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.
- Brunswik, E. & Kamiya, J. (2001). Ecological cue-validity of ‚proximity‘ and other Gestalt factors. In K. R. Hammond & T. R. Stewart (Hrsg.), *The essential Brunswik* (S. 213-225). New York, NY: Oxford University Press. (Originalausgabe 1953)
- Burge, J., Peterson, M. A. & Palmer, S. E. (2005). Ordinal configural cues combine with metric disparity in depth perception. *Journal of Vision*, 5, 534-542.
- Burton, H. E. (1945). The optics of Euclid. *Journal of the Optical Society of America*, 35 (5), 357-372.
- Busey, T. A., Brady, N. P. & Cutting, J. E. (1990). Compensation is unnecessary for the perception of faces in slanted pictures. *Perception & Psychophysics*, 48, 1-11.
- Capurro, R. & Hjørland, B. (2003). The concept of information. *Annual Review of Information Science and Technology*, 37 (1), 343-411.

- Chalmers, D. (2003). Consciousness and its place in nature. In S. P. Stich & T. A. Warfield (Hrsg.), *The Blackwell guide to philosophy of mind*. Malden, MA: Blackwell.
- Chen, Z. & Cave, K. R. (2014). Constraints on dilution from a narrow attentional zoom reveal how spatial and color cues direct selection. *Vision Research*, *101*, 125-137.
- Churchland, P. (1986). *Neurophilosophy. Toward a unified science of the mind/brain*. Cambridge, MA: MIT Press.
- Clark, J. J. & Yuille, A. L. (1990). *Data fusion for sensory information processing systems*. Boston, MA: Kluwer.
- Cole, C. (1993). Shannon revisited: Information in terms of uncertainty. *Journal of the American Society for Information Science*, *44* (4), 204-211.
- Colombo, M. & Seriès, P. (2012). Bayes in the brain: On Bayesian modeling in neuroscience. *British Journal for the Philosophy of Science*, *63* (3), 697-723.
- Coombs, C. H., Dawes, R. M. & Tversky, A. (1970). *Mathematical psychology. An elementary introduction*. Englewood Cliffs, NJ: Prentice-Hall.
- Copenhaver, R. (2014). Berkeley on the language of nature and the objects of vision. *Res Philosophica*, *91*, 29-46.
- Coren, S., Ward, L. M. & Ennis, J. T. (1999). *Sensation and perception* (5. Aufl.). Fort Worth, TX: Harcourt Brace.
- Cornsweet, T. (1962). The staircase-method in psychophysics. *American Journal of Psychology*, *75*, 485-491.
- Cutting, J. E. & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Hrsg.), *Perception of space and motion* (2. Aufl., S. 71-118). San Diego, CA: Academic Press.
- Damisch, H. (2010). *Der Ursprung der Perspektive*. Zürich: Diaphanes.
- de La Gournerie, J. (1859). *Traité de perspective linéaire contenant les tracés pour les tableaux plans et courbes, les bas-reliefs et les décorations théâtrales, avec une théorie des effets de perspective*. Zugriff auf <http://www.archive.org>
- Dennett, D. C. (1981). *Brainstorms*. Cambridge, MA: Bradford Books.
- Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT Press.
- Dennett, D. C. (1991). *Consciousness explained*. New York, NY: Back Bay Books.
- Dennett, D. C. (1996). *Content and consciousness* (2. Aufl.). London: Routledge. (Originalausgabe 1969)
- Descartes, R. (1972). *Meditationen über die Grundlagen der Philosophie mit den sämtlichen Einwänden und Erwidernungen* (4. Aufl.; A. Buchenau, Hrsg. & Übers.). Hamburg: Meiner. (Originalausgabe 1641)
- Descartes, R. (1996). *Œuvres complètes* (Bd. VII. Meditationes de prima philosophia;

- C. Adam & P. Tannery, Hrsg.). Paris: Vrin. (Originalausgabe 1641)
- Descartes, R. (2001). *Discourse on method, Optics, Geometry, and Meteorology* (P. J. Olscamp, Übers.). Indianapolis, IN: Hackett Publishing Company. (Originalausgabe 1637)
- Dretske, F. (1983). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.
- du Bois-Reymond, E. (1912). Gedächtnisrede auf Hermann von Helmholtz. Gehalten in der Leibniz-Stiftung der Akademie der Wissenschaften am 4. Juli 1895. In E. du Bois Reymond (Hrsg.), *Reden von Emil du Bois-Reymond in zwei Bänden* (2. Aufl., Bd. 2). Zugriff auf <http://www.echo.mpiwg-berlin.mpg.de>
- Eastwood, B. (1986). Alhazen, Leonardo, and late-medieval speculation on the inversion of images in the eye. *Annals of Science*, 43, 413-446.
- Eisler, R. (1904). *Wörterbuch der philosophischen Begriffe und Ausdrücke* (2. Aufl.). Zugriff auf <http://www.archive.org>
- Epstein, W. (1995). The metatheoretical context. In W. Epstein & S. Rogers (Hrsg.), *Perception of space and motion* (2. Aufl., S. 1-22). San Diego, CA: Academic Press.
- Erdmann, B. (1921). Die philosophischen Grundlagen von Helmholtz' Wahrnehmungstheorie, kritisch erläutert. *Abhandlungen der preussischen Akademie der Wissenschaften. Philosophisch-historische Klasse*, 1, 1-45. Zugriff auf <http://www.archive.org>
- Falmagne, J. (1985). *Elements of psychophysical theory*. Oxford, UK: Oxford University Press.
- Feigl, H. (1967). *The „mental“ and the „physical“*. Minneapolis, MN: University of Minnesota Press. (Originalausgabe 1958)
- Figlio, K. M. (1975). Theories of perception and the physiology of mind in the late eighteenth century. *History of Science*, 13 (3), 177-212.
- Fisher, C., Hahn, A. C., DeBruine, L. M. & Jones, B. C. (2014). Integrating shape cues of adiposity and color information when judging facial health and attractiveness. *Perception*, 43, 499-508.
- Fodor, J. A. (1975). *The language of thought*. New York, NY: Crowell.
- Fodor, J. A. (1981). The mind-body problem. *Scientific American*, 244, 114-125.
- Fodor, J. A. (1983). *RePresentations. Philosophical essays on the foundations of cognitive science*. Cambridge, MA: MIT Press.
- Fodor, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.
- Fodor, J. A. & Pylyshyn, Z. W. (1981). How direct is visual perception?: Some reflections on Gibson's „ecological approach“. *Cognition*, 9, 139-196.
- Foley, H. J. & Matlin, M. W. (2010). *Sensation and perception* (5. Aufl.). Boston,

- MA: Allyn & Bacon.
- Frisby, J. P. (1980). *Seeing: Illusion, brain and mind*. Oxford, UK: Oxford University Press.
- Fullinwider, S. P. (1990). Hermann von Helmholtz: The problem of Kantian influence. *Studies in History and Philosophy of Science*, 21 (1), 41-55.
- Gaukroger, S., Schuster, J. & Sutton, J. (Hrsg.). (2000). *Descartes' natural philosophy*. London: Routledge.
- Gelb, A. (1929). Die „Farbenkonstanz“ der Sehdinge. In A. Bethe, G. v. Bergmann, G. Embden & A. Ellinger (Hrsg.), *Handbuch der normalen und pathologischen Physiologie* (Bd. 12, S. 594-678). Berlin: Springer.
- Gibson, J. J. (1986). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin. (Originalausgabe 1979)
- Gilinsky, A. S. (1955). The effect of attitude upon the perception of size. *The American Journal of Psychology*, LXVIII (2), 173-192.
- Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 256-266.
- Goldstein, E. B. (1999). *Sensation and perception* (5. Aufl.). Pacific Grove, CA: Brooks/Cole.
- Goldstein, E. B. (Hrsg.). (2010). *Encyclopedia of perception*. Los Angeles, CA: Sage.
- Gombrich, E. H. (1960). *Art and illusion*. Princeton, NJ: Princeton University Press.
- Goodman, N. (1976). *Languages of art*. Indianapolis, IN: Hackett.
- Gordon, I. E. (1997). *Theories of visual perception* (2. Aufl.). Chichester, UK: Wiley.
- Graham, C. H. (Hrsg.). (1965). *Vision and visual perception*. Chichester, UK: Wiley.
- Gårding, J. (1992). Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision*, 2, 329-352.
- Greco, J. (1995). Reid's critique of Berkeley and Hume: What's the big idea? *Philosophy and Phenomenological Research*, 55 (2), 279-296.
- Greenberg, J. (1963). *Universals of language*. Cambridge, MA: MIT Press.
- Gregory, R. L. (1998). *Eye and brain: The psychology of seeing* (5. Aufl.). Oxford: Oxford University Press.
- Gregory, R. L. (2002). Perceptions as hypotheses. In A. Noë & E. Thompson (Hrsg.), *Vision and mind* (S. 111-134). Cambridge, MA: MIT Press.
- Grush, R. (2007). Berkeley and the spatiality of vision. *Journal of the History of Philosophy*, 45 (3), 413-442.
- Gurwitsch, A. (1966). *Studies in phenomenology and psychology*. Evanston, IL: Northwestern University Press.
- Hacker, P. M. S. (o. J.). *The relevance of Wittgenstein's philosophy of psychology to*

- the psychological sciences*. St. John's College, Oxford, UK. Zugriff auf <http://www.info.sjc.ox.ac.uk/scr/hacker/docs>
- Hacker, P. M. S. (1991). Experimental methods and conceptual confusion: An investigation into R. L. Gregory's theory of perception. *The Jerusalem Philosophical Quarterly*, 40, 289-314.
- Hacker, P. M. S. (1995). Helmholtz's theory of perception: An investigation into its conceptual framework. *International Studies in the Philosophy of Science*, 9 (3), 199-214.
- Hagen, M. A. (1976). Influence of picture surface and station point on the ability to compensate for oblique view in pictorial perception. *Developmental Psychology*, 12, 57-63.
- Hall, G. S. (1912). *Founders of modern psychology*. Zugriff auf <http://www.archive.org>
- Harper, R. S. & Boring, E. G. (1948). Cues. *American Journal of Psychology*, 61, 343-351.
- Hatfield, G. (1990). *The natural and the normative. Theories of spatial perception from Kant to Helmholtz*. Cambridge, MA: MIT Press.
- Haugeland, J. (1991). Representational genera. In W. M. Ramsey, S. P. Stich & D. E. Rumelhart (Hrsg.), *Philosophy and connectionist theory* (S. 61-90). Hillsdale, NJ: Erlbaum.
- Hecht, E. (2009). *Optik* (5. Aufl.). München: Oldenbourg.
- Heidelberger, M. (1993). Räumliches Sehen bei Helmholtz und Hering. *Philosophia Naturalis*, 30, 1-28.
- Heidelberger, M. (1995). Helmholtz als Philosoph. *Deutsche Zeitschrift für Philosophie*, 43, 835-44.
- Heidelberger, M. (1997). Beziehungen zwischen Sinnesphysiologie und Philosophie im 19. Jahrhundert. In H. J. Sandkühler (Hrsg.), *Philosophie und Wissenschaften: Formen und Prozesse ihrer Interaktion* (S. 37-58). Frankfurt a.M.: Lang.
- Heimann, P. (1974). Helmholtz and Kant: The metaphysical foundations of „Über die Erhaltung der Kraft“. *Studies in History and Philosophy of Science*, 5, 205-238.
- Helmholtz, H. L. F. v. (1998). Die Tatsachen in der Wahrnehmung. In E. Bonk (Hrsg.), *Schriften zur Erkenntnistheorie. Kommentiert von Moritz Schlick und Paul Hertz* (Bd. 2). Wien: Springer Verlag. (Originalausgabe 1896)
- Helmholtz, H. L. F. v. (2002a). Die neueren Fortschritte in der Theorie des Sehens. In J. Brüning (Hrsg.), *Gesammelte Schriften* (Bd. V. 1). Hildesheim: Olms-Weidmann. (Originalausgabe 1896)
- Helmholtz, H. L. F. v. (2002b). Die Thatsachen in der Wahrnehmung. In J. Brüning (Hrsg.), *Gesammelte Schriften* (Bd. V. 2). Hildesheim: Olms-Weidmann.

- (Originalausgabe 1896)
- Helmholtz, H. L. F. v. (2002c). Erinnerungen. In J. Brüning (Hrsg.), *Gesammelte Schriften* (Bd. V.1). Hildesheim: Olms-Weidmann. (Originalausgabe 1896)
- Helmholtz, H. L. F. v. (2002d). Ueber das Sehen des Menschen. In J. Brüning (Hrsg.), *Gesammelte Schriften* (Bd. V. 1). Hildesheim: Olms-Weidmann. (Originalausgabe 1896)
- Helmholtz, H. L. F. v. (2003a). Handbuch der Physiologischen Optik. Dritter Band. In J. Brüning (Hrsg.), *Gesammelte Schriften* (Bd. III. 3). Hildesheim: Olms-Weidmann. (Originalausgabe 1910)
- Helmholtz, H. L. F. v. (2003b). Ueber die Natur der menschlichen Sinnesempfindungen. In J. Brüning (Hrsg.), *Gesammelte Schriften* (Bd. I. 2. 2). Hildesheim: Olms-Weidmann. (Originalausgabe 1883)
- Hershenson, M. (1999). *Visual space perception*. Cambridge, MA: MIT Press.
- Hertz, H. (1894). Die Prinzipien der Mechanik in neuem Zusammenhange dargestellt. In *Gesammelte Werke* (Bd. III). Zugriff auf <http://www.archive.org>
- Hillis, J. M., Watt, S. J., Landy, M. S. & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, 4, 967-992.
- Hoffman, D. D. (2000). *Visuelle Intelligenz*. Stuttgart: Klett-Cotta.
- Hofmann, H. (1913). Untersuchungen über den Empfindungsbegriff. *Archiv für die Gesamte Psychologie*, XXVI, 1-136. Zugriff auf <http://www.archive.org>
- Howard, I. P. (2002). Depth perception. In H. Pashler (Hrsg.), *Stevens' handbook of experimental psychology* (3. Aufl., Bd. 1, S. 77-120). New York, NY: Wiley.
- Howard, I. P. (2012). *Perceiving in depth* (Bd. 1-3). New York, NY: Oxford University Press.
- Husserl, E. (1991). *Ding und Raum* (E. Marbach, Hrsg.). Hamburg: Meiner. (Originalausgabe 1973)
- Husserl, E. (2003). *Phänomenologische Psychologie* (D. Lohmar, Hrsg.). Hamburg: Meiner. (Originalausgabe 1968)
- Husserl, E. (2004). *Wahrnehmung und Aufmerksamkeit. Texte aus dem Nachlass (1893-1912)* (T. Vongehr & R. Giuliani, Hrsg.). Dordrecht, NL: Springer.
- Husserl, E. (2006). *Phantasie und Bildbewußtsein* (E. Marbach, Hrsg.). Hamburg: Meiner. (Originalausgabe 1980)
- Husserl, E. (2009). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie* (E. Marbach, Hrsg.). Hamburg: Meiner. (Originalausgabe 1913)
- Ishihara, S. (1998). *The series of plates designed as a test for colour-deficiency*. Tokyo: Kanehara.
- Jacobs, R. A. (1998). Visual cue integration for depth perception. In D. A. Rosenbaum & C. E. Collyer (Hrsg.), *Timing of behavior* (S. 61-76). Cambridge, MA: MIT

- Press.
- Jacobs, R. A. (2002). What determines cue reliability? *Trends in Cognitive Sciences*, 6 (8), 345-350.
- James, W. (2007). *The principles of psychology*. New York, NY: Cosimo. (Originalausgabe 1890)
- Johnson-Laird, P. N. (1988). *The computer and the mind*. Cambridge, MA: Harvard University Press.
- Johnston, E. B., Cumming, B. G. & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. *Vision Research*, 33 (5), 813-826.
- Kalkofen, H. (1992). Anmerkungen zur Empirismus-Nativismus-Kontroverse. *Psychologie und Geschichte*, 3, 24-45.
- Kant, I. (1998). *Kritik der reinen Vernunft* (J. Timmermann, Hrsg.). Hamburg: Meiner. (Originalausgabe 1781)
- Katz, D. (1930). *Der Aufbau der Farbwelt*. Leipzig: J. A. Barth.
- Keil, G. (2003). Über den Homunkulus-Fehlschluß. *Zeitschrift für philosophische Forschung*, 57, 1-26.
- Kemp, M. (1990). *The science of an art. Optical themes in western art from Brunelleschi to Seurat*. New Haven, CT: Yale University Press.
- Kenny, A. (1981). The homunculus fallacy. In *Interpretations of life and mind* (S. 155-165). New York, NY: Humanities Press.
- Kersten, D. & Schrater, P. R. (2002). Pattern inference theory: A probabilistic approach to vision. In D. Heyer & R. Mausfeld (Hrsg.), *Perception and the physical world* (S. 150-158). Chichester, UK: Wiley.
- Kersten, D. & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology*, 13, 150-158.
- Kitazaki, M., Kobiki, H. & Maloney, L. T. (2008). Effect of pictorial depth cues, binocular disparity cues and motion parallax depth cues on lightness perception in three-dimensional virtual scenes. *PLoS ONE*, 3 (9), e3177. Zugriff auf <http://www.plosone.org>
- Knill, D. C. (2007a). Bayesian models of sensory cue integration. In K. Doya, S. Ishii, A. Pouget & R. Rao (Hrsg.), *Bayesian brain. Probabilistic approaches to neural coding* (S. 189-206). Cambridge, MA: MIT Press.
- Knill, D. C. (2007b). Learning Bayesian priors for depth perception. *Journal of Vision*, 7 (8), 1-20.
- Knill, D. C. (2007c). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision*, 7, 1-24.
- Knill, D. C. & Richards, W. (Hrsg.). (1996). *Perception as Bayesian inference*. Cam-

- bridge, UK: Cambridge University Press.
- Koenderink, J. J., van Dorn, A. J. & Kappers, A. M. L. (1995). Depth relief. *Perception*, 24, 115-126.
- Koenigsberger, L. (1902-1903). *Hermann von Helmholtz* (Bd. 1-3). Zugriff auf <http://www.archive.org>
- Koffka, K. (2001). *Principles of gestalt psychology*. London, UK: Routledge. (Originalausgabe 1935)
- Krüger, L. (Hrsg.). (1994). *Universalgenie Helmholtz. Rückblick nach 100 Jahren*. Berlin: Akademie-Verlag.
- Kubovy, M. (1988). *The psychology of perspective and renaissance art*. Cambridge, UK: Cambridge University Press.
- Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, 35, 389-412.
- Lange, K. (1906). *Über Apperzeption*. Zugriff auf <http://www.archive.org>
- Leek, M. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, 63 (8), 1279-1292.
- Lenoir, T. (1993). The eye as mathematician. In D. Cahan (Hrsg.), *Hermann von Helmholtz and the foundations of nineteenth-century science* (S. 109-153). Berkeley, CA: University of California Press.
- Levine, M. W. (2000). *Levine and Shefner's fundamentals of sensation and perception* (3. Aufl.). Oxford, UK: Oxford University Press.
- Lindberg, D. C. (1967). Alhazen's theory of vision and its reception in the west. *Isis*, 58 (3), 321-341.
- Lindberg, D. C. (1976). *Theories of vision from al-Kindi to Kepler*. Chicago, IL: University of Chicago Press.
- Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Zugriff auf <http://www.archive.org>
- Lotze, R. H. (1881). *Grundzüge der Psychologie*. Zugriff auf <http://www.archive.org>
- Lotze, R. H. (1884). *Mikrokosmos*. Zugriff auf <http://www.archive.org>
- Luce, R. D. (2003). Whatever happened to information theory in psychology? *Review of General Psychology*, 7 (2), 183-188.
- Machlup, F. (1983). Semantic quirks in studies of information. In F. Machlup & U. Mansfield (Hrsg.), *The study of information: Interdisciplinary messages* (S. 641-671). New York, NY: Wiley.
- MacKenzie, K. J., Murray, R. F. & Wilcox, L. M. (2008). The intrinsic constraint approach to cue combination: An empirical and theoretical evaluation. *Journal*

- of Vision*, 8 (8), 1-10.
- Malebranche, N. (1968). *Von der Erforschung der Wahrheit* (A. Klemmt, Hrsg.). Hamburg: Meiner. (Originalausgabe 1674)
- Mamassian, P., Landy, M. & Maloney, L. T. (2002). Bayesian modeling of visual perception. In R. P. N. Rao, B. A. Olshausen & M. S. Lewicki (Hrsg.), *Probabilistic models of the brain* (S. 13-36). Cambridge, MA: MIT Press.
- Mamassian, P. & Landy, M. S. (1998). Observer biases in the 3d interpretation of line drawings. *Vision Research*, 38, 2817-2832.
- Marr, D. (1980). Visual information processing: The structure and creation of visual representations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 290 (1038), 199-218.
- Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.
- Marr, D. & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London B*, 204, 301-328.
- Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. . P. & Martin, K. E. (1996). Occlusion edge blur: A cue to relative visual depth. *Journal of the Optical Society of America*, 13, 681-688.
- Mather, G. (2009). *Foundations of sensation and perception* (2. Aufl.). Hove, East Sussex: Psychology Press.
- Mausfeld, R. (1994). Hermann v. Helmholtz. Die Untersuchung der Funktionsweise des Geistes als Gegenstand einer wissenschaftlichen Psychologie. *Psychologische Rundschau*, 45, 133-147.
- Mausfeld, R. (2003). Conjoint representations and the mental capacity for multiple simultaneous perspectives. In H. Hecht, R. Schwartz & M. Atherton (Hrsg.), *Looking into pictures. An interdisciplinary approach to pictorial space* (S. 17-60). Cambridge, MA: MIT Press.
- Mausfeld, R. (2010). Intrinsic multiperspectivity: On the architectural foundations of a distinctive mental capacity. In P. Frensch & R. Schwarzer (Hrsg.), *Cognition and neuropsychology: International perspectives on psychological science* (Bd. 1, S. 95-116). Sussex, UK: Psychology Press.
- Mausfeld, R. (2011). Intrinsic multiperspectivity: Conceptual forms and the functional architecture of the perceptual system. In W. Welsch, W. Singer & A. Wunder (Hrsg.), *Interdisciplinary Anthropology* (S. 19-54). Berlin: Springer.
- Maxwell, J. C. (1890). On colour vision. In W. D. Niven (Hrsg.), *The scientific papers of James Clerk Maxwell* (Bd. 2). Cambridge, UK: At the University Press. (Originalausgabe 1857)
- McGowan, W. (1982). Berkeley's doctrine of signs. In C. M. Turbayne (Hrsg.), *Berkeley: Critical and interpretive essays* (S. 231-246). Minneapolis, MN: University

- of Minnesota Press.
- Merleau-Ponty, M. (1966). *Phänomenologie der Wahrnehmung* (R. Boehm, Hrsg. & Übers.). Berlin: De Gruyter.
- Merleau-Ponty, M. (2003). Die Natur der Wahrnehmung (J. Schröder, Übers.). In L. Wiesing (Hrsg.), *Das Primat der Wahrnehmung* (S. 10-25). Frankfurt a.M.: Suhrkamp. (Originalausgabe 1934)
- Meyering, T. C. (1989). *Historical roots of cognitive science*. Dordrecht: Kluwer Academic.
- Mill, J. S. (1882). *A system of logic, ratiocinative and inductive* (8. Aufl.). Zugriff auf <http://www.archive.org>
- Mon-Williams, M., Tresilian, J. R. & Roberts, A. (2000). Vergence provides veridical depth perception from horizontal retinal image disparities. *Experimental Brain Research*, 133, 407-413.
- Muller, C. M. P., Brenner, E. & Smeets, J. B. J. (2009). Different cue weights at the same place. *Journal of Vision*, 9 (11), 1-5.
- Müller, J. P. (1837). *Handbuch der Physiologie des Menschen für Vorlesungen* (3. Aufl., Bd. 1). Zugriff auf <http://www.archive.org>
- Müller, J. P. (1840). *Handbuch der Physiologie des Menschen für Vorlesungen* (Bd. 2). Zugriff auf <http://www.archive.org>
- Nadler, J. W., Angelaki, D. E. & DeAngelis, G. C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature*, 452, 642-645.
- Neisser, U. (2014). *Cognitive psychology*. New York, NY: Psychology Press. (Originalausgabe 1967)
- Neumann, J. v. (1993). First draft of a report on the EDVAC. *IEEE Annals of the History of Computing*, 15, 27-75. (Originalausgabe 1945)
- Newell, A. (1980). Physical symbol systems. *Cognitive Science*, 4, 135-183.
- Nichols, R. (2007). *Thomas Reid's theory of perception*. New York, NY: Oxford University Press.
- Niederée, R. & Heyer, D. (2003). The dual nature of picture perception: A challenge to current general accounts of visual perception. In H. Hecht, R. Schwartz & M. Atherton (Hrsg.), *Looking into pictures. An interdisciplinary approach to pictorial space* (S. 77-98). Cambridge, MA: MIT Press.
- North, J. D. (2002). *The ambassador's secret*. London, UK: Hambledon and London.
- O'Shea, R. P., Blackburn, S. G. & Ono, H. (1994). Contrast as a depth cue. *Vision Research*, 34 (12), 1595-1604.
- Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. Lloyd (Hrsg.), *Cognition and categorization* (S. 259-303). Hillsdale, NJ: Erlbaum.

- Palmer, S. E. (1999). *Vision science*. Cambridge, MA: MIT Press.
- Panofsky, E. (1980). Die Perspektive als „symbolische Form“. In H. Oberer & E. Verheyen (Hrsg.), *Aufsätze zu Grundfragen der Kunstwissenschaft* (S. 99-167). Berlin: Verlag Volker Spiess. (Originalausgabe 1927)
- Perkins, D. N. (1973). Compensating for distortion in viewing pictures obliquely. *Perception & Psychophysics*, *14*, 13-18.
- Pester, R. (1997). *Hermann Lotze. Wege seines Denkens und Forschens*. Würzburg: Königshausen & Neumann.
- Pinker, S. (1984). Visual cognition: An introduction. *Cognition*, *18*, 1-63.
- Pirenne, M. H. (1970). *Optics, painting and photography*. Cambridge, UK: Cambridge University Press.
- Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision Research*, *41* (24), 3145-3161.
- Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology*, *47*, 44-50.
- Plümacher, M. (2004). *Wahrnehmung, Repräsentation und Wissen. Edmund Husserls und Ernst Cassirers Analysen zur Struktur des Bewusstseins*. Berlin: Parerga.
- Poggio, T. (1981). Marr's computational approach to vision. *Trends in Neurosciences*, *4* (10), 258-262.
- Polanyi, M. (1970). What is a painting? *British Journal of Aesthetics*, *10*, 225-236.
- Popper, K. & Eccles, J. C. (1987). *Das Ich und sein Gehirn*. München: Piper. (Originalausgabe 1977)
- Postman, L. & Tolman, E. C. (1958). Brunswik's probabilistic functionalism. In S. Koch (Hrsg.), *Psychology: A study of a science* (Bd. 1, S. 502-564). New York, NY: McGraw-Hill.
- Putnam, H. (1981). *Reason, truth and history*. Cambridge, UK: Cambridge University Press.
- Quine, W. V. O. (2003). Naturalisierte Erkenntnistheorie. In *Ontologische Relativität und andere Schriften* (S. 85-106). Frankfurt: Klostermann.
- Ramenahalli, S., Mihalas, S. & Niebur, E. (2014). Local spectral anisotropy is a valid cue for figure-ground organisation in natural scenes. *Vision Research*, *103*, 116-126.
- Ramsey, W. M. (2007). *Representation reconsidered*. Cambridge, UK: Cambridge University Press.
- Reid, T. (1813). An inquiry into the human mind. In D. Stewart (Hrsg.), *The works of Thomas Reid* (Bd. 1). Zugriff auf <http://www.archive.org>
- Reid, T. (1814). Essays on the intellectual powers of man. In D. Stewart (Hrsg.), *The works of Thomas Reid* (Bd. 2). Zugriff auf <http://www.archive.org>

- Richeimer, J. (2006). Familiarity and the inferential theory of perception. *Theory & Psychology*, 16 (4), 505-525.
- Richter, J. P. (1883). *The literary works of Leonardo da Vinci* (Bd. 1). Zugriff auf <http://www.archive.org>
- Riehl, A. (1879). *Der Philosophische Kriticismus und seine Bedeutung für die positive Wissenschaft* (Bd. 2). Zugriff auf <http://www.archive.org>
- Riehl, A. (1904). *Hermann von Helmholtz in seinem Verhältnis zu Kant*. Zugriff auf <http://www.archive.org>
- Rogers, S. (1995). Perceiving pictorial space. In W. Epstein & S. Rogers (Hrsg.), *Perception of space and motion* (2. Aufl., S. 119-164). San Diego, CA: Academic Press.
- Rosinski, R. R. & Farber, J. (1980). Compensating for viewing point in the perception of pictured space. In M. A. Hagen (Hrsg.), *The perception of pictures* (S. 137-176). New York, NY: Academic Press.
- Rosinski, R. R., Mulholland, T. & Degelmann, D. (1980). Picture perception: An analysis of visual compensation. *Perception & Psychophysics*, 28, 521-526.
- Roth, G. (2001). *Fühlen. Denken, Handeln. Wie das Gehirn unser Verhalten steuert*. Frankfurt a.M.: Suhrkamp.
- Schiemann, G. (1994). Die Hypothesisierung des Mechanismus bei Hermann von Helmholtz. Ein Beitrag zum Wandel der Wissenschafts- und Naturauffassung im 19. Jahrhundert. In L. Krüger (Hrsg.), *Universalgenie Helmholtz. Rückblick nach 100 Jahren* (S. 93-107). Berlin: Akademie-Verlag.
- Schiemann, G. (2008). Verlust der Welt im Bild: Ursprung und Entwicklung des Bildbegriffes bei Hermann von Helmholtz und Heinrich Hertz. In G. Wolf-schmidt (Hrsg.), *Heinrich Hertz (1857-1894) and the development of communication* (S. 31-51). Norderstedt: Nuncius Hamburgensis.
- Schiemann, G. (2014). Hermann von Helmholtz' Kantkritik. In C. Krijnen & K. W. Zeidler (Hrsg.), *Wissenschaftsphilosophie im Neukantianismus* (S. 199-232). Würzburg: Königshausen & Neumann.
- Schlick, M. (1998). Die Tatsachen in der Wahrnehmung. Erläuterungen. In E. Bonk (Hrsg.), *Hermann von Helmholtz. Schriften zur Erkenntnistheorie* (Bd. 2). Wien: Springer.
- Schlick, M. (2012). Helmholtz als Erkenntnistheoretiker. In F. Stadler & H. J. Wendel (Hrsg.), *Moritz Schlick Gesamtausgabe* (Bd. 5 Abteilung 1). Wien: Springer. (Originalausgabe 1922)
- Schlosberg, H. (1941). Stereoscopic depth from single pictures. *American Journal of Psychology*, 54, 601-605.
- Schlüter, N. & Faul, F. (2014). Are optical distortions used as a cue for material

- properties of thick transparent objects? *Journal of Vision*, 14.
- Schmitz, H. (2005). *Die Wahrnehmung. System der Philosophie III, Teil 5*. Bonn: Bouvier. (Originalausgabe 1978)
- Searle, J. R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3 (3), 417-424.
- Searle, J. R. (1982). The Chinese room revisited. *Behavioral and Brain Sciences*, 5 (2), 345-348.
- Searle, J. R. (1990). Is the brain a digital computer? *Proceedings and Addresses of the American Psychological Association*, 64, 21-37.
- Sedgwick, H. A. (1980). The geometry of spatial layout in pictorial representation. In M. A. Hagen (Hrsg.), *The perception of pictures* (S. 33-90). New York, NY: Academic Press.
- Sedgwick, H. A. (2003). Relating direct and indirect perception of spatial layout. In H. Hecht, R. Schwartz & M. Atherton (Hrsg.), *Looking into pictures. An interdisciplinary approach to pictorial space* (S. 61-75). Cambridge, MA: MIT Press.
- Seydell, A., Knill, D. C. & Trommershäuser, J. (2010). Adapting internal statistical models for interpreting visual cues to depth. *Journal of Vision*, 10 (4), 1-27.
- Shannon, C. E. & Weaver, W. (1962). *The mathematical theory of communication*. Chicago, IL: University of Illinois Press. (Originalausgabe 1949)
- Shepard, R. N. (1978). The mental image. *American Psychologist*, 33 (2), 125-137.
- Shirley, P. & Morley, R. K. (2003). *Realistic ray tracing* (2. Aufl.). Natick, MA: Peters.
- Singer, W. (2002). *Der Beobachter im Gehirn*. Frankfurt a.M.: Suhrkamp.
- Skeat, W. W. (1893). *An etymological dictionary of the English language* (2. Aufl.). Zugriff auf <http://www.archive.org>
- Skorupski, J. (1994). J. S. Mill. Logic and metaphysics. In C. Ten (Hrsg.), *Routledge history of philosophy* (Bd. VII, S. 81-100). London, UK: Routledge.
- Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, 68, 141-156.
- Smit, H. & Hacker, P. M. S. (2014). Seven misconceptions about the mereological fallacy: A compilation for the perplexed. *Erkenntnis*, 79, 1077-1097.
- Smith, A. M. (2005). The Alhacenian account of spatial perception and its epistemological implications. *Arabic Sciences and Philosophy*, 15, 219-240.
- Svarverud, E., Gilson, S. J. & Glennerster, A. (2010). Cue combination for 3D location judgements. *Journal of Vision*, 10 (1), 1-13.
- Taylor, B. (1719). *New principles of linear perspective, or the art of designing on a plane the representations of all sorts of objects, in a more general and simple method than has been done before*. Zugriff auf <http://www.archive.org>
- Thinès, G., Costall, A. & Butterworth, G. (Hrsg.). (1991). *Michotte's experimental*

- phenomenology of perception*. Hillsdale, NJ: Erlbaum.
- Tresilian, J. R., Mon-Williams, M. & Kelly, B. M. (1999). Increasing confidence in vergence as a cue to distance. *Proceedings of the Royal Society of London. Biological Sciences*, 266, 39-44.
- Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 35 (17), 2503-2522.
- Troschianko, T., Montagnon, R., Le Clerk, J., Malbert, E. & Chanteau, P. L. (1991). The role of color as a monocular depth cue. *Vision Research*, 31 (11), 1923-1930.
- Turbayne, C. M. (1963). *George Berkeley: Works on vision*. Indianapolis, IN: Bobbs-Merrill.
- Turing, A. (1950). Computing machinery and intelligence. *Mind*, LIX, 433-460.
- Turner, R. S. (1993). Vision studies in Germany: Helmholtz versus Hering. *Osiris, 2nd Series*, 8, 80-103.
- Turner, R. S. (1994). *In the eye's mind: Vision and the Helmholtz-Hering controversy*. Princeton, NJ: Princeton University Press.
- Ullman, S. (1991). Tacit assumptions in the computational study of vision. In A. Gorea (Hrsg.), *Representations of vision: Trends and tacit assumptions in vision research* (S. 305-318). Cambridge, UK: Cambridge University Press.
- Ullmann, S. (1980). Against direct perception. *Behavioural and Brain Sciences*, 3 (3), 333-381.
- Uphues, G. K. (1888). *Wahrnehmung und Empfindung. Untersuchungen zur empirischen Psychologie*. Zugriff auf <http://www.archive.org>
- Uttal, W. R. (1981). *A taxonomy of visual processes*. Hillsdale, NJ: Erlbaum.
- Vaihinger, H. (1922). *Kommentar zu Kants Kritik der reinen Vernunft* (2. Aufl.). Berlin: Union Deutsche Verlagsgesellschaft.
- van Mierlo, C. M., Brenner, E. & Smeets, J. B. J. (2007). Temporal aspects of cue combination. *Journal of Vision*, 7 (7), 1-11.
- Vishwanath, D., Girshick, A. & Banks, M. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience*, 8 (10), 1401-1410.
- Wallach, H. & Marshall, F. J. (1986). Shape constancy in pictorial representation. *Perception & Psychophysics*, 39, 233-235.
- Wee, C. (2014). Descartes and active perception. In J. F. Silva & M. Yrjönsuuri (Hrsg.), *Active perception in the history of philosophy* (S. 207-222). Heidelberg: Springer.
- Wheatstone, C. (2011). Contributions to the physiology of vision. On some remarkable, and hitherto unobserved, phenomena of binocular vision. In *The scientific papers of Sir Charles Wheatstone*. New York, NY: Cambridge University Press. (Originalausgabe 1879)

- Wiener, N. (1963). *Kybernetik* (2. Aufl.). Düsseldorf: Econ-Verlag. (Originalausgabe 1948)
- Wismeijer, D. A., Erkelens, C. J., van Ee, R. & Wexler, M. (2010). Depth cue combination in spontaneous eye movements. *Journal of Vision*, 10 (6), 1-15.
- Wittgenstein, L. (1984). Bemerkungen über die Philosophie der Psychologie. In *Werkausgabe* (Bd. 7.). Frankfurt a.M.: Suhrkamp.
- Wittgenstein, L. (1999). Philosophische Untersuchungen. In *Werkausgabe* (12. Aufl., Bd. 1.). Frankfurt a.M.: Suhrkamp. (Originalausgabe 1953)
- Wokes, D. & Palmer, P. (2008). *Perspective projection of a spheroid onto an image plane* (Bericht). Guildford, Surrey: Surrey Space Centre Research. Zugriff auf <http://surreyspacecentre.wordpress.com/2013/01/22/>
- Wolf-Devine, C. (2000). Descartes' theory of visual spatial perception. In S. Gaukroger, J. Schuster & J. Sutton (Hrsg.), *Descartes' natural philosophy* (S. 506-523). London: Routledge.
- Wolfe, J. M., Kluender, K. R. & Levi, D. M. (2009). *Sensation and perception* (2. Aufl.). Sunderland, MA: Sinauer.
- Wundt, W. M. (1862). *Beiträge zur Theorie der Sinneswahrnehmung*. Leipzig: C.F. Winter.
- Yantis, S. (2014). *Sensation and perception*. New York, NY: Worth Publishers.
- Yeung, R. W. (2002). *A first course in information theory*. New York, NY: Springer.
- Zeki, S. (1993). *A vision of the brain*. Cambridge, MA: Blackwell.
- Zhao, Y. & Heinke, D. (2014). What causes IOR? Attention or perception? – Manipulating cue and target luminance in either blocked or mixed condition. *Vision Research*, 105, 37-46.