

# Visuelle Informationsverarbeitung

*Jochen Müsseler*

<b>2.1</b>	<b>Einleitung: Fragen der visuellen Wahrnehmungsforschung – 14</b>
<b>2.2</b>	<b>Das Auge und die visuellen Verarbeitungspfade – 16</b>
2.2.1	Die Retina – 16
2.2.2	Die Bahn des Sehnervs zwischen Auge und Cortex – 17
2.2.3	Der primäre visuelle Cortex – 17
2.2.4	Weitere corticale Verarbeitungspfade – 17
<b>2.3</b>	<b>Visuelle Informationsaufnahme und -verarbeitung – 19</b>
2.3.1	Visuelle Sehschärfe und Sensitivität – 19
2.3.2	Farbwahrnehmung – 23
2.3.3	Raum- und Tiefenwahrnehmung – 24
2.3.4	Bewegungswahrnehmung – 27
2.3.5	Objektwahrnehmung – 30
<b>2.4</b>	<b>Theorien der Wahrnehmung – 35</b>
2.4.1	Die klassische Psychophysik – 35
2.4.2	Die Gestaltpsychologie – 38
2.4.3	Der wahrnehmungsökologische Ansatz von James J. Gibson – 39
2.4.4	Der computationale Ansatz von David Marr – 40
<b>2.5</b>	<b>Anwendungsbeispiele – 42</b>
<b>2.6</b>	<b>Ausblick – 43</b>
<b>2.7</b>	<b>Weiterführende Informationen – 44</b>
	<b>Literatur – 46</b>

### Eine Wahrnehmungstäuschung

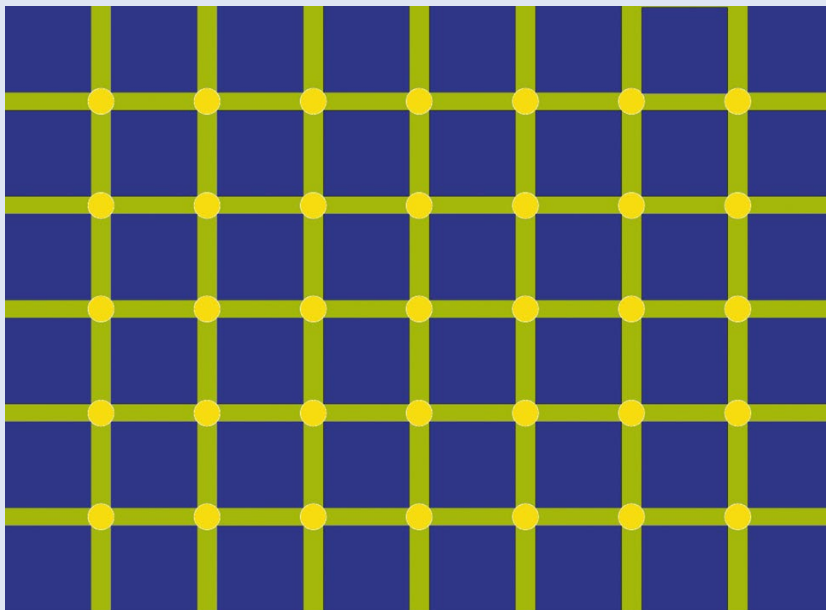
Fixiert man kurze Zeit die Bildmitte des flimmernden Gitters (Abb. 2.1), werden in den hellgelben Schnittpunkten des Gitters dunkle Punkte sichtbar, die aber fortdauernd verschwinden bzw. ihre Position zu verlagern scheinen. In unserem subjektiven Empfinden variieren die Anzahl und Position der flimmernden Punkte ständig. Ein weitere überraschende Beobachtung zu dieser Täuschung: Das Flimmern der Punkte nimmt ab, wenn man den Kopf um 45° neigt!

Derartige Sinnestäuschungen wecken unser Interesse (und nicht nur das Interesse des Wahrnehmungsforschers!), weil sie uns eindrucksvoll verdeutlichen, dass unsere

Wahrnehmungswelt keineswegs eine 1:1-Abbildung der Realität ist. Diese Erkenntnis ist in gleichem Maße faszinierend wie beunruhigend, schließt sie doch ein, dass der Wahrnehmungsinhalt auch sonst von der Realität abweichen könnte, ohne dass wir uns darüber im Klaren wären. Schlimmer noch, man muss aus der Selbstbeobachtung, dass die Anzahl der schwarzen Punkte im flimmernden Gitter mit jedem neuen Versuch variiert, schließen, dass der Zusammenhang zwischen Realität und subjektiver Wahrnehmungswelt nicht durch eine einfache – und insbesondere nicht konstante – Transformationsregel determiniert ist. Wahrnehmen ist also nicht ein passives

Aufnehmen von Information nach festgelegten Regeln, vielmehr können kurzfristige Zustände des Systems unsere momentanen Wahrnehmungsinhalte bestimmen.

Für den Wahrnehmungsforscher sind solche Sinnestäuschungen nur Mittel zum Zweck, und er bedient sich ihrer auch nur zum Teil. Sein vorrangiges Ziel ist die Analyse aller Verarbeitungsprozesse, die darüber aufklären, wie aus der reichhaltigen Information unserer Umwelt die Wahrnehmung entsteht, die schließlich unser Empfinden und Handeln leitet.



■ Abb. 2.1 Das flimmernde Gitter (Scintillating Grid). (Nach Schrauf et al. 1997)

## 2.1 Einleitung: Fragen der visuellen Wahrnehmungsforschung

Die Wissenschaft von der Wahrnehmung versucht zu ergründen, wie die Information unserer distalen Umwelt aufgenommen und verarbeitet wird, wie aus ihr die Wahrnehmung entsteht, die unser Erleben und Verhalten maßgeblich begründet. Oft findet man sogar die Vorstellung, dass alle menschlichen Erkenntnisse und Einsichten auf sensorischen Erfahrungen beruhen, also durch Wahrnehmungsprozesse erst zustande kommen. Es verwundert daher nicht, dass der Untersuchungsgegenstand „Wahrnehmung“ in der modernen Psychologie von jeher eine Sonderstellung einnahm und das Wahrnehmen als die Grundfunktion aller psychischen Funktionen angesehen wurde und zum Teil auch noch so betrachtet wird.

Diese Sonderstellung der Wahrnehmung, insbesondere auch der visuellen Wahrnehmungsmodalität, wird durch den trügeri-

schen Eindruck verstärkt, dass sich Wahrnehmung ohne unser Zutun – also weitgehend passiv – vollzieht. Oft wird visuelles Wahrnehmen mechanistisch aufgefasst und mit den Vorgängen in einer Kamera verglichen, die ja auch ein Abbild der Umwelt zu erzeugen vermag. Es hat eine lange Tradition, das Auge als Kamera zu verstehen. René Descartes war wohl Anfang des 17. Jahrhunderts einer der ersten Wissenschaftler, der am herausoperierten Rinderauge das Abbild unserer Umwelt auf der Netzhaut systematisch untersuchte (Abb. 2.2). Tatsächlich ist die Kameraanalogie bis zu einem gewissen Punkt durchaus zutreffend: Die veränderliche Brennweite der Linse, die Blende bzw. die Pupille und der belichtungsempfindliche Film bzw. die Netzhaut unterscheiden sich zwar in ihrem Aufbau, erfüllen aber ähnliche Funktionen. In beiden Fällen entsteht ein seiten- und höhenverkehrtes Abbild. Die hauptsächliche Wahrnehmungsleistung findet aber erst im Gehirn statt, in den sich anschließenden Analyse- und Verarbeitungsprozessen. Das, was wir als Wahrnehmen bezeichnen,

ist in **Abb. 2.2** daher insbesondere durch den Menschenkopf repräsentiert, der die Abbilder auf der Netzhaut beobachtet. Er repräsentiert den vom Gehirn zu leistenden Wahrnehmungsprozess – erst er weiß, die auf dem Kopf stehende Netzhautinformation „richtig“ zu interpretieren (**► Abschn. 2.3.3**).

Nicht die Projektion der Umwelt auf die Netzhaut, sondern die sich daran anschließenden Prozesse kennzeichnen den eigentlichen Wahrnehmungsvorgang. Das Auge versorgt uns mit den sensorischen Rohmaterialien; Wahrnehmen heißt vor allem aber auch, diese Rohmaterialien zu *verarbeiten*. Dabei sind zunächst einmal die Aufnahme und die Beschaffenheit der Rohdaten wichtig:

- Wie sehen die elementaren visuellen Encodierungsprozesse aus, was sind die Basismerkmale des Wahrnehmungsvorgangs?
- Wie werden die Orientierungen einfacher Linien, Winkel und Farben encodiert?

Dies sind Beispiele für Fragen, die in den folgenden Abschnitten eingehender erläutert werden und die die „frühe“ visuelle Verarbeitung (*early vision*) kennzeichnen. In dieser Phase der frühen visuellen Verarbeitung ist der Anteil an reizgetriebenen Mechanismen relativ hoch. Unter **Bottom-up-Verarbeitung** versteht man die durch einen Reiz ausgelöst und dann weitgehend automatisch ablaufenden Prozesse, die von den mehr kognitiven Funktionen (z.B. Gedächtnis, Lernen) wenig beeinflussbar sind. Mit zunehmender Verarbeitungstiefe kann man allerdings davon ausgehen, dass deren Anteil am Wahrnehmungsprozess (**Top-down-Verarbeitung**) steigt. Obwohl es auch theoretische Ansätze gibt, die eine erste – vom kognitiven Einfluss vollkommen unabhängige – visuelle Verarbeitungsstufe annehmen (z.B. Pylyshyn 1999), so ist die vorherrschende Meinung doch eher die, dass der Einfluss des Kognitiven – wenn auch in unterschiedlichem Ausmaß – auf nahezu allen Verarbeitungsstufen nachweisbar ist. Zumindest erscheint eine strikte Trennung zwischen frühen kognitiv unbeeinflussten und späteren kognitiv beeinflussten visuellen Wahrnehmungsmechanismen nicht angezeigt (z.B. Müsseler 1999a).

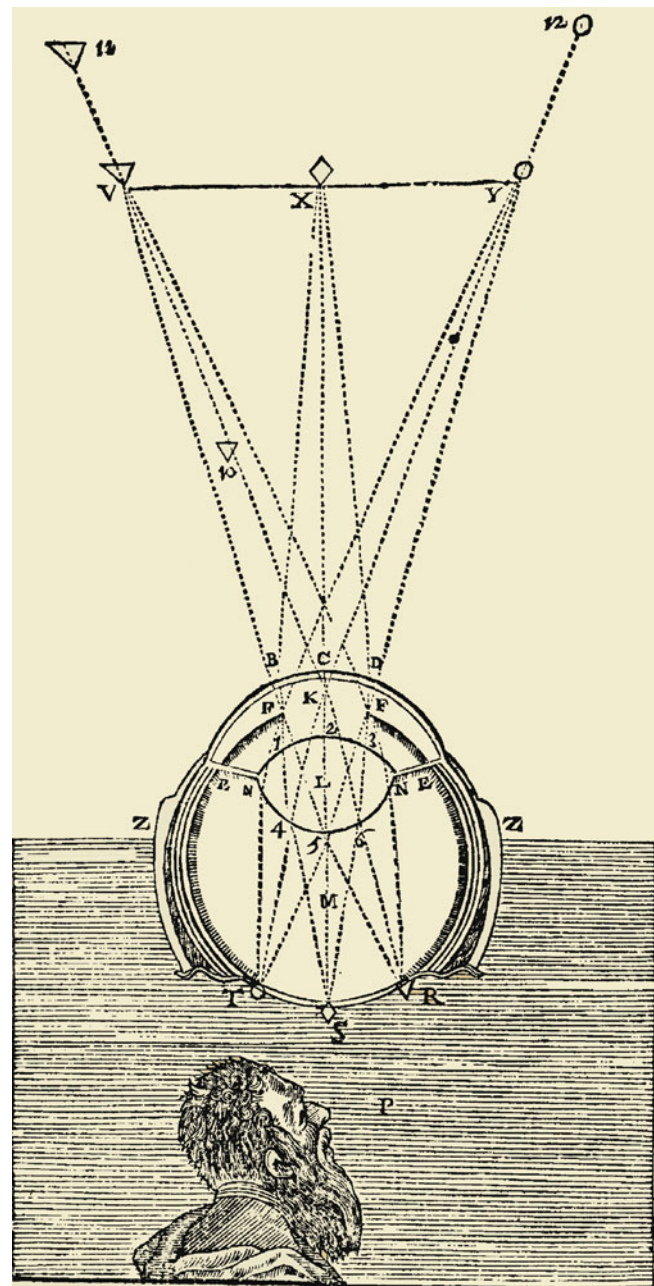
Ohne die kognitive Komponente wären in jedem Fall die folgenden, eher interpretativen Leistungen des Wahrnehmungsprozesses nicht zu erbringen:

- Wie werden Objekte voneinander abgegrenzt?
- Wie werden Objekte identifiziert und kategorisiert?

Dies sind Probleme der Objekterkennung und der Organisation in der Wahrnehmung. Eine Beantwortung dieser Fragen verdeutlicht, dass Wahrnehmen ohne den Bezug zu den bestehenden Wissensbeständen des Wahrnehmenden nicht auskommen kann. Wahrnehmen ist also auch immer eine Interpretation der Rohmaterialien. Daneben gibt es Fragen, die – je nachdem, welchen Ausschnitt man fokussiert – sowohl eine „frühe“ als auch eine „späte“ Komponente aufweisen:

- Wie werden räumliche Distanzen und Tiefe wahrgenommen?
- Wie vollzieht sich die Wahrnehmung von Bewegung?

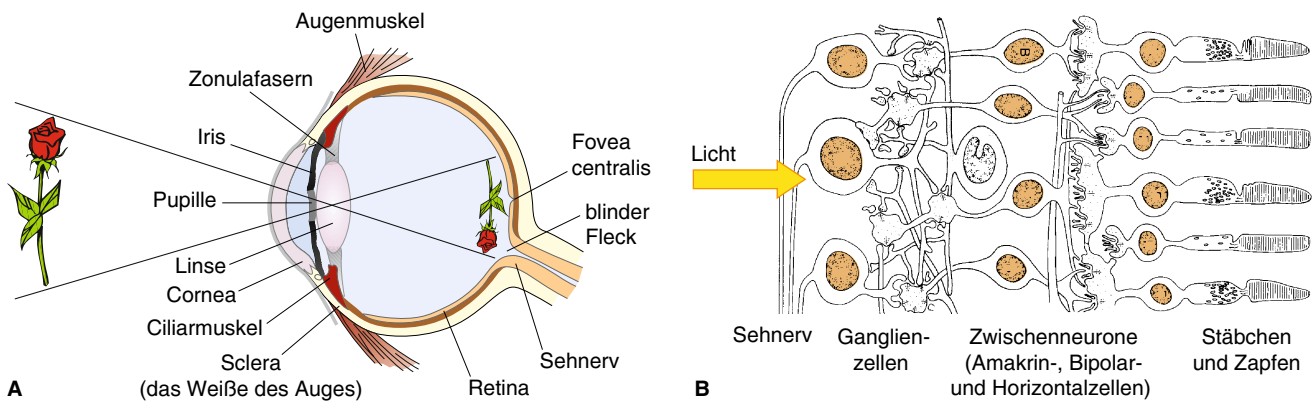
Im Folgenden wird deutlich werden, dass Wahrnehmung nicht ein passives Aufnehmen von Information ist, sondern sich aktiv



■ **Abb. 2.2** Historische Skizze zu Descartes' Beobachtungen der Abbildung unserer Umwelt auf der Netzhaut eines Rinderauges

vollzieht. Zwar senden wir keinen „Wahrnehmungsstrahl“ wie die Fledermaus aus, die mittels Echoortung ihre Umwelt räumlich „abtastet“. Dennoch erscheint die Analogie des Wahrnehmungsstrahles, von der im Übrigen der griechische Philosoph Platon überzeugt war, eher geeignet, den Wahrnehmungsprozess zu charakterisieren als die Vorstellung einer rein rezeptiv angelegten Informationsaufnahme und -verarbeitung.

Im vorliegenden Kapitel können nicht der gesamte Werdegang und Wissensstand der Wahrnehmungsforschung wiedergegeben werden – hier sei auf die speziell für diesen Zweck geschriebenen Wahrnehmungslehr- und -handbücher verwiesen (z.B. Goldstein 2015; Palmer 1999). Wir beschränken uns hier auf die Grundprobleme der Wahrnehmung und auf die zu diesem Zweck entwickelten



**Abb. 2.3** Das Auge (A) und der Aufbau der Retina (B). Man beachte, dass das Licht zunächst die (durchsichtige) Ganglien- und Zwischenneuronenschicht der Retina durchdringt, bevor es auf die lichtempfindliche Schicht der Photorezeptoren (Stäbchen und Zapfen) fällt. (Nach Dowling und Boycott 1966)

grundlegenden Wahrnehmungstheorien. Auch kann dieses Kapitel nur äußerst unzureichend über die Vielzahl und Vielfältigkeit der bekannten Wahrnehmungswahrnehmungen informieren (einen Eindruck darüber liefern entsprechende Sammelbände, z. B. Ditzinger 2006; Rodgers 1999; Seckel 2014). Dies ist insbesondere deshalb kritisch, weil in der Wahrnehmungsforschung – eher noch als in anderen Teildisziplinen der Allgemeinen Psychologie – das rein phänomenorientierte Vorgehen weit verbreitet ist. Obwohl dieses Vorgehen aus wissenschaftstheoretischen Überlegungen keineswegs dem theoriegeleiteten Vorgehen vorzuziehen ist, existieren dennoch eine Vielzahl lokaler Erklärungsansätze, die sich lediglich auf eine Erklärung einer Illusion oder einer Gruppe von Illusionen beschränken. Wir behandeln hier lediglich eine aktuelle Auswahl derartiger Phänomene und Erklärungsansätze; weitere, besonders aktuelle Phänomene, z. B. das Phänomen der Maskierung (Metakontrast), des Aufmerksamkeitsblinzels (*attentional blink*), der Wiederholungsblindheit (*repetition blindness*) oder der Veränderungsblindheit (*change blindness*), finden sich in ► Kap. 5 und 6.

## 2.2 Das Auge und die visuellen Verarbeitungspfade

Wahrnehmungspsychologische Fragestellungen werden heutzutage eng mit Fragestellungen der Sinnes- und Neurophysiologie verknüpft. Wir wissen heute, dass nahezu die Hälfte des Neocortex durch visuelle Reizung stimulierbar ist und dass diese Areale zumindest mittelbar am Wahrnehmungsprozess beteiligt sind. Dieser Abschnitt gibt deshalb zunächst einen groben Überblick über die Neurophysiologie des Sehens, den daraus abzuleitenden grundlegenden sensorischen Verarbeitungsprinzipien der Retina und den nachfolgenden (corticalen) Verarbeitungspfaden. Kenntnisse der neurophysiologischen Grundlagen werden dabei vorausgesetzt (z. B. Bear et al. 2009).

### 2.2.1 Die Retina

Wie in ► Abschn. 2.1 angedeutet, gleicht das Auge in seinen Funktionsmerkmalen einer Kamera. Einfache Störungen (z. B. Weit- und Kurzsichtigkeit) können mit einfachen physikalischen Ge-

setzmäßigkeiten der Lichtbrechung erklärt und korrigiert werden. Diese Vergleichbarkeit hat natürlich aufgrund der unterschiedlichen Hardwarerealisierungen der optischen Apparate ihre Grenzen, sie hört aber spätestens bei der Projektionsfläche auf: Die Retina (Abb. 2.3B) ist keineswegs eine dem analogen Film oder dem digitalen Bildsensor ähnliche homogene, lichtempfindliche Schicht. Die im gesunden Auge (Abb. 2.3A) angesiedelten ca. 126 Mio. Photorezeptoren sind unterschiedlich licht- und farbempfindlich. Weil die 120 Mio. Stäbchen (*rods*) eine deutlich höhere Lichtempfindlichkeit aufweisen als die 6 Mio. Zapfen (*cones*), nehmen wir beispielsweise in der Dunkelheit nur die *achromatischen* Farben Weiß, Grau oder Schwarz wahr („Nachts sind alle Katzen grau!“, skotopisches Sehen), während am helllichten Tag das *chromatische* Farbsehen der Zapfen mit den Farben Blau, Rot, Grün und Gelb vorherrscht (photopisches Sehen; zum Farbsehen s. auch ► Abschn. 2.3.2). Außerdem ist die Verteilung der Stäbchen und Zapfen im Gesichtsfeld sehr unterschiedlich. Die Zapfendichte ist in der Fovea centralis – einem Gebiet, das nur etwa 2° des gesamten Gesichtsfeldes ausmacht – am höchsten. Sie nimmt mit zunehmender Exzentrizität vom Fixationspunkt deutlich ab, und die relative Dichte der Stäbchen steigt. Stäbchen sind zwar durch eine bessere Lichtsensitivität ausgewiesen, das Zapfensystem übermittelt aber neben dem Farbsehen eine höhere räumliche und zeitliche Auflösung (Überblick in Skavenski 1990).

Zum Teil sind die letztgenannten Befunde auch dadurch bedingt, dass bereits auf der Retina erste verschiedenartige neuronale Verschaltungen zwischen den Stäbchen bzw. Zapfen einerseits und den Zwischenneuronen und Ganglienzellen andererseits wirksam werden. Dementsprechend überlappen sich zum einen die **rezeptiven Felder** der Ganglienzellen, zum anderen findet man in den aus dem Auge austretenden Zellen auch eine erste funktionale Spezialisierung in *M-Typ*-(magnozellular-) und *P-Typ*-(parvozellulär-)ähnliche Zellen. Die *M-Typ*-ähnlichen Ganglienzellen haben vergleichsweise große rezeptive Felder und antworten auf grobe achromatische Reizung mit einer schnellen transienten Aktivierung. Die *P-Typ*-ähnlichen Ganglienzellen haben dagegen kleine rezeptive Felder und antworten auf eine fein achromatische und auf chromatische Reizung mit einer langsamen tonischen Aktivierung. Diese prinzipielle Zellunterscheidung wird bis zu den corticalen Projektionsarealen beibehalten (Überblick in Bear et al. 2009; Kandel 1996).



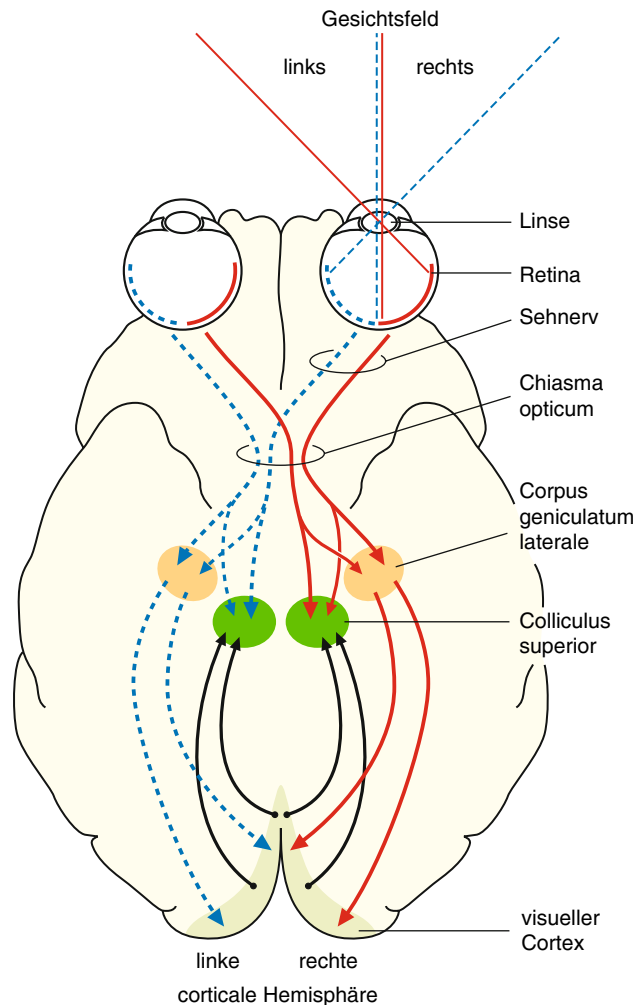
### 2.2.2 Die Bahn des Sehnervs zwischen Auge und Cortex

Während die Sehbahnen der temporalen Retinahälften nach dem Austritt aus dem Auge ungekreuzt verlaufen, kreuzen die jeweils nasalen Retinahälften im Chiasma opticum die Hemisphäre, sodass das jeweilige Gesichtsfeld kontralateral projiziert wird (Abb. 2.4). Bevor allerdings corticale Regionen erreicht werden, ziehen etwa 10 % des Sehnervs in den Colliculus superior (CS) – eine Region, die maßgeblich an der Steuerung von Blickbewegungen beteiligt ist. Die übrigen Anteile des Sehnervs entsenden hauptsächlich Fasern zum Corpus geniculatum laterale (CGL), einer Struktur unterhalb des eigentlichen Thalamus, die sich in zwei retinotop organisierte magno- und vier parvozelluläre Zellschichten gliedert. Eine selektive Zerstörung der magnozellularen Schichten führt beim Affen zu starken Einbußen in der Bewegungswahrnehmung, während parvozelluläre Läsionen die Wahrnehmung von Farbe, feinen Texturen, Formen und räumlicher Tiefe verhindern (Schiller et al. 1990).

### 2.2.3 Der primäre visuelle Cortex

Die Neurone aller Schichten des CGL projizieren zum visuellen Cortex im Okzipitallappen. Der primäre visuelle Cortex wird aufgrund seines gestreiften Aufbaus auch als Area striata (oder als Brodmann-Areal 17 oder einfach als V1) bezeichnet. Er besteht aus neun Schichten (1 bis 6, wobei die vierte Schicht in 4A, 4B, 4Ca und 4Cb unterteilt ist). Wie im CGL liegt in V1 eine räumlich retinotop Organisation der Zellen vor, wobei allerdings den fovealen Regionen weit mehr Platz eingeräumt wird als den peripheren Regionen (**corticaler Vergrößerungsfaktor**; Cowey und Rolls 1974; Drasdo 1977). Die magnozellularen CGL-Schichten projizieren größtenteils zur Schicht 4Ca, die wiederum Fasern zur Schicht 4B entsendet. Auch die *einfachen Zellen* der Schicht 4B sind zum Teil bewegungssensitiv, reagieren aber insbesondere auf Orientierung. Die parvozellulären Schichten des CGL entsenden Fasern zur Schicht 4Cb, die wiederum die Schichten 2 und 3 des primären visuellen Cortex innervieren (Abb. 2.5). Die dort zu findenden *komplexen Zellen* des V1, insbesondere die Zellen der parvozellulären Blob-Bahn (so benannt nach ihrer äußeren Erscheinungsform nach einer Einfärbung; *blob* = „Tropfen“), sind farbsensitiv. In den ebenfalls komplexen Zellen der parvozellulären Interblob-Bahnen finden sich dagegen richtungs- und formsensitive Zellen für Reize einer bestimmten Orientierung. Sie sind weniger positionsrestriktiv als die einfachen Zellen, d. h., ihr Zell-tuning ist relativ unabhängig davon, wo der Reiz das receptive Feld durchstreicht. – Neben den einfachen und komplexen Zellen sind in V1 außerdem die *endinhierten Zellen* (hyperkomplexen Zellen) isoliert worden, die sensitiv auf sich bewegende Ecken und Winkel einer bestimmten Länge reagieren.

Diese äußerst grobe Beschreibung der Zellen in V1 wird bei Weitem nicht der anzunehmenden Verarbeitungskomplexität gerecht, die dieses Areal leistet. Man kann jedoch festhalten: Je weiter man sich von der Retina entfernt, desto spezialisierter reagieren die einzelnen Neurone auf spezifische Reizeigenschaften.



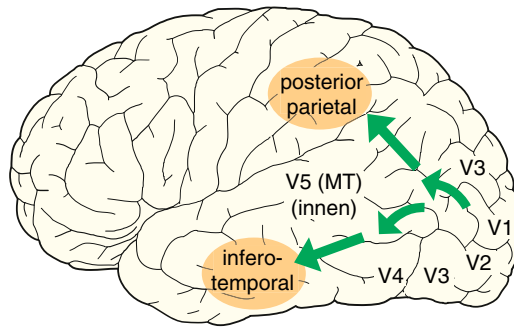
■ **Abb. 2.4** Die Bahn des Sehnervs zwischen Auge und primärem visuellen Cortex. Dabei kreuzen die Neuronenbahnen aus den nasalen Retinahälften im Chiasma opticum die Hemisphäre, sodass das jeweilige Gesichtsfeld kontralateral projiziert wird. Die Information des linken Gesichtsfeldes wird in die rechte corticale Hemisphäre (durchgezogene Linie) projiziert und umgekehrt (gestrichelte Linie)

ten. Seit Hubel und Wiesel (z. B. 1959, 1968) bahnbrechenden Einzelzellableitungen an Katzen und Affen in diesem Areal geht man deshalb davon aus, dass die Zellen von V1 als erste Merkmalsdetektoren fungieren.

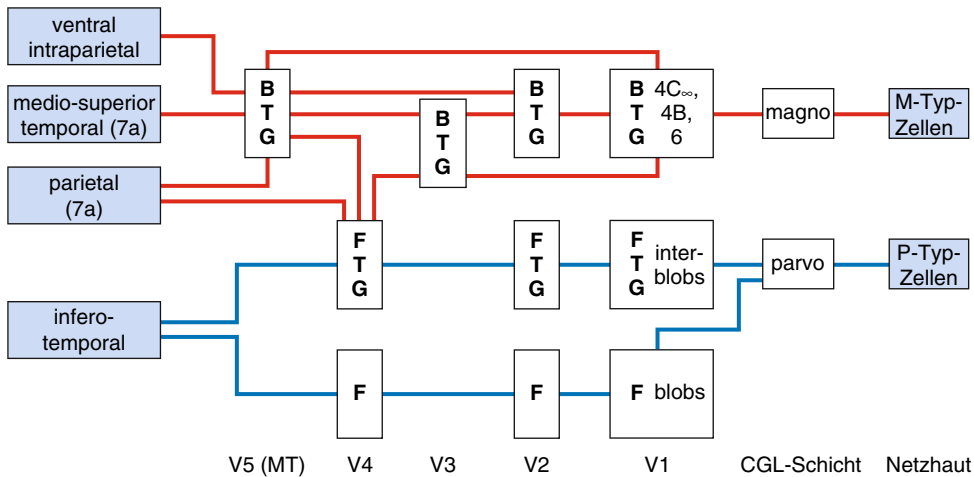
### 2.2.4 Weitere corticale Verarbeitungspfade

Die weitere Verarbeitung im extrastriaten visuellen Cortex verläuft durch die Areale V2 bis V5, wobei letzteres auch als mediotemporales Areal (MT) bezeichnet wird. Die Areale sind feiner differenzierbar (z. B. in V3a und V5a; z. B. Felleman und Van Essen 1991), wichtig ist aber der Befund, dass der Verarbeitungsfluss nicht – wie man früher annahm – strikt seriell verläuft. Spätestens ab V3 bzw. V4 teilt sich der parvo- und magnozellular-visuelle Informationsfluss in einen temporalen (ventralen) und einen parietalen (dorsalen) Verarbeitungspfad (Abb. 2.5).

Diese unterschiedlichen Verarbeitungspfade wurden vor allem von Ungerleider und Mishkin (Mishkin et al. 1983; Unger-



■ **Abb. 2.5** Der temporale und parietale Verarbeitungspfad des visuellen Systems mit den dort hauptsächlich übermittelten Reizeigenschaften. F = Farbe, T = räumliche Tiefe, G = Gestalt, B = Bewegung. (Nach DeYoe und Van Essen 1988; Kandel 1996)



leider und Mishkin 1982) beschrieben, wobei sie dem *temporalen Pfad* (dem Was-Pfad) die Funktion der Objekterkennung und dem *parietalen Pfad* (dem Wo-Pfad) die Funktion der Objektllokalisierung zuschrieben. Mittlerweile stellt sich der Zusammenhang zwischen Wahrnehmungsfunktionen und den beteiligten visuellen Arealen deutlich differenzierter dar (einen Eindruck vermittelt ■ Abb. 2.5). Außerdem ist die Was-Wo-Unterscheidung auch zugunsten einer Was-Wie-Unterscheidung diskutiert worden. Während Ungerleider und Mishkin ihre Konzeption von der Frage her entwickelten, *welche* Information in den beiden Pfaden verarbeitet wird, ist die Leitfrage von Milner und Goodale (1995), *wozu* die Information in den beiden Strängen dient. Ähnlich wie bei Ungerleider und Mishkin (1982) findet im Was-Pfad die neuronale Codierung von Objektmerkmalen statt. Die Neurone im inferotemporalen Cortex reagieren besonders auf komplexe visuelle Reize wie Hände oder Gesichter, aber auch Farbe (insbesondere in V4; Zeki et al. 1991) wird vorrangig im ventralen Pfad verarbeitet. Nach Milner und Goodale (1995) ist dieser Pfad daher vor allem für die bewusste Objekterkennung zuständig. Inwieweit bewusste Wahrnehmungsvorgänge dem ventralen Verarbeitungspfad zuzuschreiben sind, wird in ► Kap. 6 eingehender erörtert (dort findet sich auch eine Darstellung von verschiedenen Wahrnehmungsbeeinträchtigungen nach pathologisch bedingten Läsionen der hier besprochenen Areale; z. B. Blindsight, visuelle Agnosien, visuospatialer Neglect).

Gegenüber dem Was-Pfad wird der dorsale Pfad für die (nicht notwendigerweise bewusste) visuelle Steuerung von Handlungen verantwortlich gemacht. Im Gegensatz zum ventralen System stellt der dorsale Pfad sozusagen online diejenigen visuellen Informationen zur Verfügung, die für die visuelle Kontrolle von

Handlungen benötigt werden. Das betrifft sowohl die räumliche Position von Objekten als auch Informationen über die Richtung und Geschwindigkeit von bewegten Objekten (insbesondere im MT bzw. in V5; Zeki et al. 1991). Es wird also angenommen, dass die Analyse der Objektmerkmale, die für die Identifikation und das Wiedererkennen von Objekten notwendig ist, unabhängig von der Analyse derjenigen Objektmerkmale stattfindet, die es einer Person erlauben, ein Objekt zu greifen. Dieser Ansatz ist allerdings nicht unumstritten. So gibt es beispielsweise Befunde, die schon frühe Interaktionen zwischen dem dorsalen und dem ventralen Pfad belegen (z. B. Merigan und Maunsell 1993; Van Essen und DeYoe 1995). Folgt man Rossetti und Pisella (2002), repräsentieren die beiden Verarbeitungspfade ein sensomotorisches und ein kognitiv-sensorisches System, die je nach Stimulus- und Responsetyp und deren zeitlicher Verkopplung zum Einsatz kommen. Festzuhalten bleibt, dass die vermuteten unterschiedlichen Funktionen der beiden Verarbeitungspfade und damit der Zusammenhang zwischen Wahrnehmung und Handlung in den letzten Jahrzehnten zu einem enormen Anstieg auch psychophysischer Experimente geführt haben. Die Diskussion darüber erscheint keineswegs abgeschlossen.

Fassen wir zusammen: Mit zunehmender (corticaler) Verarbeitung findet man in den entsprechenden Hirnregionen eine zunehmende funktionale Neuronenspezialisierung (► Zur Vertiefung 2.1). Diese geht umgekehrt proportional einher mit der Größe der zugehörigen rezeptiven Felder. Die Ganglienzellen der Retina antworten auf eine unspezifische Lichtreizung innerhalb einer eng umgrenzten Netzhautregion, während die Zellen in V1 bei größer werdenden rezeptiven Feldern z. B. nur auf Lichtstreifen einer bestimmten Orientierung antworten. Im inferotempo-

## Zur Vertiefung 2.1

**Bildgebende Verfahren zur Untersuchung von Hirnprozessen**

Insbesondere in den letzten Jahrzehnten sind **bildgebende Verfahren** entwickelt worden, die es ermöglichen, am wachen Menschen neurale Prozesse nahezu nichtinvasiv zu erfassen. So können Teilaspekte der Wahrnehmung und anderer kognitiver Funktionen mit den corticalen Regionen in Verbindung gebracht werden, die bei der Darbietung von Reizen oder beim Lösen bestimmter Aufgaben aktiv sind (detaillierte Beschreibung der Verfahren z. B. in Jäncke 2005).

Die *Positronenemissionstomografie* (PET) basiert auf der Messung von radioaktiven Markersubstanzen, die zuvor in den Blutkreislauf injiziert wurden und Positronen emittieren. Die radioaktiv markierten Stoffwechselsubstanzen werden verstärkt in metabolisch aktiven Zellen verbraucht, also den Zellen, die an dem jeweiligen kognitiven Prozess beteiligt sind. Ein Detektor, der in einer Vorrichtung drehbar um den Kopf geführt wird, zählt die Photonen, die bei der Annihilation von Positronen mit Elektronen im Körper emittiert werden, sodass der Computer die Orte stärkster und schwächster Strahlung bestimmen kann. Die zeitliche Auflösung beim PET liegt allerdings nur im Bereich von 10 s, die räumliche Auflösung erlaubt dagegen die Lokalisation mit Millimeterpräzision.

Bei der *funktionellen Kernspintomografie* (bzw. funktionellen Magnetresonanztomografie, fMRT; *functional magnetic resonance imaging*, fMRI) werden nicht die Veränderungen radioaktiver Strahlungen, sondern magnetischer Felder erfasst. Dazu werden Wasserstoffatome durch ein starkes statisches Magnetfeld parallel zu den Feldlinien des Magnetfeldes ausgerichtet. Daneben wird ein kleines, radiofrequentes Feld induziert. Wenn die Neurone aktiv werden, ändert sich der Sauerstoffverbrauch im Bluthämoglobin, was sich wiederum auf die Eigenschaften des magnetischen Feldes auswirkt. Diese Veränderungen können mithilfe der fMRI-Technik gemessen werden. Aufgrund der Trägheit

des Wechsels im Blutfluss ist allerdings die zeitliche Auflösungsfähigkeit auf ungefähr 1 s beschränkt. Allerdings lässt sich dieser Wert aufgrund von autoregressiven Modellierungen (und gleichzeitiger Gültigkeit der entsprechenden Modellierungsannahmen) erheblich verbessern (z. B. Goebel et al. 2003). Die hohe räumliche Auflösung im Millimeterbereich wird meist durch eine zusätzliche Magnetresonanztomografie (MRT; *magnet resonance imaging*, MRI) sichergestellt, wodurch individuelle anatomische Lageinformationen berücksichtigt werden.

Methoden, die auf jeden Fall eine hinreichend gute zeitliche Auflösung gewährleisten, um die in kognitiven Prozessen ablaufenden physiologischen Aktivitäten im Millisekundenbereich zu erfassen, sind das *ereigniskorrelierte Potenzial* (EKP; *event-related potential*, ERP) und die *Magnetencephalografie* (MEG). Das EKP erfasst alle elektrocorticalen Potenziale, die vor, während und nach einem sensorischen, motorischen oder psychischen Reiz- oder Reaktionsereignis im Elektroencephalogramm (EEG) auf der Kopfoberfläche messbar sind. Das EKP stellt die summierte Aktivität dieser elektrischen Prozesse dar. Die wesentliche Annahme bei der EKP-Messung liegt darin, dass es geordnete Aktivierungsmuster gibt, die mit psychischen Vorgängen eng zusammenhängen. Die Muster werden nach ihrer Ausrichtung, negativ oder positiv, und ihrem zeitlichen Auftreten klassifiziert. Zum Beispiel wird eine Positivierung des Potenzials im Bereich von 95–140 ms nach der Reizdarbietung als P1 bezeichnet, die mit frühen reizbezogenen Verarbeitungsprozessen in Verbindung gebracht wird.

Allerdings verfügt das EKP neben der guten zeitlichen über eine schlechte räumliche Auflösung. Dieses Manko hat die MEG nicht, wenngleich die hohe räumliche Auflösungsfähigkeit der MEG mit zunehmender Tiefe etwas abnimmt. Auch die MEG nutzt die Tatsache aus, dass elektrische Ströme im Gehirn magnetische Felder erzeugen. Diese Felder kann man

mithilfe sogenannter SQUIDs (*superconducting quantum interference devices*) messen. Allerdings sind die durch die magnetischen Felder verursachten Ausschläge extrem klein, und der Messort muss entsprechend aufwendig und kostenintensiv abgeschirmt werden.

Zusammenfassend weisen PET und fMRI eine hohe räumliche Auflösung auf, was allerdings zu Lasten der zeitlichen Auflösung geht. In zeitlicher Hinsicht (allerdings nicht in räumlicher) ist die Ableitung des EKP informativer. Wenn man in seinen Fragestellungen sowohl auf eine hohe räumliche als auch auf eine hohe zeitliche Auflösung angewiesen ist, ist zurzeit die MEG das Maß der Dinge.

Neben den bisher genannten reaktiven Verfahren zur Messung von Gehirnaktivitäten, die nach der Präsentation sensorischer Reizinformation oder bei der Bearbeitung kognitiver Aufgaben auftreten, kommt in den letzten Jahrzehnten vermehrt auch ein Verfahren zum Einsatz, das transkraniell – also durch den Schädel hindurch – von außen auf die neuronalen Aktionspotenziale einwirkt. Mithilfe von elektrisch erzeugten Magnetfeldern gelingt es, gezielt und fast schmerzlos durch die Haut und den Knochen hindurch die Nervenzellen einzelner Hirngebiete zu stimulieren. Dieses Verfahren wird als *Transkranielle Magnetstimulation* (TMS) bezeichnet. Dabei wird unterschieden zwischen der Stimulation durch einzelne Magnetfeldpulse und der Stimulation durch Impulssalven (von bis zu 100 Hz), der sogenannten *repetitiven Transkraniellen Magnetstimulation* (rTMS). Durch Magnetstimulation über dem motorischen Cortex kann man beispielsweise Muskelzuckungen auslösen. Über dem visuellen Cortex kann eine derartige Stimulation zu Phosphenen (Lichterscheinungen), aber auch zu Skotomen (Wahrnehmungsausfällen innerhalb des Gesichtsfeldes) führen. Eine rTMS von Hirnregionen, die für Sprache zuständig sind, kann für einige Minuten eine Verschlechterung der sprachlichen Ausdrucksfähigkeit bewirken.

ralen Cortex präferieren die Zellen sogar höchst komplexe Reizeigenschaften (z. B. Gesichter), die relativ unabhängig vom Ort der retinalen Stimulation sind (z. B. Tanaka 1996). Eine Folge dieser zunehmenden Spezialisierung ist eine verteilte Codierung von Reizeigenschaften in verschiedenen Hirnregionen. Wir werden darauf zurückkommen (► Abschn. 2.3.5) und ein damit verbundenes Problem, das Bindungsproblem (*binding problem*), erörtern.

## 2.3 Visuelle Informationsaufnahme und -verarbeitung

In diesem Abschnitt geht es darum, die Wahrnehmungsleistungen zu erörtern, zu interpretieren und diese im Zusammenhang mit neurophysiologischen Befunden zu diskutieren. Dabei stehen

die psychophysischen Leistungen des Wahrnehmungsapparats im Vordergrund, also die Leistungen, die durch Einsatz entsprechend objektivierbarer Methoden validiert worden sind (► Zur Vertiefung 2.2 und ► Zur Vertiefung 2.4).

### 2.3.1 Visuelle Sehschärfe und Sensitivität

Die räumliche Auflösungsfähigkeit des visuellen Systems gilt als recht hoch. Sie steigt von 5 min arc bei 10° retinaler Exzentrizität bis über 1 min arc in der Fovea (z. B. Badcock und Westheimer 1985; Westheimer 1981). Die Abhängigkeit der Sehschärfe vom retinalen Ort der Reizung kann man sich leicht selbst klarmachen, indem man einen Buchstaben in einer Zeilenmitte dieses Buches fixiert und bei fixiertem Blick die Buchstaben am

## Zur Vertiefung 2.2

## Psychophysische Methoden

Durch Gustav Theodor Fechner (1860) wurden in der klassischen Psychophysik Messmethoden etabliert, die zur Erfassung psychologischer Phänomene dienen. Die für den Wahrnehmungspsychologen wichtigsten Verfahren sind bis heute die *Grenzmethode*, die *Konstanzmethode* und die *Herstellungsmethode*. Daneben existieren weitere Verfahren, die aber meist auf einer Variante oder einer Kombination dieser Methoden beruhen. Die grundlegenden Messmethoden werden im Folgenden anhand einer absoluten **Wahrnehmungsschwelle** („Ab welcher Lichtintensität ist ein Reiz überhaupt wahrnehmbar?“) und der zugehörigen **Unterschiedsschwelle** (die Schwelle des eben merklichen Unterschieds, „Ab welchem Intensitätsunterschied werden zwei Lichtreize als unterschiedlich empfunden?“) beispielhaft erläutert.

**Grenzmethode**

Zur Messung der absoluten Wahrnehmungsschwelle steigert der Versuchsleiter bei dieser Messmethode die Intensität eines zunächst deutlich unterschwelligen Reizes in festgelegten Schritten, bis der Beobachter angibt, diesen Reiz wahrzunehmen (aufsteigendes Verfahren). Daraufhin wird die Messreihe abgebrochen, und der Versuchsleiter startet eine weitere Messreihe, in der er die Lichtintensität eines deutlich überschwelligen Reizes reduziert, bis der Beobachter diesen Reiz nicht mehr wahrnimmt (absteigendes Verfahren). Die Schwelle ergibt sich aus dem Mittelwert mehrerer ab- und aufsteigender Messreihen. Zur Messung der Unterschiedsschwelle fragt man analog in ab- und aufsteigenden Messreihen danach, ob sich ein Reiz von einem Vergleichsreiz unterscheidet oder nicht.

In der *Staircase-Variante* dieses Verfahrens werden ab- und aufsteigende Darbietungen zufällig gemischt, der Versuchsleiter behält aber die Kontrolle über beide Messreihen. Der wesentliche Unterschied zur ursprünglichen Grenzmethode besteht darin, dass – sobald der Beobachter den Reiz in der aufsteigenden Messreihe wahrgenommen hat – die Reizintensität zunächst wieder verringert wird und man sich so der Schwelle iterativ nähert. Analog verfährt man in der absteigenden Messreihe (adaptives Staircase-Verfahren; ■ Abb. 2.6). Die Schwelle ergibt sich dann als Mittelwert aus den beobachteten Umkehrpunkten. Eine besonders effektive Variante der Staircase-Methode ist bei Pentland (1980) und Lieberman und Pentland (1982) als Best-PEST-Verfahren beschrieben.

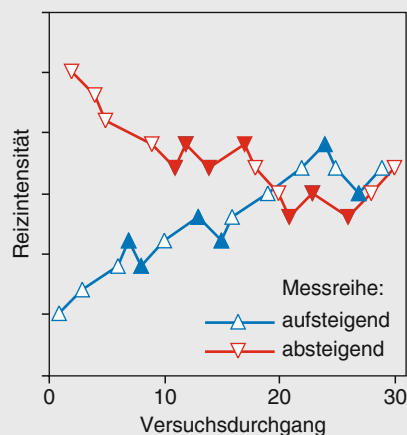
**Konstanzmethode**

Im Gegensatz zur Grenzmethode, in der die Messreihe abgebrochen wird, sobald der Beobachter signalisiert, einen Reiz erkannt zu haben (oder nicht), werden bei der Konstanzmethode in zufälliger Reihenfolge immer alle vom Experimentator ausgewählten Reizausprägungen dargeboten. Zur Auswertung wird dann den Urteilswahrscheinlichkeiten der Beobachter eine Funktion angepasst (z. B. unter Zuhilfenahme einer *Logit-* oder *Probit-Analyse*; Finney 1971; Lieberman 1983), die es erlaubt, die absolute Schwelle als entsprechenden Wert der Lichtintensität auf der x-Achse zu bestimmen. Den Konventionen entsprechend ist dieser Wert durch die 50 %-Urteilswahrscheinlichkeit festgelegt, also bei dem Wert der x-Achse, an dem der Beobachter mit gleicher Wahrscheinlichkeit angibt, den Reiz erkannt zu haben oder nicht (durchgezogene rote Linie in ■ Abb. 2.7).

Handelt es sich um überschwellige Reizintensitäten, so kann man auf diese Weise auch den „Punkt subjektiver Gleichheit“ zweier sensorischer Empfindungen ermitteln (*point of subjective equality*, PSE-Wert). Soll ein Beobachter beispielsweise einen Reiz danach beurteilen, ob er in seiner Lichtintensität mit einem Vergleichsreiz, der aber vor einem andersfarbigen Hintergrund präsentiert wird, übereinstimmt oder nicht, so gibt die 50 %ige Urteilswahrscheinlichkeit den PSE-Wert an. Einer weiteren Konvention folgend werden in diesem Fall die untere und obere Unterschiedsschwelle bei 25 %iger bzw. 75 %iger Urteilswahrscheinlichkeit angesiedelt (bisweilen auch bei 20 %iger bzw. 80 %iger Urteilswahrscheinlichkeit; punktierte blaue Linien in ■ Abb. 2.7). – Die Konstanzmethode ist naturgemäß aufwendiger als die Grenzmethode, wird aber im Allgemeinen als sehr genau angesehen.

**Herstellungsmethode**

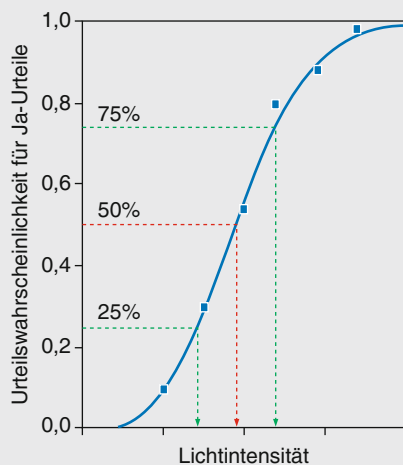
Zur Bestimmung der absoluten Schwelle stellt der Beobachter bei dieser Methode selbst (z. B. mittels eines Leuchtdichtereglers) die Reizintensität ein, an der er einen Reiz gerade noch wahrnimmt bzw. ihn nicht mehr wahrnimmt (absolute Schwelle). Analog dazu stellt er bei der Unterschiedsschwelle den gerade merklichen Unterschied zu einem Vergleichsreiz ein. Der Schwellenwert ergibt sich aus dem Mittelwert mehrerer Einstellungen. – Die Herstellungsmethode ist die einfachste der genannten psychophysischen Methoden. Sie benötigt nur eine geringe Anzahl von Versuchsdurchgängen, gilt allerdings wegen ihrer Anfälligkeit im Hinblick auf mögliche Antworttendenzen als weniger genau.



■ Abb. 2.6 Adaptives Staircase-Verfahren mit einer auf- und einer absteigenden Messreihe. Die ausgefüllten Dreiecke markieren die für die Schwellenbestimmung wichtigen Umkehrpunkte



## Zur Vertiefung 2.2 (Fortsetzung)



**Abb. 2.7** Ein Beobachter beurteilt wiederholt die Wahrnehmbarkeit eines Reizes mit einem Ja-nein-Urteil (hier beispielsweise bei Darbietung verschiedener Lichtintensitäten). Abgetragen sind hier die Urteilswahrscheinlichkeiten der Ja-Urteile (schwarze Quadrate), denen eine Funktion angepasst wird (hier z.B. eine kumulative Normalverteilung, Probit-Funktion; durchgezogene Linie). Der Abszissenwert bei 50 %-Urteilswahrscheinlichkeit determiniert die absolute Wahrnehmungsschwelle oder – beim Vergleich überschwelliger Reizintensitäten – den Punkt subjektiver Gleichheit (PSE-Wert; gestrichelte Linie). Die 25 %- und 75 %-Urteilswahrscheinlichkeit sind bisweilen als untere bzw. obere Unterschiedsschwelle (punktstrich Linien) definiert



A



B

**Abb. 2.8** Die Sehschärfe-Illusion. **A** Schematisch wird veranschaulicht, dass die Buchstaben in der retinalen Peripherie größer dargestellt werden müssten, um bei Fixation in der Bildmitte gleich gut identifizierbar zu sein. **B** Unten ist die zunehmende retinale Unschärfe einer Szene bildhaft dargestellt. Unser alltägliches Sehen vermittelt uns dagegen den Eindruck, immer über ein „bis in alle Ecken“ scharfes Abbild zu verfügen (oben). (Der Abdruck erfolgt mit freundlicher Genehmigung durch den Karlsruher-Dombauverein Aachen. Das Foto wurde der Webseite ► <http://www.karlsruher-dombauverein.de/die-perfekte-geometrie-des-aachener-doms/> entnommen.)

rechten und linken Rand zu identifizieren versucht. Dies wird misslingen. **Abb. 2.8A** veranschaulicht schematisch, dass die Buchstaben in der retinalen Peripherie größer sein müssten, um sie fehlerfrei identifizieren zu können. Merkwürdigerweise wird uns diese zunehmende periphere Unschärfe nur selten bewusst. Stattdessen nehmen wir vor unserem geistigen Auge ein überall gleich scharfes „Foto“ wahr – eine Täuschung, die man auch als *Sehschärfe-Illusion* bezeichnet (**Abb. 2.8B**).

Die Überlegenheit des fovealen Sehens lässt sich schon auf der Retina mit der unterschiedlichen Verteilung von Stäbchen und Zapfen und deren neuronalen Verschaltungen in Verbindung bringen. Da in der Peripherie immer mehrere Rezeptoren

auf eine Ganglienzelle konvergieren (die rezeptiven Felder dieser Ganglienzellen sind entsprechend groß), hingegen die Rezeptoren in fovealen Arealen weit dichter und weniger verschaltet sind (mit entsprechend kleineren rezeptiven Feldern), resultiert schon daraus eine höhere foveale Sehschärfe.

Der foveale Vorteil wird besonders deutlich bei Verwendung von Standardmethoden. Zum Beispiel ist bei der *Landolt'schen Sehprobe* die Öffnungsrichtung einer rotierten, dem Buchstaben C ähnlichen Figur unter ansonsten optimalen Bedingungen anzugeben (Größe und Öffnung des Zeichens sind in der Europäischen Norm EN ISO 8596 festgeschrieben); auch bei der *Vernier-Sehprobe* hat der Beobachter unter optimalen Bedingun-



**Abb. 2.9** Testtafeln zur Ermittlung der Sehschärfe. (A und B) Die alphanumerischen Zeichen haben den Nachteil, dass die Zeichen individuell mehr oder weniger vertraut sind und deswegen unterschiedlich gut berichtet werden (Stichwort Analphabetismus). Dies ist bei den Landolt-Ringen (C) und bei dem Vernier-Sehschärfetest (Noniussehschärfe; D) nicht der Fall

gen die Versetzung zweier übereinander angeordneter Striche zu erkennen – optimal heißt hier jeweils, dass die Reize lang andauernd, gleichzeitig, stationär und mit hohem Kontrast präsentiert werden (Abb. 2.9).

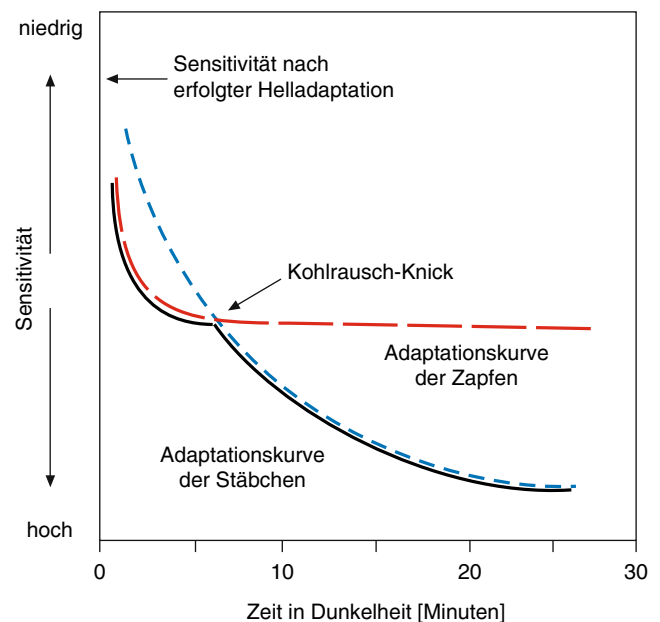
Bereits Skavenski (1990) hat darauf hingewiesen, dass die mit Standardmethoden ermittelte Sehschärfe keineswegs allen Aufgaben zugutekommen muss. So kann in Textursegmentierungsaufgaben (Abb. 2.21) unter bestimmten Bedingungen ein Zielreiz in der Fovea schlechter erkannt werden als in der Peripherie (Meinecke und Kehler 1994; Kehler 1997). Die Sensitivität des visuellen Systems hängt noch von weiteren Faktoren ab, z. B. von der Reizdauer und der Leuchtdichte. Dabei bleibt dem *Bloch'schen Gesetz* folgend die Sehschärfe konstant, wenn mit zunehmender Reizdauer die Leuchtdichte reduziert wird. Das reziproke Verhältnis zwischen Reizdauer und Leuchtdichte gilt bei der Landolt'schen Sehprobe bis etwa 150 ms Reizdauer (Kahneman et al. 1967). Darüber lässt sich eine Abnahme der Leuchtdichte nicht mehr mit einer Zunahme der Reizdauer kompensieren. Das Bloch'sche Gesetz ist offensichtlich darauf zurückzuführen, dass die Reizenergie an den Rezeptoren bis zu einem bestimmten Intervall zeitlich aufsummiert wird.

Die räumliche Auflösungsfähigkeit des visuellen Apparats wird auch durch eine Sukzessivität in der Darbietung negativ beeinflusst. Haben Versuchspersonen beispielsweise die Aufgabe, einen peripher dargebotenen Zielreiz danach zu beurteilen, ob er sich links oder rechts von der Mitte eines Vergleichsreizes befindet, so nehmen sie die relative Position des Zielreizes selbst dann relativ genau wahr, wenn beide Reize sehr kurzzeitig, aber simultan dargeboten werden. Trennt man die Darbietung beider Reize zeitlich nur um 100 ms, so kommt es zu einer systematisch wahrgenommenen Verschiebung des Zielreizes nach außen und zwar in einer Größenordnung, die mit 0,3–0,5° weit über der aufgrund der Standardmethoden zu erwartenden räumlichen Auflösungsfähigkeit liegt (Müsseler et al. 1999; Müsseler und Van der Heijden 2004; Bocianski et al. 2008, 2010).

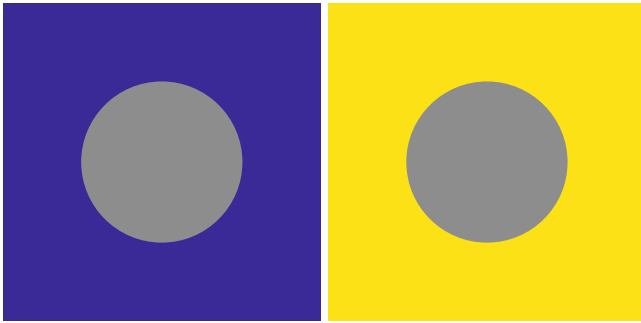
Eine weitere Variable, die das visuelle Auflösungsvermögen nachhaltig beeinflusst, ist der Adaptationszustand des visuellen Systems. Die Flamme einer Kerze wird in klarer Nacht erst nach erfolgter (Dunkel-)Adaptation aus mehreren Kilometern wahr-

nehmbar. Ebenso braucht das System zur (Hell-)Adaptation einige Zeit, wenn man plötzlich in einen hellen Raum eintritt. Die optimale Sensitivität des Systems wird durch Veränderungen der Pupillengröße und durch die veränderte photochemische Lichtempfindlichkeit der Zapfen und Stäbchen erreicht. Da beide Rezeptorsysteme – wie erwähnt – unterschiedlich lichtempfindlich sind, ist die Dunkeladaptationskurve (Abb. 2.10) durch den sogenannten Kohlrausch-Knick gekennzeichnet, der den Übergang vom Zapfen- zum Stäbchensehen markiert.

Zumindest das letztgenannte Beispiel bezieht sich auf Sensitivitätsveränderungen, die im absoluten Schwellenbereich beobachtbar sind. Man kann aber auch Sensitivitätsveränderungen im überschwelligen Bereich bei konstanter physikalischer Reizgröße beobachten. Abb. 2.11 verdeutlicht dies anhand der je nach Hintergrund unterschiedlich empfundenen Helligkeit von



**Abb. 2.10** Dunkeladaptationskurven der Stäbchen und Zapfen. Die durchgezogene Linie zeigt die aus beiden Kurven resultierende Adaptationskurve mit dem Kohlrausch-Knick



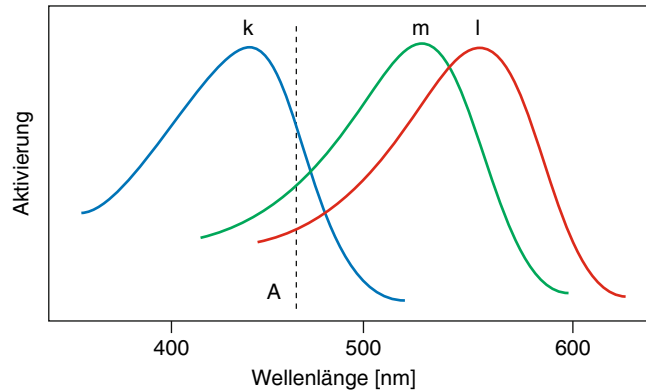
■ **Abb. 2.11** Simultankontrast. Fixieren Sie in der Mitte zwischen den beiden Quadraten und beurteilen Sie die beiden grauen Scheiben. Obwohl sie identisch sind, erscheint die Scheibe vor dem hellen gelben Hintergrund deutlich dunkler

Oberflächen, des *Simultankontrasts*. Obwohl die beiden grauen Scheiben sich nicht unterscheiden, wirkt das Feld vor dem hellen Hintergrund deutlich dunkler. Erklärt wurde dies zunächst mit einem eher peripheren Mechanismus, der als *laterale Inhibition* bezeichnet werden kann: Die Lichtreizung von Rezeptoren hemmt die Entladung benachbarter Rezeptoren – mit der Konsequenz, dass der hellere Hintergrund die Rezeptoren der rechten grauen Scheibe stärker hemmt als der dunkle Hintergrund. Neurophysiologisch ließ sich dieser Mechanismus tatsächlich bereits auf der Retina nachweisen (Hartline et al. 1956). Neuere Befunde legen allerdings auch nahe, dass am Simultankontrast zusätzlich corticale Faktoren beteiligt sind (neuerer Überblick zur Kontrastverarbeitung in Wirtz 2001; vgl. auch das Phänomen der Helligkeitskonstanz in ► Abschn. 2.3.5).

Laterale Inhibition wird im Übrigen auch als Erklärung für das Kontrastgitter (Hermann'sches Gitter; eine Variante, das Scintillating Grid, ist in ■ Abb. 2.1 dargestellt) herangezogen. Die Annahme ist, dass die Kreuzungspunkte von vier Seiten laterale Inhibition erhalten, während die übrigen Linien nur von zwei Seiten lateral gehemmt werden. Die deshalb ausgeprägtere laterale Inhibition lässt in der Peripherie an den Kreuzungspunkten dunkle Punkte entstehen (vgl. auch Spillmann 1994). – Der Simultankontrast ist übrigens nicht auf Helligkeitsstufen beschränkt. Ein auf ähnlichen Mechanismen beruhender Farbsimultankontrast lässt sich beispielsweise erzeugen, wenn ein graues Feld von einer roten Umrandung umgeben wird. Es erscheint dann leicht grünlich.

### 2.3.2 Farbwahrnehmung

Wie viele Farben können wir sehen? Wenn weißes Licht mithilfe eines Prismas gebrochen wird, zeigt sich, dass das menschliche Auge für Wellenlängen zwischen 400 und 700 nm empfindlich ist. „Farben“ existieren in der physikalischen Außenwelt nicht, das Licht wird nur mit unterschiedlicher Wellenlänge von den Oberflächen reflektiert. Die Farbwahrnehmung bestimmt sich neben der Wellenlänge aus der *Intensität* eines Farbreizes (also der Helligkeit) und dessen *Sättigung*. Trifft z. B. hauptsächlich Licht der Wellenlänge 500–570 nm auf die Retina, nehmen wir ein gesättigtes Grün wahr. Werden andere Wellenlängen in gleichen Anteilen hinzugemischt (nimmt also der Weißanteil einer Farbe



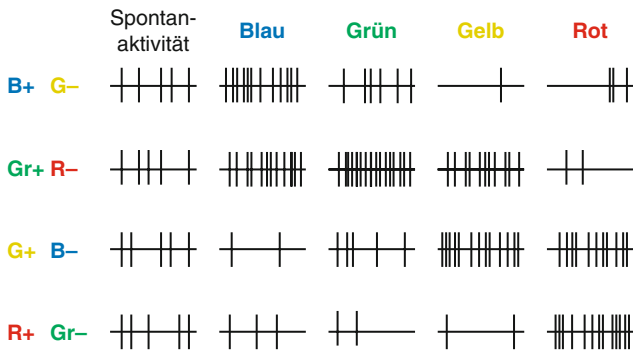
■ **Abb. 2.12** Aktivierungskurven von relativ kurz- (k, blauwellige Zapfen, Maxima ca. 420 nm), mittel- (m, grünwellige Zapfen, ca. 530 nm) und langwelligem (l, rotwellige Zapfen, ca. 560 nm) Rezeptorsystemen. Reiz A stimuliert die drei Rezeptortypen in einem bestimmten Verhältnis

zu), nimmt die Sättigung ab. Man schätzt, dass das menschliche Auge auf diese Weise zwischen 2 und 7 Mio. Farbabstufungen differenzieren kann (Tessier-Lavigne und Gouras 1996).

Wie sieht aber der sensorische Code aus, der diese differenzierte Farbwahrnehmung ermöglicht? Betrachtet man die ungeheure Anzahl der wahrnehmbaren Farbabstufungen auf jedem Punkt der Retina, ist klar, dass nicht jede Farbnuance durch ihren eigenen Rezeptor und einen zugehörigen neuronalen Verarbeitungsstrang übermittelt wird. Vielmehr ist anzunehmen, dass eine begrenzte Anzahl von Neuronenpopulationen ein spezifisches Aktivierungsmuster generiert, das zur Farbwahrnehmung führt. Diese Vorstellung hat schon früh zu psychophysischen Farbmischungsexperimenten und Theorien der Farbwahrnehmung geführt. So stellten beispielsweise Thomas Young und Hermann von Helmholtz im 19. Jahrhundert fest, dass jede mögliche Farwahrnehmung durch eine Mischung der drei Grundfarben Blau, Rot und Grün erzeugt werden kann. In ihrer *Dreifarbentheorie* beruht Farbwahrnehmung daher auf drei Rezeptorsystemen mit jeweils unterschiedlicher spektraler Empfindlichkeit (Helmholtz 1910). In ■ Abb. 2.12 hat z. B. der Reiz A einen relativ hohen Anteil an einer kurzwelligen Aktivierungskomponente und niedrige Anteile an der mittel- und langwelligen Aktivierungskomponente, was in diesem Fall zu einem blauen Farwahrnehmen führt. Werden die Rezeptorsysteme in einem anderen Verhältnis aktiviert, entsteht eine andere Farwahrnehmung. Man beachte aber, dass trotz eines unterschiedlich zusammengesetzten physikalischen Frequenzspektrums die Empfindung gleich bleiben kann. In diesem Fall spricht man von *metameren Farben*.

Erst in den 1980er Jahren konnte man beim Menschen mit den sich eröffnenden neuen physiologischen Registrierungsmöglichkeiten drei verschiedene Zapfensysteme mit Pigmentabsorptionskurven ermitteln, die ähnliche Maxima und Verläufe aufweisen wie in ■ Abb. 2.12 (z. B. Dartnall et al. 1983). Die Dreifarben Theorie erhielt auf diese Weise ihre physiologische Bestätigung.

Zur gleichen Zeit, als die Dreifarben Theorie entwickelt wurde, machte man auch folgende Beobachtungen: Betrachtet man über mehrere Minuten ein blaues Feld und schaut danach auf eine weiße Fläche, so stellt sich ein gelbes *Nachbild* ein (so wie sich nach dem Betrachten eines gelben Feldes ein blaues Nachbild



**Abb. 2.13** Mögliche Entladungsraten von Gegenfarbenzellen im CGL des Rhesusaffen über die Zeit. Die Zelle B+G- reagiert z. B. mit einer vermehrten Aktivität auf eine Reizung mit blauem Licht, während sie auf eine Reizung mit gelbem Licht eher inhibiert wird. (Nach DeValois und Jacobs 1968)

einstellt). Ähnliches gilt für die Farben Rot und Grün. Aufgrund solcher Feststellungen und aufgrund der Beobachtung, dass Rotblinde immer auch grünblind bzw. Blaublinde immer auch gelbblind sind, verfasste Ewald Hering seine *Gegenfarbentheorie* (Überblick in Hurvich und Jameson 2001). Danach sind Rot/Grün, Blau/Gelb und auch Schwarz/Weiß antagonistisch wirkende Farbpaare, d. h., sie lösen jeweils einen Gegenfarbenmechanismus aus, der dazu führt, dass auf eine Farbe positiv und auf die andere negativ reagiert wird.

Eine neurophysiologische Bestätigung fand die Gegenfarbentheorie in den 1960er Jahren in den sogenannten Gegenfarbenzellen des CGL des Rhesusaffen (zur Lokalisation des CGL beim Menschen s. **Abb. 2.4**). Dort fand man beispielsweise (B+G-)-Zellen, die mit einer vermehrten Aktivität auf eine Reizung mit blauem Licht reagierten, während sie auf eine Reizung mit gelbem Licht eher inhibiert wurden (DeValois und Jacobs 1968; **Abb. 2.13**).

Dreifarbentheorie und Gegenfarbentheorie schließen sich nicht – wie man zunächst glaubte – aus. Beide Mechanismen ergänzen sich. Neuronale Schaltkreise der Eingangsimpulse von den k-, l- und m-Zapfen erzeugen die aktivierenden und hemmenden Wirkungen der nachgeschalteten Gegenfarbenzellen, die sich u. a. auch in den Ganglienzellen der Retina finden. Auf diese Weise ist man in der Lage, verschiedenartige Farbphänomene zu erklären, unter ihnen die Effekte des Farbnachbildes oder des Farbsimultankontrasts (s. oben; neuere Überblicke zur Farbwahrnehmung bei Gegenfurtner 2012; Gegenfurtner und Sharpe 2000). Unabhängig davon hat gerade die Erforschung der Farbwahrnehmung gezeigt, dass bisweilen aus psychophysischen Experimenten gewonnene Theorien weit vor ihrer physiologischen Bestätigung vorliegen können. Im vorliegenden Fall ist zwischen der Entwicklung der Dreifarben- bzw. der Gegenfarbentheorie und ihren physiologischen Bestätigungen mehr als ein halbes Jahrhundert vergangen.

### 2.3.3 Raum- und Tiefenwahrnehmung

Visuelle Raum- und Tiefenwahrnehmung entsteht aus den zweidimensionalen Projektionen unserer dreidimensionalen Umwelt auf die Retina. Wir befassen uns zunächst mit zweidimensionalen

len Vorlagen, die einen Raum- und Tiefeneindruck entstehen lassen, und gehen dann zu den dreidimensionalen Anordnungen über.

#### Raum- und Tiefenwahrnehmung bei zweidimensionalen Vorlagen

**Relative Ortsinformation ohne Tiefe** Betrachten wir zunächst die relative Ortsinformation ohne Tiefe, die bei der Betrachtung einer einfachen zweidimensionalen Fläche entsteht. Zum Beispiel stehen die Abbildungen und Absätze auf einer Buchseite in einer räumlichen Relation zueinander, die es zu identifizieren und zu repräsentieren gilt. Dies allein ist alles andere als trivial. An dieser Stelle gilt es, zwei Trugschlüsse auszuräumen, die in diesem Zusammenhang in populärwissenschaftlichen Darstellungen immer wieder auftauchen. Beide stehen in Verbindung mit einer zu wortwörtlich genommenen Kameraanalogie (**► Abschn. 2.1**).

Der erste Trugschluss entsteht dadurch, dass unsere Umwelt auf der Retina seiten- und höhenverkehrt abgebildet ist, d. h., die oberste Zeile einer Seite ist auf der Retina unten projiziert, die unterste Zeile oben. Da wir unsere Welt aber nicht auf dem Kopf stehend wahrnehmen, scheint eine Leistung des visuellen Verarbeitungsapparats darin zu bestehen, die Welt wieder „richtig“ herum zu drehen. Dies ist natürlich irreführend. Es gibt keinen Verarbeitungsprozess, und es besteht auch keine Notwendigkeit für einen Verarbeitungsprozess, der das seiten- und höhenverkehrte Abbild „umdreht“. Was das System bestenfalls während der ontogenetischen Entwicklung lernt, ist die Konsistenz und die Etikettierung dessen, was in unserem Sprachgebrauch als „oben“ und „unten“ bezeichnet wird. Oben auf der Retina ist der Ort, zu dem sich die Objekte bewegen, wenn man sie loslässt – also, unserem Sprachgebrauch folgend, nach unten (vgl. in diesem Zusammenhang die sogenannten Umkehrbrillenexperimente, die oben und unten vertauschen; Kohler 1962).

Der zweite Trugschluss entsteht, wenn man versucht, Raumwahrnehmung durch die Existenz einer der retinalen Abbildung ähnlichen topografischen Gehirnkarte zu erklären. In der heutigen Zeit ist diese Gefahr recht groß, kann man doch durch geeignete neurophysiologische Methoden derartige topografische Karten sogar an verschiedenen Stellen des visuellen Verarbeitungspfad nachweisen. Beispielsweise sind Strukturen im CGL (dort insbesondere in Schicht 6), aber auch in V1 und späteren visuellen Arealen retinotop organisiert. Diese Karten bilden die räumlichen Retinaverhältnisse neuronal nach, wenn auch meist erheblich verzerrt. Dennoch liegt die Annahme nahe, dass der visuell wahrgenommene Raum mit der Existenz einer solchen Karte repräsentiert und Raumwahrnehmung als solche hinreichend geklärt ist (*position as a code for position*-Trugschluss; Van der Heijden et al. 1999). In Wirklichkeit handelt es sich lediglich um eine Verschiebung des Problems, weil man zur Interpretation einer topografischen Karte einen Homunculus („einen kleinen Mann im Kopf“) benötigt, der die räumlichen Informationen auszulesen weiß (**Homunculus-Problem**; nicht zu verwechseln mit dem Homunculus im primär-motorischen Cortex; **► Kap. 20**). Es nützt zum Verständnis des Wahrnehmungsprozesses und hier insbesondere der Raumwahrnehmung nichts, wenn man die Retina in **Abb. 2.2** durch eine retinotop Karte des Gehirns ersetzt. Erklärungsbedürftig bleibt dann, wie der



Menschenkopf bzw. der Homunculus die räumliche Karte ausliest und damit zu einem Raumeindruck gelangt.

So muss man grundsätzlich bezweifeln, dass die retinotopische Organisation neuronaler Karten unmittelbar metrische Informationen über die konkreten Abstände von Objekten enthält. Zwar liefert die retinale Stimulation die Substanz für die räumliche Struktur, indem durch sie die lokalen Nachbarschaftsbeziehungen der Objekte (*local receptive field structure*) festgelegt wird, es bedarf aber eines zusätzlichen global operierenden Codes (*multi-local code*; Koenderink 1990), um die wahrgenommene Position eines Objekts zu bestimmen. Einige Autoren vermuten, dass sich dieser Code erst durch verifizierbare Blickbewegungen etabliert (Van der Heijden et al. 1999; Wolff 1999; Van der Heijden 2004, Kap. 8; für eine frühe Version dieser Idee vgl. Helmholtz 1866).

**Tiefenwahrnehmung bei zweidimensionalen Vorlagen** Wenden wir uns nun dem Tiefenaspekt zu. Das Betrachten eines zweidimensionalen Fotos oder eines Gemäldes ist meist mit einem mehr oder weniger starken Tiefeneindruck verbunden. Dies ist selbst dann der Fall, wenn man ein Auge abdeckt, weswegen diese **Tiefenkriterien** auch als *monokulare Tiefenkriterien* bezeichnet werden. Es sind dies u. a.:

- **Linearperspektive und Texturgradient:** Für die Linearperspektive ist kennzeichnend, dass parallele Linien in der Tiefe in einem Fluchtpunkt zusammenlaufen, für den Texturgradienten, dass z. B. frontal parallele Linien, die gleich weit voneinander entfernt sind, in der Tiefe dichter gepackt erscheinen. Diese Tiefenkriterien wurden in der Malerei der Frührenaissance perfektioniert, entstammen also nicht der neuzeitlichen Psychologie. So führte auch Leonardo da Vinci (1452–1519) neben der Linearperspektive die *Farbperspektive* (d. h. die Verblassung bzw. Verblauung von Farben mit zunehmender Tiefe) und die *Luftperspektive* (auch atmosphärische Perspektive oder *Sfumato*, d. h. das Verschwimmen der Konturen bei sehr weiten Distanzen) ein.
- **(Teil-)Verdeckung:** Verdeckt ein Objekt ein anderes, so wird dies als näher empfunden.
- **Relative Größe und Bekanntheitsgrad von Objekten:** Größere Objekte erscheinen näher als kleinere Objekte, allerdings wird dies durch die Lage des Horizonts modifiziert (sogenannte Mondtäuschung). Ist die Größe eines Objekts bekannt, wirkt sich dies ebenfalls auf die wahrgenommene Größe aus.

Schon diese Tiefenkriterien erlauben in den meisten Fällen eine ausgezeichnete Abschätzung der tatsächlichen Tiefe. An der Tiefenwahrnehmung sind aber noch weitere Faktoren beteiligt.

#### ■ **Raum- und Tiefenwahrnehmung in dreidimensionalen Anordnungen**

**Querdisparation** Betrachten wir in einer realen (oder simulierten) dreidimensionalen Anordnung von Objekten zunächst einen stationären Beobachter bei fixiertem Blick. Schaltet man in dieser Anordnung die o. g. monokularen Tiefenkriterien weitgehend aus, entsteht dort Tiefenwahrnehmung aufgrund der paarigen Anlegung unserer Augen – genau genommen aufgrund der beiden geringfügig unterschiedlichen Augenblickwinkel. In

sogenannten Stereoskopen oder 3-D-Filmen macht man sich diese Eigenschaft des visuellen Apparats zunutze, indem man unterschiedliche Bilder zu den beiden Augen projiziert. Wie entsteht dadurch der Tiefeneindruck?

Jeder Punkt auf einer Netzhaut korrespondiert mit einem Punkt auf der anderen Netzhaut (vgl. aber das Korrespondenzproblem; s. unten). Fixiert man einen Punkt *F* in einer räumlichen Anordnung, so wird er auf korrespondierende Netzhautstellen abgebildet, in diesem Fall in der Fovea. Auf korrespondierende Netzhautstellen fallen aber auch alle Punkte eines virtuellen Kreises, der durch den fixierten Punkt *F* und dem optischen Mittelpunkt beider Augen verläuft, den sogenannten theoretischen **Horopter** (der empirische Horopter divergiert von dieser Idealform insbesondere im äußeren Gesichtsfeld; Palmer 1999, S. 206 ff.). Genau genommen ist der Horopter keine eng umgrenzte Linie, sondern umfasst ein mehr oder weniger breites Areal um den Horopter, in dem die Objekte der beiden Netzhäute noch zu einem Objekt fusioniert werden (*Panum-Areal*). Ansonsten würden Doppelbilder wahrgenommen. In jedem Fall fallen alle anderen Punkte unseres Gesichtsfeldes außerhalb des Horopters auf nicht korrespondierende Netzhautstellen, d. h., die Orte ihrer Projektionen weichen im linken und rechten Auge ein wenig voneinander ab, und es entsteht die **Querdisparation**. Das unterschiedliche Ausmaß der Querdisparation lässt die unterschiedlich wahrgenommenen Entfernungen zu den Objekten entstehen (Abb. 2.14; s. auch ► Zur Vertiefung 2.3).

In V1 finden sich die ersten binokularen Zellen, die nur auf Stimulation in beiden Augen reagieren. Zellen, die ein optimales Tuning bei einem bestimmten Ausmaß an Querdisparation zeigen, lassen sich bei der Katze vor allem in V1 und V2 isolieren (LeVay und Voigt 1988). Allerdings scheinen querdisparationsempfindliche Zellen zusätzlich auf der gesamten parietalen Bahn und zum Teil auch auf der temporalen Bahn zu existieren (Tyler 1990).

Ungeklärt ist dabei bis heute weitgehend, wie das visuelle System das **Korrespondenzproblem** löst. Es muss ja zunächst die korrespondierenden Informationen der beiden Netzhauthälften zueinander in Beziehung setzen, bevor es die Disparität feststellen kann. Dies ist ein besonders auffälliges Problem, wenn man Stereogramme aus Zufallspunktmustern betrachtet (*random-dot stereograms*; Julesz 1971). Dort ist die Zuordnung der Punkte ja keineswegs trivial (neuerer Überblick zum stereoskopischen Sehen in DeAngelis 2000; Marr 1982; computationale Lösungsvorschläge dieses Problems in Marr und Poggio 1979).

#### **Okulomotorische Faktoren der Raum- und Tiefenwahrnehmung**

Querdisparation ist ein binokulares Tiefenkriterium. Weitere binokulare Tiefenkriterien sind die Konvergenz und die Akkommodation. Der *Konvergenzwinkel* der Augen, also die Stellung der Augen zueinander, gemessen als Winkel am Fixationspunkt, ist bei Fixation eines nahen Objekts größer als bei Fixation eines weit entfernten Objekts. Die *Akkommodation* ist die Anpassung der Form der Linse, die bei nahen Objekten bauchiger und bei entfernten Objekten flacher ist. In beiden Fällen berücksichtigt das visuelle System u. a. die propriozeptive Rückmeldung von den Augenmuskeln bei der Tiefenberechnung, sodass diese Tiefenkriterien auch als okulomotorische Tiefenkriterien bezeichnet werden.

Allgemeine Psychologie

Müsseler, J.; Rieger, M. (Hrsg.)

2017, XXIV, 850 S. 357 Abb. in Farbe., Hardcover

ISBN: 978-3-642-53897-1