e-teaching.org

Lernen mit Text und Bild

Autoren: Uwe Oestermeier & Alexander Eitel
Portalbereich: Didaktisches Design
Stand: 05. Mai 2014

Inhaltsverzeichnis

Einleitung und Übersicht 1
1. Bild und Sprache 2
2. Eine Taxonomie bildlicher und grafischer Codes 4
3. Kulturelle und kognitive Voraussetzung des Verstehens von Bildern 10
4. Psychologische Theorien des Bildverstehens 11
5. Potenziale des Bildgebrauchs 14
6. Probleme des Bildgebrauchs 17
7. Theorien und Befunde zum Lernen mit Text-Bild-Kombinationen 19
8. Richtlinien für Text-Bild-Kombinationen 24
Literatur 30
Glossar 34
Die Autoren 35

Begriffe ohne Anschauung sind blind.
Immanuel Kant, Kritik der reinen Vernunft, 1781

Einleitung und Übersicht

Ein Lernen nur mit Bildern und Grafiken gibt es nicht, jedenfalls nicht in nennenswertem Umfang. Fast alle Formen der Bildkommunikation, die zu Lernzwecken tagtäglich in Familie, Kindergarten, Schule und Weiterbildungseinrichtungen stattfinden, sind Mischformen, bei denen immer geschrieben oder gesprochen Erläuterungen zu den Grafiken und Bildern hinzukommen. Sei es, dass Eltern Dinge in Bilderbüchern benennen, Lehrer erzählen, auf was zu achten ist und Dozenten einschränkend bemerken, was auf den Bildern nicht zu sehen ist. Es ist deshalb die Frage, ob es in solchen Lernkontexten überhaupt bildliche Informationen gibt, die ohne eine sprachliche Einbettung verständlich sind. Selbst in Kunstmuseen, in denen Bilder als künstlerische Ausdrucksform noch am ehesten für sich selbst stehen, gleitet der Blick des Betrachters in der Regel zuerst auf die kleinen Schildchen neben den Bildern, um anhand der sprachlichen Titel- und der Herkunftsangaben das Bild und seinen Inhalt einordnen zu können. Bildverstehen setzt deshalb in der Praxis regelmäßig Sprachverständnis voraus.

Dieses Abhängigkeitsverhältnis von Bild und Sprache ist der rote Faden dieses Textes: Im ersten Kapitel wird das Verhältnis von Bild und Sprache eingeführt und erläutert. Nach einigen unabdingbaren Begriffsbestimmungen im zweiten Kapitel, in denen verschiedene bildliche und grafische Formen unterschieden werden, wird dann im dritten Kapitel versucht zu verdeutlichen, dass Bilder und Grafiken zwar als Selbstverständlichkeiten angesehen werden, sie aber tatsächlich nur selbstverständlich sind, weil zahlreiche kulturelle und individuelle Voraussetzungen wie z.B. das Lesen von Schrift, Mathematikkenntnisse u.v.a.m. stillschweigend vorausgesetzt werden.

Der Hauptteil dieses Textes ist psychologisch geprägt und problematisiert das Verhältnis von bildlichem und sprachlichem Denken anhand der sogenannten „Imagery-Debatte", in der Kognitionspsychologen vor allem in den siebziger und achtziger Jahren den Stellenwert von bildlichen Vorstellungen diskutierten haben. In diesem Abschnitt werden auch aktuelle Theorien des Text-Bildverstehens und deren Erklärungskraft problematisiert. Zwei Kapitel über Potenziale und Probleme des Bildgebrauchs leiten schließlich zu praktischen Konsequenzen im letzten Kapitel über. Auch ein Großteil dieser praktischen Hinweise ergibt sich aus dem Zwang zur sprachlichen Einbettung bildlicher Informationen: Wer didaktisch gut aufbereitetes Bildmaterial erstellen möchte, sollte sein Hauptaugenmerk nicht primär auf die Gestaltung des Bildmaterials richten, sondern auf die Text-Bild-Bezüge und die Einbettung der Bilder in den Lernkontext. Die zentrale These lautet also: Wer Lernen durch den Einsatz von Grafiken und Bildern fördern will, sollte nicht nur Grafiken und Bilder gestalten, sondern vor allem Text-Bild- bzw. Audio-Bild-Kombinationen.

1. Bild und Sprache

Die kognitive Funktion von Bildern ist bereits in der klassischen griechischen Philosophie intensiv diskutiert worden. Die damals angesprochenen Probleme sind auch in den erkenntnistheoretischen Debatten der folgenden Jahrhunderte immer wieder aufgegriffen worden und bestimmen bis heute das Nachdenken über das Denken in der modernen Kognitionspsychologie und Philosophie des Geistes.

Viele Menschen halten das bildliche Denken subjektiv für wichtig und unentbehrlich, darunter auch viele Wissenschaftler (wie u.a. Albert Einstein), die von sich behaupten, dass sie ihre Entdeckungen vor allem bildlichen Vorstellungen verdanken (Miller, 1984). Auch in der Alltagssprache gelten „Anschaulichkeit" und „Bildlichkeit" als Synonyme für Verständlichkeit. Aber letztlich geht es in der psychologischen Diskussion auch nicht um diese subjektiven Einschätzungen, sondern um die Frage, ob es ungeachtet dieser introspektiven Berichte und sprachlichen Redewendungen überhaupt mentale Repräsentationen gibt, die in einem nachvollziehbaren Sinne bildhaft sind und ob diese mentalen Repräsentationen eine echte erkenntnisfördernde Funktion haben (Oestermeier, 1998). Dass externe Bilder solch eine Funktion haben können, ist wohl unbestritten, wenn es auch immer wieder Stimmen gibt, die davor warnen, die Bedeutung von Bildern zu hoch zu bewerten (Pörksen, 1997) und die Bedeutung der Sprache zu unterschätzen.

Das Verhältnis von Sprache und Bild ist von einer großen Vielfalt an Bezugsmöglichkeiten geprägt. Dies gilt umso mehr, je stärker konkrete Bilder von fiktiven und realen Personen, Tieren und anderen körperlichen Gegenständen zugunsten abstrakter Grafiken und Diagramme zurücktreten, die keine sichtbaren Einzeldinge sondern unsichtbare Gegenstände und allgemeine Gesetzmäßigkeiten illustrieren sollen. Die folgende Tabelle vermittelt einen ersten groben Überblick über die verschiedenen kognitiven Funktionen von Bildern und Grafiken und ihren Bezug zu den jeweils notwendigen sprachlichen Zusatzinformationen:

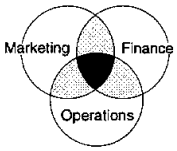
Kognitive Funktion	*Beispiel*	*Gegenstandsbereich*	*Bildtypen*	*Sprachliche Ergänzungen*
Wahrnehmen und Wiedererkennen		sichtbare körperliche Gegenstände	Fotos, Abbilder, Malerei	Namen, Orts- und Zeitangaben, Beschreibungen, Behauptungen
Klassifikation und Veranschaulichung		unsichtbare Gegenstände, Klassen von Gegenständen	Typisierte Illustrationen, Planskizzen, Modellzeichnungen	Definitionen, Erläuterungen, Erklärungen
Schlussfolgerungen		abstrakte mathematische Gegenstände, quantitative Zusammenhänge, Naturgesetze	geometrische Planfiguren, logische Diagramme	Negationen, Definitionen, Beweise, Formeln, Gesetze, Argumente

Tabelle 1. Kognitive Funktionen von Grafiken und Bildern

Diese Tabelle geht davon aus, dass in der Praxis immer sprachliche Zusatzinformationen benötigt werden, um die kognitive Funktion der Grafiken nutzen zu können.

Blickbewegungsstudien belegen, dass fast alle Leser zuerst die Bilder einer Seite überfliegen, um dann im Wesentlichen bei dem Text zu bleiben (Frey, 1999; Rayner, Rotello, Stewart, Keir, & Duffy, 2001). Auch Bilder, die einen im Verhältnis zum Text echten Mehrwert an Informationen enthalten, werden oftmals nicht länger angeschaut (Folker, Ritter, & Sichelschmidt, 2005; Hannus & Hyönä, 1999). Die Erklärung dafür ist vermutlich in der Strategie der Lernenden zu suchen, ihre kognitiven Ressourcen möglichst ökonomisch auf die wesentlichen und schwierigen Teile zu konzentrieren, wobei sie davon ausgehen, dass Bilder im Vergleich zu Texten leicht zu verstehen sind und deshalb kaum Aufmerksamkeit beanspruchen (Weidenmann, 1988). Eine Ursache dafür könnte in der Lesesozialisation liegen, die die nichtsprachlichen Anteile im Medium Buch nicht eigens thematisiert und als weitgehend unproblematisch ansieht. Untersuchungen zur Visual Literacy, d.h. zu der Frage inwieweit Menschen Bilder und andere visuelle Codes wie Tabellen und Listen verstehen bzw. „lesen" können, zeigen jedoch, dass diese Annahme falsch ist. Selbst in so klassischen Industrienationen wie den USA, Kanada und Deutschland haben mehr als 40 % der Bevölkerung Probleme Bilder, Grafiken und Tabellen in Bezug zu Texten und anderen Informationsquellen zu setzen und damit zu verstehen (OECD & Statistics Canada, 1995).

Bei der Gestaltung von Lernmaterialien kommt es also darauf an, dass der Text-Bild-Integration ein besonderes Augenmerk geschenkt wird. Es bringt wenig, wenn aufwändig erstellte Bilder und Grafiken von den Lesern als bloße Dekoration oder triviale Zutat abgetan werden. Es empfiehlt sich gegebenenfalls im Text explizite Lesehilfen für die Bilder und Grafiken zu formulieren und auf wesentliche Elemente und Inhalte direkt Bezug zu nehmen und zu erläutern wie sich Text und Bild ergänzen. Um den Leser bzw. Betrachter die

Text-Bild-Integration zu erleichtern sollte auch beachtet werden, dass die aufeinander bezogenen Bild- und Textteile möglichst nah beieinander auf derselben Seite angeordnet sind. Aus drucktechnischen Gründen werden Bilder oft auf gesonderte Teile des Buches beschränkt, allerdings gibt es kaum ein effektiveres Mittel als ständiges Blättern, um die Text-Bild-Integration zu unterbinden.

2. Eine Taxonomie bildlicher und grafischer Codes

Im Gegensatz zu Texten sind die syntaktischen Möglichkeiten von Bildern viel unbestimmter und letztlich auch vielfältiger. Die Spannbreite reicht von abstrakten Liniendiagrammen, die mit einigen klar definierten Strichen und Ziffern auskommen, bis zu konkreten Aquarellen und Fotografien, bei denen es auf feinste Farb- und Formnuancen ankommen kann. Diese Unterschiede in der Darstellungsform sind für die Verarbeitung von Bildern schon allein deshalb relevant, weil es fraglich ist, ob bei der Verarbeitung von realistischen Bildern überhaupt dieselben kognitiven Prozesse ablaufen wie bei der Interpretation abstrakter Darstellungen. Bei realistischen Bildern ist es plausibel, sie als Stellvertreter für die konkreten Gegenstände selbst zu betrachten, bei denen dieselben hoch automatisierten und extrem schnellen Wahrnehmungs- und Objekterkennungsprozesse ablaufen, wie bei der Beobachtung der Gegenstände der normalen Umwelt selbst. Weidenmann spricht deshalb in diesem Zusammenhang von einem ökologischen Verstehensmodus (Weidenmann, 1988).

Unter normalen Umständen und bei vertrauten Gegenständen laufen diese Prozesse ohne weiteres Nachdenken ab. Demgegenüber ist es bei vielen abstrakten Grafiken von vornherein klar, dass sie nicht von sich aus verständlich sind und Verstehensleistungen erfordern, bei denen bewusst und kontrolliert Zusatzinformationen gesucht bzw. Vorwissen abgerufen werden muss, um durch weitere Schlussfolgerungen den vollen Gehalt der Darstellung zu erschließen. Weidenmann spricht hier von einem indikatorischen Verstehensmodus, der nur aktiviert wird, wenn das ökologische Bildverständnis allein nicht greift und es zu Irritationen und einem Normalisierungsbedarf kommt.

Bei näherer Betrachtung zeigt sich jedoch, dass diese unterschiedlichen vom Rezipienten zu erbringenden Verständnisleistungen nicht zwangsläufig an unterschiedlichen Darstellungsformen festgemacht werden können. Auch Grafiken können naiv angeschaut werden, wenn z.B. ein Kind einen Graphen als Berg und Tal bezeichnet, und auch realistische Bilder können Stoff zu längerem Nachdenken liefern (wie z.B. die Wehrmachtsausstellung gezeigt hat).

Der Unterschied zwischen ökologischem und indikatorischem Bildverständnis ist natürlich trotzdem wichtig, denn die vertiefte Reflektion über einen Sachverhalt ist die eigentliche didaktische Stärke von statischen Bildern. Animationen, Filme und andere dynamische Bilder nehmen unser auf Bewegungserkennung getrimmtes visuelles System dermaßen in Anspruch, dass man buchstäblich keinen klaren Gedanken fassen kann, solange sich im Blickfeld etwas bewegt. Nicht vorhersagbare Bewegungen im Bild führen zu einem neuronalen Gewitter im visuellen Cortex, das andauernde unwillkürliche Orientierungsreaktio-

nen nach sich zieht. Dies erklärt auch, warum viele Menschen zur Entspannung so gerne fernsehen (Kubey & Csikszentmihalyi, 2002). Fernsehen befreit vom selbstbestimmten Nachdenken. Bei einem unbewegten statischen Bild hingegen bestimmt der Rezipient, wann und wie oft er sich mit einem Teilaspekt befassen will. Er kann dabei ganz seinen eigenen Interessen und seinem Vorwissen gemäß die Rezeptionsweise und -geschwindigkeit festlegen.

Auch in Zeiten der zunehmenden Verbreitung von bewegten und interaktiven Bildern lohnt es, sich auf die besonderen didaktischen Funktionen statischer Abbildungen zu konzentrieren und deren strukturellen Merkmale gesondert zu betrachten. Wir lassen hier also bewusst dynamische und interaktive Bilder und Grafiken außen vor. Die im Folgenden vorgeschlagene Taxonomie lehnt sich an den in der Psychologie dominierenden angelsächsischen Sprachgebrauch an (vgl. Vekiri, 2002). Die Unterschiede ergeben sich im wesentlichen aus den folgenden Aspekten, die für Visualisierungen zentral sind:

- Die Hauptgattungen "darstellende Bilder", "logische Bilder" und "Mischformen" unterscheiden sich vor allem dadurch, welchen Gegenstandsdimensionen (z.B. Raum, Zeit, Statusunterschiede) die räumlichen Dimensionen des Papiers oder des Bildschirms bezeichnen.
- Weitere Binnendifferenzierungen ergeben sich durch den typischen Gegenstandsbereich und die unterschiedliche kognitive Verwendung der Darstellungen, sowie
- die Nutzung nicht-bildlicher Codes und anderer oft nicht genau abzugrenzender Dimensionen wie z.B. "konkret – abstrakt ", "dokumentarisch – fiktiv", " realistisch – metaphorisch".

Darstellende (gegenständliche) Bilder. Mit der Erfindung der Fotographie konnten auch zeichnerisch Unbegabte erstmals getreue Abbildungen von Personen und Gegenständen anfertigen. Der Ausdruck „Fotorealismus“ ist geradezu zu einem Synonym für Abbildungstreue geworden, obwohl natürlich auch Fotos durch Brennweite, Auflösung, Ausschnitt, Farbabweichungen und nicht zuletzt durch die Flächigkeit des Mediums selbst nie völlig gegenstandsgetreu sein können.

Das Grundprinzip der gemeinten Abbildungstreue ist jedoch klar: Zwischen der räumlichen Konfiguration der abgebildeten Gegenstände und der Darstellung selbst besteht eine klar definierte geometrische Gesetzmäßigkeit, bei der benachbarte Punkte im Raum zu benachbarten Punkten in der Abbildung werden. Bei der Interpretation der Abbildung dominiert die Ähnlichkeit zwischen Gegenstand und Zeichen, so dass man sagen kann, dass es kennzeichnend für darstellende Bilder ist, dass sie räumlich-körperliche Gegenstände räumlich ähnlich repräsentieren. Diese Ähnlichkeit kann natürlich mehr oder weniger ausgeprägt sein:

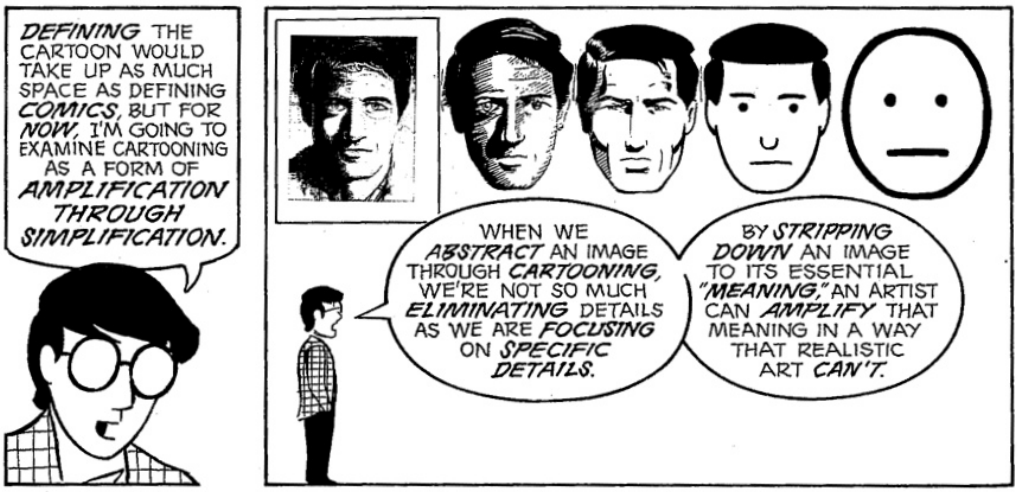


Abb. 1. Darstellende Bilder unterschiedlichen Abstraktionsgrades. Die räumliche Konfiguration der Bildelemente (der Farben bzw. Linien auf dem Papier) entspricht immer der tatsächlichen räumlichen Konfiguration der dargestellten Elemente (hier den Elementen Augen, Mund und Kopfform) (aus McCloud, 1993, S. 30).

- *Unterarten:* Ausgehend von dieser Grundbestimmung lassen sich mehrere Unterarten benennen, die entlang der Dimensionen „realistisch – schematisch" bzw. „konkret – abstrakt" angeordnet werden können, wie z.B. Fotos, Portraits, Zeichnungen, Skizzen, Karten, Explosions- und Aufrisszeichnungen, geometrische Planfiguren.
- *Gegenstandsbereiche*: Durch die Isomorphie sind solche Darstellungen auf räumliche Gegenstände beschränkt. Typische Einsatzbereiche für solche Darstellungen sind das Vermessungs- und Ingenieurwesen, die Botanik, Medizin und Architektur, um nur einige Bereiche zu nennen.
- *Kognitive Funktionen.* Dabei lassen sich mindestens zwei typische didaktische und kognitive Funktionen unterscheiden. Zum einen *veranschaulichen* darstellende Bilder Gegenstände, die dem bloßen Auge nicht zugänglich sind, bzw. nicht in ihrer Gänze überschaut werden können. Beispiele dafür sind Mikroskopaufnahmen, Röntgenbilder und Wetterkarten. Eine andere Funktion besteht in der *Erklärung* des Aufbaus und der Funktionsweise von komplexen konkreten Gegenständen, wie bei Bauplänen und Explosionszeichnungen technischer Apparate.

Logische Bilder. Im Gegensatz zu darstellenden Bildern zeigen logische Bilder (Schnotz, 2002) auch nicht-räumliche Gegenstände und Eigenschaften und sind somit inhärent abstrakt.

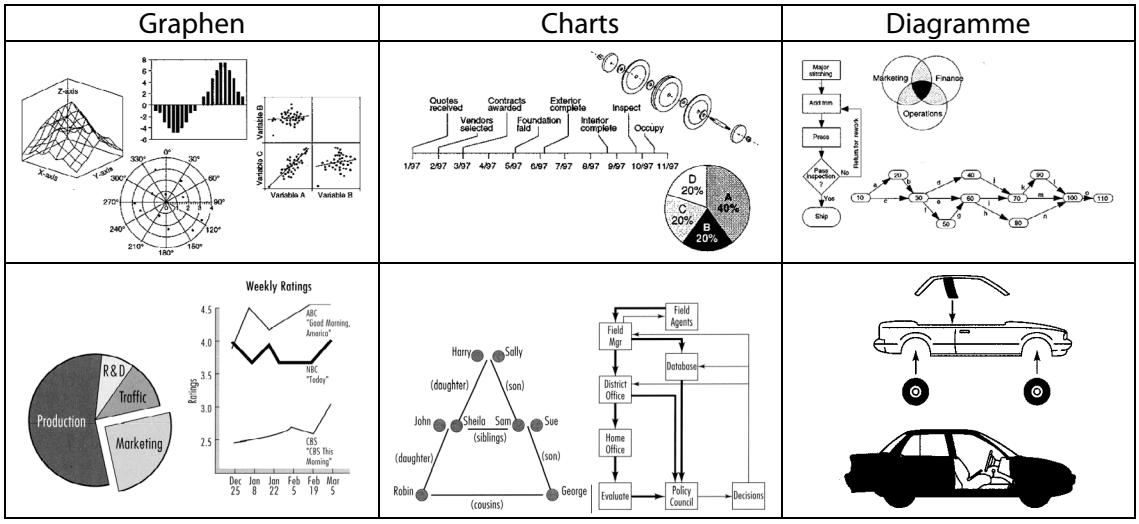


Tabelle 2. Unterschiedliche Beispiele für Graphen, Charts und Diagramme. Die Beispiele in der oberen Hälfte stammen aus Harris (1999, Titelseite). Die unteren Beispiele stammen aus Kosslyn (1994, S. 107, 141, 243, 254). Tortendiagramme werden z.B. von Kosslyn als „pie graphs“ und von Harris als „pie charts“ bezeichnet. Während bei Harris Diagramme im Wesentlichen abstrakt sind, gehören für Kosslyn neben symbolischen auch immer bildlich-konkrete Elemente zu einem Diagramm.

- *Unterarten:* Viele Autoren unterscheiden bei logischen Bildern Graphen, Charts und Diagramme. Leider ist der Sprachgebrauch hier sehr uneinheitlich, wie Tabelle 2 zeigt. Die folgende Unterscheidung ist aber relativ weit verbreitet.

  Von *Graphen* wird in der Regel gesprochen, wenn quantitative Werte in einem Koordinatensystem dargestellt werden, wobei die Achsen im Gegensatz zu Karten in der Regel nicht räumlich zu interpretieren sind. Die X- und Y-Achsen stellen z.B. Zeit, Leistungs- oder Verbrauchswerte dar.

  Von *Charts* wird gesprochen wenn es nicht um kontinuierlich variierende numerische Werte, sondern um Relationen zwischen distinkten Objekten geht. Beispiele sind Organigramme, Ablaufdiagramme u.ä. Auch bei Charts sind die räumlichen Relationen der Abbildungsfläche in der Regel nicht räumlich zu interpretieren. Die Relationen, die oft durch Linien oder Pfeile gekennzeichnet sind, stehen für Teil-Ganzes-, Ursache-Wirkungs- oder Vorher-Nachher-Relationen o. ä.

  Der Ausdruck *Diagramm* schließlich wird von manchen Autoren (z.B. Kosslyn, 1989) auf schematisierte Abbildungen angewendet, bei denen Teile der Abbildung Gegenstände darstellen und darüber hinaus noch abstrakte Konzepte (wie z.B. Kraft und Energiefluss) visualisieren. Diagramm wird aber oft auch als Oberbegriff verwendet, so dass alle logischen Bilder als Diagramme bezeichnet werden können. Welche Bedeutung letztlich gemeint ist, muss in der Regel aus dem Kontext erschlossen werden, und auch wir werden im Folgenden nicht von „Kausalcharts“ oder „Flusscharts“, sondern wie allgemein üblich, von „Kausaldiagrammen“ und „Flussdiagrammen“ sprechen.

- *Gegenstandsbereiche:* Prinzipiell lassen sich qualitative und quantitative Relationen unterscheiden, die in logischen Bildern dargestellt werden. Sofern numerische Daten und Koordinaten der Abbildung zugrunde liegen, spricht man von quantitativen Graphen, die insbesondere in den Naturwissenschaften zu den wichtigsten visuellen Kommunikations- und Darstellungsmitteln überhaupt gehören, aber natürlich auch in allen anderen quantitativ arbeitenden Disziplinen verwendet werden, wie z.B. der Ökonomie, der experimentellen Psychologie und vielen Ingenieursdisziplinen. Bei qualitativen Darstellungen hingegen kommt es nicht auf graduelle sondern auf kategoriale Unterschiede und Relationen an. Beispiele sind Venn-Diagramme, die in der elementaren Logik eingesetzt werden, oder Flussdiagramme, die oft in der Informatik zur Visualisierung von Programmabläufen eingesetzt werden. Auch in der Psychologie werden solche Kästchendiagramme häufig zur Visualisierung von Modellen und Theorien eingesetzt.
- *Kognitive Funktionen:* Bei quantitativen Zusammenhängen kommt es oftmals darauf an, Trends und Unterschiede möglichst übersichtlich aufzubereiten. Durch diese Übersetzung der Rohdaten – die ursprünglich meistens in Tabellen vorliegen – in Verlaufslinien und Größen sollen „auf einen Blick" Prognosen und Vergleiche durchgeführt werden können. Bei qualitativen Darstellungen liegt die Ausgangsbeschreibung meist verbal vor und die Visualisierung dient eher dem Erkennen von Strukturen und Relationen zwischen einzelnen Elementen, die einem linearen Text nur schwer zu entnehmen sind.
- *Symbole:* Da quantitative Darstellungen letztlich Zahlen visualisieren, gehören Ziffern und Variablennamen zu den üblichen Bestandteilen von Grafiken, während Charts und Diagramme oft mit kurzen verbalen Erläuterungen versehen sind. Die numerischen Werte dienen dabei oftmals der Erhöhung der Ablesegenauigkeit, während die anderen Symbole – wie bei den qualitativen Diagrammen auch – vor allem der Bezugnahme und damit der Text-Bild-Kohärenz dienen.

Mischformen. Über die genannten Gattungen und Arten hinaus ergeben sich weitere Möglichkeiten aus der Mischung der verschiedenen Bild- und Grafiktypen. Es ist keineswegs so, dass die beiden Interpretationen der räumlichen Fläche (Raum als Raum und Raum als nicht-räumliche Größe) nicht in einer Fläche kombiniert werden können. So können in einem Bild gleichzeitig Größenunterschiede von Gebäuden Nähe und Ferne signalisieren, während die Größe der Person Statusunterschiede signalisiert, indem ein groß dargestellter Herrscher aus einer Menge klein dargestellter Untertanen hervorgehoben wird. Für diese Mischung, bei denen die räumliche Größe der Abbildung mal perspektivisch und mal metaphorisch zu interpretieren ist, gibt es jedoch keinen eigenen Oberbegriff.

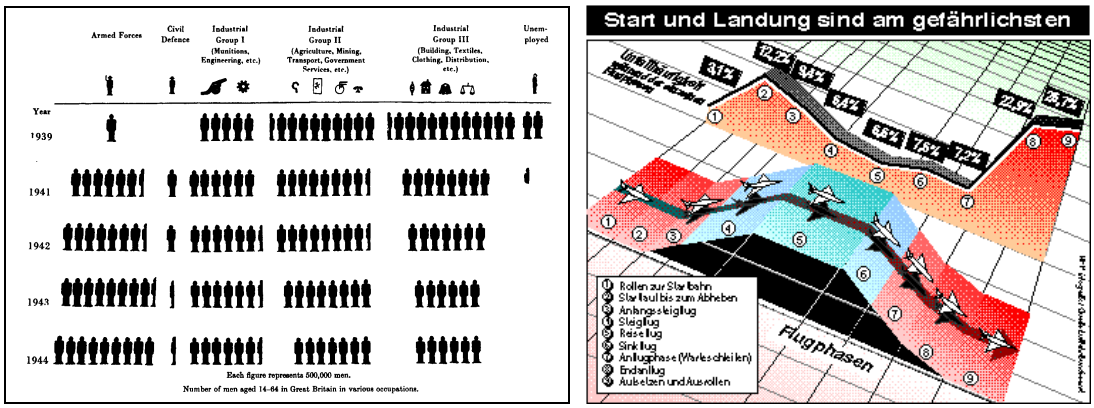


Abb. 2. ISOYTPE (links; aus Macdonald-Ross, 1977, S. 389) und Infografik (rechts; Quelle AFP) als Mischformen. In der ISOTYPE sind lediglich die Piktogramme dem Gegenstand ähnlich, die räumlichen Größen selbst geben vor allem Mengenverhältnisse wieder. In der Infografik wird hingegen die räumliche Höhe des Fluges räumlich analog wiedergegeben und mit einer räumlichen Darstellung nicht-räumlicher Häufigkeiten verglichen.

- *Unterarten: Infografiken* beispielsweise kombinieren räumliche und nicht-räumliche Relationen oftmals direkt nebeneinander, wenn sie z.B. vor dem Hintergrund einer Karte Schuldenberge oder Arbeitslosenziffern der einzelnen Länder zeigen.

  Bei *Comics* liegt hingegen eine klare Trennung der Rauminterpretation innerhalb einer Zeichnung und zwischen der Abfolge der Zeichnungen vor. Innerhalb einer einzelnen Zeichnung ist die Lesart gegenständlich und auf räumliche Kohärenz angelegt, während die lineare Abfolge der Zeichnungen zeitlich zu interpretieren ist.

  Eine weitere Mischform stellen *ISOTYPE* dar, die räumlich-tabellarische Arrangements von quantitativen Größen mit schematischen Piktogrammen kombinieren, die die Bezugnahme der einzelnen Elemente auf konkrete Gegenstände symbolisieren soll. Diese Darstellungsform wurde zwar nicht von Otto Neurath (1991) erfunden, es ist aber sein Verdienst, die Gestaltungsprinzipien dieser Visualisierungsform systematisiert und expliziert zu haben.

  Neben Infografiken, Comics und ISOTYPEN gehören auch metaphorische Bilder und Bildarrangements wie Aneinanderreihungen und Gegenüberstellungen mehrerer Bilder und Fotos zu den Darstellungsmitteln, die unterschiedliche Rauminterpretationen verwenden.

- *Einsatzgebiete:* Während Infografiken und ISOTYPE vor allem im Journalismus und in der Didaktik von professionellen Grafikern erstellt werden, stehen Gegenüberstellungen und Aneinanderreihungen allen Bildproduzenten als zusätzliche Gestaltungsmittel zur Verfügung. Oftmals liegt der didaktische Nutzen von Bildern nicht so sehr in den einzelnen Bildern selbst, sondern in dem Vergleich, zu dem Gegenüberstellungen implizit aufrufen.

- *Kognitive Funktionen:* Während ISOTYPE und Gegenüberstellungen vor allem Vergleiche evozieren, zielen Infografiken oft auf die Erklärung komplexer Zusammenhänge, bei denen es insbesondere auf Ursache-Wirkungs-Beziehungen ankommt.

3. Kulturelle und kognitive Voraussetzung des Verstehens von Bildern

Die Verfügbarkeit einer Schriftsprache und noch mehr deren Verbreitung durch den Buchdruck hat dramatische Auswirkungen für die einzelnen Mitglieder einer Kultur und auch insbesondere darauf, wie sie intellektuelle Probleme formulieren, angehen und lösen (Lindberg, 1992; Donald, 1993). Ohne Schrift, Zahlen, Grafiken und Instrumente wäre die moderne Wissenschaft und Technologie schlichtweg unmöglich. Die kulturell vermittelten Symbolsysteme und Produktionstechniken sind eng miteinander verzahnt und in vielfältiger Weise voneinander abhängig. So wie Spielfilme auf einem Drehbuch beruhen, setzen viele Bild- und Grafiktypen die Kenntnis und Beherrschung anderer nicht-bildlicher Symbolsysteme sowie ein Grundverständnis der Produktionstechniken voraus.

Bilder sind der Ursprung der Schrift und über die Kulturen hinweg weiter verbreitet als die Schrift. Es gibt nur wenige bildlose Kulturen. Aber selbst bei bildlosen Kulturen hat sich gezeigt, dass ihre Mitglieder ohne Training in der Lage sind, Bilder zu rezipieren und zu produzieren. Die neuere Forschung scheint zu bestätigen, dass Bilder bei weitem nicht so konventionell und arbiträr sind wie Sprachen.

In seinem Überblicksartikel berichtet Deregowski (1989) von Mitgliedern bilderloser Kulturen (die Tallensi in Ghana), die auf Aufforderung Strichzeichnungen erstellten, die zwar schematisch erscheinen, die dargestellten Gegenstände jedoch recht gut erkennen lassen.

Obwohl ein Bildverständnis im vollen Sinne wohl nicht von vorneherein gegeben ist (wie Untersuchungen von Hochberg & Brooks, 1962, bzw. DeLoache & Burns, 1994, belegen), unterscheiden sich gegenständliche Bilder hinsichtlich des Erlernens grundlegend von der Schrift. Das naive Bildverständnis entwickelt sich allem Anschein nach spontan und ohne spezifische Ausbildung. Darüber hinaus gibt es natürlich noch zahlreiche andere Deutungsebenen von Bildern, die weit über die bloße Erkennbarkeit und Identifizierbarkeit von Gegenständen und das Verständnis des Bild-Gegenstand-Bezuges hinausgehen.

Dass trotz eines universellen Bildverständnisses nicht in allen Kulturen Bilder vorhanden sind, liegt vermutlich daran, dass Bilder im Gegensatz zur gesprochenen Sprache ein geeignetes Trägermedium und Farben sowie Gravur bzw. Maltechniken voraussetzen. Papier und Bleistift sind für uns Alltagsgegenstände, über die wir nicht mehr groß nachdenken, historisch gesehen sind leicht handhabbare, universell verfügbare Medien jedoch keineswegs selbstverständlich.

Die Bindung an technische Trägermedien und Erzeugungsmittel gilt noch mehr für logische Diagramme und Karten, sofern sie genau und zuverlässig sein sollen. Im Gegensatz zu darstellenden Bildern sind logische Bilder jedoch viel voraussetzungsreicher: Sie setzen viele Kulturtechniken und lange individuelle Lerngeschichten voraus. Viele Graphen verlangen z.B. ein Grundverständnis der analytischen Geometrie. Johann Heinrich Lambert war der erste Physiker, der Graphen zur Datenanalyse einsetzte (Tilling, 1975). Diese Darstellungsform ist heute ein so selbstverständlicher Bestandteil der Schulbücher und Massenmedien, dass es schwer fällt, sich in die Zeit zurück zu versetzen, in der Forscher ohne sie auskommen mussten. Vor und auch noch lange nach Lambert publizierten Naturwissenschaftler ihre Daten fast ausschließlich in Tabellen und nicht in Grafiken. Auf ganzer

Linie setzten sich Graphen erst im 19. Jahrhundert durch.

Dieses Beispiel zeigt, wie stark der Einsatz und die Verwendungsweise von logischen Bildern historischen Entwicklungen unterworfen ist und es eben vom Einfallsreichtum einzelner Personen abhängt, ob Bildern und Grafiken neue Einsatzgebiete erschlossen werden können. Aus medienpsychologischer Perspektive ist daran aber vor allem interessant, dass in dieser historischen Entwicklung bildliches Denken und Vorstellen mit sprachlich-symbolischen Codes untrennbar verwoben ist. Ohne Lesefertigkeiten und mathematische Kenntnisse kann man Graphen schlichtweg nicht verstehen. Sie werden deshalb auch erst nach längerem Training zu einer Selbstverständlichkeit. Wie Studien zur Visual Literacy zeigen, gelingt der Schritt zur routinierten Beherrschung nur einer Minderheit (OECD and Statistics Canada, 1995)

Generell lässt sich also festhalten, dass es eine stark variierende Rezeptionskompetenz gibt und schon allein deshalb nicht jede Abbildung einen Sachverhalt anschaulicher im Sinne von verständlicher macht. Die Produktionskompetenz ist vermutlich noch deutlich schwächer ausgebildet, da im Gegensatz zur Schrift, die in vielen alltäglichen und beruflichen Kontexten benötigt wird, Bilder (mit Ausnahme von Fotos) regelmäßig nur von einigen Spezialisten produziert werden.

4. Psychologische Theorien des Bildverstehens

Psychologische Theorien des Bildverstehens versuchen zu erklären, wie Bilder verarbeitet, verstanden und letztlich mental repräsentiert werden. Eine traditionelle, aber naive Vermutung ist, dass Bildverstehen darin besteht, dass externe Bilder in interne bildliche Vorstellungen überführt werden und als Bilder im Gedächtnis gespeichert werden. Es ist wohl unbestritten, dass Bilder memotechnische Funktionen erfüllen können. Trotzdem ist diese Auffassung in mehrfacher Hinsicht naiv:

- Auswendiglernen und exakte Wiedergabe hat nur wenig mit echtem Verstehen zu tun, wie allein die pathologischen Fälle von Eidetikern belegen, die sich zwar an alle Details erinnern können, aber das Wichtige nicht vom Unwichtigen unterscheiden können. Bildliche Wahrnehmungen und Vorstellungen gehören sicherlich zu Bestandteilen des Bildverstehens, echtes Verstehen verlangt aber mehr als bloße exakte Wiedergabe. Für ein echtes Verstehen von Bildern gelten dieselben Kriterien wie für das Verstehen von Texten, z.B. die Fähigkeit Bezüge zu anderem Wissen herstellen zu können, das erworbene Wissen auf andere Problemstellungen anwenden zu können und anderen Menschen den Gehalt des Dargestellten mit eigenen Worten erklären zu können.
- Wir haben schon darauf hingewiesen, dass die meisten Bilder ohne die kontextuelle – und das heißt zumeist sprachliche Einbettung – in ihrem eigentlichen Bedeutungsgehalt gar nicht verstanden werden können. Eine Theorie des Bildverstehens müsste also jeweils in eine Theorie des Textverstehens eingebettet sein, die auf jeden Fall die rhetorische Funktion der Bilder mit berücksichtigt. Dazu liegen bislang nur wenige Ansätze vor (Bonsiepe, 1968; Oestermeier & Hesse, 2000; Schnotz, 2001).

- Zum dritten ist fraglich, ob die aktuelle Psychologie überhaupt über Methoden verfügt, empirisch zu entscheiden, ob die naiv vermutete Korrespondenz zwischen externen und internen Bildern richtig oder falsch ist.
- In der sogenannten Imagery-Debatte wurde auch darüber gestritten, ob überhaupt sinnvoll von Bildern im Kopf geredet werden kann. Es ist immer wieder darauf hingewiesen worden, dass man wohl kaum davon ausgehen könne, dass aus der metaphorischen Redeweise von Bildern im Geiste oder Bildern im Kopf, die wir im Alltag benutzen, auch wirklich folgt, dass es bildhafte mentale Repräsentationen gibt.

Imagery-Debatte I: Bilder als Datenstrukturen. Ob Menschen bildhaft denken ist eine alte Streitfrage der Psychologie und sie ist es bisher geblieben. Die vielleicht prominentesten Gegenpole in dieser Debatte waren die beiden amerikanischen Psychologen Zenon Pylyshyn und Stephen M. Kosslyn. Unstrittig auch in der neueren Imagery-Debatte war, dass Menschen ihre Vorstellungen als bildhaft erleben. Strittig war vielmehr, ob es mentale Repräsentationen geben könne, die in einem klar definierbaren Sinne bildhaft sind und wirklich für diejenigen Problemlösungen benötigt werden, die subjektiv als bildliches Denken erlebt werden. Diese Diskussion wurde vor allem vor dem Hintergrund geführt dass mit dem Computer eine leistungsfähige Metapher des Denkens zur Verfügung stand, mit der kognitive Vorgänge als Symbolverarbeitung gedeutet wurden.

Eine wichtige Messgröße der vor diesem Hintergrund diskutierten Experimente waren Reaktionszeiten, in die man seit den Rotationsexperimenten von Shepard und Metzler (1971) große Hoffnung gesetzt hatte, weil sich ein beinahe linearer Zusammenhang zwischen Drehwinkel und Dauer einer vorgestellten Rotation nachweisen ließ. Ähnliche Befunde konnte Stephen Kossyln vorweisen, der umso längere Antwortzeiten maß, je größer die Entfernungen auf einer vorgestellten Karte waren. Kosslyn (1980) vertrat aufgrund solcher Befunde die Auffassung, dass Menschen über einen visuellen Puffer bzw. Kurzzeitspeicher verfügen, der mit Vorstellungsbildern gefüllt wird, die von einer Art innerem Auge abgesucht werden.

Wie die Imagery-Debatte letztlich zeigte, sind solche Verhaltensdaten prinzipiell nicht geeignet, zu entscheiden, ob mentale Repräsentationen bildhaft oder symbolisch sind. Denn es lassen sich für bildhafte mentale Repräsentation äquivalente propositionale Repräsentationen postulieren, die zu denselben Antwortmustern und Verarbeitungsgeschwindigkeiten führen (Anderson, 1978).

Imagery-Debatte II: Bilder im Gehirn. Viele Psychologen hatten die Imagery-Debatte deshalb auch schon als empirisch unentscheidbar ad acta gelegt. Der Boom der Neuropsychologie und die starke Verbreitung bildgebender Verfahren in den letzten beiden Jahrzehnten hat die Thematik jedoch wieder neu belebt. Ganz platt gesagt wurde und wird versucht, Bilder im Gehirn zu fotografieren. Diese liefern Auskunft darüber, welche Hirnareale bei welchen Vorstellungstätigkeiten aktiviert werden und lassen sich so mit den zahlreichen älteren Befunden über die Funktion bestimmter Areale in Beziehung setzen.

Tatsächlich glaubt Kosslyn (1994), damit die Imagery-Debatte endgültig für sich entschieden zu haben. Nach seiner Theorie besitzen wir einen visuellen Puffer, der erstens ikonisch d.h. bildhaft organisiert ist und zweitens sowohl vom Wahrnehmungs- als auch vom Vor-

stellungsvermögen benutzt wird. Beide Punkte sind nicht wirklich neu. Schon im 19. Jahrhundert war bekannt, dass mit Gehirnverletzungen im visuellen Cortex bildliche Wahrnehmungen und Vorstellungen oft gleichermaßen beeinträchtigt werden (Oestermeier, 1998). Bis heute sind mindestens 15 sogenannter retinotoper (d.h. analog zur Netzhaut strukturierten) Regionen mittels Messung des Stoffwechselumsatzes im Cortex identifiziert worden, wobei aber im Einzelnen umstritten ist, welche Regionen welche Funktionen haben. Es gibt also tatsächlich ikonische Repräsentationen im Gehirn und damit in einem nicht-metaphorischen Sinne Bilder im Kopf.

Stephen Kossyln hat diese retinotopen Regionen, die sowohl an Wahrnehmungs- als auch an Vorstellungsaufgaben beteiligt sind, mit seinem visuellen Puffer identifiziert. Trotzdem bleiben Zweifel angebracht. Es mag zwar sein, dass bei visuellen Vorstellungsaufgaben in denselben Regionen viel Energie benötigt wird, wie bei visuellen Wahrnehmungen. Der Energieumsatz allein sagt aber nicht, wie wichtig diese Areale für die Gesamtfunktion sind. In einem Ballsaal setzten die Tänzerinnen und Tänzer auch mehr Energie als die Musiker um, und trotzdem bestimmt die Musik das Geschehen. In gleicher Weise könnten Regionen, die für sprachliches oder begriffliches Denken verantwortlich sind, mit geringem Energieaufwand andere Regionen steuern und dominieren.

Ohne detaillierteres Wissen über die Arbeitsteilung der verschiedenen Hirnregionen wird man diese Möglichkeit nicht ausschließen können.

Theorien des Text-Bild-Verstehens. Hinzu kommt, dass in realen Lernszenarien nicht isolierte Bilder und Grafiken sondern Text-Bild-Kombinationen untersucht werden müssen. Deshalb haben bildgebende Verfahren nur einen sehr begrenzten Aussagewert für die eigentliche Problematik des Lernens mit Bildern. Die Vorstellungsaufgaben, die mit den neuen bildgebenden Verfahren untersucht wurden, sind sehr elementar und bestehen zumeist darin, sich bestimmte konkrete Gegenstände vorzustellen. Da Kernspintomographien in einem engen Tunnel stattfinden, lassen sich Problemlöseaufgaben, bei denen man neben mentalen Vorstellungen auch externe Unterstützungsmedien wie z.B. Bücher und Notizzettel benötigt, schlecht durchführen.

Da der aktuelle Wissensstand aufgrund der genannten methodischen Probleme begrenzt ist, ist es schwierig, Prozesse, die beim Lernen mit Text-Bild Kombinationen stattfinden, auf kortikaler Ebene zu untersuchen. Die Überprüfung theoretischer Annahmen des Lernens mit Text-Bild Kombinationen erfolgt daher meist sehr indirekt, nämlich über die ermittelten Leistungen in nachfolgenden Wissenstests (vgl. Mayer, 2009). Jedoch werden in der neueren Forschung zum Lernen mit Text-Bild Kombinationen immer häufiger prozessorientierte Verfahren wie Lautes-Denken oder Blickbewegungsanalysen eingesetzt, die in Kombinationen mit den gemessenen Testleistungen neue Aufschlüsse über die zugrundeliegenden kognitiven Prozesse liefern (Van Gog & Scheiter, 2010). Hierzu befindet sich ein Überblick über die aktuellen Theorien und Befunde des Lernens mit Text-Bild Kombinationen in Kapitel 7.

5. Potenziale des Bildgebrauchs

Wie genau Texte und Bilder im Gehirn verarbeitet und integriert werden, wird vermutlich noch länger unklar bleiben. Alle gängigen „Kästchenmodelle“, die die internen Verarbeitungs- und Verstehensprozesse in handhabbare Untermodule aufbrechen sollen, gehen in der Regel von der Idealisierung aus, dass die zentralen Prozesse unabhängig von perzeptuellen und motorischen Prozessen behandelt werden können und deshalb als Input- und Outputmodule keiner weiteren Spezifikation bedürfen. Die Wahrnehmungsorgane wie Augen und Ohr liefern die Rohinformationen, dann findet Kognition, d.h. Denken und Beurteilen statt, und der Körper schließlich führt das Gedachte aus.

Diese Idealisierung, wonach Kognition allein im Kopf stattfindet, geht jedoch von einem zu simplen Schema der Interaktion eines Organismus mit seiner Umwelt aus und verkennt geradezu den Witz des Gebrauchs von Bildern und Grafiken, nämlich die Tatsache, dass der senso-motorische Apparat durch den Gebrauch von Medien in einem nicht-metaphorischen Sinne kognitive Funktionen übernimmt. Wenn man die Medien von vorneherein mit in die Bearbeitungsschleifen mit einbezieht, wird deutlich, dass die externen Medien mehr als bloße Informationsträger sind. Sie werden zu speziellen Werkzeugen, die auf das Innere von Menschen gerichtet sind (Wygotsky, 1934, S. 55). Anspruchsvolle Beispiele sind der Gebrauch von Formeln und Graphen in den mathematisierten Wissenschaften sowie die stetig wachsenden Anwendungsfelder von Computersimulationen.

Dabei fungieren die externen Repräsentationen zum einen natürlich als permanent verfügbare Speicher. Da ihr Einfluss auf die Gedächtnisleistung so einschneidend und ihre Verfügbarkeit so weitgehend ist, haben einige Psychologen Zweifel geäußert, ob es in Gedächtnisexperimenten überhaupt sinnvoll ist, die Versuchspersonen nur kurz mit einem Stimulus zu konfrontieren um dann die Speicherdauer im Gedächtnis zu messen (Norman, 1993, S. 77). Denn sobald man Notizen machen kann, spielt eine limitierte Speicherdauer des Kurzzeitgedächtnisses in der Praxis überhaupt keine nennenswerte Rolle mehr: Der Stimulus bleibt schlichtweg bestehen und das Gedächtnisproblem stellt sich erst gar nicht.

Die Externalisierung von Gedanken und Erinnerungen ist aber mehr als nur eine Auslagerung von Information, die sonst das Kurzzeitgedächtnis überlasten würde. Die Externalisierung von Gedanken strukturiert und ordnet Elemente, so dass man Zusammenhänge "sehen" kann, auf die man durch Denken und Vorstellen allein nicht kommt.

Seit einigen Jahren wird deshalb versucht, genauer zu verstehen, wie Menschen innerhalb ihrer Umwelt Zeichen einsetzen, um ihren kognitiven Apparat zu steuern (Norman, 1993; Hutchins, 1995). Für die speziell zu diesem Zweck hergestellten physikalischen Gegenstände wurde der Ausdruck „kognitive Artefakte“ geprägt. Mit diesen Werkzeugen versetzt sich der Mensch in die Lage, mittels basaler Kulturtechniken wie Lesen, Rechnen und Schreiben komplexe Schlussfolgerungen und Gedächtnisleistungen zu ziehen, die ohne den kulturellen Hintergrund und die innerhalb einer Kultur verfügbaren externen kognitiven Artefakte nicht möglich wären.

Kognitive Funktionen. Welche kognitiven Funktionen übernehmen nun diese Artefakte? Eine Klassifikation kognitiver Grundfunktionen kann sich an den üblichen Einteilungen von

Standardwerken zur Kognitionspsychologie orientieren, die typischerweise entlang solcher Stichworte wie "Wahrnehmung", "Aufmerksamkeit", "Gedächtnis", "Kategorisierung", "Kommunizieren", "Problemlösen" und "Denken" organisiert sind. Die folgende Tabelle zeigt diese Stufung von oben (= basale Fähigkeiten) nach unten (= höhere Funktionen):

Grundvermögen	**Funktion der Artefakte**	**Beispiele für visuelle Medien**
Wahrnehmung	• erweitern • quantifizieren und präzisieren	• Röntgenaufnahmen • Messinstrumente wie Tachometer, Druckmesser, Thermometer samt Abbildung der Messergebnisse
Aufmerksamkeit	• sequenzieren und steuern	• Bildfolgen, Hervorhebungen in Abbildungen, sprachliche Hinführung
Gedächtnis	• externalisieren • kollektivieren	• Skizzenblock • Fotos historischer Ereignisse
Kategorisierung	• definieren • hierarchisieren • klassifizieren • identifizieren • idealisieren, schematisieren	• Mengendiagramme • Baum des Porphyrius, Stammbäume • Zeichnungen in Bestimmungsbüchern • Passfotos • Stilisierte Zeichnungen, Karikaturen
Kommunikation	• sequenzieren und steuern • dokumentieren • verbreiten • Sender markieren	• Ablaufdiagramme • Beweisfotos • Zeitung, Telefon, Computerkonferenzen, Radio, TV • Firmenlogos
Problemlösen und Denken	• probieren und simulieren • deduktiv beweisen • induktiv denken • orientieren und planen o im Raum o in der Zeit o im sozialen Raum	• Simulationsprogramme • geometrische Planskizzen, diagrammatische Logiken • Visualisierungen Statistikpakete o Karten o Zeitachsen, Ablaufdiagramme o Organigramme

Tabelle 3: Kognitive Funktionen und ihre kognitiven Artefakte.

Selbstverständlich können Artefakte multifunktional und multimedial sein, wie die steigende Zahl der Funktionen und Medien zeigt, die vom Computer übernommen werden und ihn zu einem universellen kognitiven Artefakt machen. Doch auch traditionelle Artefakte sind nach dem obigen Schema multi-funktional. Eine Straßenkarte kann gleichzeitig der individuellen räumlichen Planung einer Route und der Kommunikation räumlicher Informationen dienen.

Die obige Beispielliste soll deutlich machen, dass moderne arbeitsteilige Gesellschaften so hohe Anforderungen an den kognitiven Apparat und dessen Genauigkeit, Differenzierung, Koordination und Reliabilität stellen, dass diese Anforderungen nur mit Hilfe externer Zeichen erbracht werden können. Diese Beispiele zeigen auch, wie radikal der Zeichenge-

brauch die kognitiven Anforderungen umstrukturiert. Es ist einfach Karten, Wegweiser und Organigramme zu lesen (wenn man es einmal gelernt hat). Es ist alles andere als trivial, sich in einer unbekannten Stadt oder einem undurchschaubaren Großkonzern durchzufragen und sich anhand lückenhafter sprachlicher Berichte und Gerüchte zu orientieren.

Ein Großteil der kognitiven Artefakte funktioniert dabei nach dem *Prinzip der kognitiven Umverteilung von oben nach unten*: Anspruchsvolle Inferenz- und Problemlöseschritte (die sogenannten höheren kognitiven Funktionen) im unteren Teil der Tabelle werden auf einfache Wahrnehmungsprobleme (die sogenannten niedrigeren kognitiven Funktionen) im oberen Teil der Tabelle reduziert, wobei allerdings eine erhebliche Lernzeit für das Training spezialisierter Wahrnehmungsfähigkeiten, wie z.B. das Lesen spezieller Instrumente, aufgebracht werden muss (Hutchins,1995).

Wie die Beispiele ferner zeigen, erfolgt menschliche Informationsverarbeitung nicht isoliert als ausschließlich interner mentaler Prozess, sondern *situiert* durch eine flexible Nutzung der in der aktuellen Situation verfügbaren informationellen Ressourcen (Lave, 1988). Oder einfacher ausgedrückt: Beim Denken benutzt man nicht nur seinen Kopf sondern auch die Gegenstände um einen herum. Der Rückgriff auf externe Informationsverarbeitungsangebote ermöglicht eine hohe kognitive Ökonomie, denn die Zweckmäßigkeit der Artefakte unterliegt einem kulturellen Ausleseprozess.

Vergleichsstudien. So gesehen sind die verschiedenen Grafik- und Bildtypen hochgradig spezialisierte Werkzeuge, die ganz unterschiedlichen Zwecken dienen und diese speziellen Zwecke oft nachweisbar schneller und effektiver erfüllen als das Universalinstrument Sprache, das zwar enorm flexibel einsetzbar ist, aber eben die Effizienz und das hohe Auflösungsvermögen des visuellen Apparates nicht ausnutzen kann. Aufgrund dieser Spezialisierungen und Differenzierungen ist es auch sinnlos, das Lernen mit Bildern generell über ein Lernen mit Texten zu stellen. Bei einer solchen Fülle von kognitiven Funktionen, die ja beim Lernen im umfassenden Sinne alle eine Rolle spielen, lassen sich immer Fälle finden, bei denen der Bildgebrauch die Leistung verbessert, es lassen sich aber auch immer Fälle finden, in denen Bilder nicht ihren speziellen Stärken gemäß eingesetzt werden und sich deshalb keine Leistungsverbesserungen finden lassen.

Es ist stets genauer zu fragen, welche der obigen Funktionen durch sprachliche Mittel beinahe genauso gut erledigt werden können, und bei welchen Funktionen Texte einfach schlechter abschneiden müssen. In den letzten hundert Jahren sind viele solcher differenzierten funktionalen Analysen empirisch und theoretisch durchgeführt worden, es fehlen aber bislang theoretische Ansätze, die diese Befunde zu einem verständlichen Gesamtbild zusammen führen.

So verglich Washburne (1927a, 1927b) die Darstellung quantitativer Werte in verschiedenen Graphen, Tabellen und Texten. Er kam zu dem Ergebnis, dass *Tabellen* am besten geeignet sind, wenn es auf eine *hohe Ablesegenauigkeit* ankommt, *Säulendiagrame* bei *einfachen Vergleichen von Werten* und *Liniendiagramme* beim *Erkennen dynamischer Verläufe*. Texte schnitten bei diesen relativ elementaren Aufgaben am schlechtesten ab. Larkin und Simon (1987) haben u.a. Diagramme von Flaschenzügen untersucht, die Probanden spontan aufgrund von schriftlichen Beschreibungen angefertigt haben. Diese Diagramme explizierten alle wichtigen beschriebenen geometrischen und topologischen Beziehungen,

wobei neue Relationen schon allein deshalb leichter im visuellen Modus gezogen werden konnten, weil sich alle relevanten Informationen zu einem Objekt auch räumlich nebeneinander befanden und nicht erst aus einer Liste von Sätzen extrahiert und in Bezug zueinander gesetzt werden mussten. Die *visuellen Schlussfolgerungen* wurden also ohne große Aufmerksamkeitssteuerung und Suche quasi in direkt ablesbarer Form mitgeliefert.

Andere spezifische Vorteile von Abbildungen sind in der enorm hohen Auflösung zu sehen, die es erlaubt, gewaltige Datenmengen in einem Bild zusammen zu fassen. Dieser Vorteil besteht vor allem bei Papier. Während gute Bücher mit 1200 dpi und mehr gedruckt werden, kommen viele Computerbildschirme nach wie vor nur auf 72 dpi. Da das unbewaffnete Auge bis zu O.1 mm und mehr auflösen kann (Tufte, 1983) nutzt diese Pixelgröße des Bildschirms das visuelle Auflösungsvermögen bei weitem nicht aus. Karten setzten z.B. oft schon eine Auflösung von 250 dpi und mehr voraus. Jede gute gedruckte Verkehrskarte, die neben eine Karte eines Routenplaners auf dem Computerbildschirm gehalten wird, belegt diese These.

Diese Überblicksfunktion kann sprachlich nur unzureichend durch Zusammenfassungen simuliert werden, da jede Hervorhebung eines Aspektes die Vernachlässigung anderer Aspekte bedeutet. Jeder Rezipient, der dieses Interesse an spezifischen Aspekten nicht teilt, ist deshalb mit einer visuellen Gesamtdarstellung besser bedient. Obwohl diese Funktion der Informationsballung auf engem Raum bei gleichzeitiger einfacher Lesbarkeit natürlich nicht unbekannt ist, ist sie in der Praxis jedoch nur sehr schwer für einen einzelnen Autor zu realisieren.

Auch rhetorische Funktionen können von Bildern effizient übernommen werden. Ob Beispiele, Aufhänger, Metaphern, Analogien und Argumente besser bildhaft als sprachlich vermittelt werden, ist oft schwer zu ermitteln, da es dabei oft um verschiedene Verstehensebenen geht und Bilder eben – wie mehrfach angeführt – diese Funktion auch nur innerhalb der rhetorischen Gesamtstruktur der Text-Bild-Kombination übernehmen können.

6. Probleme des Bildgebrauchs

Bei Medien- und Darstellungsvergleichen kommt es immer wieder zu Ergebnissen, deren Relevanz für die hier angesprochene Thematik – vorsichtig ausgedrückt – eher gering ist. Zum Teil resultiert dies aus isolierten Vergleichen, bei denen die tatsächlichen Einsatzbedingungen von Grafiken und Bildern völlig außen vor gelassen werden.

Solch ein künstliches Problem, dessen Relevanz sich in der Praxis leicht in Luft auflöst, ist z.B. in der Debatte um die Verwendung von Torten- und Säulendiagrammen zu sehen (Macdonald-Ross, 1977). Es gab Autoren, die gegen Tortendiagramme argumentierten, und darauf hinwiesen, dass Längenvergleiche dem Auge leichter fallen als Flächenvergleiche. Experimentelle Tests zeigten, dass Probanden jedoch Tortendiagramme mindestens genauso korrekt in Prozentwerte umsetzen können, wie Balken- oder Säulendiagramme. Worauf andere wiederum argumentierten, dass es nicht auf das Ablesen einzelner Werte ankomme, sondern auf Vergleiche, die entlang einer linearen Skala einfacher fallen als bei unterschiedlichen Tortenstückgrößen. Diese Ergebnisse sind m.E. deshalb irrelevant, weil

die jeweiligen Defizite der Darstellungen durch sprachliche Erläuterungen ausgeglichen werden können.

Trotz der methodischen Schwierigkeiten kann man echte Probleme des Text-Bild-Verstehens identifizieren, die aber weniger aus dem Leistungsvermögen von Darstellungsformaten resultieren, als aus den unterschiedlichen Einstellung der Rezipienten gegenüber den verschiedenen Formaten.

Illusion des Verstehens. Generell scheint zu gelten, dass Bilder nur kurz angeschaut werden und in aller Regel nur eine geringe Verarbeitungstiefe nach sich ziehen. Bilder gelten als leicht und schnell zu erfassen (innerhalb von 300 ms hat man das Grundmotiv erfasst) und letztlich nicht der Mühe wert (Peeck, 1993). Es gibt außerdem oft kein klares Kriterium, anhand dessen der Rezipient entscheiden könnte, ob er das Bild oder die Grafik wirklich verstanden hat. Während man bei einem Text oftmals an schwierigen Stellen hakt, weil die Bedeutung einzelner Wörter unklar ist oder eine Passage inkohärent wirkt, stellt sich ein vergleichbares Gefühl bei Grafiken viel seltener ein. Die visuellen Elemente sind in der Regel bekannt und die Figuren als solches meist klar zu erkennen. Bilder und Grafiken vermitteln also leicht das trügerische Gefühl, etwas verstanden zu haben („illusion of full understanding", vgl. Peeck, 1993).

Kognitive Belastung. Diese Illusion des Verstehens verhindert kognitive Anstrengungen, obwohl gerade bei komplexen Bild-Text-Kombinationen ein erheblicher Codierungs- und Integrationsaufwand erforderlich ist. Schlecht gestalte Text-Bild-Kombinationen ziehen dabei Verarbeitungskapazität ab, indem sie den Leser zwingen, zwischen Text und Bild hin und her zu springen. An diesem Punkt setzt die *cognitive load theory* (Sweller, 1994; Kirschner, 2002) ein, die im Grunde auf eine Trivialität hinausläuft. Lehr-Lernmaterial sollte möglichst so gestaltet werden, dass unnötiger Aufwand („extrinsic cognitive load") vermieden wird um so Kapazität für die eigentliche Lernaufgabe zu schaffen („germane [passend, relevant] cognitive load"). Erreicht wird dies durch eine enge *räumliche* Verzahnung von Grafiken und erklärenden Texten oder einer engen *zeitlichen* Verzahnung von Grafiken und auditiven Erläuterungen. Obwohl dieser Hinweis banal ist, ist er in der Praxis nicht ohne Auswirkungen, da aus drucktechnischen Gründen oder schlicht mangelnder Sorgfalt, Bild und Text oft schlecht aufeinander abgestimmt sind.

Verblöden durch Bilder? Eine andere Problematik, die sich aus einer weit verbreiteten Vorliebe für oberflächlich verarbeitete Bilder ergeben könnte, ist die Zurückdrängung der sprachlich argumentativen Auseinandersetzung mit dem Lerngegenstand. Kritiker des Bildzeitalters wie Neil Postman (1985) oder Uwe Pörksen (1997) sehen vor allem in der beschränkten argumentativen Ausdrucksstärke von Bildern ein Hauptmanko der visuellen Kommunikation. Verkürzt ausgedrückt bestehen die Bedenken darin, dass Bildern die syntaktischen Mittel fehlen um Behauptungen, Begründungen, Negationen und Differenzierungen auszudrücken, die Grundbestandteile jeder ernst zu nehmenden kritischen Aneignung eines Stoffes sind. Die Sprache verfüge dazu über ein reiches argumentatives Vokabular: „weil", „da", „somit", „dagegen", „zwar ... aber ..." u.v.m. Schon allein deshalb sei es unmöglich, sich rein visuell mit komplexen Sachverhalten angemessen auseinander zu setzen, weshalb eine Dominanz der Bilder letztlich zu einem Sinken des intellektuellen Niveaus breiter Bevölkerungsschichten führe.

An dieser Kritik ist zweifellos richtig, dass sich verbale und visuelle Mittel auf der Ebene der Syntax bzw. Zeichenform grundlegend unterscheiden: Bei den meisten Bildformen gibt es weder ein definiertes Vokabular noch Regeln für Wohlgeformtheit noch syntaktische Äquivalente für Tempi und Modi. Doch auch hier greift das Kompensationsargument: Es stimmt einfach nicht, dass Bilder die Hauptlast der Argumentation tragen, in fast allen argumentativen Kontexten wie z. B. Vorträgen, Artikeln, Werbespots etc. sind Sprechakte nach wie vor allgegenwärtig. Es ist richtig, dass die Bildkommunikation oft auf einem niedrigem Niveau geführt wird und viele Schaubilder und Infografiken ihre Inhalte nicht als das kennzeichnen, was sie sind – bloße Spekulationen und Vermutungen. Aber es ist doch sehr fraglich, ob die rein sprachlichen Diskurse nicht genauso Stereotypen, unzuverlässigen Vereinfachungen und Fehlinterpretationen aufsitzen wie Bilder.

Der richtige Aspekt an dieser Kritik des Bildgebrauchs könnte sein, dass der argumentative Gehalt von Bildern selten problematisiert wird, weil es dazu an Gelegenheiten und Vorbildern fehlt. Es gibt kaum Diskussionsrunden im Fernsehen oder in der Schule, die sich kritisch mit Bildern befassen, aber zahllose Runden, bei denen rein sprachlich diskutiert wird. Und wenn es stimmt, dass Bilder tatsächlich besser zu erinnern sind als sprachliche Aussagen, führt dies tatsächlich dazu, dass sie sich leichter im kollektiven Gedächtnis ungeprüft festsetzen können als sprachliche Thesen.

7. Theorien und Befunde zum Lernen mit Text-Bild-Kombinationen

Psychologische Theorien des Lernens mit Texten und Bildern lassen sich grob in zwei Sichtweisen unterschieden: in eine ergebnisorientierte und eine prozessorientierte Sichtweise (vgl. Schmidt-Weigand & Scheiter, 2011). Die ergebnisorientierte Sichtweise entstammt der Gedächtnisforschung. Die prozessorientierte Sichtweise stammt vorzugsweise aus der Forschung zum Textverstehen bzw. zu Verständnisprozessen im Allgemeinen. Aus beiden Sichtweisen lassen sich Vorhersagen für effektives Lernen mit Text-Bild Kombinationen ableiten.

Ergebnisorientierte Sichtweise. Wie der Name schon sagt, liegt der Fokus der ergebnisorientierten Sichtweise auf den *Ergebnissen* des Lernens mit Texten und Bildern. Sie basiert vorwiegend auf den Theorien der Dualen Kodierung (Paivio, 1986) sowie der „conjoint retention“ (Kulhavy, Lee, & Caterino, 1985). Kurz gesagt nehmen beide Theorien an, dass es vorteilhaft für den späteren Abruf von Informationen, also für die Gedächtnisleistung ist, wenn Informationen gemeinsam bzw. miteinander verbunden abgespeichert (gelernt) wurden. Lernt man mit Text und Bild, dann ist es vorteilhaft, wenn die Informationen, die über Text und Bild vermittelt werden, in enger räumlicher, zeitlicher und mit ausreichend thematischer Nähe präsentiert werden. Dann werden sie im Gedächtnis eng miteinander verbunden abgespeichert, obwohl die Informationen aus Text und Bild in zwei verschiedenen Kodierungen im Gedächtnis vorliegen. Möchte man Informationen aus Text und Bild zu einem späteren Zeitpunkt abrufen, dann reicht der erfolgreiche Abruf der Informationen z.B. aus dem Text aus, um auch an alle Informationen über das Bild (und andersherum) zu gelangen. Die gemeinsame Speicherung von Text- und Bildinformation führt zum gemeinsamen Abruf dieser Informationen. Dadurch ist die Erinnerungsleistung verbessert.

Diese Annahmen übernahm Richard E. Mayer (2009) für seine Kognitive Theorie des Multimedialen Lernens (kurz: CTML), die sehr einflussreich im Bereich des Lernens mit Text und Bild ist. Etwas seltsam ist dabei, dass sich die CTML auf das Prinzip der dualen Kodierung bezieht, um damit bessere *Verständnis*leistungen vorherzusagen, obwohl die Theorie der Dualen Kodierung (Paivio, 1986) ursprünglich bessere *Erinnerungs*leistungen vorhersagte. Grob gesagt geht die CTML davon aus, dass Texte und Bilder in getrennten verbalen und piktorialen Kanälen verarbeitet werden und im Arbeitsgedächtnis mithilfe von Vorwissen aus dem Langzeitgedächtnis integriert werden (siehe Abb. 3).

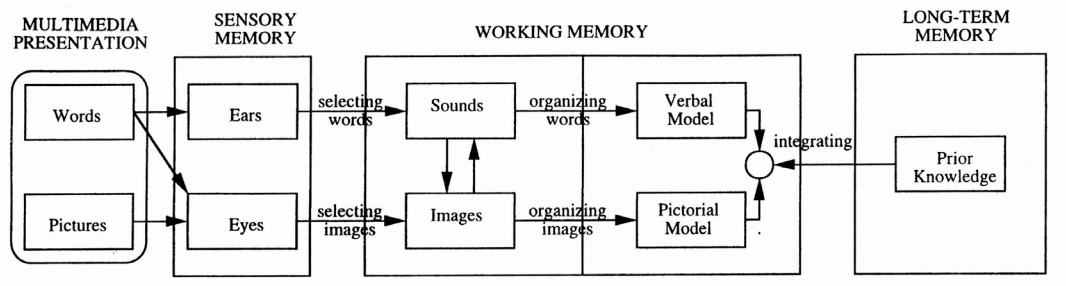


Abb. 3. Die Kognitive Theorie des Multimedialen Lernens (Mayer, 2009, S. 61).

Das Modell ist insofern problematisch, als das die Pfeile mehr oder weniger alle Kästchen irgendwie mit allen Kästchen verbinden und somit eigentlich nicht mehr sagen als "Alles hängt mit allem zusammen“, was bei komplexen psychologischen Problemen nie falsch sein kann. Das Modell verletzt daher ein zentrales Gütekriterium von Theorien im Allgemeinen – nämlich, dass sie falsifizierbar sind. Überdies beinhaltet sie damit zu viele Annahmen über mentale Repräsentationen (siehe viele Kästchen in Abb. 3), die praktisch nicht nachweisbar sind und auch keine schärferen Vorhersagen erlauben als wenn es sie nicht geben würde. Das heißt, sie erklärt Prozesse beim Lernen mit Text-Bild Kombinationen weniger sparsam als nötig (vgl. Eitel, Scheiter, Schüler, Nyström, & Holmqvist, 2013). Andererseits könnte man sogar noch entgegnen, dass zentrale Pfeile fehlen, und zwar Pfeile, die vom Langzeitgedächtnis auf die Wahrnehmungsprozesse zurückgehen. Denn es ist erwiesen, dass die Auswahlprozesse, die Mayer zwischen den sensorischen und dem Arbeitsgedächtnis ansiedelt, hochgradig vorwissensabhängig sind und insofern vom Langzeitgedächtnis gesteuert werden (z.B. Canham & Hegarty, 2010). Daher werden die Annahmen der CTML in der neueren Forschung zum Lernen mit Texten und Bildern zu Recht kritisch betrachtet (z.B. Eitel, Scheiter, & Schüler, 2013; Schmidt-Weigand & Scheiter, 2011; Rummer, Schweppe, Scheiter, & Gerjets, 2008).

Nichtsdestotrotz lässt sich aus der ergebnisorientierten Sichtweise auf das Lernen mit Text und Bild im Allgemeinen ableiten, dass Bilder vor allem dann die Erinnerungsleistung erhöhen, wenn sie einen hohen raumzeitlichen und thematischen Bezug zum entsprechenden Text aufweisen – eine Schlussfolgerung, die sich aus den Theorien zur dualen Kodierung (Paivio, 1991) und zur „conjoint retention“ (Kulhavy et al., 1985) ableiten lässt, und die mehrfach empirisch bestätigt wurde. So zeigten sich in Studien Erinnerungsvorteile vor allem für Informationen, die sowohl im Text standen als auch im entsprechenden Bild illustriert wurden (Levie & Lentz, 1982). Durch den hohen thematischen Bezug dieser Informa-

tionen wurden sie gemeinsam abgespeichert und konnten somit gemeinsam abgerufen werden.

Informationen aus Texten und Bildern werden besser erinnert, wenn sie „gemeinsam" – also mit hoher raumzeitlicher und thematischer Überlappung - abgespeichert wurden (Kulhavy et al., 1985; Paivio, 1986).

Prozessorientierte Sichtweise. Wie der Name hier auch schon sagt, liegt der Fokus der prozessorientierten Sichtweise auf den Prozessen, die dem (erfolgreichen) Lernen mit Texten und Bildern zugrunde liegen. Das Ziel dieser Sichtweise ist es, lernförderliche Prozesse zu identifizieren und zu fördern. Lernerfolg wird vor allem über ein besseres *Verständnis* der über Text und Bild vermittelten Sachverhalte definiert (Schmidt-Weigand & Scheiter, 2011).

Eine Theorie, die der prozessorientierten Sichtweise zuzuordnen ist, ist das integrative Modell des Text- und Bildverstehens von Wolfgang Schnotz (2001; siehe auch Schnotz & Bannert, 2003). Dieses Modell bezieht sich – anders als die zuvor dargestellte Theorie von Mayer (2009) – vorwiegend auf Theorien zum Textverständnis (z.B. van Dijk & Kintsch, 1983), um zu erklären, welchen Einfluss die kombinierte Darstellung von Texten und Bildern auf das Verständnis hat. Dabei nimmt Schnotz (2001) an, dass das Verständnis sowohl von Texten als auch von Bildern in *einem mentalen Modell* reflektiert ist (siehe auch Abb. 4). Denn sowohl beim Lesen eines Textes als auch beim Ansehen von Bildern macht man sich, neben dem tatsächlich vorhandenen Informationsgehalt der Wörter und Abbildungen, eigene Vorstellungen über deren Bedeutung, die dann mental abgelegt werden. Das ist meistens auch nötig, um einen Text überhaupt verstehen oder richtig einordnen zu können. Treffenderweise spricht man im Volksmund daher auch von „man liest zwischen den Zeilen" oder „man macht sich sein eigenes Bild davon". In die Fachsprache übersetzt heißt das, man erstellt sich üblicherweise sein eigenes mentales Modell eines Sachverhalts unter Zuhilfenahme jeglicher verfügbarer Information (vgl. van Dijk & Kintsch, 1983). Das können Text- und Bildinformationen sein, aber auch eigene Vorstellungs- und Schlussfolgerungsprozesse sowie Informationen, die aus dem eigenen Vorwissen abgerufen werden. Das resultierende mentale Modell ist damit eine Art Sammlung von Informationen über einen Sachverhalt (aus Text, aus Bild etc.). Die Informationen sind darin so verknüpft, dass man in der Lage dazu ist, Schlussfolgerungen auf Basis dieser Informationen zu ziehen. Daher wird das Verständnis für einen Sachverhalt häufig mit dem Aufbau eines kohärenten mentalen Modells für diesen gleichgesetzt.

Möchte man nun ein Verständnis für einen Sachverhalt, also ein kohärentes mentales Modell, aufbauen, dann unterscheiden sich Texte und Bilder in ihrer Effektivität. Zum Beispiel besitzen Texte eine höhere Ausdrucksmöglichkeit als Bilder. Texte erlauben es, allgemeine Negationen oder Disjunktionen wie z.B. „Haustiere nicht gestattet" oder „Sitzplatz für Ältere oder für Mütter mit Kleinkindern" effizient und gut verständlich auszudrücken (Beispiel aus Schnotz, 2001). Versuchen Sie das mit einem Bild auszudrücken. Sie würden wohl Schwierigkeiten dabei haben, ein Bild zu finden, welches stellvertretend für alle Haustiere und nur für Haustiere steht. Ebenfalls müsste ein intuitives und allgemein verständliches Bild für die Disjunktion „oder" wohl auch erst noch erfunden werden.

Die eigentliche argumentative Last trägt immer der Begleittext (Oestermeier & Hesse, 2000).

Auf der anderen Seite sind Bilder besser geeignet, um Schlussfolgerungen zu ziehen, da der gesuchte Inhalt und dessen Beziehung zu anderen Inhalten direkt aus der Abbildung abgelesen werden kann (Ainsworth, 2006). Stellen Sie sich beispielsweise vor, Sie lesen die Beschreibung „Die Figur besteht aus zwei symmetrischen diagonalen Linien, die sich oben an der Spitze treffen und ungefähr in der Mitte durch eine horizontale Linie verbunden sind". Werden Sie anschließend gefragt, ob die Figur eine geschlossene Region enthält und, wenn ja, welche Form diese Region hat, dann fällt Ihnen die Antwort eventuell schwer. Zeigt man hingegen das entsprechende Bild, nämlich lediglich das eines großgeschriebenen Buchstabens „A", dann fällt die Antwort denkbar leicht – ein Dreieck (Beispiel aus Schnotz, 2001).

Das Beispiel macht deutlich, dass Bilder förderlich für das Verständnis sein können, weil sie automatisierte Wahrnehmungsprozesse ausnutzen, um Schlussfolgerungen zu ermöglichen (siehe auch die Vertiefung Gestaltgesetze auf e-teaching.org). Das heißt, über direkt wahrgenommene räumliche Relationen in Bildern lassen sich Schlussfolgerungen über Bedeutungszusammenhänge ableiten (Schnotz & Bannert, 2003). Man kann sein mentales Modell, also sein Verständnis für einen Sachverhalt, auf den räumlichen Informationen aus dem Text aufbauen. Räumliche Informationen aus dem Bild dienen als Gerüst für das entsprechende mentale Modell (vgl. Eitel et al., Scheiter, Schüler, Nyström et al., 2013). Schnotz (2001) bezeichnet den Prozess des Aufbaus einer mentalen Modellstruktur aus dem Ablesen der räumlichen Struktur eines Bildes als Strukturabbildungsprozess. Die visuelle Wahrnehmung eines Bildes resultiert somit direkt in der Struktur des mentalen Modells (siehe Abb. 4, rechter Teil).

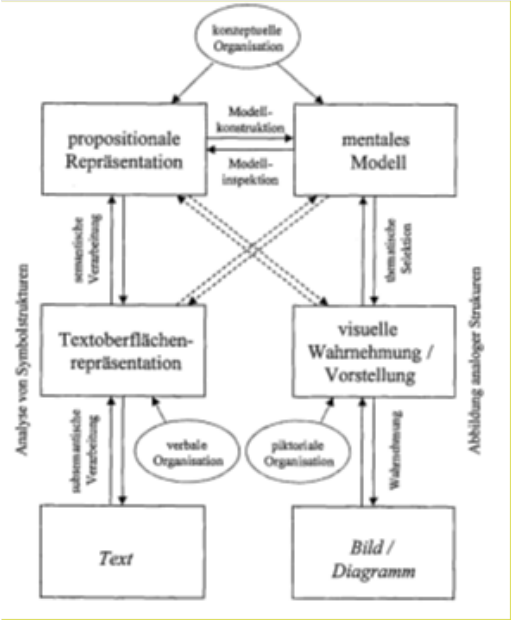


Abb. 4. Das integrative Modell des Text- und Bildverstehens (Schnotz, 2001, S. 306).

Müsste man sich hingegen sein mentales Modell nur auf Basis eines Textes aufbauen, so müsste man laut Schnotz (2001) einen Umweg zum mentalen Modell gehen, nämlich über die Textoberflächenrepräsentation, aus der zunächst eine propositionale Repräsentation des Textes gebildet wird (vgl. „text base"; van Dijk & Kintsch, 1983). Erst durch zusätzliche Vorstellungsprozesse und durch Zuhilfenahme von Vorwissen etc. würde das mentale Modell erstellt werden (siehe Abb. 4, linker Teil). Das kann zum Aufbau eines mentalen Modells führen, dass die Textinhalte inadäquat reflektiert, weil Fehlvorstellungen bzw. Fehlinterpretationen der Textinhalte ins mentale Modell einfließen können. Man läuft Gefahr, den Text falsch zu verstehen. Diese Gefahr ist reduziert, wenn zusätzlich ein Bild präsentiert wird, da es die Struktur für ein adäquates mentales Modell bereits vorgibt. So können Bilder das Verständnis für Texte fördern.

Das ist jedoch nur dann der Fall, wenn (1) das Bild die Textinhalte adäquat reflektiert, wenn (2) das Bild bereits vor oder gleichzeitig mit dem Text verarbeitet werden kann, so dass nicht zunächst ein (inadäquates) mentales Modell auf Basis des Textes aufgebaut wird, welches dann mit dem nachfolgenden, aus dem Bild abgelesenen, mentalen Modell interferiert, und wenn (3) sich Text und Bild auf den gleichen Sachverhalt beziehen, so dass das Bild Fehlinterpretationen des Textes verhindert. Ein einfaches Beispiel: Liest man in einem Text, dass sich „der Kasten mit der Aufschrift *Text* neben dem Kasten mit der Aufschrift *Bild* befindet", so könnte man sich vorstellen, dass sich der Kasten mit der Aufschrift *Text* sowohl links als auch rechts neben dem Kasten mit der Aufschrift *Bild* befindet. Wird das entsprechende Bild dazu präsentiert, dann sind solche Ambiguitäten schnell aufgelöst und mögliche Fehlinterpretationen werden verhindert. Denn Bilder stellen räumliche Relationen gemeinhin ohne Interpretationsspielraum dar (siehe Abb. 4, unterer Teil).

Gleichzeitig würde eine reine Bildpräsentation in den meisten Fällen ebenfalls zu einem mentalen Modell führen, das den Sachverhalt inakkurat bzw. unzureichend reflektiert, denn ähnlich wie Texte benötigen Bilder häufig zusätzliche Erklärungen, um verstanden zu werden; bzw. damit man überhaupt weiß, wie sie „gelesen" werden sollen. Ein statisches Bild kann leicht wahrgenommen werden. Um ein Bild jedoch wirklich zu verstehen, benötigt man häufig einen zusätzlichen Text (vgl. Hegarty & Just, 1993). Dementsprechend wird eine verständnisförderliche Wirkung durch das Hinzufügen von Bildern zu Texten im Rahmen des Schnotz-Modells und im Rahmen der prozessorientierten Sichtweise im Allgemeinen hauptsächlich darauf zurückgeführt, dass sich Texte und Bilder beim effektiven Aufbau eines mentalen Modells ergänzen. Bilder können Prozesse der mentalen Modellkonstruktion aus dem Text ersetzen, zumindest wenn sie sich auf dasselbe Thema beziehen.

Bilder und Texte ergänzen sich im Aufbau von Verständnis. Ein Bild erleichtert die (semantische) Interpretation eines Textes und im Gegenzug erleichtert das Lesen eines Textes die Auswahl (Selektion) von relevanten Informationen in dem entsprechenden Bild (vgl. Eitel & Scheiter, 2014).

8. Richtlinien für Text-Bild-Kombinationen

Aus den gerade genannten Theorien und Befunden des Lernens mit Text-Bild Kombinationen ergeben sich einige Richtlinien, die für die effektive Gestaltung von Text-Bild Kombinationen sinnvoll sind. Für die sprachliche Einbettung von Grafiken und Bildern folgt zunächst eine einfache Faustregel:

Syntaktische Unterschiede hinsichtlich der Farbwahl, Linienstärke, Kontrastbildung, Komplexität des Aufbaus von Bildern und Grafiken etc. mögen Auswirkungen haben, sie sind aber zweitrangig gegenüber der oftmals viel größeren Gefahr, dass ein Lerngegenstand inhaltlich und didaktisch inadäquat dargestellt wird.

Doch was heißt inhaltlich und didaktisch adäquat? Es ist naheliegend, dazu an die linguistischen Standards anzuknüpfen. Die von dem Sprachphilosophen Grice formulierten Konversationsmaximen für rationale sprachliche Diskurse lassen sich leicht auf die visuelle Kommunikation übertragen (Grice, 1975; Oestermeier et al., 2001). Die vier Griceschen Maximen, die ein kooperativ agierender Sprecher einhalten muss, können wie folgt für Visualisierungen umformuliert werden:

- *Qualität*: Ein visueller Beleg für eine Behauptung muss wahr bzw. gut bestätigt sein.
- *Quantität*: Die Visualisierung muss in Bezug auf das Vorwissen des Rezipienten informativ sein.
- *Relevanz*: Die Visualisierung muss relevant für die Hauptthese des Textes sein.
- *Modalität*: Die Visualisierung muss klar und verständlich sein.

Diese Liste zeigt, dass die Verständlichkeit nur ein Aspekt von mehreren ist, obwohl sie in vielen psychologischen Vergleichsstudien im Vordergrund steht. Doch was folgt nun aus diesen Maximen im Einzelnen für die Praxis?

Qualität im Sinne von Beweiskraft und Vertrauenswürdigkeit einer Visualisierung hängt zu einem großen Teil vom Produktionsprozess der Grafiken und Bilder ab; von der verwendeten Technologie, den beteiligten Institutionen und den involvierten Interessen. Gedruckte Landkarten sind vertrauenswürdiger als von Hand gezeichnete Skizzen, Abbildungen in wissenschaftlichen Aufsätzen vertrauenswürdiger als Abbildungen auf einer provisorischen Website. Doch selbst bei seriösen Quellen bleibt Vorsicht geboten: Die Untersuchungskommission der DFG, die die Veröffentlichungen des Krebsforschers Herrmann überprüfte, stellte z.B. fest: *"Die Hinweise auf Datenmanipulation und wissenschaftliches Fehlverhalten ergaben sich vorrangig aus der Analyse der Abbildungen der Publikationen."* (DFG, 2000).

Die daraus folgende Maxime ist einfach:

Lieber auf beeindruckende Visualisierungen verzichten, wenn die Belege für die veranschaulichte These zweifelhaft sind!

Damit wird z.B. die folgende Visualisierung (Abb. 5) hinfällig, die immer wieder im Zusammenhang mit dem Lernen mit Multimedia als zentrales Argument angeführt wird.

Abb. 5. Ein visuelles Argument für den Einsatz von Multimedia. Die Zeit, Nr. 48, 21.11.2002, S. 37.

Diese Grafik existiert in vielen Varianten und wird meistens ohne Quellenangaben übernommen (dies ist in Abb. 5 zwar anders, siehe Fußzeile). Wer aber in dem genannten Buch nachschlägt, wird dort die Zahlen auf Seite 46 finden, aber nur mit dem lapidaren Hinweis „Die Lernforschung hat nachgewiesen …". Weder werden an dieser Stelle überprüfbare experimentelle Daten angeführt noch die im Textabschnitt unterhalb der Grafik genannte "American Audiovisual Society" erwähnt (bei der sich die Zahlen allerdings so auch nicht finden!). Es bleibt also völlig unklar, woher die Zahlen wirklich stammen. Es wäre auch verwunderlich, wenn sich in einem Experiment die Effekte kombinierter Faktoren so schön aus den Bestandteilen aufsummieren ließen. Kritische Anmerkungen dazu finden sich bei Weidenmann (2002).

Nicht immer lassen sich haltlose Behauptungen so leicht entlarven, denn einige Bildfälschungen haben ihren Weg über wissenschaftliche Artikel in moderne Schulbücher gefunden. Das folgende Foto in Abb. 6 findet sich in vielen Biologie-Büchern als Beleg für die Darwinsche Evolutionstheorie. Es zeigt zwei Varianten der englischen Motte (Biston betularia), eine helle und eine dunkle. Angeblich soll sich die dunkle Motte mit der einsetzenden Industrialisierung und dem verstärkten Ruß stark vermehrt haben, während die helle Motte ihre Tarnung verlor und verstärkt von Vögeln gefressen wurde.

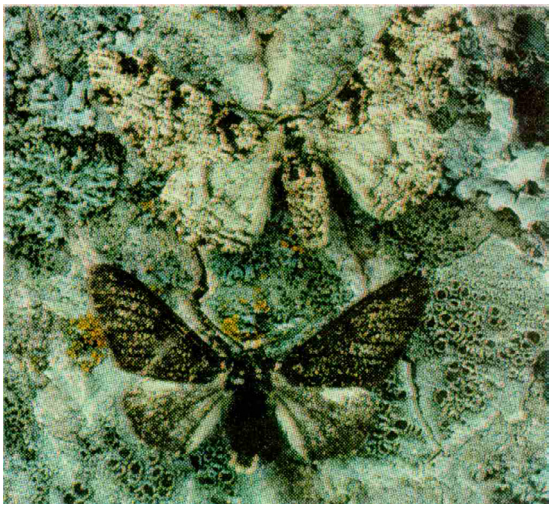


Abb. 6. Der "Beweis" für die Darwinsche Evolutionstheorie. Die Zeit, Nr. 46, 7.11.2002, S. 35

Auf den ersten Blick ist an dem Foto nichts besonders, doch hat sich eine bemerkenswerte Debatte darum entfacht, als bekannt wurde, dass die ursprüngliche Untersuchung zu den Mottenpopulationen vermutlich "frisiert" und das Foto gestellt war.

Eigentlich ist es nicht verwunderlich, dass solche Bilder gestellt werden, denn so direkt vergleichbar wird man diese zwei Motten kaum finden. Vermutlich wurde dieses Foto sogar mit gutem Gewissen eingesetzt, denn es soll lediglich verdeutlichen, wieso die helle Motte verschwand. Der Witz des Fotos besteht nun darin, sich den hellen Hintergrund wegzudenken um so den Evolutionsvorteil der dunklen Motte zu sehen. D.h. das Foto wird eigentlich erst verständlich, wenn man eine Änderung hineinprojiziert. Aber auch das macht das Foto noch nicht irreführend, denn diese zwar nicht direkt gezeigte aber nahegelegte Erklärung könnte ja trotzdem im Prinzip richtig sein. Als wirklich irreführend wurde daran kritisiert, dass die helle Motte sich als nachtaktives Tier tagsüber praktisch nie auf Baumrinden aufhalte, sondern im feinen Geäst der Kronen. Das Foto suggeriere einen Auslesemechanismus, der so nicht stattfinde. Oder anders ausgedrückt: Gäbe es Hinweise auf solch einen Mechanismus, wäre auch das gestellte Foto als Illustration durchaus legitim.

Dieses auf den ersten Blick so einfache Beispiel zeigt, wie vertrackt die Beziehung von visuellen und sprachlichen Argumenten sein kann. Es gehört somit eigentlich zu den Pflichten eines Autors, deutlich zu machen, was die Visualisierung überhaupt zeigen soll: Soll sie eine Behauptung beweisen oder einen anderweitig belegten Sachverhalt erläutern? Dies unterbleibt oft, da anscheinend die Meinung herrscht, dass gut ausgewählte Bilder für sich sprechen. Als Produzent sollte man sich also immer den genauen argumentativen Gehalt klar machen, d.h. letztlich ihn zu verbalisieren versuchen. Ergeben sich dann Zweifel, ob diese sprachliche Übersetzung haltbar ist, sollte man diese Zweifel benennen und es nicht dem Leser überlassen, darauf zu kommen, da es keine Garantie gibt, dass er diese Visualisierung von vorneherein kritisch betrachtet.

Dieser Maxime entsprechend gehört es eigentlich auch zu den Pflichten, bei Graphen über die Qualität der Daten Auskunft zu geben, indem man Scheinpräzisionen als solche kenn-

zeichnet (z.B. durch eingezeichnete Streuungs- und Messfehlerbereiche) und Verzerrungen vermeidet (z.B. Skala vom Nullpunkt aus und nicht von selbstgewählter Basislinie). Auch das Auftragen zweier Skalen auf der Y-Achse kann hochgradig irreführend sein, da es im Belieben des Produzenten steht, wie stark er den Zusammenhang darstellen möchte:

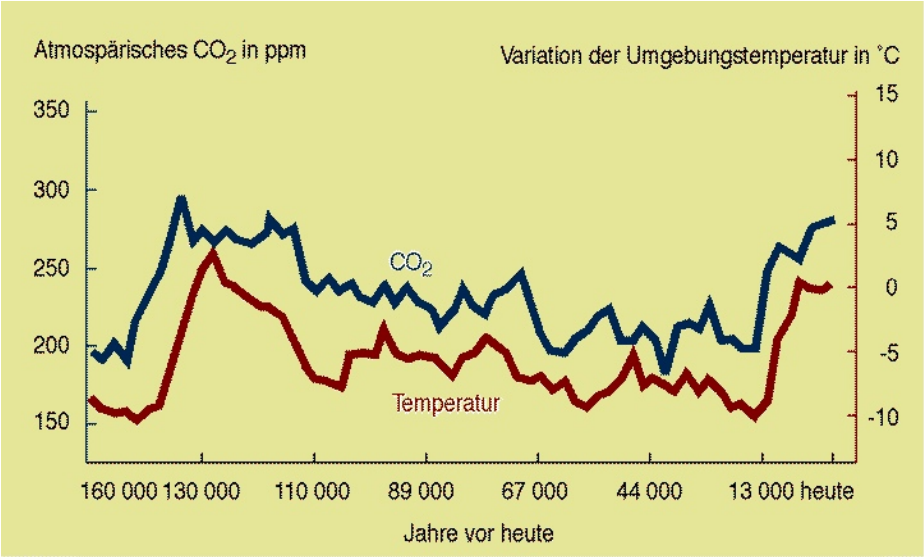


Abb. 7. Zwei verschiedene Skalen auf einer Dimension. Durch Reskalierung einer Größe kann der Zusammenhang praktisch als nichtexistent dargestellt werden, da die Schwankungen gegen Null gehen können, oder aber als beinahe perfekt wie hier. Welche Aussage ist richtig? Aus: von Weizsäcker, E. U., Lovins, A. B., Lovins, L. H. (1995). Faktor vier: Doppelter Wohlstand – halbierter Naturverbrauch, München: Droemer Knaur, S. 251.

Quantität. In einem kooperativen Gespräch versuchen Sprecher ihren Gesprächspartnern weder zu wenig noch zu viele Informationen zu geben. Die Hauptschwierigkeit bei der Operationalisierung dieser Maxime besteht darin, den semantischen Informationsgehalt von Visualisierungen zu messen (und nicht etwa Bits). Dieser semantische Gehalt hängt u. a. von dem Vorwissen des Rezipienten, den vorhandenen Kontextinformationen und den aus der Darstellung erschließbaren Informationen ab, alles Faktoren die praktisch nie genau bestimmt werden können. Als einfache Faustregel für Visualisierungen kann jedoch Tuftes (1983) "data-ink-ratio" gelten:

Eine gute Visualisierung präsentiert ein Maximum an Daten mit einem Minimum an Druckerschwärze.

Dieses Prinzip ist natürlich zu einfach, denn gerade in Lehr-Lern-Kontexten kommt es nicht immer auf ein Maximum an Informationen sondern eher darauf an, Wichtiges von Unwichtigem zu trennen und Wichtiges dabei in Text und Bild so zu präsentieren, dass es gut gelernt werden kann. Drei grobe Richtlinien für die effektive Gestaltung von Text-Bild Kombinationen lassen sich ableiten.

(1) Visuelle Sparsamkeit: Dazu gibt es zahlreiche grafische Mittel um für den Adressaten Wichtiges gegenüber Unwichtigem hervorzuheben (z. B. durch Überzeichnung, Einfär-

bung, Unterlegung). Die einfachste Möglichkeit ist jedoch, irrelevante Information einfach wegzulassen. So haben sich z. B. nüchterne Strichzeichnungen und schematische Abbildungen als effektiver erwiesen als detailreiche Realbilder (Dwyer, 1978; Scheiter, Gerjets, Huk, Imhof, & Kammerer, 2009). Trotzdem gehört der Verstoß gegen die Relevanzmaxime zu den häufigsten Sünden, da viele Grafiker sich vermutlich von Haus aus eher als Kreative denn als nüchterne Nachzeichner verstehen. Die Abbildung 8 (links) setzt sich aus zwei Säulendiagrammen zusammen, die einen statistischen Zusammenhang zwischen Investitionen und Arbeitsplätzen veranschaulichen. Leider wird durch unnötige Piktogramme und Schattierungen die relevante Information verdeckt. Eine visuell abgespeckte Darstellung der Abbildung (rechts) zeigt auf einen Blick, dass Investitionen – meist mit einer Verzögerung –auf den Arbeitsmarkt durchschlagen.

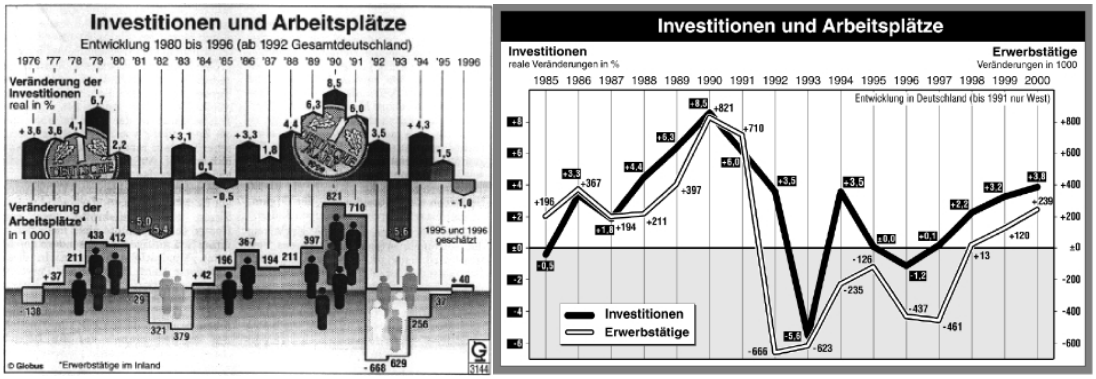


Abb. 8. Links: Der Zusammenhang von Investitionen und Arbeitsplätzen in einer Infografik aus dem Wirtschaftsteil der Frankfurter Rundschau am 23.3.1996. Rechts: Diese Visualisierung aktuellerer Daten ist von zahlreichen unnötigen Informationen entschlackt. Die beiden Kurven demonstrieren auf einen Blick einen Zusammenhang zwischen den beiden Variablen (Redesign von A. Bengsch).

(2) Text-Bild Verweise: Autoren sind jedoch oft auf vorgegebenes Grafikmaterial angewiesen, sodass auch hier nur die Sprache als Kompensation weiterhelfen kann. Bei unübersichtlichen und komplexen Gegenständen muss die Aufmerksamkeit sprachlich gesteuert werden, wobei prinzipiell zwei Möglichkeiten bestehen. Zum einen kann man im laufenden Text immer wieder auf relevante Bildelemente verweisen, was zu einem häufigen Springen zwischen Text und Bild führt. Solche Verweise zwischen Text und Bild werden auch als „Signals“ bezeichnet. Diese können explizite sprachliche Verweise sein („siehe Abbildung“) oder sie können thematische Zusammenhänge zwischen Text- und Bildinformationen etwas subtiler, nämlich über entsprechende Farbmarkierungen hervorheben („color coding“). Studien haben gezeigt, dass das Aufzeigen von Zusammenhängen in Text und Bild sowohl über explizite Verweise als auch über color coding zu besseren Lernleistungen führte (Mautone & Mayer, 2001; Ozcelik, Arslan-Ari, & Cagiltay, 2010; Ozcelik, Karakus, Kursun, & Cagiltay, 2009; Scheiter & Eitel, 2010). Blickbewegungsanalysen zeigten dabei, dass bessere Lernleistungen über schnelleres Auffinden von Informationen im Bild (Ozcelik et al., 2009) sowie über eine höhere Bildaufmerksamkeit (Scheiter & Eitel, 2010) erklärt werden konnten. Demnach wurde die Integration von Text- und Bildinformationen

durch das Hinzufügen von Verweisen in Text und Bild verbessert. Ein Nachteil solcher Verweise in Text und Bild ist jedoch, dass der Lesefluss eventuell unterbrochen wird. Eine andere Möglichkeit zur Förderung von Integration besteht daher darin, in Überschriften auf bestimmte Bildelemente zu verweisen. Dann werden die entsprechenden Bildelemente in der Regel erst am Ende des entsprechenden Absatzes angeschaut. Bei einer Untersuchung (Bausmith, 2000) dieser beiden Möglichkeiten haben sich keine nennenswerten Unterschiede gezeigt, es wurde aber deutlich, dass zumindest bei einem hohen Vorwissen der Probanden diese sprachlichen Steuerungen positive Effekte hatten. Leider trat dies nicht bei Probanden mit niedrigem Vorwissen auf; wieder ein Beleg dafür, dass es keine Patentrezepte gibt und selbst die Formel „Text + Bild + Integration sind besser als Text + Bild" nicht generell gilt.

(3) Raumzeitliche Nähe: Eine weitere Maßnahme, um die Integration von Text- und Bildinformationen und somit den Lernerfolg, zu verbessern, ist die räumlich und zeitliche nahe Präsentation von Text und Bild. Eine Meta-Analyse von Ginns (2006) hat gezeigt, dass eine gleichzeitige Präsentation von Text- und Bildinformationen zu besseren Lernleistungen führte als eine sequentielle (nicht-gleichzeitige) Text-Bild Darbietung. Dies war vor allem dann der Fall, wenn die Informationen komplex waren. Werden Texte und Bilder gleichzeitig, also auf derselben Seite dargeboten, dann sollte man ferner die räumliche Distanz zwischen Text und Bild reduzieren, um die Integration zu fördern und somit den Lernerfolg zu verbessern (vgl. Johnson & Mayer, 2012). So ist die Wahrscheinlichkeit höher, dass Informationen gemeinsam abgespeichert werden (Kulhavy et al., 1985; Paivio, 1986). Außerdem ist die Wahrscheinlichkeit höher, dass Texte und Bilder ergänzende Funktionen für den Verständnisaufbau ausüben können (Schnotz, 2001).

Relevanz. Diese Maxime fordert:

Visualisierungen müssen einen Bezug zum Thema bzw. zum Kommunikationsziel haben.

Es ist nahezu trivial, anzunehmen, dass das Hinzufügen von Bildern zu einem Text nur dann eine lernförderliche Wirkung hat, wenn sich Text und Bild auf ein gemeinsames Thema beziehen. Nichtsdestotrotz ist es wichtig, dies zu erwähnen. Besonders darum, weil sich diese Richtlinie noch etwas spezifizieren lässt. Texte und Bilder sollten sich nämlich nicht nur auf ein gemeinsames Thema beziehen, sondern sie sollten auch jeweils den Teil an Informationen darbieten, für den sie als Medium besser geeignet sind. So ist ein zentraler Vorteil von Bildern, dass sie kaum Interpretationsspielraum beim Vermitteln von räumlichen Relationen bieten. Man kann in einer Abbildung direkt ablesen, wo sich drei Kästchen A, B und C in Relation zueinander im Raum befinden. Würde man das allein sprachlich erklären müssen, dann bräuchte man Zeit und kognitive Ressourcen (Larkin & Simon, 1987). Das heißt: Möchte man räumliche Informationen vermitteln oder sind räumliche Relationen zentral um einen Sachverhalt verstehen zu können, dann sollte man diese Informationen mithilfe eines Bildes – und nicht mithilfe eines Textes – darstellen (vgl. Schmidt-Weigand & Scheiter, 2011). Andererseits ist Text besser geeignet, um Verallgemeinerungen, Disjunktionen oder Negationen auszudrücken (vgl. „Haustiere nicht gestattet"). Text ist damit sehr viel expressiver, und daher besser geeignet um argumentative Abwägungen, Erörterungen oder abstrakte Funktionsweisen auszudrücken, und sollte vorzugsweise auch dafür eingesetzt werden.

Modalität. Diese Maxime fordert:

Gestalte die Visualisierung so, dass sie den Betrachtenden keine unnötigen Schwierigkeiten bei der kognitiven Verarbeitung bereitet.

Die Umsetzung dieser Maxime ist natürlich ebenfalls je nach Kontext keineswegs trivial, aber es ist klar, dass Rezipienten genau diese Dienstleistung von Visualisierungen erwarten. Während zum eindeutigen und klaren Sprechen und Schreiben unzählige Untersuchungen vorliegen und daraus bewährte Richtlinien abgeleitet sind, ist die verständliche Gestaltung von Bildern noch kaum systematisch ausgearbeitet (Ballstaedt, 1999).

Soweit zur Diskussion der einzelnen Maximen. Ohne den theoretischen Rahmen der Griceschen Maximen hat Tufte (1997) vergleichbare Qualitätsstandards für Visualisierungen formuliert:

> "For information displays, design reasoning must correspond to scientific reasoning. Clear and precise seeing becomes as one with clear and precise thinking. For example, the scientific principle, make controlled comparisons, also guides the construction of data displays, prescribing that the ink or the pixels of graphics should be arranged so as to depict comparisons and context. … design quality grows from intellectual quality. Such … principles … include (1) documenting the sources and the characteristics of the data, (2) insistently enforcing appropriate comparisons, (3) demonstrating mechanisms of cause and effect, (4) expressing those mechanims quantitatively, (5) recognizing the inherently multivariate nature of analytic problems, and (6) inspecting and evaluating alternative explanations." (p. 53)

Mit anderen Worten:

Gute Bilder und Grafiken sollten vor allem relevant und problemadäquat sein, informativ und ehrlich. Alles andere ist zweitrangig.

Literatur

Ainsworth, S. (2006). DeFT: a conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198.

Anderson, J. R. (1978). Arguments Concerning Representations for Mental Imagery. In: *Psychological Review, 85*, 4, 249–277.

Baddeley, A. (1990). *Human Memory: theory and practice*. Hove: Erlbaum.

Ballstaedt, St.-P. (1999). Verständlichkeit von technischen Bildern. In J. Hennig & M. Tjarks-Sobhani (Hg.), *Verständlichkeit und Nutzungsfreundlichkeit von technischer Dokumentation*. Lübeck: Schmidt-Römhild, 89–100 .

Bausmith, J. L. (2000). The effects of prior knowledge and explicit cueing on students' understanding of scientific systems. *Dissertation Abstracts International Section A: Humanities & Social Sciences, Vol 61(1-A)*, Jul 2000, 67ff.

Bonsiepe, G. (1968). Visuell/verbale Rhetorik, *Format* IV/5.

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*, 155–166.

DeLoache, J. S. & Burns, N. M. (1994). Early Understanding of the Representational Function of Pictures. In: *Cognition, 52*, 83–110.

Deregowski, J. B. (1989). Real Space and Represented Space: Cross-Cultural Perspectives. In: *Behavioural and Brain Sciences, 12*, 51–119.

DFG Pressemitteilung Nr. 26: Task Force legt Abschlußbericht vor. 19. Juni 2000: http://dfg.de/service/presse/pressemitteilungen/2000/pressemitteilung_nr_26/index.html. Zuletzt abgerufen am 05.05.2014).

Donald, M. (1993). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA.: Harvard University Press.

Dwyer, F.M. (1978). *Strategies for improving visual learning. A handbook for the effective selection, design, and use of visualized materials*. State College, PA: Learning Services.

Eitel, A., Scheiter, K., & Schüler, A. (2013). How inspecting a picture affects processing of text in multimedia learning. *Applied Cognitive Psychology, 27*, 451–461.

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction, 28*, 48–63.

e-teaching.org Redaktion (2012). Gestaltgesetze. Zuletzt geändert am 25.08.2012. Leibniz-Institut für Wissensmedien: http://www.e-teaching.org/didaktik/gestaltung/visualisierung/gestaltgesetze/. Zuletzt abgerufen am 05.05.2014.

Folker, S., Ritter, H., & Sichelschmidt, L. (2005). Processing and integrating multimodal material: The influence of color-coding. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the cognitive science society*). Mahwah, NJ: Erlbaum, 690–695.

Frey, S. (1999). *Die Macht des Bildes: der Einfluß der nonverbalen Kommunikation auf Kultur und Politik*. Bern: Huber.

Ginns, P. (2006). Integrating information: a meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*, 511–525.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts*. New York: Academic Press, 41–58.

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low-and high-ability children. *Contemporary Educational Psychology, 24*, 95-123.

Harris, L. R. (1999). *Information Graphics*. Oxford: Oxford University Press.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*, 717–742.

Hochberg, J. E. & Brooks, V. (1962). Pictorial Recognition as an Unlearned Ability. In: *American Journal of Psychology*, 75, 624–628.

Hutchins, E. (1995). How a Cockpit Remembers its Speed. *Cognitive Science 19*, 265–288.

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied, 18*, 178.

Kirschner P.A. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction, 12,* 1–10.

Kosslyn, S. M (1980). *Image and Mind*. Cambridge, Mass.: Harvard University Press.

Kosslyn, S. (1989). Understanding Charts and Graphs. *Applied Cognitive Psychology*, Vol. 3. 185–226.

Kosslyn, S. M (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, Mass.: MIT Press.

Kulhavy, R. W., Lee, J. B., & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology, 10*, 28–37.

Kubey, R. & Csikszentmihalyi, M. (2002). Fernsehsucht. *Spektrum der Wissenschaft, 5/2002*, 70–75.

Larkin, J.H. & Simon, H.A. (1987). Why a diagramm is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-–99.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: a review of research. *Educational Communication and Technology, 30*, 195–232.

Lindberg, D. C. (1992). *The Beginnings of Western Science*. Chicago: University of Chicago Press.

Macdonald-Ross, M. (1977). Graphics in texts. In L.S. Shulman (Ed.), *Review of research in education*, Vol. 5, Itasca, Ill.: Peacock.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*, 377–389.

McCloud, S. (1993). Understanding Comics: The Invisible Art. Kitchen Sink Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge: Cambridge University Press.

Messaris, P. (1997). *Visual Persuasion: The Role of Images in Advertising.* Thousands Oaks, CA: Sage Publications.

Miller, A. I. (1984). *Imagery in Scientific Thought: Creating 20th-Century Physics*. Boston: Birkhäuser.

Neurath, O. (1991). *Gesammelte bildpädagogische Schriften*. Wien: Hölder-Pichler-Tempsky.

Norman, D. A. (1993). *Things that make us Smart,* Reading: Addison-Wesley.

OECD and statistics Canada (1995), *Literacy, Economy and Society: Results of the First International Adult Literacy Survey*, Paris and Ottawa.

Oestermeier, U. & Hesse, F. W. (2000). Verbal and visual causal arguments. *Cognition*, *75*, 65–104.

Oestermeier, U. (1998). Bildliches und logisches Denken: Eine Kritik der Computertheorie des Geistes. In: C. Habel & G. Rickheit (Hrsg.), *Studien zur Kognitionswissenschaft*, Wiesbaden: Deutscher Universitäts-Verlag.

Oestermeier, U., Reinhard-Hauck, P. & Ballstaedt, St.-P. (2001). Gelten die Griceschen Maximen auch für visuell Argumente? In K. Sachs-Hombach (Hrsg.), *Bildhandeln*. Magdeburg: Scriptum Verlag, 207-222.

Ozcelik Arslan- Ari, I., Cagiltay, K., E. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*, 110–117.

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers and Education, 53*, 445–453.

Paivio, A. (1986). *Mental Representations*. New York: Oxford University Press.

Peeck, J. (1993). Increasing Picture Effects in Learning From Illustrated Text. *Learning and instruction 3* (3), 227-238.

Pörksen, U. (1997). *Weltmarkt der Bilder*. Eine Philosophie der Visiotype. Stuttgart: Klett-Cotta.

Postman, N. (1985). *Amusing Ourselves to Death. Public Discourse in the Age of Show Business*, New York: Viking-Penguin (Elisabeth Sifton Books).

Pylyshyn, Z.W. (1984). *Computation and Cognition: Towards a Foundation for Cognitive Science*. Cambridge, Mass.: MIT Press.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied, 7*, 219–226.

Scheiter, K., & Eitel, A. (2010). The effects of signals on learning from text and diagrams: how looking at diagrams earlier and more frequently improves understanding. In A. K. Goel, M. Jamnik, & N. H. Narayanan (Eds.), *Diagrammatic representation and inference–6th International Conference, Diagrams 2010 (LNAI 6170)*. Heidelberg: Springer, 264–270.

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction, 19*, 481-494.

Schmidt-Weigand, F., & Scheiter, K. (2011). The role of spatial descriptions in learning from multimedia. *Computers in Human Behavior, 27*, 22–28.

Schnotz, W. (2001). Wissenserwerb mit Multimedia. *Unterrichtswissenschaft 29*, 292–318.

Schnotz, W. (2002). Wissenserwerb mit Texten, Bildern und Diagrammen. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia*. Weinheim: Psychologie Verlags Union, 65–82.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction, 13*, 141–156.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4,* 295-312.

Shepard, R. N. & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects, in: *Science, vol. 171*, 701–703.

Tilling, L. (1975). Early experimental graphs. *The British Journal for the History of Science*, Vol. 8, No. 30, 193–-213.

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.

Tufte, E.R. (1997). *Visual explanations: images and quantities, evidence and narrative*. Cheshire, CT: Graphic Press.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic.

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*, 95–99.

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14,* (3), 261–312.

Weidenmann, B. (1988). *Psychische Prozesse beim Verstehen von Bildern*. Bern: Huber.

Weidenmann, B. (1994). Wissenserwerb mit Bildern. Instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen. Bern, Göttingen, Toronto, Seattle: Huber.

Weidenmann, B. (2002). Multicodierung und Multimodalität im Lernprozeß. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia*. Weinheim: Psychologie Verlags Union, 45–64.

Washburne, J.N. (1927a). An experimental study of various graphic, tabular and textual methods of presenting quantitative material. *Journal of Educational Psychology, 18,* 361–376.

Washburne, J.N. (1927b). An experimental study of various graphic, tabular and textual methods of presenting quantitative material. *Journal of Educational Psychology, 18*, 465–476.

Wygotsky, L. (1934/1969). *Denken und Sprechen*. Frankfurt a.M.: Fischer.

Glossar

Blickbewegungsstudien: Eine wichtige Methode der Medienpsychologie, bei der die Augenbewegungen (Sakkaden) erfasst und in Bezug zu den fixierten Bereichen gesetzt werden. Praktiker berichten, dass die Auswertung der dabei anfallenden Datenmengen regelmäßig Probleme bereitet.

Imagery-Debatte: In der Imagery-Debatte behaupten die Piktorialisten auf der einen Seite die Existenz eines eigenständigen bildhaften bzw. analogen Codes, der die Besonderheiten des bildlichen Vorstellens und Erinners erklären soll, während auf der anderen Seite die Propositionalisten argumentieren, diese Besonderheiten seien auf der Basis eines einheitlichen mentalen propositional-symbolischen Codes erklärbar.

Indikatorischer Verstehensmodus: Ein an der Kommunikationsabsicht geschulter Verstehensmodus (Was will uns der Autor damit sagen?), bei dem Vorwissen aktiviert wird und andere Informationsquellen, wie z.B. der Begleittext, hinzugezogen werden, um den vollen Gehalt der Abbildung zu erschließen. Didaktische Bilder und Grafiken erfordern oft diesen weitergehenden Verstehensmodus, obwohl der flüchtige Leser zum *ökologischen Verstehensmodus* neigt.

Kognitive Artefakte: Werkzeuge, mit denen kognitive Leistungen erbracht werden. In diesem weiten Sinne sind natürlich auch Grafiken und Bilder kognitive Werkzeuge. Sie führen etwas vor Augen, was sonst im Gedächtnis behalten werden müsste. Über die gedächtnisentlastende Funktion hinaus dienen Grafiken aber vor allem der Sichtbarmachung abstrakter Sachverhalte, die damit überhaupt erst zu einem dauerhaften Reflexionsgegenstand werden können. Kulturhistorisch betrachtet sind Schrift und Bild somit unabdingbare Voraussetzung für technologische und wissenschaftliche Leistungen.

Ökologischer Verstehensmodus: Ein an der natürlichen Umwelt geschulter Verstehensmodus, bei dem die Interpretation der Bilder durch hochgradig automatisierte und extrem schnelle Interpretationsmechanismen erfolgt. Dieser Verstehensmodus reicht in der Regel nicht aus, um den didaktischen Gehalt von Abbildern und Grafiken zu erschließen. Die meisten Bilder werden in diesem oberflächlichen Modus behandelt (vgl. Indikatorischer Verstehensmodus).

Text-Bild-Integration: Ein Verständnis komplexer Grafiken und Bilder setzt zumeist voraus, das die Lernenden aktiv Bezüge zwischen Bild- und Textinhalten herstellen, diese Inhalte also integrieren. Aus Produzentensicht kann diese Aktivität des Rezipienten durch eine gute Gestaltung gefördert werden, bei der die Leser möglichst wenig zwischen verschiedenen Seiten blättern, bzw. zwischen weit auseinander liegenden Bildelementen und Legenden hin und her springen müssen.

Dual-Coding: Eine Theorie der mentalen Repräsentation, die davon ausgeht, dass innerhalb des kognitiven Systems sprachlich-begriffliche und bildhafte Gedächtnisstrukturen existieren. Diese Annahme wurde innerhalb der Imagery-Debatte kontrovers diskutiert, wobei insbesondere die Frage im Mittelpunkt stand, ob die These der dualen Codierung und die Gegenthese eines einheitlichen mentalen Codes überhaupt entscheidbare empirische Implikationen haben.

Die Autoren

Name: Uwe Oestermeier

Internet: http://www.iwm-kmrc.de/www/de/mitarbeiter/

E-Mail: u.oestermeier@iwm-kmrc.de

Dr. Uwe Oestermeier leitet die Entwicklungsgruppe innerhalb der Medientechnik des Instituts für Wissensmedien. Er studierte Philosophie, Psychologie und Religionswissenschaften an den Universitäten Bonn und Tübingen (Magister 1989) und promovierte an der Universität Tübingen zum Thema "Bildliches und logisches Denken". Von 1989 bis 1993 war er wissenschaftlicher Angestellter am Wilhelm-Schickard-Institut für Informatik der Universität Tübingen. 1994 wechselte er als wissenschaftlicher Angestellter zum Deutschen Institut für Fernstudienforschung (DIFF) in die Abteilung für Angewandte Kognitionswissenschaft. Seit 2003 ist er Leiter der Softwareentwicklungsgruppe des Instituts für Wissensmedien.

Wichtigste Veröffentlichungen: Begriffliche und empirische Fragen der Kausalkognition. Kognitionswissenschaft, 6, 1997, 2, 70-85. Bildliches und logisches Denken. Eine Kritik der Computertheorie des Geistes. Wiesbaden: DUV 1998; (mit F.W. Hesse): Verbal and Visual Causal Arguments, in: Cognition 75, 2000, 65-104.

Name: Alexander Eitel

Internet: http://www.iwm-kmrc.de/www/de/mitarbeiter/

E-Mail: a.eitel@iwm-kmrc.de

Dr. Alexander Eitel ist wissenschaftlicher Mitarbeiter am Leibniz-Institut für Wissensmedien. Er studierte Psychologie an der Universität Tübingen (Diplom 2009). Von April 2009 bis April 2013 promovierte er bei Prof. Dr. Katharina Scheiter am Leibniz-Institut für Wissensmedien. Im Rahmen seiner Promotion befasste er sich damit, wie räumliche Informationen, die bereits aus einer frühen Phase der Bildverarbeitung entnommen werden, Prozesse der Textverarbeitung beeinflussen. Im April 2013 schloss er seine Promotion ab. Seither arbeitet er als Postdoktorand in der Arbeitsgruppe „Wissenserwerb mit Multimedia“ unter der Leitung von Prof. Dr. Katharina Scheiter.
Wichtigste Veröffentlichungen:

Eitel, A., Kühl, T, Scheiter, K., & Gerjets, P. (2014). Disfluency meets cognitive load in multimedia learning: Does harder-to-read mean better-to-understand? *Applied Cognitive Psychology*. doi:10.1002/acp.3004

Eitel, A., & Scheiter, K. (2014). Picture or text first? Explaining sequence effects when learning with pictures and text. *Educational Psychology Review*. doi:10.1007/s10648-014-9264-4

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: Evidence for scaffolding. *Learning and Instruction, 28*, 48-63. doi:10.1016/j.learninstruc.2013.05.002

Eitel, A., Scheiter, K., & Schüler, A. (2013). How inspecting a picture affects processing of text in multimedia learning. *Applied Cognitive Psychology*. doi:10.1002/acp.2922