

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



UNIVERSITÄT TÜBINGEN

LEHRSTUHL FÜR KOGNITIVE NEUROWISSENSCHAFTEN

BACHELORARBEIT
KOGNITIONSWISSENSCHAFT

Veränderungen der wahrgenommenen
Identität in der Gesichtserkennung

Vorgelegt von:
Esther Kutter

Betreuer:
Prof. Dr. Hanspeter A. Mallot
Zweitgutachter:
Prof. Dr. Martin V. Butz

30. August 2012

Erklärung

Ich versichere, die vorliegende Bachelorarbeit selbstständig und unter ausschließlicher Verwendung der angegebenen Quellen angefertigt zu haben.

Tübingen, 30. August 2012

(Esther Kutter)

Danksagung

Ein herzliches Dankeschön gilt an erster Stelle Wolfgang Röhrich für seine kritische Unterstützung und Anleitung sowie Professor Dr. Mallot für seine zahlreichen Anregungen und Ratschläge. Darüber hinaus gilt mein Dank allen, die mich unterstützt und mir geholfen haben, diese Arbeit zu realisieren.

Zusammenfassung

Die kategoriale Wahrnehmung unserer Umwelt ist ein faszinierendes Phänomen, das Wissenschaftler schon seit vielen Jahren in seinen Bann zieht. Ein häufig untersuchter Aspekt ist dabei die Wahrnehmung und Verarbeitung künstlicher Kontinua, wie sie beispielsweise beim Ineinandermorphen zweier Gesichter entstehen. In früheren Studien konnte insbesondere über Kategorisierungs- und Diskriminationsaufgaben nachgewiesen werden, dass auch diese unter bestimmten Bedingungen kategorial wahrgenommen werden. Ziel der vorliegenden Studie war es nun, eine neue Methode zu testen, bei der die Kategoriengrenze innerhalb einer Morphsequenz direkt und unmittelbar von den Probanden bestimmt werden sollte. Dabei konnte herausgefunden werden, dass die Probanden unter den hier gegebenen Bedingungen im Allgemeinen auf einfache Bildervergleiche zurückgriffen. Zusätzlich wurde die empfundene Ähnlichkeit und Markanz der verwendeten Gesichter erfasst, um deren Einfluss auf die Lösungsfindung zu untersuchen. Dabei konnten geringe Korrelationen nachgewiesen werden. Hauptdiskussionspunkte waren daher zum einen die grundsätzliche Anwendbarkeit der verwendeten Methode sowie das Zusammenspiel der untersuchten Faktoren.

Inhaltsverzeichnis

1	Einleitung	1
2	Material und Methoden	5
2.1	Versuchspersonen	5
2.2	Stimulusmaterial	5
2.3	Apparatur	7
2.4	Prozedur	8
2.4.1	Lernaufgabe	8
2.4.2	Drehregleraufgabe	9
2.4.3	Bewertungsaufgabe	11
3	Ergebnisse	13
3.1	Lernaufgabe	13
3.2	Analyse der Drehregleraufgabe	13
3.3	Bezug zu Markanz und Ähnlichkeit der Gesichter	17
3.4	Hysterese-Effekte	20
4	Diskussion	22
4.1	Interpretation der Ergebnisse	22
4.1.1	Analyse der Drehregleraufgabe	22
4.1.2	Bezug zu Markanz und Ähnlichkeit der Gesichter	23
4.1.3	Hysterese-Effekte	24
4.1.4	Lösungsstrategien und Lerneffekte	25
4.2	Kritikpunkte	26
4.2.1	Inhaltliche Kritikpunkte	26

4.2.2	Methodische Kritikpunkte	27
4.3	Zusammenfassung und Ausblick	29
	Literaturverzeichnis	31
	Weitere Abbildungen	33

Abbildungsverzeichnis

2.1	Grundprinzip des Morphalgorithmus	7
2.2	Lernaufgabe	9
2.3	Drehregleraufgabe	10
2.4	Bewertungsaufgabe	11
3.1	Verteilungsfunktionen für die weiblichen Gesichterpaare bei einer einzelnen Versuchsperson	14
3.2	Verteilungsfunktionen für Gesichterpaare bei allen Versuchspersonen	15
3.3	Verschiebungsrichtungen zwischen den Stimuli	17
3.4	Korrelation zwischen Markanz und Verschiebungsstärke	18
3.5	Korrelation zwischen Antwortvarianz und Ähnlichkeit	19
3.6	Hysterese-Effekte für die Gesichterpaare bei allen Versuchspersonen	20
3.7	Beispiele für das typische Klickverhalten der Versuchspersonen	21
1	Verwendete Stimuli	33
2	Bewertung der Einzelgesichter hinsichtlich ihrer Markanz	34
3	Bewertung der Gesichterpaare hinsichtlich ihrer Ähnlichkeit . .	34
4	Klickverhalten und benötigte Zeit über die Versuchsblöcke hinweg	35

Kapitel 1

Einleitung

In jedem Augenblick unseres Lebens werden wir geflutet von einer unendlichen Fülle verschiedenster Sinneseindrücke. Um diese Flut von Informationen aufzunehmen und zu verarbeiten, hat unser Gehirn im Laufe der Evolution verschiedenste Mechanismen entwickelt. Einer davon ist die so genannte „kategoriale Wahrnehmung“. Die vielen, kontinuierlichen Sinneseindrücke werden nach bestimmten Kriterien in eine überschaubare Menge diskreter Kategorien eingeordnet. Dabei ist diese Klassifizierung ab einem bestimmten Punkt weitestgehend unabhängig von der tatsächlichen, physikalischen Ähnlichkeit: Objekte derselben Gruppe werden als gleich oder ähnlich wahrgenommen, solche verschiedener Gruppen als unterschiedlich, selbst wenn die physikalischen Differenzen zwischen diesen Objekten in allen Fällen gleich sind (Harnad, 1990).

Anschaulichstes Beispiel hierfür ist wohl der Regenbogen, der zwar alle Wellenlängen des Lichts enthält (also theoretisch auch unendlich viele Farben), in dem wir im Allgemeinen aber trotzdem nur sechs Grundfarben wahrnehmen. Und die Unterscheidung einzelner Farben fällt uns an den Kategorien Grenzen (grün-gelb) leichter als innerhalb einer Kategorie (grün-grün), selbst wenn der Wellenlängenunterschied zwischen den beiden Grüntönen derselbe ist wie zwischen dem Grün- und Gelbton (Bornstein & Korda, 1984). Weitere Beispiele kategorialer Wahrnehmung sind unter anderem auch in der Wahrnehmung und Verarbeitung von Sprache zu finden (Lieberman, Harris, Hoff-

1. EINLEITUNG

man & Griffith, 1957; Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967).

Dieser Effekt scheint im Allgemeinen immer dann aufzutreten, wenn wir mit verschiedenen Reizen aus einem natürlichen physikalischen Kontinuum konfrontiert werden. Fraglich ist nur, ob es sich dabei um einen angeborenen oder erlernten Prozess handelt und auf welcher Verarbeitungsebene die Kategorisierung stattfindet. Für beide Sichtweisen gibt es Evidenz: Beispielsweise verfügen wir Menschen über Neuronen, die selektiv auf die vier Grundfarben blau, grün, gelb und rot reagieren (DeValois & DeValois, 1975). Dies sowie die Verwendung kulturübergreifend gleicher Farbkategorien legen nahe, dass es sich – zumindest bei der Farbwahrnehmung – eher um angeborene Mechanismen handelt. Bei der Verarbeitung von Sprache hingegen scheinen sowohl angeborene Prozesse (beispielsweise Sensitivität für bestimmte akustische Merkmale) als auch Erfahrungen eine wichtige Rolle zu spielen (Rosen & Howell, 1987).

Wahrscheinlich ist also, dass bei der kategorialen Wahrnehmung beide Faktoren zusammenwirken. Um diese Hypothese zu überprüfen, schlug Harnad (1990) die Verwendung „künstlicher Kontinua“ vor, also solcher, die in dieser Form nicht natürlich vorkommen und die gegebenenfalls nicht nur in einem Faktor (beispielsweise der Veränderung der Wellenlänge) variieren (Harnad, 1990). Darauf basierend wurden in den vergangenen Jahrzehnten zahlreiche Studien zur Wahrnehmung von Gesichtern durchgeführt. Diese sind für die Untersuchung kategorialer Wahrnehmung geradezu prädestiniert: Sie gehören nicht nur zu einer Objektklasse, die uns Menschen quasi von Geburt an sehr vertraut ist (Diamond & Carey, 1977), sondern sie sind darüber hinaus auch kontinuierlich natürlichen Transformationen (wie Mimik, Beleuchtung, Perspektivwechseln oder Alterungsprozessen) unterworfen. Nichtsdestotrotz haben wir im Laufe unseres Lebens gelernt, vertraute Gesichter wiederzuerkennen. Wie aber verändert sich unsere Wahrnehmung, wenn die Veränderungen nicht innerhalb eines Gesichts stattfinden, sondern zwischen den (an sich statischen) Gesichtern zwei verschiedener Personen?

Solch eine Transformation lässt sich durch so genanntes Morphen realisieren. Dabei werden zwischen zwei Einzelbildern Zwischenübergänge berechnet, so

dass sie letztendlich in fließenden Übergängen ineinander überführt werden können (siehe dazu auch Kapitel 2.2). Spielt man diese Bildersequenz anschließend als Film ab, erhält man den Eindruck einer stetigen Transformation.

Zumeist fanden diese Morphsequenzen in Kategorisierungsaufgaben Anwendung, in denen die Probanden ein Bild der Sequenz einem der beiden Ausgangsgesichter zuordnen mussten. Aus den Antworten wurden anschließend psychometrische Funktionen berechnet, anhand derer erkennbar war, ob die Gesichter kategorial wahrgenommen wurden oder nicht. Beale und Keil (1995) konnten mittels dieser Methode zeigen, dass es zwischen Gesichtern, die entlang eines solchen künstlichen Kontinuums präsentiert wurden, in der Tat eine scharfe Kategoriengrenze gibt, wenn die Personen den Probanden hinreichend bekannt waren (Beale & Keil, 1995). Kircher, Senior, Phillips und weitere (2001) generierten später Kontinua aus dem eigenen Gesicht beziehungsweise dem des Partners (also einer hochvertrauten Person!) und unbekanntem Gesichtern. Auch diese Bilder wurden kategorial wahrgenommen. Außerdem beobachteten sie, dass es bei einer seriellen (im Gegensatz zur randomisierten) Präsentation der Morphbilder systematische Verschiebungen der Kategoriengrenze in Richtung der zu Beginn präsentierten Person hin gab (Kircher et al., 2001). Diese Ergebnisse zeigen deutlich, dass es unter bestimmten Bedingungen in der Tat möglich ist, Kategorien auch innerhalb künstlicher Kontinua zu erlernen.

Der bisherige methodische Ansatz beruhte im Allgemeinen auf der Berechnung der psychometrischen Funktion, aus der anschließend indirekt Rückschlüsse auf die Kategoriengrenze gezogen wurden. In dieser Studie sollte nun näher untersucht werden, ob diese Effekte ebenfalls auftreten, wenn die Probanden explizit angewiesen werden, die subjektive Mitte (also die Kategoriengrenze) innerhalb einer Morphsequenz zu finden. Dazu wurden fein abgestufte Morphsequenzen generiert, in denen die Unterschiede zwischen zwei benachbarten Bildern möglichst gering waren. Die Probanden konnten dann über eine Art „Drehregler“ die Sequenz entlang scrollen, so dass sie die Transformation vom einen zum anderen Gesicht in einer fließenden Bewegung betrachten konnten. Nun erhielten sie die Aufgabe, jenes Bild der Morphse-

1. EINLEITUNG

quenz zu finden, in dem ihrer Meinung nach von beiden Originalgesichtern gleich viele Elemente enthalten sind (also die subjektive Mitte). Dieses Bild entspricht gerade der Kategoriengrenze, also dem Umschlagpunkt vom ersten zum zweiten Gesicht. Sollte es auch hier zur kategorialen Wahrnehmung kommen, so müsste diese subjektive Mitte stets bei demselben Bild festgelegt werden.

Zusätzlich wurde die empfundene Ähnlichkeit und Markanz der verwendeten Stimuli erfasst, um deren Einfluss auf die Ergebnisse zu analysieren. Ersteres könnte die Unterscheidung der einzelnen Bilder erschweren, was zu größeren Varianzen in den Antworten führen würde und die Kategorisierung beeinträchtigen könnte. Letzteres stellt vor allem auch einen möglichen Indikator für die Verarbeitungsebene dar: Wird ein markantes Gesicht als dominanter innerhalb der Sequenz empfunden, könnte sich dies auf die Lokalisation der Kategoriengrenze auswirken. Dies würde bedeuten, dass es sich nicht mehr ausschließlich um einen simplen Abgleich einzelner Merkmale handelt, sondern durchaus höhere kognitive Prozesse die Kategorisierung beeinflussen.

Kapitel 2

Material und Methoden

2.1 Versuchspersonen

An dem folgenden Versuch nahmen insgesamt 20 Personen (10 männlich, 10 weiblich) im Alter von 18 bis 33 Jahren teil. Die meisten von ihnen waren Studenten der Eberhard-Karls-Universität Tübingen, Deutschland. Alle Probanden verfügten über eine normale oder korrigierte Sehstärke und waren naiv hinsichtlich des Nutzens und der Fragestellung des durchgeführten Experiments. Vor Beginn wurden sie über den Ablauf des Versuchs sowie den Verbleib der erhobenen Daten aufgeklärt.

2.2 Stimulusmaterial

Den Ausgangspunkt bildeten insgesamt 10 qualitativ hochwertige Farbfotografien von menschlichen Gesichtern aus der Frontalperspektive (je fünf Männer und fünf Frauen). Hierbei handelte es sich weder um berühmte Persönlichkeiten noch um die Probanden selbst oder ihnen bekannte Personen. Um Frisuren als Hinweisreiz zur Identifizierung auszuschließen, wurden bärtige Gesichter von vornherein aussortiert und die Haare in allen Fällen bestmöglich nach hinten gebunden, so dass beide Ohren sichtbar waren. Alle Personen zeigten einen möglichst neutralen Gesichtsausdruck.

Mit dem Bildbearbeitungsprogramm *Adobe Photoshop* (Adobe Systems, San

2.2. STIMULUSMATERIAL

Jose, USA) wurden die Bilder weiter bearbeitet. So wurden auf den Bildern Kopf und Hals freigestellt sowie die Haare bis zum Ansatz entfernt und durch einen kahlen Schädel ersetzt. Der Hintergrund wurde mit neutralem Schwarz gefüllt. Zuletzt wurden alle Bilder auf 795x795 Pixel skaliert, mit einer Punktdichte von 96 dpi und einer Farbtiefe von 24-bit (True Color).

Die im Anhang auf Seite 33 befindliche Abbildung 1 (links) zeigt die bearbeiteten und im folgenden Versuch verwendeten Gesichter (100%-Morphs, im Weiteren auch Zielgesichter genannt). Um die Anonymität der Personen zu wahren, wurden die Frauen mit A-E codiert, die Männer mit W-Z. Für den ersten Teil des Experiments wurden zusätzlich je fünf weitere männliche und weibliche Gesichter auf dieselbe Weise bearbeitet (siehe Abbildung 1, rechts). Diese dienten jedoch lediglich als Distraktorreize und fanden nur in diesem Teil des Versuchs Verwendung.

Anschließend wurden die Gesichter zu Paaren zusammengefasst, indem jedes Bild mit jedem anderen Bild desselben Geschlechts kombiniert wurde. Bei je fünf männlichen und weiblichen Gesichtern ergaben sich so insgesamt 10 männliche (VW, VX, VY, VZ, WX, WY, WZ, XY, XZ, YZ) und 10 weibliche (AB, AC, AD, AE, BC, BD, BE, CD, CE, DE) Gesichterpaare.

Die Bilder jedes Paares wurden dann ineinander gemorpht, also ein lineares Bilderkontinuum zwischen diesen zwei Gesichtern mit Hilfe des Morphprogramms *FotoMorph* (Digital Photo Software, Norwegen) erzeugt.

Bei diesem Verfahren werden zuerst auf den beiden Gesichtern etwa 150 bis 300 korrespondierende Punkte markiert (beispielsweise Konturen, Augen, Ohren, Nase und so weiter). Aus dieser Punktmenge wird als nächstes mittels Delaunay-Triangulation ein Dreiecksnetz erstellt, indem je drei benachbarte Punkte zu Dreiecken ohne sich überkreuzende Linien verbunden werden.

Anschließend werden die beiden Bilder durch „Warping“ (Verzerren) und „Fading“ (Überblenden) ineinander überführt. Beim Warping werden die Kontrollpunkte des einen Gesichts mit gleicher Schrittweite geradlinig in Richtung der jeweils korrespondierenden Punkte im anderen Gesicht verschoben (mittels so genannter linearer Interpolation). Dabei werden alle Punkte gleichmäßig verzerrt, so dass die Relationen untereinander bestehen bleiben

2.3. APPARATUR

(Wolberg, 1990). Dadurch werden die Konturen beider Bilder sukzessive in Übereinstimmung gebracht. Analog werden beim Fading die RGB-Werte linear interpoliert und vermischt. Mit jedem Interpolationsschritt nimmt der Anteil des einen Bildes ab und der des anderen zu. Das jeweilige Verhältnis beider Gesichter wird im Folgenden Morphlevel genannt werden (siehe auch Abbildung 2.1).

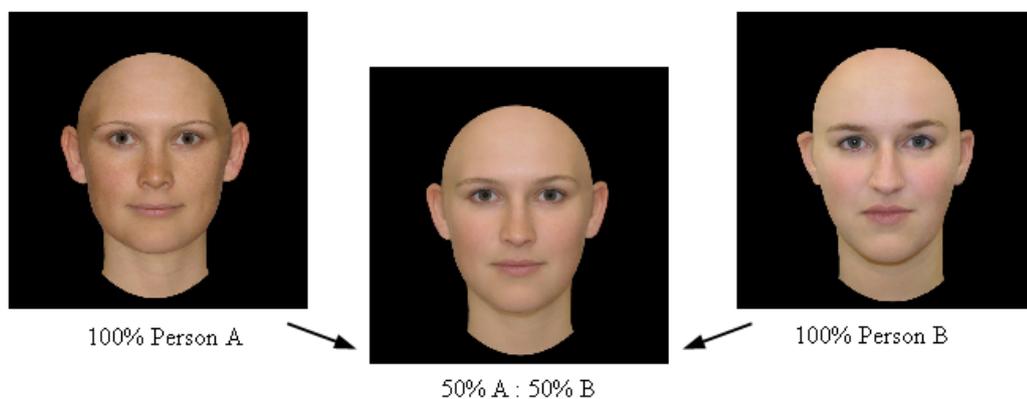


Abbildung 2.1: Grundprinzip des Morphalgorithmus. Eingangs auf beiden Gesichtern definierte Referenzpunkte werden trianguliert und anschließend durch Warping in Übereinstimmung gebracht. Die Texturwerte der Dreiecke werden durch Fading einander angeglichen. Dadurch können Mischgesichter (Mitte) mit Anteilen beider Originalgesichter (rechts, links) in beliebigem Verhältnis (Morphlevel) generiert werden.

Auf diese Weise wurde pro Paar eine Bildersequenz aus insgesamt 26 Bildern erstellt, in der das eine Gesicht jeweils in 4%-Schritten (also 24 „künstlichen“ Bildern) in das andere Gesicht überführt wird. Die Schrittweite wurde relativ gering gewählt, um die Unterscheidung von benachbarten Bildern zu erschweren und um fließendere Übergänge zu erhalten.

2.3 Apparatur

Die Probanden saßen in einem Abstand von 50 cm vor einem Standard-Monitor (Auflösung 1280x1024 Pixel). Das Experiment wurde mit der Psych-toolBox in *MatLab R2012a* (MathWorks, Natick, USA) programmiert und

ausgeführt. Zum Antworten und Reagieren während des Experiments standen den Versuchspersonen jeweils bestimmte Tasten einer Standard-Tastatur zur Verfügung (siehe auch Kapitel 2.4).

2.4 Prozedur

Das Experiment bestand insgesamt aus drei verschiedenen Abschnitten: eine kurze Lernaufgabe, gefolgt vom Hauptexperiment (im Folgenden auch Drehregleraufgabe genannt) und zuletzt eine Bewertung der Stimuli.

2.4.1 Lernaufgabe

Der erste Teil des Versuchs umfasste eine einfache Lernaufgabe und diente dazu, sich mit dem verwendeten Stimulusmaterial vertraut zu machen. Den Probanden wurden erst laminierte A5-Karten vorgelegt, auf die jeweils die Zielgesichter gedruckt waren. Sie bekamen die Anweisung, sich diese anzuschauen und so gut einzuprägen, dass sie sie später in einer Menge von vielen Gesichtern wiederfinden könnten. Hierfür war ihnen kein Zeitlimit gesetzt, aber im Mittel wurden etwa zwei bis drei Minuten benötigt.

Wenn die Versuchspersonen sich sicher genug fühlten, wurden die Karten wieder eingesammelt und der erste Programmteil begann (siehe Abbildung 2.2). Auf dem Bildschirm erschien ein Fixationskreuz und nach 300 Millisekunden eines von 20 möglichen Gesichtern: 10 Ziel- und 10 Distraktorgesichter (siehe Abbildung 1 auf Seite 33 im Anhang). Die Aufgabe war nun, zu entscheiden, ob das präsentierte Gesicht bei den zuvor gelernten Bildern dabei gewesen war. Bejahten sie die Frage – handelte es sich also um eines der Zielgesichter – so sollten sie auf dem Zahlenblock der Tastatur die Nummerntaste „1“ drücken; handelte es sich hingegen um eines der Distraktorgesichter, sollte die Nummerntaste „2“ gedrückt werden. Für diese Entscheidung war ihnen kein Zeitlimit gesetzt. Im Anschluss erhielten die Probanden eine Rückmeldung, ob die gegebene Antwort richtig war. Danach folgte der nächste Durchgang. Die Versuchspersonen mussten jedes der Bilder fünfmal korrekt klassifizieren. Bei einer falschen Entscheidung wurde das Gesicht noch einmal mehr

2.4. PROZEDUR

präsentiert. Die Anzahl der Durchgänge variierte daher je nach Leistung der Probanden, insgesamt ergaben sich aber mindestens 60 Durchgänge (20 Gesichter x 3-mal richtig) in komplett randomisierter Reihenfolge. Im Mittel benötigten die Versuchspersonen für diesen Teil des Experiments etwa fünf Minuten; eine Pause war nicht eingeplant.

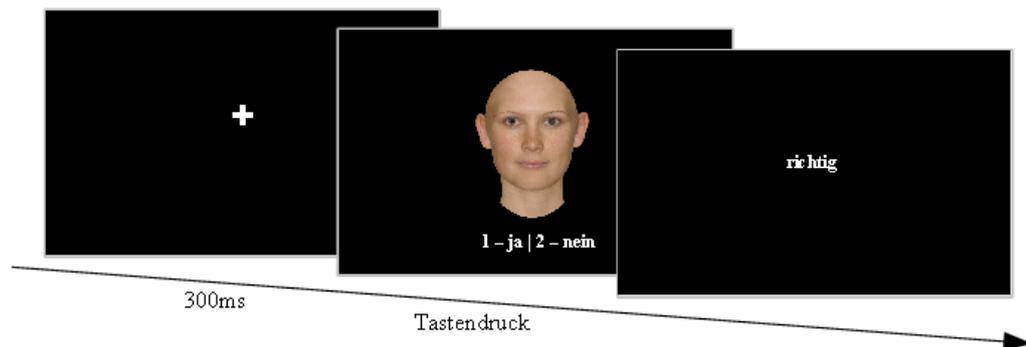


Abbildung 2.2: Lernaufgabe. Präsentation eines Fixationskrezes, gefolgt von einem Ziel- oder Distraktorgesicht nach 300 Millisekunden. Die Entscheidung, ob es sich um eines der zu lernenden Bilder handelte, erfolgte über die Nummerntasten, anschließend gab es eine Rückmeldung über die Korrektheit der Antwort.

2.4.2 Drehregleraufgabe

Im zweiten Teil des Experiments wurde den Probanden nach einem Fixationskruz (300 Millisekunden) jeweils ein Gesichterpaar sowie ein variables Bild aus der zugehörigen Morphsequenz präsentiert. Das Morphbild erschien in der Bildschirmmitte, die beiden Gesichter des Paares in zufälliger Anordnung auf der rechten und linken Seite (siehe Abbildung 2.3). Die Probanden erhielten nun die Instruktion, das Morphbild in der Mitte so einzustellen, dass sie es nicht mehr eindeutig einem der Ursprungsbilder zuordnen könnten, also den individuellen Umschlagpunkt zwischen den beiden Originalgesichtern innerhalb der Sequenz zu finden. Dazu konnten sie das Bild über die Pfeiltasten verändern, indem sie quasi wie mit einem Drehregler die Morphsequenz entlang scrollten: Bei einmaligem Drücken der rechten beziehungsweise linken Pfeiltaste wurde das mittlere Bild ersetzt durch das nächste Bild der Sequenz

2.4. PROZEDUR

in Richtung des Gesichts auf der rechten beziehungsweise linken Seite. Hielt man die Pfeiltaste gedrückt, konnte man in fließenden Übergängen bis zum Ende der Sequenz (also einem Originalbild) scrollen. Dabei gab es keine Skala oder ähnliches, die den Probanden explizit angab, auf welchem Morphlevel sie sich gerade befanden, die beiden 100%-Morphs waren aber zur Orientierung während der ganzen Zeit sichtbar. Die Probanden konnten beliebig oft hin- und herscrollen; hatten sie ihr gewünschtes Endbild eingestellt, so konnten sie die Auswahl mittels der Leertaste bestätigen. Danach folgte der nächste Durchgang.

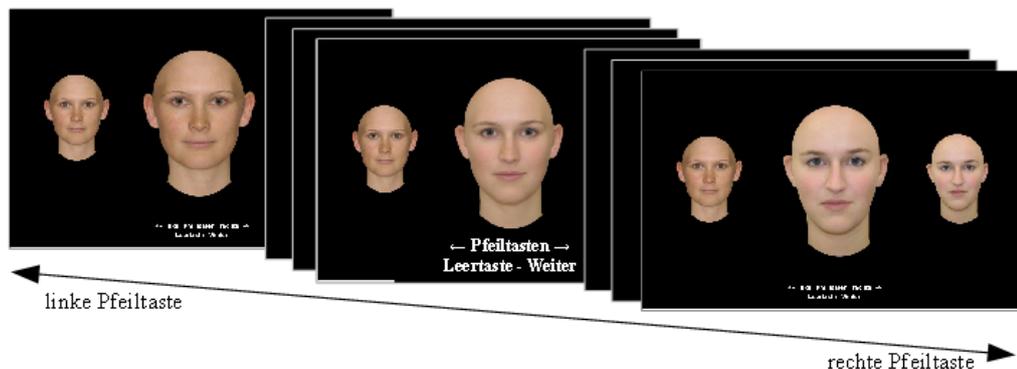


Abbildung 2.3: Drehregleraufgabe. Präsentation zweier Originalgesichter (auf der rechten und linken Seite) sowie eines Morphbildes in der Bildschirmmitte. Über die Pfeiltasten (rechts/links) wurde dieses Gesicht verändert, so dass die Morphsequenz entlang gescrollt werden konnte. Auf diese Weise sollte die subjektive Mitte der Sequenz eingestellt und anschließend über die Leertaste bestätigt werden. Die 100%-Morphs waren stets sichtbar.

In diesem Abschnitt des Versuchs wurden die Faktoren Bilderpaar und Anfangsbild variiert. Insgesamt gab es 20 verschiedene Gesichterpaare beziehungsweise -sequenzen (je 10 pro Geschlecht). Pro Sequenz kamen als Anfangsbilder jeweils sechs verschiedene Bilder auf unterschiedlichem Morphlevel in Frage: Sie waren entweder 26%, 30% oder 34% vom 50:50-Morph entfernt. Auf diese Weise sollte gewährleistet werden, dass die Probanden nicht einfach abzählen konnten, wie viele Schritte zur Mitte benötigt würden. Jedes dieser möglichen Anfangsbilder wurde zweimal präsentiert. Insgesamt ergaben sich auf diese Weise 240 Durchgänge (20 Paare x 2x6 Anfangsbilder).

2.4. PROZEDUR

Alle Bedingungen wurden den Probanden in vollständig randomisierter Reihenfolge präsentiert. Der Versuch war in acht Blöcke zu je 30 Durchgängen unterteilt; zwischen den einzelnen Blöcken konnten die Probanden Pausen beliebiger Länge einlegen. Für diesen Teil des Experiments benötigten die Versuchspersonen durchschnittlich 45 bis 50 Minuten.

2.4.3 Bewertungsaufgabe

Im letzten Teil des Experiments wurden die einzelnen Gesichter beziehungsweise die Gesichterpaare noch einmal bewertet. Dazu sollten die Probanden für jedes der einzeln präsentierten Zielgesichter auf einer 7-stufigen Skala angeben, als wie durchschnittlich (Stufe 1) beziehungsweise markant (Stufe 7) sie dieses Bild empfanden. Nach der Präsentation und Bewertung aller Einzelbilder wurden ihnen die Gesichterpaare auf gleiche Weise dargeboten, mit der Anweisung, die Paare hinsichtlich ihrer Ähnlichkeit (Stufe 1 – sehr ähnlich beziehungsweise Stufe 7 – sehr verschieden) zu bewerten. Die Eingabe erfolgte über die Nummerntasten „1“ bis „7“ des Zahlenblocks der Tastatur.



Abbildung 2.4: Bewertungsaufgabe. Präsentation eines Einzelgesichts (links), welches hinsichtlich seiner Markanz (durchschnittlich – sehr markant) bewertet werden sollte. Im zweiten Teil wurde ein Gesichterpaar (rechts) präsentiert, dessen Ähnlichkeit (sehr ähnlich – sehr verschieden) bestimmt werden sollte. Die Einschätzung erfolgte in beiden Fällen über die Nummerntasten 1 bis 7.

Jedes Bild beziehungsweise Bilderpaar wurde genau einmal präsentiert, das heißt, es gab insgesamt 10 randomisierte Durchgänge bei den Einzelbildern sowie 20 randomisierte Durchgänge bei den Bilderpaaren. Für diesen Teil

2.4. PROZEDUR

benötigten die Probanden noch einmal fünf bis sieben Minuten.

Das gesamte Experiment war als Within-Subject-Studie gestaltet, das heißt, jede Versuchsperson hatte stets alle Bedingungen zu bewältigen. Männer- und Frauengesichter wurden in allen Teilen des Experiments gemeinsam präsentiert, es gab also keine reinen Männer- beziehungsweise Frauenblöcke. Insgesamt dauerte der Versuch etwa 60 Minuten.

Kapitel 3

Ergebnisse

3.1 Lernaufgabe

Im ersten Teil des Experiments hatten die Probanden jeweils zu entscheiden, ob es sich bei dem präsentierten Bild um eines der Zielgesichter oder um ein Distraktorgesicht handelte. Dieser Abschnitt wurde erst abgeschlossen, wenn jedes Gesicht fünfmal korrekt klassifiziert worden war. Zwar benötigten die Versuchspersonen hierfür unterschiedlich viele Durchgänge, dennoch war aufgrund der obigen Restriktion gewährleistet, dass sie mit den Stimuli hinreichend vertraut waren. Daher wurde in diesem Abschnitt auf eine detailliertere Analyse verzichtet.

3.2 Analyse der Drehregleraufgabe

In der Drehregleraufgabe sollten die Versuchspersonen das Bild in der Morphsequenz auswählen, welches für sie die Mitte zwischen den beiden Gesichtern eines Paares darstellte. Hier wurden im ersten Schritt alle Daten nach den verschiedenen Gesichterpaaren sortiert und für jede Versuchsperson einzeln ausgewertet, um einen groben Eindruck von der Art der gewonnenen Daten zu bekommen. Dabei wurde untersucht, wie oft jedes der 26 Bilder der Sequenz als Endbild eingestellt wurde. Abbildung 3.1 zeigt exemplarisch die resultierenden Verteilungsfunktionen bei weiblichen Gesichterpaaren, wie sie

3.2. ANALYSE DER DREHREGLERAUFGABE

von der zehnten Versuchsperson gewählt wurden.

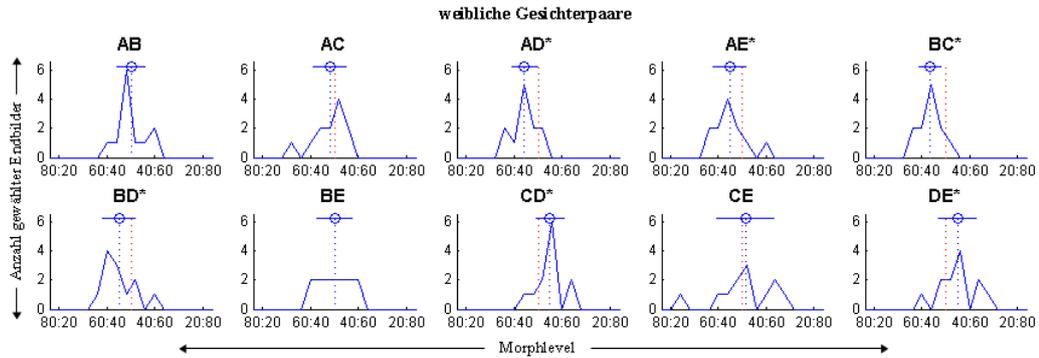


Abbildung 3.1: Verteilungsfunktionen für die weiblichen Gesichterpaare bei einer einzelnen Versuchsperson. Für jedes Gesichterpaar (Bezeichnung über den einzelnen Diagrammen) ist neben der Anzahl der eingestellten Endbilder pro Morphlevel (von insgesamt 12) zusätzlich in der oberen Hälfte der Mittelwert und die Standardabweichung dargestellt. Die senkrechte rote Linie entspricht der objektiven Mitte (50:50-Morphlevel) und die senkrechte blaue Linie der subjektiven Mitte. Fälle, in denen beide Werte signifikant voneinander abweichen, sind mit einem * markiert.

Dieser Abbildung kann man zum einen entnehmen, dass die Varianz innerhalb der einzelnen Gesichterpaare zumeist sehr gering ist (mit Ausnahme des Paares CE). Dies bedeutet, dass die Endbilder im Allgemeinen stets etwa gleich gewählt wurden. Zum anderen ergab ein Einstichproben-t-Test, dass die eingestellte subjektive Mitte in mehreren Fällen deutlich von der objektiven Mitte abwich: AD: $T(11) = -3,7422, p < 0,01$; AE: $T(11) = -2,7222, p < 0,05$; BC: $T(11) = -4,8378, p < 0,010$; BD: $T(11) = -2,5289, p < 0,05$; CD: $T(11) = 3,0446, p < 0,05$ und DE: $T(11) = 2,2094, p < 0,05$. In diesem Fall sind die Paarnamen in Abbildung 3.1 mit einem * markiert.

Um zu untersuchen, ob diese Verschiebungen systematisch auftreten, wurden für die einzelnen Gesichterpaare jeweils jene Versuchspersonen zusammengefasst, welche dieselbe Verschiebungsrichtung, also eine signifikante Abweichung vom 50:50-Morphlevel zum gleichen Bild hin, aufwiesen (siehe Abbildung 3.2 auf Seite 15).

3.2. ANALYSE DER DREHREGLERAUFGABE

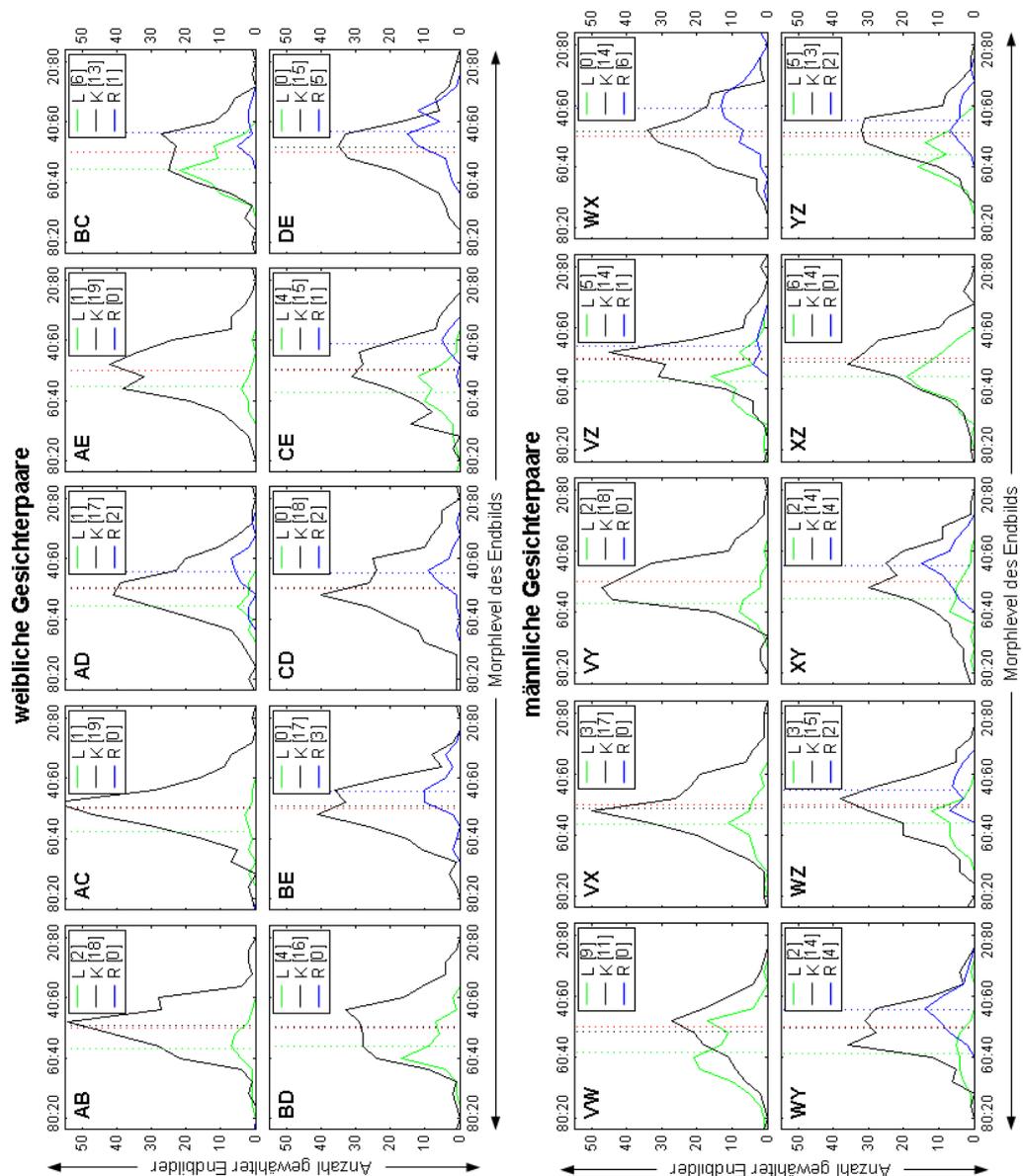


Abbildung 3.2: Verteilungsfunktionen für die weiblichen (oben) und die männlichen (unten) Gesichterpaare bei allen Versuchspersonen. Für jedes Gesichterpaar (Bezeichnung oben links in den einzelnen Diagrammen) ist die Anzahl der eingestellten Endbilder pro Morphlevel (von insgesamt 240) dargestellt. Diejenigen Versuchspersonen, welche jeweils signifikante Verschiebungen in die gleiche Richtung (L-links, K-keine, R-rechts; siehe hierzu auch die Erklärung im Text) aufwiesen, wurden zu Gruppen zusammengefasst; die jeweilige Gruppengröße ist durch den Wert in Klammern in der Legende angegeben. Die senkrechte rote Linie entspricht der objektiven Mitte (50:50-Morphlevel) und die senkrechte grüne, schwarze beziehungsweise blaue Linie der subjektiven Mitte.

3.2. ANALYSE DER DREHREGLERAUFGABE

Zu beachten ist hierbei, dass die an dieser Stelle verwendeten Bezeichnungen „links/rechts“ *nicht* einer Verschiebung zu dem auf der linken beziehungsweise rechten Seite präsentierten Gesicht entsprechen; sie dienen lediglich der anschaulicheren Beschreibung der Diagramme! Inhaltlich verbirgt sich hinter einer Verschiebung nach „links“ die Einstellungstendenz zum zuerst im Paar genannten Gesicht (beispielsweise zu A im Paar AB) beziehungsweise bei „rechts“ die zum zuletzt genannten Gesicht des Paares (also zu B im Paar AB).

Dieser Abbildung kann man zuerst einmal entnehmen, dass es über alle Versuchspersonen hinweg nur in verhältnismäßig wenigen Fällen zu signifikanten Abweichungen vom 50:50-Morphlevel kommt. Das bedeutet, dass die Probanden im Allgemeinen in der Lage sind, die tatsächliche, objektive Mitte innerhalb der Morphsequenzen zu finden. Der zumeist sehr steile Kurvenverlauf zeigt zudem, dass diese Bestimmung auch noch sehr präzise ist. Bei den Paaren AB, AC, VX und VY beispielsweise ist dies besonders gut zu beobachten.

Im Folgenden sollen nun die Fälle näher betrachtet werden, in denen es signifikante Verschiebungen gab. Hier zeigte sich, dass die subjektiv wahrgenommene Mitte zwischen den Bildern im Mittel zwischen 5% und 9% vom 50:50-Morphlevel abwich. Besonders interessant sind vor allem die Paare VW und WX: Bei diesen gab es vergleichsweise starke Verschiebungen (8,6% und 9%) weg von Gesicht W, die zudem von verhältnismäßig vielen Personen (9 und 6) eingestellt wurden.

In der folgenden Abbildung 3.3 wurde anschließend zusammengefasst, bei wie vielen Personen es Abweichungen zu welchen Gesichtern hin gab. Dabei wurde jeweils die Differenz aus den beiden Richtungen gebildet, um die Gesamtverschiebung zu bestimmen. Die Fälle, in denen subjektive und objektive Mitte übereinstimmten, wurden an dieser Stelle nicht mit betrachtet. Die Stärke eines Pfeils in der Grafik spiegelt dabei die Anzahl der Personen wider, bei denen es eine Verschiebung in Richtung der Orientierung des Pfeils gab.

3.3. BEZUG ZU MARKANZ UND ÄHNLICHKEIT DER GESICHTER

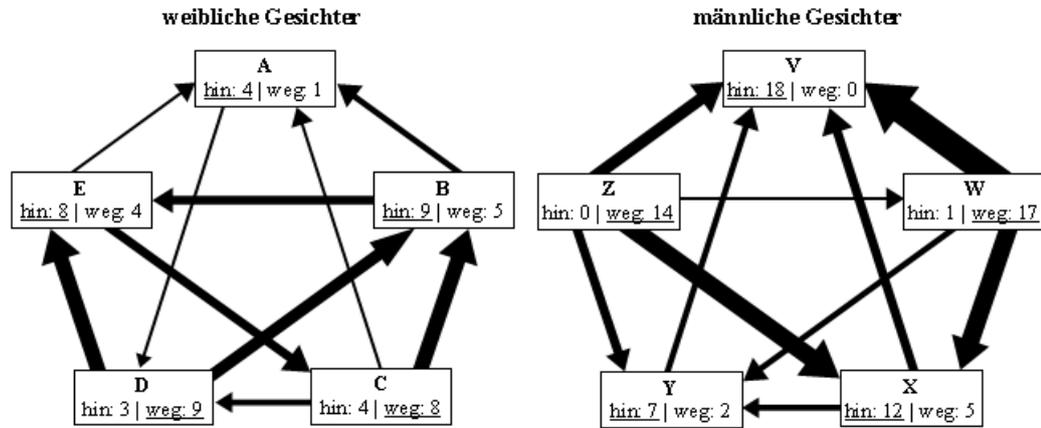


Abbildung 3.3: Verschiebungsrichtungen zwischen den Stimuli. Dargestellt ist jeweils die Gesamtverschiebung zwischen den einzelnen männlichen (rechts) beziehungsweise weiblichen (links) Gesichtern. Die Stärke eines Pfeils spiegelt dabei die Anzahl der Personen wider, bei denen es eine Verschiebung in Richtung der Orientierung des Pfeils gab. Unter den Namen ist jeweils die Gesamtzahl der Personen mit einer Verschiebung zu diesem Bild hin beziehungsweise vom ihm weg (unabhängig von den einzelnen Paaren) angegeben.

Insbesondere bei den Männern (siehe Abbildung 3.3, rechts) ist deutlich zu erkennen, dass es einzelne Gesichter (V, X, Y) gibt, zu denen die Verschiebungen stets hin tendieren, während sie bei anderen Gesichtern (W, Z) stets weg tendieren (entsprechend den Zahlen unter dem Gesichternamen). Bei den Frauengesichtern (siehe Abbildung 3.3, links) ist dieser Effekt ebenfalls zu beobachten, wenn auch nicht in so ausgeprägtem Maße.

3.3 Bezug zu Markanz und Ähnlichkeit der Gesichter

Anschließend wurde analysiert, ob diese Verschiebungsrichtungen mit der empfundenen Markanz der einzelnen Gesichter korrelierten und ob es einen Zusammenhang zwischen der wahrgenommenen Ähnlichkeit und der Varianz der Gesichterpaare gibt.

Bei der Beurteilung der Markanz gab es bei den männlichen Bildern zwei deutliche Tendenzen: Die Gesichter V und W wurden als sehr durchschnitt-

3.3. BEZUG ZU MARKANZ UND ÄHNLICHKEIT DER GESICHTER

lich bewertet (hier gab es kaum signifikanten Unterschiede), die Gesichter X, Y und Z hingegen als sehr markant (auch hier wurden wieder keine Signifikanzen beobachtet). Zwischen diesen beiden Gruppen allerdings waren die Differenzen hoch signifikant ($T[19] = -6,7836, p < 0,01$). Eine solche Tendenz (A, C, D versus B, E) ließ sich ebenfalls – wenn auch in schwächerem Maße – bei den weiblichen Gesichtern beobachten ($T[19] = -3,0403, p < 0,05$ zwischen den Gruppen). Bei diesen waren auch insgesamt die Unterschiede zwischen den einzelnen Gesichtern kleiner (siehe hierzu auch Abbildung 2 auf Seite 34 im Anhang).

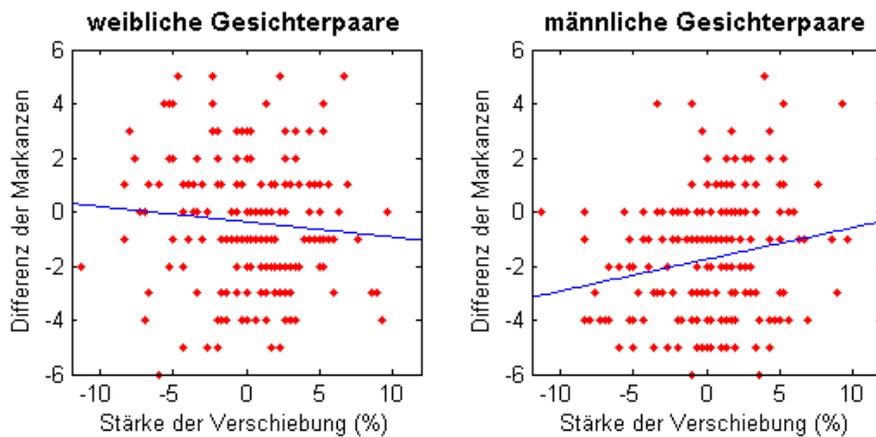


Abbildung 3.4: Korrelation zwischen Markanz und Verschiebungsstärke. Für jede Versuchsperson und jedes Gesichterpaar wurde die Differenz aus den Markanzwerten der jeweiligen Einzelgesichter gebildet und die Korrelation zur Stärke der Abweichung vom 50:50-Morphlevel bestimmt. Bei den Männern (rechts) wurde ein schwacher, aber hoch signifikanter Effekt gefunden ($r = 0,1898$); die Korrelation bei den Frauen (links) ergab keine Signifikanz ($r = -0,0898$).

Zur Bestimmung der Korrelation (siehe Abbildung 3.4) wurde nun die Differenz aus den Markanzwerten der beiden Gesichter eines Paares gebildet. Bei den männlichen Gesichtern wurde hierbei ein schwacher, aber hoch signifikanter Effekt ($r = 0,1898; p < 0,01$) gefunden. Dies bedeutet, dass eine markantere Person schon früh im Morphbild wiedererkannt wird, selbst wenn noch mehr Elemente der anderen Person enthalten sind. Bei den weiblichen Gesichtern konnte zwar keine signifikante Korrelation beobachtet wer-

3.3. BEZUG ZU MARKANZ UND ÄHNLICHKEIT DER GESICHTER

den ($r = -0,0898; p > 0,20$), dennoch ist eine schwache Tendenz erkennbar. Auch bei der Bewertung der Ähnlichkeit zeigten sich deutliche Unterschiede zwischen männlichen und weiblichen Gesichtern (siehe hierzu auch Abbildung 3 auf Seite 34 im Anhang). So ergab eine einfaktorielle Varianzanalyse, dass die Varianzen für die einzelnen Paare bei den Frauen ($MS = 2,1376$) wesentlich höher als bei den Männern ($MS = 1,7547$) waren, während die Unterschiede zwischen den Paaren vergleichsweise gering ausfielen ($MS = 17,8583$ versus $MS = 25,0356$). Zudem wurden die weiblichen Paare insgesamt als viel ähnlicher eingestuft: Sie erhielten Bewertungen zwischen 2,3 und 5,2; die Männer hingegen zwischen 2,8 und 6,0.

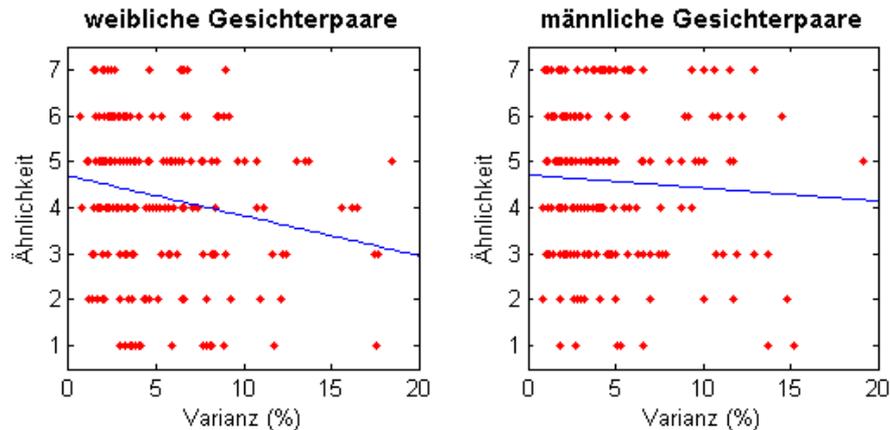


Abbildung 3.5: Korrelation zwischen Antwortvarianz und Ähnlichkeit. Für jede Versuchsperson und jedes Gesichterpaar wurde die Korrelation zwischen der Ähnlichkeit des Paares und der Varianz innerhalb der Antworten bestimmt. Bei den Frauen (links) wurde ein schwacher, aber hoch signifikanter Effekt gefunden ($r = -0,2168$); die Korrelation bei den Männern (rechts) ergab keine Signifikanz ($r = -0,0591$).

Für beide Geschlechter konnte im Folgenden ein negativer Zusammenhang zur Varianz innerhalb eines Paares gefunden werden (siehe Abbildung 3.5). In diesem Fall zeigte sich bei den Frauen ein ebenfalls schwacher, aber hoch signifikanter Effekt ($r = -0,2168; p < 0,01$), während bei den Männern nur eine sehr schwache Tendenz – aber in die gleiche Richtung – beobachtet wurde ($r = -0,0591; p > 0,40$). Dies bedeutet, dass es bei sehr ähnlichen Gesichtern schwerer ist, stets dasselbe Endbild einzustellen. Je verschiedener die

3.4. HYSTERESE-EFFEKTE

Gesichter hingegen sind, desto einfacher fällt es den Probanden, das Endbild wiederzufinden.

3.4 Hysterese-Effekte

Darüber hinaus wurde untersucht, ob das zu Beginn des Durchgangs präsentierte Morphbild einen Einfluss darauf hatte, welches Endbild später eingestellt wurde. Vorstellbar wäre hier, dass systematische Verschiebungseffekte weg vom oder hin zum Anfangsbild auftreten (so genannte Hysterese-Effekte). Die folgende Abbildung 3.6 zeigt die über alle Probanden gemittelten Werte des gewählten Endmorphs in Abhängigkeit vom Morphlevel des Startbildes.

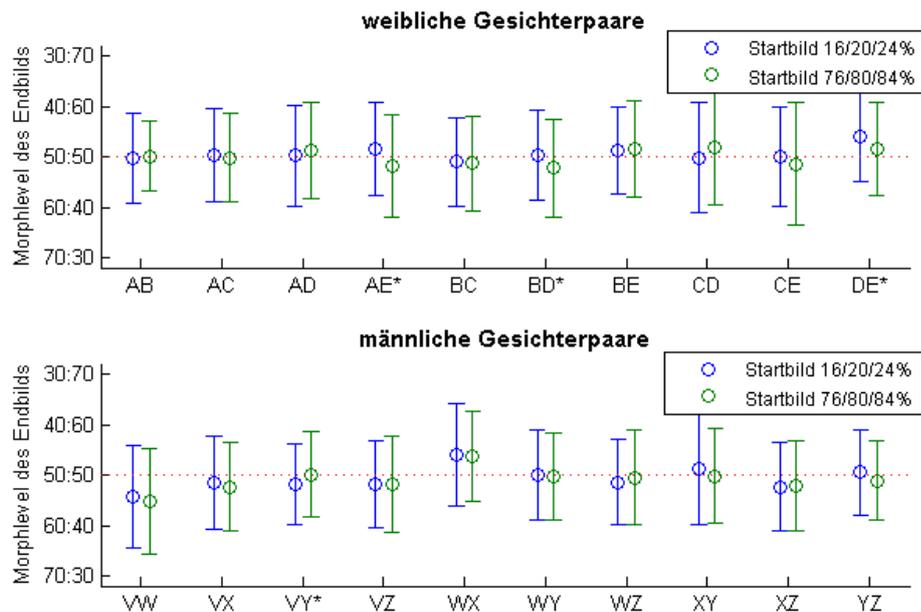


Abbildung 3.6: Hysterese-Effekte für die weiblichen (oben) und die männlichen (unten) Gesichterpaare bei allen Versuchspersonen. Dargestellt sind die bei den einzelnen Gesichterpaaren eingestellten Endbilder in Abhängigkeit vom zu Beginn des Durchgangs präsentierten Anfangsbild (angegeben ist der prozentuale Anteil des ersten Bildes im Morph). Die wenigen Fälle, in denen es signifikante Unterschiede gab, sind mit einem * markiert.

3.4. HYSTERESE-EFFEKTE

Ein Zweistichproben-t-Test ergab hier, dass das Anfangsbild nur in einzelnen Fällen zu signifikanten Unterschieden im Endbild führte: AE: $T(119) = 2,8401$, $p < 0,01$; BD: $T(119) = 2,1147$, $p < 0,05$; DE: $T(119) = 2,1649$, $p < 0,05$ und VY: $T(119) = -2,0530$, $p < 0,05$.

Diese Differenzen standen im Allgemeinen in enger Verbindung mit dem Klickverhalten der Probanden (siehe Abbildung 3.7):

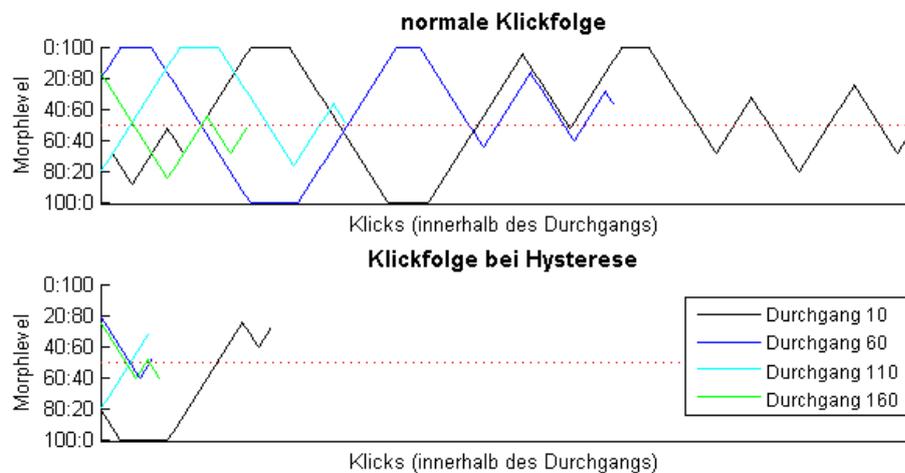


Abbildung 3.7: Beispiele für das typische Klickverhalten der Versuchspersonen. Im Allgemeinen scrollten die Probanden vielfach hin und her, bis sie sich auf ein Endbild festlegten (oben). Versuchspersonen, bei denen Hysterese-Effekte auftraten, scrollten aber zumeist nur in eine Richtung und bestätigten das Endbild dann ohne große Korrekturen (unten).

Die Versuchspersonen, bei denen eine Hysterese beobachtet wurde, scrollten häufig nur in eine Richtung und bestätigten das eingestellte Bild dann häufig ohne weitere Korrekturen (siehe Abbildung 3.7, unten). Dabei sind die beiden Tendenzen, über die Mitte hinaus zu scrollen oder bereits vor der Mitte anzuhalten, etwa gleich häufig zu beobachten. Zumeist scrollten die Probanden aber mehrfach hin und her, bis sie sich nach vielen Korrekturen auf ein Endbild festlegten (siehe Abbildung 3.7, oben).

Aufgrund dieser Ergebnisse wurde in der vorangegangenen Analyse auf die Differenzierung zwischen den Anfangsbildern verzichtet.

Kapitel 4

Diskussion

4.1 Interpretation der Ergebnisse

4.1.1 Analyse der Drehregleraufgabe

In diesem Experiment hatten die Probanden die Aufgabe, die subjektive Mitte innerhalb einer Morphsequenz zu finden. Eine Analyse der Verteilungsfunktion für das jeweils eingestellte Endbild ergab, dass dieses in den meisten Fällen sehr gut mit der tatsächlichen, objektiven Mitte (50:50-Morphlevel) übereinstimmte. Die zumeist recht kleine Varianz war zudem ein Indikator für die Präzision, mit der diese Bestimmung erfolgte.

Eine mögliche Erklärung für diesen Befund wäre, dass die Probanden die Mitte durch den systematischen Vergleich einzelner Gesichtsmerkmale ermittelten. Bei so hochkomplexen Objekten wie Gesichtern wäre dies zwar eine schwierige Aufgabe, da beim bloßen Gegenüberstellen der beiden Gesichter viele Unterschiede gar nicht oder nur schwer sichtbar wären. Dadurch, dass es aber möglich war, die Sequenz beliebig schnell entlang zu scrollen, wurden die Veränderungen beziehungsweise das einander Angleichen in fließenden Bewegungen sichtbar. Konzentrierte man sich nun auf ein stark bewegtes Gesichtsmerkmal, erleichterte dies das Finden der Mitte. So berichteten einzelne Personen, sie hätten sich bei ihrer Entscheidung beispielsweise am „Auf- und Abwandern“ der Ohren orientiert und dann einfach auf halber Strecke angehalten.

4.1. INTERPRETATION DER ERGEBNISSE

Die über mehrere Probanden hinweg beobachteten Verschiebungen der subjektiven Mitte weisen allerdings darauf hin, dass die Festlegung dieser Kategoriengrenze nicht ausschließlich auf niedrigerer Verarbeitungsebene durch einen einfachen Bildervergleich erfolgte. Hier scheinen durchaus auch höhere kognitive Prozesse eine Rolle zu spielen: Gibt es eine Verschiebung von einem bestimmten Gesicht weg, so bedeutet das, dass diese Person schon wiedererkannt wird, selbst wenn noch mehr Elemente der anderen Person im Morphbild enthalten sind. Sie dominiert quasi die Morphsequenz.

Interessant ist nun vor allem, dass diese Dominanz auch über eine Morphsequenz hinaus zu gehen scheint. Wie der Abbildung 3.3 zu entnehmen ist, gibt es insbesondere bei den Männern einzelne Gesichter, zu denen die Verschiebungen immer hin tendieren, während sie bei anderen stets weg tendieren. Auch bei den weiblichen Gesichtern ist dieser Effekt tendenziell zu erkennen. Das heißt also: Ein Gesicht, das in der einen Sequenz dominant ist, ist es wahrscheinlich auch in einer anderen. Es stellt quasi einen salienteren Reiz dar, was wiederum Einfluss auf die wahrgenommene Kategoriengrenze hat.

4.1.2 Bezug zu Markanz und Ähnlichkeit der Gesichter

Um zu untersuchen, welche Faktoren diese Salienz nun beeinflussen, wurden in dem vorliegenden Experiment die Ähnlichkeit der Gesichterpaare sowie die Markanz des einzelnen Gesichts näher untersucht.

Die subjektive Mitte einer Morphsequenz entspricht dem Punkt, an dem keines der beiden Originalgesichter mehr eindeutig im Morphbild identifiziert werden beziehungsweise an dem man die Merkmale beider Gesichter mit gleicher Stärke erkennen kann. Nun konnte in verschiedenen Studien gezeigt werden, dass markante Gesichter nicht nur schneller aus dem Gedächtnis abgerufen werden können (Light, Kayra-Stuart & Hollander, 1979), sondern zudem auch schneller identifiziert werden als prototypische Gesichter (Valentine, Bruce et al., 1986). In einem Morphbild, das Elemente zweier Gesichter enthält, könnte man daher erwarten, dass die Merkmale eines charakteristischen Gesichts ebenfalls eher wiedererkannt werden, was zu einer Verschiebung der Kategoriengrenze führen würde. Dies spiegelte sich in der

Tat auch in einer zwar schwachen, aber dennoch hoch signifikanten Korrelation zwischen Markanz und beobachteter Verschiebungsstärke wider.

Hinsichtlich der Ähnlichkeit hätte man erwartet, dass dieser Faktor vor allem die Varianz der subjektiven Mitte beeinflussen würde: Bei sehr verschiedenen Gesichtern sollte die empfundene Unterschiedlichkeit an der Kategoriengrenze stärker sein als bei sehr ähnlichen Gesichtern. Dadurch würde es einem leichter fallen, diese Grenze auch in späteren Durchgängen wiederzufinden, was sich in einer kleineren Varianz widerspiegeln würde. Diesbezüglich wurde insbesondere bei den Frauen eine signifikante Korrelation beobachtet, was die obige Annahme bestätigt.

Beide Faktoren – Markanz und Ähnlichkeit – scheinen also die Lokation und die Präzision der Kategoriengrenze bis zu einem gewissen Grade zu beeinflussen.

4.1.3 Hysterese-Effekte

Darüber hinaus wurden mögliche Hysterese-Effekte untersucht, wie sie beispielsweise von Kircher et al. (2001) beobachtet wurden. Sie hatten in ihrer Studie herausgefunden, dass es bei einer seriellen Präsentation der Stimuli während der Kategorisierungsaufgabe zu systematischen Verschiebungen der Kategoriengrenze in Richtung der zu Beginn präsentierten Person gab. Eine mögliche Ursache für diesen Effekt sahen die Autoren in einer Art Antwortbias, wodurch die Probanden die Kategoriengrenze gewissermaßen „antizipieren“ konnten (Kircher et al., 2001).

Die serielle Präsentation lässt sich in gewisser Weise mit dem in diesem Experiment verwendeten „Drehregler“ vergleichen – auch hier wurden die Bilder stets in aufeinander folgender Reihenfolge dargeboten. Ein wesentlicher Unterschied besteht allerdings darin, dass die Probanden über die Pfeiltasten die Präsentationsreihenfolge und -zeit selbst bestimmen konnten. Dies würde aber eine Antizipation der Kategoriengrenze unmöglich und eventuelle Hysterese-Effekte zunichte machen.

Genau dies spiegelt sich auch in den Ergebnissen wider. Einige Versuchspersonen scrollten stets nur in eine Richtung und bestätigten das eingestellte

Bild dann meist ohne weitere Korrekturen (siehe Abbildung 3.7, unten, in Kapitel 3.6). Eben diese Probanden sind es, bei denen Hysterese-Effekte beobachtet werden konnten. Allerdings gab es nicht nur Verschiebungen zum Anfangsbild hin (wie bei Kircher et al., 2001), sondern ebenso häufig auch Verschiebungen von diesem weg. Zumeist aber scrollten die Probanden mehrfach hin und her, bis sie sich auf ein Endbild festlegten (siehe Abbildung 3.6, oben). Dadurch relativierte sich der Einfluss des Anfangsbildes wieder. Die wenigen beobachteten Hysterese-Effekte ließen sich im Allgemeinen also nicht durch die Verschiedenheit des Stimulusmaterials oder irgendwelche kognitiven Prozesse erklären, sondern sie wurden vor allem durch das individuelle Klickverhalten der Probanden hervorgerufen. Was die Versuchspersonen letztendlich zu ihrer jeweiligen Vorgehensweise bewog, ist nicht bekannt.

4.1.4 Lösungsstrategien und Lerneffekte

Eine weitere Beobachtung, die aus Abbildung 3.7 hervorgeht, ist, dass die Probanden über die Durchgänge hinweg immer weniger Klicks durchführten, bis sie sich festlegten. Im ersten Block drückten sie im Durchschnitt etwa 67 Mal die rechte oder linke Pfeiltaste, in den letzten drei Blöcken hingegen nur noch etwa halb so oft (man bedenke, dass zum einmaligen Durchscrollen der gesamten Sequenz nur 25 Klicks notwendig sind!). Dies könnte sowohl auf Müdigkeits- als auch auf Lerneffekte zurückzuführen sein. Wahrscheinlicher sind hier allerdings Lerneffekte, da die Klickrate anfangs sehr schnell zurückgeht und sich bereits ab dem zweiten oder dritten Block auf einem relativ konstanten Level einpendelt. Dieser Effekt spiegelt sich auch in der pro Durchgang benötigten Zeit wider (siehe hierzu auch Abbildung 4 auf Seite 35 im Anhang).

Eine mögliche Strategie, die tatsächliche Mitte relativ präzise zu finden, wäre, die Zeit zu „erlernen“, die zum kompletten Durchscrollen der Sequenz benötigt wird, und anschließend einfach nach der Hälfte dieser Zeit anzuhalten. Eine Anwendung dieser Strategie kann in Einzelfällen zwar nicht ausgeschlossen werden, ist aber aus mehreren Gründen sehr unwahrscheinlich: Zum einen kann das Einpendeln bei einem Endbild, also das mehrfache nach rechts

und links Korrigieren, häufig noch selbst bei recht späten Durchgängen beobachtet werden (siehe Abbildung 3.7). Ein solcher Klickverlauf wäre bei der betrachteten Strategie aber eher nicht zu erwarten. Außerdem wären dann Hysterese-Effekte in viel ausgeprägterem Maße zu erwarten, die jedoch, wie schon erwähnt, kaum zu beobachten waren. Entscheidendster Punkt ist aber, dass sich in diesem Fall keinerlei Verschiebungseffekte einstellen dürften beziehungsweise die Ergebnisse über die verschiedenen Gesichterpaare hinweg ähnlich sein müssten, da diese Strategie allein die Zeit, nicht aber das Stimulusmaterial, berücksichtigt. Dies war aber de facto nicht der Fall (siehe auch Abbildung 3.2). Die Anwendung einer solchen materialunabhängigen Strategie kann also als sehr unwahrscheinlich angenommen werden.

4.2 Kritikpunkte

4.2.1 Inhaltliche Kritikpunkte

Aus den gewonnenen Ergebnissen wurde die Schlussfolgerung gezogen, dass sowohl die Markanz als auch die Ähnlichkeit Einfluss auf die Bestimmung der Kategoriengrenze haben (siehe Abschnitt 4.1.2). Dennoch waren die beobachteten Korrelationen für beide Faktoren eher schwach.

Dies ließe sich beispielsweise damit erklären, dass neben der Markanz auch noch viele weitere Bedingungen, beispielsweise Sympathie oder Attraktivität, die Salienz eines Gesichts beeinflussen. So gaben einige Probanden an, dass sie sich bei der Einstellung des Endbildes unter anderem davon leiten ließen, wie sympathisch sie die präsentierten Personen fanden. Dieser Faktor wurde aber nicht weiter systematisch untersucht. Darüber hinaus ist die Markanz eines Gesichts schon an sich ein kompliziertes, mehrdimensionales Konzept, welches beispielsweise durch Gesichtsform und -farbe, wahrgenommenes Alter, die Relationen einzelner Merkmale und viele weitere Faktoren untereinander beeinflusst wird (Bruce, Burton & Dench, 1994; Bruce, Burton & Hancock, 1995). Auch die Vertrautheit der Stimuli – insbesondere auch im Zusammenspiel mit der Markanz – spielt eine wichtige Rolle. So beobachteten beispielsweise Beale und Keil (1995) nur dann Kategorisierungseffekte,

4.2. KRITIKPUNKTE

wenn sie hochvertraute, berühmte Persönlichkeiten (statt eher unbekanntem Personen) präsentierten (Beale & Keil, 1995). Die hier verwendeten Stimuli waren den Probanden zwar alle gleichermaßen vertraut, aber es handelte sich bei keinem um übermäßig bekannte oder gar berühmte Persönlichkeiten. Daher konnte dieser Faktor nicht weiter analysiert werden.

Auch die Ähnlichkeit zweier Gesichter ist ein sehr allgemeines und diffuses Konzept. Zum einen muss zwischen subjektiver (empfundener) und objektiver (tatsächlicher) Ähnlichkeit unterschieden werden. Zwar gibt es verschiedene Theorien und Modelle zu deren Bestimmung (Reed & Friedman, 1973; Tversky, 1977), diese sind häufig aber sehr abstrahiert formuliert und berücksichtigen nur selten auch subtilere Elemente oder Texturen. Und die subjektive Ähnlichkeit lässt sich (genau wie die Markanz) nur schwer kontrollieren, da zumeist unklar ist, welche Faktoren (beispielsweise Form oder Textur des Gesichts oder bestimmter Gesichtspartien) diese Bewertung beeinflussen. Eine anschließende Nachbefragung der Probanden hätten diesbezüglich vermutlich weiter Aufschluss geben können. Dies hätte sich aber aufgrund der Vielzahl an möglichen Faktoren als sehr schwierig erwiesen und den Rahmen dieser Arbeit bei Weitem gesprengt.

Alles in allem kann man also sagen, dass ein gewisser Einfluss der beiden untersuchten Faktoren zwar durchaus nachgewiesen werden konnte, diese aber – unter anderem aufgrund ihrer Komplexität – nicht ausreichend kontrolliert wurden, um tiefer gehende Schlussfolgerungen zu ziehen. Auch sollte in weiteren Studien überprüft werden, inwieweit die sich die vermutete „globale Dominanz“, die bei den hier verwendeten Gesichtern gefunden wurde, verallgemeinern lässt.

4.2.2 Methodische Kritikpunkte

Im vorangegangenen Abschnitt 4.1.2 wurde argumentiert, dass die Markanz einen gewissen Einfluss auf die Salienz eines Gesichts hat, was sich wiederum auf die Lokation der Kategoriengrenze auswirkt. Dennoch weisen die Ergebnisse der weiblichen Gesichter nur tendenziell in diese Richtung; die Korrelation von Markanz und Verschiebungsstärke war hier nur marginal.

4.2. KRITIKPUNKTE

Dies lässt sich teilweise durch das verwendete Stimulusmaterial erklären. Bei der Zusammenstellung der Bilder wurde vom Versuchsleiter eine Vorauswahl aus bereits vorhandenen Gesichtern getroffen, was zu einem Bias hinsichtlich der Markanz führte. Während bei den Männern das Spektrum durchschnittlich–markant weitestgehend vollständig abgedeckt wurde, bewegten sich die Frauen alle eher in der Mitte des Spektrums; besonders durchschnittliche und sehr markante Gesichter waren nicht dabei. Dies ist zwar nur teilweise überraschend, da Frauengesichter insgesamt seltener als markant empfunden werden; sinnvoller wäre dennoch ein separater Vortest zur Markanz mit mehr Gesichtern und anderen Versuchspersonen gewesen, um für beide Geschlechter ein gleichmäßigeres Spektrum zu gewährleisten und diesen Faktor besser kontrollieren zu können.

Zwar konnte durch die zusätzliche Analyse der Zeit- und Klickdaten die Verwendung einer stimulusunabhängigen, rein zeitbasierten Strategie weitestgehend ausgeschlossen werden (siehe Abschnitt 4.1.4). Es ist natürlich trotzdem möglich, dass einige Probanden eine andere, individuelle Strategie angewandt haben. So könnte es beispielsweise sein, dass sich manche Versuchspersonen stets auf bestimmte Gesichtspartien (wie Mund, Ohren oder Nase) konzentrierten, während andere sich in ihrer Entscheidung eher vom Gesamteindruck oder anderen hier nicht erwähnten Aspekten leiten ließen.

Ein wesentlicher Diskussionspunkt ist natürlich auch die Wahl der Messmethode. Wie schon in Kapitel 1 erwähnt, finden in den meisten Studien zu dieser Thematik Diskriminations- oder Kategorisierungsaufgaben Anwendung. Die Kategoriengrenze – definiert als Punkt subjektiver Gleichheit – wird in der Regel indirekt über psychometrische Funktionen bestimmt. Im Gegensatz dazu erhielten die Probanden in diesem Experiment die Anweisung, die subjektive Mitte direkt einzustellen. Dies hat den Vorteil, dass die Aufgabenstellung in diesem Experiment einen weniger ausgeprägten Anforderungscharakter hat. Da die Probanden in der Kategorisierungsaufgabe explizit angewiesen wurden, „in Schubladen zu denken“, ist eine kategoriale Wahrnehmung der Bilder kaum überraschend – erstaunlich ist hierbei nur die Schärfe der Kategoriengrenze. Aus diesem Grund wurde auch kein klassisches Staircase-Verfahren angewendet. (Dies soll die gewonnenen Erkenntnis-

se nicht in Frage stellen oder relativieren; sie wurden ja insbesondere durch die Ergebnisse der Diskriminationsaufgaben mehrfach bestätigt!) Darin liegt allerdings auch ein entscheidender Nachteil der gewählten Methode: Im Grunde genommen können aus den hier gewonnenen Ergebnissen auch nur indirekt Rückschlüsse auf eine kategoriale Wahrnehmung gezogen werden. Wie bereits im vorangegangenen Abschnitt 4.1.1 beschrieben könnte die Aufgabe auch allein durch das Vergleichen der beiden Originalgesichter gelöst worden sein – dies würde nicht zwangsläufig eine Kategorisierung erfordern. Aufgrund der zahlreichen vorangegangenen Studien kann allerdings angenommen werden, dass auch hier die Stimuli kategorial wahrgenommen wurden. Daher ist diese Methode sehr gut geeignet, die Bestimmung der Kategoriengrenze näher zu untersuchen. Dennoch sollte sie noch einmal in einem Folgeexperiment validiert werden, indem sie der Kategorisierungsaufgabe in einem Experiment gegenüber gestellt und anschließend die jeweiligen Ergebnisse (insbesondere beispielsweise das Morphlevel, bei dem die Kategoriengrenze festgelegt wird) verglichen würden.

4.3 Zusammenfassung und Ausblick

In dieser Studie wurde untersucht, wie beziehungsweise wo die Probanden die Kategoriengrenze zwischen zwei Gesichtern innerhalb einer Morphsequenz festlegten. Dazu konnten sie mittels einer Art Drehregler die Sequenz beliebig entlang scrollen, bis sie sich für ein Bild entschieden hatten. Diese Methode steht im Gegensatz zu der klassischen Kategorisierungsaufgabe, in der diese Grenze nur indirekt bestimmt wird.

Es konnte gezeigt werden, dass die Versuchspersonen im Allgemeinen ziemlich präzise die objektive Mitte (50:50-Morphlevel) einstellten. Dabei orientierten sich viele Probanden an den beim Durchscrollen auftretenden Bewegungseffekten, die durch die Transformationen innerhalb der Gesichter entstanden. Dies legt die Vermutung nahe, dass die Lösung der Aufgabe unter anderem auf einer eher niedrigen Verarbeitungsebene durch den Vergleich der Bilder erfolgte. Daneben konnten allerdings auch systematische Verschiebungen der Kategoriengrenzen beobachtet werden. Besonders interessant war hier-

4.3. ZUSAMMENFASSUNG UND AUSBLICK

bei, dass sich die daraus abgeleitete Dominanz eines Gesichts auch in analogen Verschiebungsrichtungen bei anderen Morphsequenzen widerspiegelte. Dies deutet darauf hin, dass durchaus auch höhere kognitive Prozesse – wie die Bewertung eines Gesichts – eine wichtige Rolle spielen. Diese Salienz wird durch multiple Faktoren bedingt, deren Zusammenspiel allerdings noch nicht vollständig untersucht wurde. Einen möglichen Faktor stellt dabei die Markanz eines Gesichts dar, für die in diesem Experiment eine gewisse Korrelation ermittelt werden konnte. Ebenso konnte ansatzweise gezeigt werden, dass eine hohe Ähnlichkeit die Kategorisierung erschwerte.

Die Wahrnehmung von Gesichtern ist ein faszinierender Prozess, der nach wie vor nicht in seiner ganzen Komplexität verstanden wird. Durch die Generierung künstlicher Kontinua zwischen verschiedenen Gesichtern konnte nachgewiesen werden, dass diese unter bestimmten Bedingungen kategorial wahrgenommen werden (Beale & Keil, 1995; Kircher et al., 2001; Bühlhoff & Newell, 2004). In verschiedenen prototyp- oder exemplarbasierten Modellen wurde außerdem versucht, die mentale Repräsentation von Gesichtern zu erklären (Valentine, 1991b, 1991a). Darin werden Gesichter als Vektoren eines Euklidischen Gesichtsraums gespeichert, dessen Dimensionen über verschiedene Gesichtsmerkmale definiert werden. Je nach Typikalität beziehungsweise Markanz befinden sich die Gesichter eher zentrumsnah oder weiter entfernt. In dieser Studie wurde ein erster Versuch unternommen, diese beiden Bereiche zu verbinden. So konnte gezeigt werden, dass die unterschiedlichen Positionen im Gesichtsraum Einfluss auf die Kategorisierung haben. Dennoch sind noch viele Fragen offen: Welche Dimensionen bestimmen letztendlich den Gesichtsraum? Wie lässt sich daraus die Salienz eines Gesichts ermitteln? Vor allem aber: Wie werden künstliche Kontinua in diesem Raum repräsentiert? Verstehen wir diese Fragen, könnten wir nicht nur weitere Vorhersagen für die Wahrnehmung von Gesichtern ableiten, sondern wir würden auch den grundlegenden Mechanismus der Kategorisierung besser verstehen.

Literaturverzeichnis

- Beale, J. & Keil, F. (1995). Categorical effects in the perception of faces. *Cognition*, 57 (3), 217–239.
- Bornstein, M. & Korda, N. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological research*, 46 (3), 207–222.
- Bruce, V., Burton, A. & Hancock, P. (1995). Missing dimensions of distinctiveness. *T. Valentine, Cognitive and computational aspects of face recognition*, 138–158.
- Bruce, V., Burton, M. & Dench, N. (1994). What's distinctive about a distinctive face? *The Quarterly Journal of Experimental Psychology*, 47 (1), 119–141.
- Bülthoff, I. & Newell, F. (2004). Categorical perception of sex occurs in familiar but not unfamiliar faces. *Visual cognition*, 11 (7), 823–855.
- DeValois, R. & DeValois, K. (1975). Neural coding of color. In *Handbook of perception*.
- Diamond, R. & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology*, 23 (1), 1–22.
- Harnad, S. (1990). Introduction: psychological and cognitive aspects of categorical perception: a critical overview. In S. Harnad (Hrsg.), *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.
- Kircher, T., Senior, C., Phillips, M., Rabe-Hesketh, S., Benson, P., Bullmore, E., ... David, A. (2001). Recognizing one's own face. *Cognition*, 78 (1), B1–B15.

- Liberman, A., Cooper, F., Shankweiler, D. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, 74 (6), 431.
- Liberman, A., Harris, K., Hoffman, H. & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology; Journal of Experimental Psychology*, 54 (5), 358.
- Light, L., Kayra-Stuart, F. & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory; Journal of Experimental Psychology: Human Learning and Memory*, 5 (3), 212.
- Reed, S. & Friedman, M. (1973). Perceptual vs conceptual categorization. *Memory & Cognition*, 1 (2), 157–163.
- Rosen, S. & Howell, P. (1987). Auditory, articulatory, and learning explanatory explanations of categorical perception in speech. In S. Harnad (Hrsg.), *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.
- Tversky, A. (1977). Features of similarity. *Psychological review*, 84 (4), 327.
- Valentine, T. (1991a). Representation and process in face recognition. In *Pattern recognition by man and machine*.
- Valentine, T. (1991b). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, 43 (2), 161–204.
- Valentine, T., Bruce, V. et al. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception*, 15 (5), 525–535.
- Wolberg, G. (1990). *Digital image warping* (Bd. 10662). IEEE computer society press California.

Weitere Abbildungen

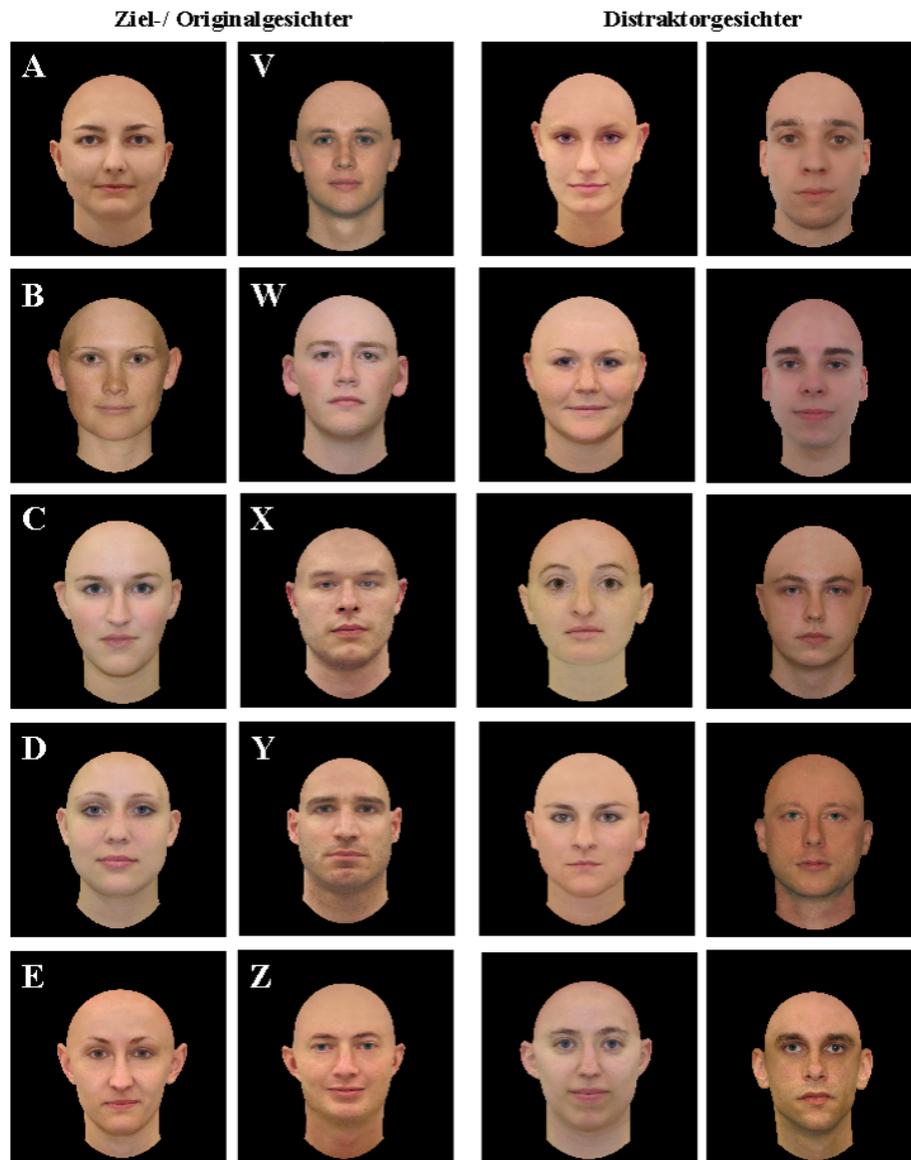


Abbildung 1: Verwendete Stimuli. Kopf und Hals der Gesichter wurden freigestellt, die Haare entfernt und der Hintergrund mit Schwarz gefüllt. Die weiblichen Zielgesichter (links) wurden mit A-E codiert, die Männer mit W-Z. Da die Distraktorgesichter (rechts) nur im ersten Abschnitt des Experiments verwendet wurden, wurde auf eine Codierung verzichtet.

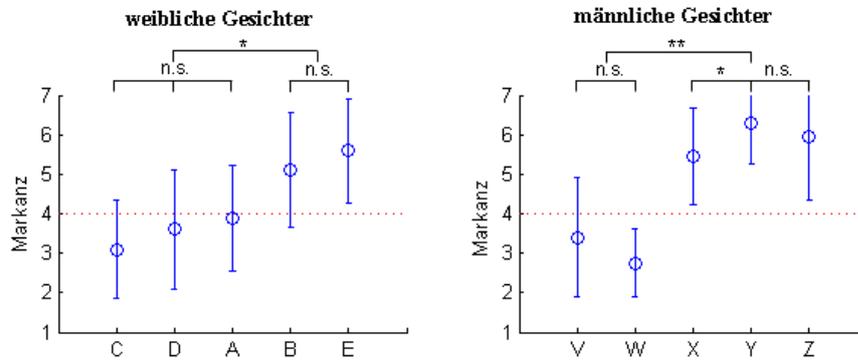


Abbildung 2: Bewertung der Einzelgesichter hinsichtlich ihrer Markanz. Niedrige Bewertungen bedeuten, dass ein Gesicht eher durchschnittlich ist, bei hohen Werten ist es sehr markant. Die Unterschiede zwischen den einzelnen Frauen (links) sind geringer als bei den Männern (rechts). Letztere werden insgesamt häufiger als markant angesehen.

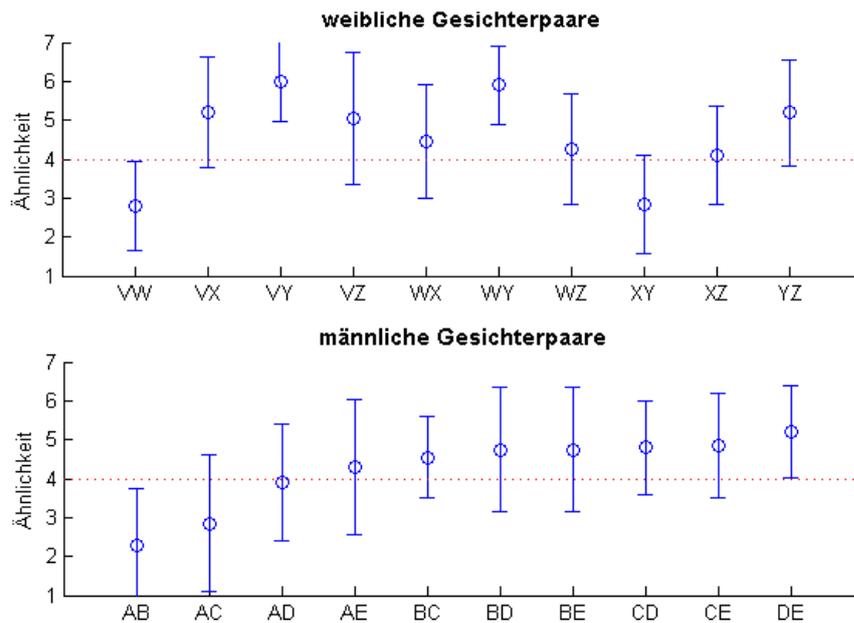


Abbildung 3: Bewertung der Gesichterpaare hinsichtlich ihrer Ähnlichkeit. Niedrige Bewertungen bedeuten, dass sich beide Gesichter sehr ähnlich sind, bei hohen Werten sind sie sehr verschieden. Bei den Frauen (oben) waren die Varianzen innerhalb der einzelnen Paares wesentlich höher als bei den Männern (unten). Zudem gab es weniger ausgeprägte Unterschiede zwischen den weiblichen Paaren. Aus Gründen der Übersichtlichkeit wurde darauf verzichtet, Signifikanzen zwischen den einzelnen Paaren einzutragen.

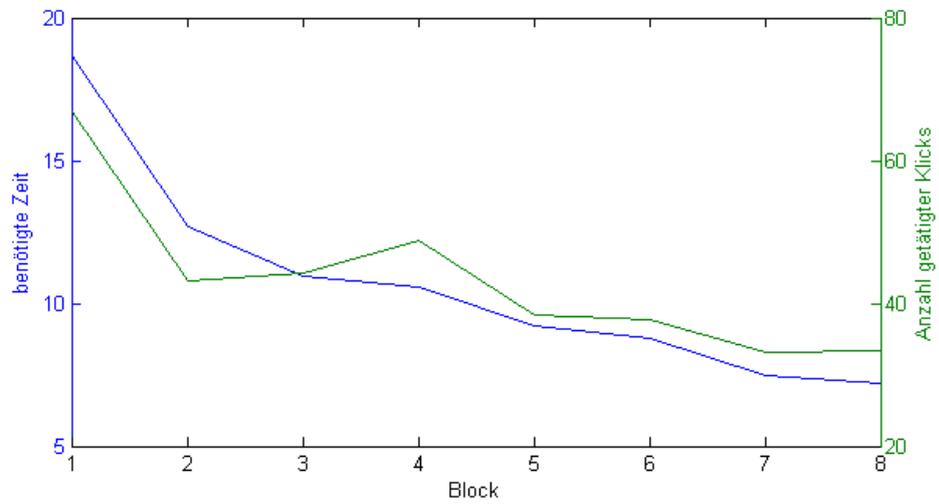


Abbildung 4: Klickverhalten und benötigte Zeit über die Versuchsblöcke hinweg. Zu Beginn wird oft hin- und hergescrollt, wofür viel Zeit benötigt wird. Aber schon nach dem ersten Block bleiben beide Faktoren auf etwa konstantem Level.