Universitätsklinikum Ulm
Klinik für Psychosomatische Medizin und Psychotherapie
Ärztlicher Direktor: Prof. Dr. H. Gündel
Sektion Medizinische Psychologie
Leitung: Prof. Dr. H. C. Traue

Die standardisierte Erfassung der Emotionserkennungsfähigkeit aus Gesichtern – Erstellung und Anwendung der „Pictures of Facial Affect-Ulm“

Dissertation
zur Erlangung des Doktorgrades der Humanbiologie
der Medizinischen Fakultät der
Universität Ulm

Vorgelegt von
Dipl.-Psych. Kerstin Limbrecht, geboren in Stuttgart

Ulm, 2012

Amtierender Dekan: Prof. Dr. Thomas Wirth

1. Berichterstatter: Prof. Dr. Harald C. Traue

2. Berichterstatter: Prof. Dr. Ferdinand Keller

Tag der Promotion: 20.07.2012

Für HE, UL, ML, JML, KD

INHALTSVERZEICHNIS

	INHALTSVERZEICHNIS	**i**
	ABKÜRZUNGSVERZEICHNIS	**iii**
1	**EINLEITUNG**	**1**
2	**THEORETISCHER HINTERGRUND UND ABLEITUNG DER FORSCHUNGSZIELE**	**3**
2.1	**Theoretische Grundlagen der Emotionspsychologie**	**3**
2.2	**Neurobiologische Grundlagen der Emotionserkennung**	**10**
2.3	**Empirische Arbeiten aus dem Bereich Emotionserkennung**	13
2.3.1	Interindividuelle Unterschiede der Emotionserkennung	13
2.3.2	Emotionserkennung und klinische Phänomene	18
2.3.3	Entwicklung von Stimulusmaterial zur Überprüfung der Emotions-erkennungsfähigkeit	20
2.4	**Die Überprüfung der psychometrischen Qualität eines Testverfahrens**	**23**
2.5	**Ziele der vorliegenden Arbeit**	**25**
3	**MATERIAL UND METHODIK**	**27**
3.1	**Das Facial Action Coding System**	**27**
3.2	**Der FEEL-Test**	**30**
3.3	**Erstellung Bildmaterial PFA-U**	**34**
3.3.1	Rekrutierung der Darsteller	34
3.3.2	Training der Darsteller	35
3.3.3	Erstellung der Aufnahmen und Weiterverarbeitung des Bildmaterials	37
3.3.4	Datenauswertung und statistische Analyse	40
4	**TESTTHEORETISCHE ÜBERPRÜFUNG DER FRONTALBILDER**	**41**
4.1	**Studie 1: Überprüfung der Emotionserkennungsraten zur ersten Selektion**	**42**
4.2	**Studie 2: Überprüfung der Kriteriumsvalidität**	**50**
4.3	**Studie 3: Beurteilung der Intensität und Natürlichkeit des Emotionsausdrucks und der Attraktivität der fotografierten Person**	**58**

4.4	**Studie 4: Auswahl des endgültigen Bildersatzes und Normierung der Erkennungsleistung**	**70**
4.4.1	Überprüfung von Intensität, Natürlichkeit und Attraktivität des Stimulusmaterials am endgültigen Bilderset	85
4.4.2	Bestimmung der Kriteriumsvalidität am endgültigen Bilderset	91
4.4.3	Zusammenfassende Diskussion der Ergebnisse aus der Normierungsstudie	92
4.5	**Studie 5: Vergleich der Erkennungsleistung über zwei Zeitpunkte (Retest-Reliabilität)**	**93**
5	**EMOTIONSERKENNUNG AUS UNTERSCHIEDLICHEN PERSPEKTIVEN**	**98**
5.1	**Studie 6: Variation der Stimuli über Emotion und Perspektive**	**101**
5.2	**Studie 7: Konstanthaltung der Stimuli über die Perspektive**	**111**
5.3	**Diskussion der Emotionserkennung über verschiedene Perspektiven**	**119**
6	**ALLGEMEINE DISKUSSION**	**121**
7	**AUSBLICK**	**126**
8	**ZUSAMMENFASSUNG**	**129**
	LITERATURVERZEICHNIS	**131**
	TABELLENVERZEICHNIS	**143**
	ABBILDUNGSVERZEICHNIS	**145**
	ANHANG	**146**
	DANKSAGUNG	**147**
	Lebenslauf	**148**
	Publikationen und Posterbeiträge	**149**

ABKÜRZUNGSVERZEICHNIS

Abkürzung	**Beschreibung**
adk	Akademie der Künste
AU/AUs	Action Unit/Action Units
CAFE	California Facial Expressions dataset
CAS-PEAL	Chinese Academy of Sciences – pose, expression, accessory, lightning
d	Durchschnittliche Effektstärke
EmFACS	Emotional Facial Action Coding System
ERQ	Emotion Regulation Questionnaire
FACS	Facial Action Coding System
FEEL	Facially Expressed Emotion Labeling
FEEL-R	Revision des FEEL
FEEL-Rpersp	FEEL-Test zur Emotionserkennungsfähigkeit aus fünf Perspektiven
JACFEE/JACNeuF	Japanese and Caucasian Facial Expressions of Emotion and Neutral Faces
KDEF	Karolinska Directed Emotional Faces
KFDB	Korean Faces Database
M	Mittelwert
ms	Milisekunden
NEO-FFI	Neuroticism, Extraversion, Openness-Five Factor Invetory
PFA-U	Pictures of Facial Affect Ulm
PoFA	Pictures of Facial Affect
RaFD	Radboud Face Database
s	Sekunden
SA	Standardabweichung
TeiQue	Trait emotional intelligence questionnaire
THYMOS	tachistoskopische Analyse emotionaler Sensitivität
TIPI	Ten Item Personality Inventory
UCDSEE	University of California, Davis Set of Emotion Expressions
VERT	Vienna Emotion Recognition Task
VIEMER	Vienna Memory of Emotion Recognition Task

1. EINLEITUNG

„Das menschliche Gehirn enthält rund zehn Milliarden Neurone, die auf ungeheuer komplizierte Weise untereinander verdrahtet sind. Es ist schon erstaunlich und verblüffend genug, welche Dinge die elektrischen Funken in diesen Zellen und die chemischen Wechselwirkungen zwischen ihnen bewirken, doch die Erzeugung unserer Emotionen gehört zu ihren erstaunlichsten und verblüffendsten Meisterleistungen.“ (LeDoux, 2003, S. 25)

Das Erleben von Emotionen ist für jeden Menschen vollkommen alltäglich, Emotionen prägen die Auseinandersetzung des Menschen mit sich selbst und seiner Umwelt. Emotionen werden in zwischenmenschlichen Begegnungen nicht nur durch die reichhaltige Sprachsemantik oder Gestik, sondern auch unmittelbar durch emotional-expressives Verhalten der Mimik kommuniziert. Die Fähigkeit, Emotionen auszudrücken, zu erkennen und zu interpretieren ist für die soziale Interaktion bedeutsam und ein wichtiges Element in den Konstrukten wie Empathie, Emotionsregulation und Emotionaler Intelligenz (Ortony & Turner, 1990). In einer Interaktion wird die Mimik oder ein anderes Signale wie die Gestik oder die Sprache im Erkennungsprozess nach innen gespiegelt und verarbeitet (Buck, 1984; Ekman, 1992; Ekman, 1993). Partner einer zwischenmenschlichen Interaktion analysieren gegenseitig die emotionalen Hinweisreize, um über den emotionalen Zustand des jeweils anderen informiert zu sein. Gleichzeitig werden die Informationen auch intraindividuell verarbeitet und das Individuum über seine eigenen emotionalen Reaktionen informiert („facial fedback“-Hypothese). Das mimisch-expressive Verhalten ist insofern Hinweis für einen bestimmten emotionalen Zustand im Individuum sowie zwischen den Individuen. Es wird angenommen, dass gerade die Mimik entstanden ist, um eine Vermittlung der gerade in der Person vorherrschenden Emotion rasch und sicher zu ermöglichen. Das bedeutet: Emotionen haben einen korrespondierenden Gesichtsausdruck beziehungsweise eine mimische Entsprechung (**Emotionsausdruck**). Im Rahmen der zwischenmenschlichen Interaktion ist es wichtig, diese mimischen Entsprechungen korrekt zu „dechiffrieren“, um Interaktionen zufriedenstellend zu gestalten (**Emotionserkennung**). Aus Informationen wie der Stellung der Augenbrauen, der Nase und des Mundes lassen sich Informationen ablesen, aus denen ein spezifischer emotionaler Ausdruck geschlussfolgert werden kann.

Erfolgreiches Zuordnen dieser emotionalen Gesichtsausdrücke zur zugrundeliegenden Emotion erscheint insofern sinnvoll, als dass individuell auf den Gemütszustand einer Person eingegangen werden kann und die Interaktion in der Folge störungsfreier verlaufen kann. Die Fähigkeit, emotionale Gesichtsausdrücke adäquat zuzuordnen scheint allerdings nicht grundsätzlich bei allen Personen gleichermaßen ausgeprägt zu sein. Es sind sowohl individuelle Unterschiede als auch gruppenspezifische Unterschiede zu beobachten. So können beispielsweise Personen mit einer Angststörung Emotionen schlechter aus Gesichtern erkennen als Personen ohne Angststörung (Kessler, Roth, Deighton & Traue, 2005) und Frauen Emotionen schon bei einer viel geringeren Intensität des mimischen Ausdrucks korrekt zuordnen als Männer (Hoffmann, Kessler, Eppel, Rukavina & Traue, 2010). Ein weit verbreitetes Paradigma zur Erforschung der Fähigkeit zur Emotionserkennung (**Emotionserkennungsleistung**) ist das Präsentieren von Fotografien mit verschiedenen Emotionsausdrücken, meist aufgenommen zum Zeitpunkt stärkster Expressivität. Überlegung hierbei ist, dass Emotionen bestimmten Emotionsausdrücken zuzuordnen sind. Die Verwendung mimischen Bildmaterials erfasst die „Rückübersetzung“ des emotionalen Gesichtsausdrucks zur basalen Emotion. Zur standardisierten Erfassung dieser Erkennungsleistung, wie sie im Bereich der Diagnostik und Intervention vorausgesetzt wird, wird Stimulusmaterial benötigt, welches die emotional-expressiven Gesichter valide darstellt. Kernstück der vorliegenden Arbeit stellt die Konzeption und testtheoretische Überprüfung des neuen Bildersets Pictures of Facial Affect-Ulm (kurz: PFA-U) dar. Im Vordergrund steht einerseits der Einbezug der ökologischen Validität. Im Unterschied zu bisherigen Stimulussets wird eine Erweiterung des Bildmaterials vorgenommen. Bildmaterial aus fünf unterschiedlichen Perspektiven soll der Wahrnehmung von mimischen Ausdrücken im Alltag stärker entsprechen als es bei Frontalbildern der Fall ist. Außerdem wird die Natürlichkeit und Intensität des Gesichtsausdrucks und die Attraktivität des abgebildeten Gesichts bei der Auswahl berücksichtigt. Es wird angenommen, dass diese Kriterien eine wichtige Rolle bei der Emotionserkennung spielen. Andererseits sollte die Überprüfung bestimmter psychometrischer Gütekriterien erfolgen, welche für die Verwendung psychologischer Testverfahren definiert wurden. Die bisherigen Bildersets wurden einem solchen Prozedere nicht unterzogen. Die Bereitstellung des Stimulusmaterials soll langfristig durch die Einbettung in ein bereits bestehendes und gut validiertes Testsystem geschehen, dem FEEL-Test (Kessler, Bayerl, Deighton & Traue, 2002).

2. THEORETISCHER HINTERGRUND UND ABLEITUNG DER FORSCHUNGSZIELE

Emotionen sind vielschichtige – sowohl psychische als auch physische Prozesse sind beteiligt. Aufgrund dieser Komplexität der teilnehmenden Prozesse ist es zum momentanen Zeitpunkt noch nicht möglich, mit wissenschaftlichen Methoden eine einheitliche und objektive Definition des Emotionsbegriffes zu erzielen. Dies spiegelt sich in einer Fülle teils widersprüchlicher Emotionstheorien wider. Emotionen haben unter anderem eine mimische Entsprechung, welche es ermöglicht, diese schnell und effektiv in einer sozialen Interaktion zu kommunizieren. Somit erscheint es bedeutsam, an dieser Stelle die Theorie der Basisemotionen vorzustellen, welche für die mimische Emotionserkennung von besonderer Bedeutung ist. Theoretische Überlegungen und empirische Befunde sollen zusammenfassend dargestellt werden, um hieraus Ziele und Fragestellungen der vorliegenden Arbeit ableiten zu können.

2.1 Theoretische Grundlagen der Emotionspsychologie

Betrachtet man gängige Emotionsdefinitionen auf Übereinstimmungen hin, so lässt sich folgendes Übereinkommen ableiten:

Emotionen bestehen aus vier verschiedenen Komponenten, einer subjektiven Komponente (individuelles inneres Erleben), der physiologischen Aktivität bzw. Reaktivität des Körpers, einer kognitiven und bewertenden Komponente und dem motorisch-expressiven Ausdrucksverhalten (Traue & Kessler, 2003). Sie alle dienen der Reizbewertung, Systemregulation, Handlungsvorbereitung, Kommunikation von Reaktion und Intention und Reflexion und Kontrolle (Merten, 2003).

Es handelt sich hierbei nicht um eine Definition, eher um einen Konsens zwischen verschiedenen Standpunkten innerhalb der Emotionsforschung. Ein häufig auftretendes Missverständnis scheint zu sein, dass wissenschaftlich orientierte Emotionsforschung mit einer exakten Gegenstandsbestimmung zu beginnen habe (Otto, Euler & Mandl, 2000; Meyer, Reisenzein & Schützwohl, 2003). Vielmehr ist es so, dass, läge so eine Definition bereits vor, das Phänomen bereits in all seinen Einzelheiten bekannt wäre. Somit kann man

eine exakt zutreffende Definition des Emotionsbegriffes an dieser Stelle noch nicht voraussetzen. Daraus resultiert, dass die oben genannten Komponenten jeweils unterschiedlich gewichtet und interpretiert werden können, was die Ausbildung verschiedener theoretischer Schwerpunkte ermöglicht (Scherer, 2005). Grundsätzlich können hierbei diskrete oder kategoriale von dimensionalen Emotionstheorien unterschieden werden.

Bei der **dimensionalen Klassifikation** von Emotionen wird angenommen, dass eine Emotion das Resultat einer bestimmten Ausprägung von unterschiedlichen Dimensionen ist. Verschiedene Dimensionen werden und wurden diskutiert. Wilhelm Wundt (1905) nahm erstmals folgende drei Dimensionen an: Spannung versus Lösung, Lust versus Unlust und Erregung versus Beruhigung. Jede Emotion setzt sich in einer bestimmten Ausprägung aus diesen drei Dimensionen zusammen. Aktuellere theoretische Überlegungen stammen von Lang und Kollegen (Lang, Greenwald, Bradley & Hamm, 1993; Bradley & Lang, 1994). Bei ihnen heißen die drei Dimensionen Valenz, Erregung und Dominanz. Ein großer Vorteil dieses Ansatzes ist die Möglichkeit der feinen Abstufung. Er ermöglicht nicht nur eine qualitative Unterscheidung, sondern auch eine quantitative Abstufung. Dies kommt unserem Alltagserleben sehr entgegen, wir spüren eine Emotion manchmal weniger, in anderen Situationen dafür sehr stark. Innerhalb des dimensionalen Ansatzes werden die alltäglich angewendeten Emotionswörter wie Angst, Ärger oder Freude als Kernaffekt berücksichtigt, aber nicht differenzierend verwendet. Diese Zuordnung ist Kernpunkt der kategorialen Ansätze. Von enormer Bedeutung für die vorliegende Arbeit ist die Theorie der Basisemotionen, welches nachfolgend als Beispiel des kategorialen Ansatzes herangezogen werden soll.

Forscher des **kategorialen Ansatzes** von Basisemotionen postulieren distinkte angeborene und elementare Emotionen als Ergebnis evolutionärer Prozesse. Sie beziehen sich hierbei auf die evolutionstheoretischen Überlegungen von Charles Darwin, der mit seinem 1872 herausgebrachten Buch „The Expression of Emotions in Man and Animals“ als „Urvater“ dieses theoretischen Ansatzes bezeichnet werden kann. Erkannt werden sollen diese distinkten „Basis“-Emotionen anhand einer universellen Mimik – also mittels eines Emotionsausdrucks, der in den unterschiedlichen Kulturen der Welt in etwa dieselbe Aussagekraft hat. Darwin begann seine Beobachtungen an Tieren, da er von einer

kontinuierlichen Entwicklung des Emotionsausdrucks vom Tier zum Mensch ausging. Er beobachtete ebenfalls den Gesichtsausdruck von Kindern, bei denen der Ausdruck noch weitestgehend ungehemmt ist, legte verschiedenen Beurteilern fotografierte Gesichtsausdrücke zur Einschätzung vor, stellte schriftliche Befragungen für interkulturelle Vergleiche an, beobachtete den emotionalen Ausdruck von Haustieren und verglich ihn mit dem menschlichen Emotionsausdruck (Meyer, Schützwohl & Reisenzein, 2003; Ekman, Sorensen & Friesen, 1969; Izard, 1971). Darwin kam in seinen Überlegungen zu dem Ergebnis, dass Emotionen durch kognitive Einschätzungen ausgelöst werden, die ihrerseits den emotionalen Ausdruck in Form von Mimik, Gestik, Stimme und physiologischen Veränderungen verursachen. Außerdem ging auch er schon davon aus, dass es eine fest determinierte Anzahl an Emotionsausdrücken mit organismischer und kommunikativer Funktion gibt, die für das Überleben und die Fortpflanzung von großer Bedeutung sind. Viele dieser Überlegungen bilden die Grundannahmen der Theorie der Basisemotionen. Emotionen werden dabei als Anpassungsprozesse der Individuum-Umwelt-Schnittstelle gesehen, die kommunizierbare Handlungs- und Reaktionsmuster ermöglichen (Traue, 1998). Angenommen wird eine Art „Notfallsystem“, welches die kognitive Verarbeitung umgehen und schnell physiologische Reaktionen auslösen kann, die nötig sind, um die Situation bestmöglich zu meistern (Cacioppo, 2004). Jeder der Basisemotionen liegt vermutlich ein spezifisches psychobiologisches Reaktionsmuster zugrunde (Ekman, 2003; Ekman & Cordaro, 2011). Durch den mimischen Ausdruck werden beispielsweise in der zwischenmenschlichen Kommunikation Situationsbewertungen und Handlungsbereitschaften sichtbar.

Zwar stimmen Forscher in ihren theoretischen Überlegungen in weiten Teilen überein, die Anzahl der angenommenen Grundemotionen variiert jedoch stark und reicht von 2-18 Grundemotionen. „Mut“ wird ganz selten als eine elementare Emotion beschrieben, Emotionen wie Ärger, Furcht, Traurigkeit und Freude lassen sich hingegen bei vielen Forschern wiederfinden. Silvan Tomkins beispielsweise ging von acht elementaren Emotionen aus: Überraschung, Interesse, Freude, Wut, Furcht, Ekel, Scham und Angst (Tomkins, 1984). Carroll Izard hingegen beschrieb zehn Basisemotionen: Interesse, Leid, Widerwillen, Freude, Zorn, Überraschung, Scham, Furcht, Verachtung und Schuldgefühl (Izard, 1971). Angst, Ärger, Ekel, Freude, Trauer und Überraschung nahmt auch Robert Plutchik als Basisemotionen an, ergänzte sie in seiner Theorie aber um die Emotionen

Akzeptanz, Erwartung und Erstaunen (Plutchik, 1980). Frijda wiederum führte ganz andere Emotionen als Basisemotionen auf: Begierde, Freude, Interesse, Sorge, Überraschung und Verwunderung (Frijda, 1986; Frijda, 1988). Diese Aufzählung umfasst nur die bekanntesten Vertreter und ihre Konzepte. Doch auch viele andere Forscher(gruppen) haben eigene Sets von Emotionen zusammengestellt, die ihrer Meinung nach als elementar anzusehen sind. Philip Johnson-Laird und Keith Oatley haben sich beispielsweise angeschaut, welche Wörter der Mensch hat, um Emotionen auszudrücken und diese dann zusammengefasst. Sie gingen von fünf Basisemotionen aus: Ärger, Angst, Ekel, Freude, Trauer (Oatley & Johnson-Laird, 1987). Jaak Panksepp stimulierte Rattengehirne elektrisch und folgerte aus den Verhaltensreaktionen seiner Versuchstiere, dass es vier elementare Emotionen geben müsste: Panik, Wut, Erwartung und Furcht (Panksepp, 1982).

Paul Ekmans Aufzählung umfasst sieben Basisemotionen, die da lauten: Freude, Überraschung, Ärger, Trauer, Ekel, Furcht und Verachtung (Ekman, 1982). Ekman definierte in Zusammenarbeit mit verschiedenen Forschern dreizehn Aspekte, die eine Emotion erfüllen muss, um eine Basisemotion zu sein. Aufgrund der Tatsache, dass diese bisher nur auf Englisch veröffentlicht wurden, werden sie hier, um Ungenauigkeiten in der Übersetzung zu vermeiden, in der Originalsprache aufgeführt. Sie lauten (Ekman & Cordaro, 2011, S. 365):

1. distinctive universal signals
2. distinctive physiology
3. automatic appraisal
4. distinctive universals in antecedent events
5. presence in other primates
6. capable of quick onset
7. can be of brief duration
8. unbidden occurence
9. distinctive thoughts, memories, and images
10. distinctive subjective experience
11. refractory period filters information available to what supports the emotion
12. target of emotion unconstrained
13. the emotion can be enacted in either a constructive or destructive fashion

Diese Aspekte sieht er in „seinen“ Basisemotionen als erfüllt an. „Verachtung“ als Basisemotion zu bezeichnen, ist dennoch nicht unumstritten (Ortony & Turner, 1990;

Wagner, 1997; Wagner, 2000), so dass häufig auch nur von sechs Basisemotionen die Rede ist (siehe Abbildung 1). Beispielsweise führte Robert Plutchik an, dass „Verachtung“ eine Emotion darstelle, bei der eine Bewertung des Gegenübers vorgenommen werden müsse, also höhere kognitive Prozesse eine Rolle spielten. Er plädierte außerdem dafür, dass „Verachtung“ mehr eine Mischung von Ekel und Ärger darstelle als eine reine und konkret abgrenzbare Emotion. Bei Verachtung spielt vermutlich die kognitive Komponente eine wesentlich ausgeprägtere Rolle als bei den anderen postulierten Basisemotionen. Hinweise hierfür ergeben sich aus Beobachtung, dass Tiere nicht in der Lage zu sein scheinen, Verachtung zu empfinden oder zu äußern. Wagner (2000) zufolge kann „Verachtung“ außerdem nur in sozialen Situationen empfunden werden, da nur Subjekte, aber keine Objekte oder Situationen verachtet werden können. Aufgrund dieser Kontroverse hat „Verachtung“ bei der vorliegenden Arbeit keine Beachtung erfahren.

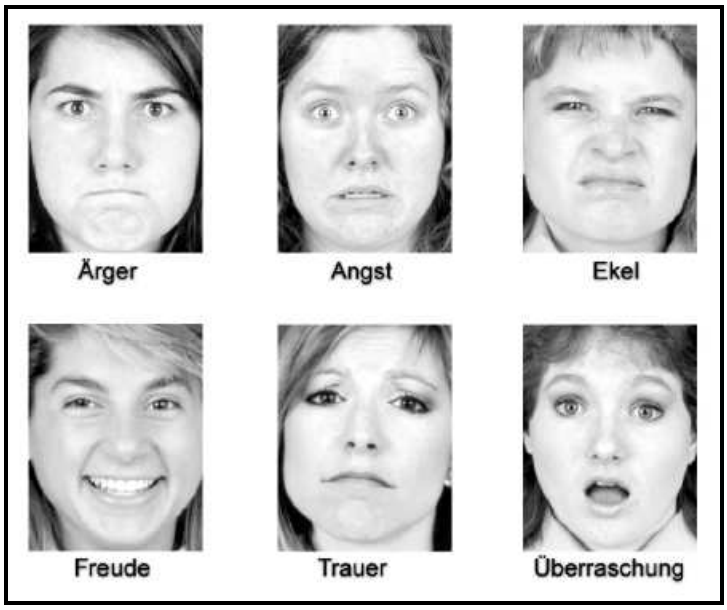


Abb. 1: Die sechs Basisemotionen nach Ekman (aus dem JACFEE-Bildersatz: Matsumoto & Ekman, 1988)

Eine unvollständige Übersicht über die verschiedenen Vertreter der Theorie der Basisemotionen ist in der nachfolgenden Tabelle 1 zusammengefasst.

Tab. 1: Auszug der von verschiedenen Forschern postulierten Basisemotionen (nach Ortony & Turner, 1990)

Autoren und Jahr	**Basisemotionen**
Ekman (1982)	Ärger, Freude, Trauer, Überraschung, Angst, Ekel, Verachtung
Frijda (1986)	Begierde, Freude, Interesse, Sorge, Überraschung, Verwunderung
Izard (1971)	Ärger, Angst, Ekel, Freude, Interesse, Scham, Schuld, Überraschung, Verachtung, Verzweiflung
Oatley & Johnson-Laird (1987)	Ärger, Angst, Ekel, Freude, Trauer
Panksepp (1982)	Angst, Erwartung, Panik, Wut
Plutchik (1980)	Ärger, Anerkennung, Ekel, Erwartung, Freude, Angst, Trauer, Überraschung
Tomkins (1984)	Ärger, Angst, Ekel, Freude, Interesse, Scham, Überraschung, Verachtung, Verzweiflung

Anhänger der Theorie der Basisemotionen postulierten anfänglich, dass die kommunikativen Signale des Gesichtsausdrucks nicht willentlich zu steuern wären. Diese Behauptung ist nach heutigem Kenntnisstand nicht mehr haltbar. Verschiedene Studien (zum Beispiel Kappas, 2003) haben gezeigt, dass es nicht notwendigerweise einen direkten Zusammenhang zwischen Emotion und dem gezeigten Gesichtsausdruck geben muss. Der Mensch ist in der Lage, die Mimik auch als bewussten Aspekt in der Kommunikation einzusetzen. Beispiele hierfür sind unter anderem der traurige Gesichtsausdruck, etwa um einer Person sein Mitgefühl auszudrücken oder das so genannte Konversationslächeln, um bei einem ersten Gespräch einen freundlichen und netten Eindruck zu hinterlassen. Anhand spezifischer Merkmale des Gesichts kann ihm eine bestimmte Emotion zugeordnet werden.

Die Überlegungen, auf denen die Theorie der Basisemotionen fußt, sind nicht unumstritten. Kritisiert werden vor allem zwei Punkte: Erstens die Bezeichnung einer bestimmten Anzahl von Emotionen als elementar oder als „Basis“ und zweitens die Universalität dieser Emotionen. Der erste Kritikpunkt kommt in erster Linie dadurch zustande, dass die verschiedenen Forscher in diesem Bereich unterschiedlich viele Emotionen als Basisemotionen postuliert haben (Meyer, Schützwohl & Reisenzein, 2003). Ortony und Turner (1990) entgegneten in diesem Zusammenhang, dass, gäbe es eine einheitliche empirische Basis, die Anzahl der Basisemotionen eindeutig sein müsste. Die Kritik der Universalität der Basisemotionen wurde in erster Linie vom James Russell (Russell, 1991; Russell, 1994) formuliert. Russell führte Daten an, die zeigten, dass bestimmte isolierte Kulturen bei der mimischen Emotionserkennung schlechter abschnitten als westliche Vergleichsgruppen. Russell führte dies auf den Einfluss der Massenmedien und methodische

Probleme zurück. Ob dieses Argument richtig ist, kann derzeit noch nicht beantwortet werden.

Abschließend soll erwähnt werden, dass im aktuellen Diskurs zu Emotionen die Theorie der Basisemotionen mit dem dimensionalen Ansatz vereint wird. Es sollen die Vorteile beider Ausrichtungen miteinander verbunden werden. Roseman (2011) beispielsweise stellt einen theoretischen Rahmen vor, der diskrete Emotionen in einen Raum aus vier Dimensionen („Appraisal“, „Circumstances“, „Valence“ und „Control“) einordnet (siehe Abbildung 2). Mit diesem Hybrid-Modell können feine Abstufungen auf den jeweiligen Dimensionen vorgenommen werden, die gebräuchlichen Emotionswörter können ebenfalls beibehalten werden. „Distress“ (englisch für: Leid, Not, aber auch Bedrängnis) kann in diesem Modell beispielsweise folgendermaßen eingeordnet werden: Das Erleben einer solchen Emotion geht mit geringer Kontrolle (Dimension „Control“), Ungewissheit (Dimension „Circumstances“), Aversion (Dimension „Valence“) und dem Wunsch dieser Situation zu entkommen, (Dimension „Appraisal“) einher. Das Modell bietet zudem die Möglichkeit, nicht nur die Basisemotionen, sondern auch komplexere emotionale Zustände zu definieren.

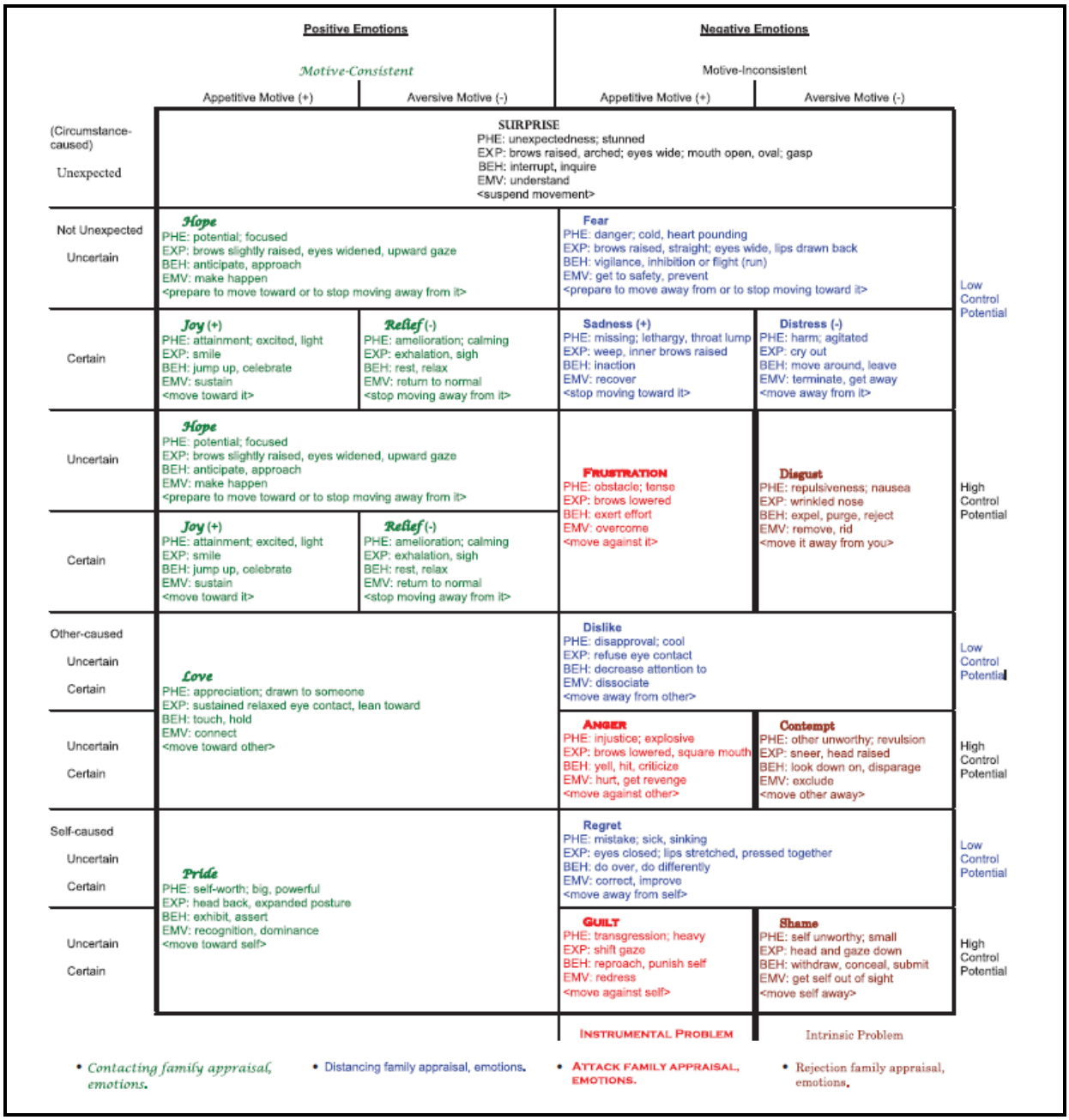


Abb. 2: Übersicht Hybrid-Modell von Roseman (2011, S. 438)

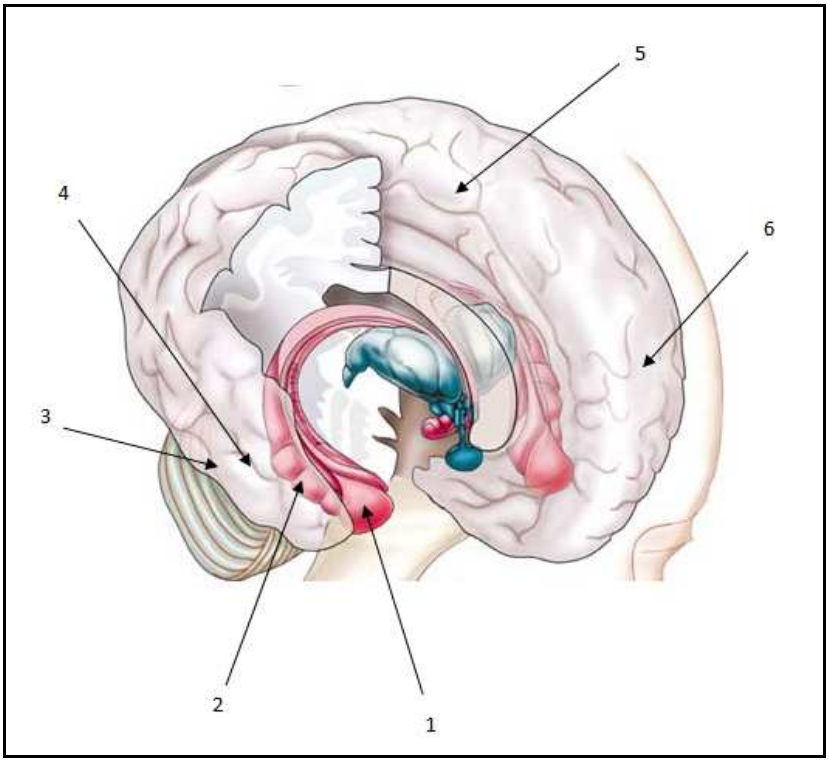
2.2 Neurobiologische Grundlagen der Emotionserkennung

Ziel der neurowissenschaftlichen Emotionsforschung ist es, den „Sitz der Emotionen im Gehirn“ zu finden. Im Mittelpunkt stehen das limbische System und der präfrontale Kortex (Gur, Schroeder, Turner, McGrath, Chan, Turetsky, Alsop, Maldjian & Gur, 2002). Die emotionale Funktion des präfrontalen Kortex kann in drei verschiedene Bereiche unterteilt werden. Der **ventro-mediale präfrontaler Kortex** ist für die Repräsentation elementarer positiver und negativer affektiver Zustände in Abwesenheit unmittelbar verfügbarer Anreize

zuständig. Der **dorsolaterale präfrontale Kortex** leitet den Handlungsentwurf, die Repräsentation von Zielzuständen der positiven und negativen Affekte ein. Der **orbitofrontale Kortex** bestimmt die emotionale Intensität (zum Beispiel bei Schmerz), unterstützt schnelles Verstärkungslernen und das Ändern von Assoziationen, wenn Kontingenzen sich geändert haben. Unterschieden werden kognitive Bereiche wie Aufmerksamkeitssteuerung, Entwurf von Handlungsplänen oder sprachlichen Äußerungen von Bereichen, die diese „kognitiven" Gebiete mit emotionalen Aspekten anreichern. Wichtig ist, darauf hinzuweisen, dass keine Trennung dieser beiden beschriebenen Bereiche im eigentlichen Sinne anzunehmen ist (Sprengelmeyer, Rausch, Eysel & Przuntek, 1998), sondern dass Emotionen alle Prozesse beeinflussen. Untersuchungen mit bildgebenden Verfahren zeigen, dass eine Vielzahl neuronaler Areale beim Erleben und Wahrnehmen von Emotionen aktiviert werden (Murphy, Nimmo-Smith & Lawrence, 2003). Bezüglich des Empfindens von Emotionen im Gehirn ist die Lateralisierungshypothese weit verbreitet (Otto, Euler & Mandl, 2000; Traue, 1988; Singer & Deutsch,). Hierzu gibt es drei unterschiedliche Überlegungen. Eine Überlegung ist, dass Emotionen in der rechten Hemisphäre lokalisiert sind, nicht in der linken. Einer anderen Überlegung nach sind in der linken Hemisphäre eher positive und in der rechten Hemisphäre eher negative Emotionen lokalisiert. Eine dritte Überlegung ist, dass so genannte Annäherungsemotionen links und Vermeidungsemotionen rechts lokalisiert sind. Für keine der drei Überlegungen gibt es bislang überzeugende empirische Befunde.

Vermutet wird, dass Gedächtnisinhalte in der Amygdala emotional angereichert werden (siehe Abbildung 3). Fällt diese aus, führt dies zu einer Unfähigkeit, die affektive Valenz von Stimuli wahrzunehmen. So kann es nach einer Läsion vorkommen, dass auch zuvor als eklige wahrgenommene Dinge verspeist werden. Eine Furchtkonditionierung ist nicht mehr möglich. Die Amygdala hat Ein- und Ausgänge zum motorischen (zentrales Höhlengrau), autonomen und endokrinen System (Hypothalamus), so dass hier zahlreiche Interaktionen angenommen werden können. Eine Verbindung zum Neokortex über den Basalkern garantiert eine Assoziation der Sinne, eine Verbindung zum Hippocampus gewährleistet den Einbezug von Gedächtnis und räumlichem Denken. Nach LeDoux (1995) werden in der Amygdala implizite und im Hippocampus explizite Anteile einer Situation gespeichert. Zwei Wege der Stimulusverarbeitung liegen vor: ein schneller und direkter Weg über den Thalamus zur Amygdala, welcher als Frühwarnsystem fungiert und ein zweiter langsamerer

Weg über den Neokortex, welcher eine Relevanzunterscheidung ermöglicht. Beispielsweise kann es sein, dass wir uns vor einem Gartenschlauch erschrecken, da darin bei einer ersten Sondierung eine Schlange erkannt wurde. Der Körper wird angeregt, um eine schnelle Reaktion zu gewährleisten. Auf dem langsamen Weg wird in der Zwischenzeit differenziert und der Gartenschlauch erkannt. Ein Flucht- oder Kampfreaktion ist nicht notwendig. Dieses Zwei-Wege-System bietet große Vorteile bezüglich einer schnellen Reaktion versus dem Einsparen von Energie.



1 = Amygdala, 2 = Hippocampus, 3 = Gyrus fusiformis, 4 = Sulcus temporalis, 5 = Neokortex, 6 = Präfrontalkortex

Abb. 3: Übersicht über die für die Emotionsverarbeitung relevanten Hirnareale (Copyright: www.julius-ecke.de)

Für die mimische Emotionserkennung von besonderer Bedeutung sind der **Gyrus fusiformis** und der **Sulcus temporalis superior**. Erst genanntes Hirnareal ist für die Gesichtererkennung und für die Wahrnehmung unveränderlicher Gesichtsmerkmale zuständig. Verschiedene Studien (zum Beispiel Rhodes, Brake & Atkinson, 1993; Calder, Young, Keane & Dean, 2000) zeigen, dass spezifische Merkmale des Gesichts und ihre räumliche Anordnung beziehungsweise Beziehung zueinander eine wichtige Rolle spielen. Der Sulcus temporalis superior hingegen ist für die Wahrnehmung sich zeitlich verändernder Gesichtsmerkmale zuständig und detektiert mimische Veränderungen.

Eine Metastudie von Murphy, Nimmo-Smith und Lawrence (2003) zeigte, dass bei der Emotionserkennung für einzelne Emotionen unterschiedliche Gehirnareale aktiviert werden. Dies erklärt, warum bei einigen psychischen/neurologischen Störungen nur ganz bestimmte Emotionen nicht mehr ausreichend gut erkannt werden. Bei der Erkennung von „Angst" spielt die Aktivierung der Amygdala eine bedeutende Rolle, so die Ergebnisse der Metastudie. Die Verarbeitung der Emotion „Ekel" hingehen geht einher mit einer Zunahme der neuronalen Aktivität der Insula und des Pallidums. Für die anderen Basisemotionen konnten noch keine einheitlichen Ergebnisse abgeleitet werden. Ärgerverarbeitung könnte aber vornehmlich in orbitofrontalen Bereichen, die Verarbeitung von Freude und Trauer im medialen präfrontalen Kortex stattfinden (Murphy, Nimmo-Smith & Lawrence, 2003).

2.3 Empirische Arbeiten aus dem Bereich der Emotionserkennung

Auf dem Konzept der Basisemotionen beruht eine ganze Reihe wissenschaftlicher Arbeiten, die Aspekte des emotionalen Verhaltens und Erlebens vielfältig untersucht haben. Bei der Erstellung eines standardisierten Bildersets ist es von Bedeutung, diese Ergebnisse in angemessener Weise zu berücksichtigen. Daher sollen zunächst einige Befunde vorgestellt werden, die Einfluss auf die Erstellung der „Pictures of Facial Affect-Ulm" hatten.

2.3.1 Interindividuelle Unterschiede der Emotionserkennung

Die bisherige Forschung legt nahe, dass die Fähigkeit zur Emotionserkennung zu den Basisfertigkeiten des Menschen gehört. So zeigt sich einerseits, dass „Freude" und „Überraschung" in der Regel sehr gut und „Angst" häufig relativ schlecht erkannt wird (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & Knippenberg, 2010; Tracy, Robins & Schriber, 2009; Goeleven, De Raedt, Leyman & Verschuere, 2008). „Freude" wird möglicherweise so gut erkannt, da sie schon seit Entstehung der Menschheit existiert. In einer Studie von Ross, Owren und Zimmermann (2009) konnte nachgewiesen werden, dass durch Kitzeln induziertes Lachen in seiner Form stark dem von Primaten ähnelt. Die Autoren ziehen den Schluss, dass gemeinsame Vorfahren schon Freude ausdrücken konnten. Trotzdem können individuelle Unterschiede bei der mimischen Emotionserkennung beobachtet werden. Einige Aspekte sollen im Folgenden aufgeführt werden.

a) **Alter**: Das Alter des Beurteilers eines emotionalen Stimulus und das Alter der Person, deren emotionaler Ausdruck beurteilt werden soll, beeinflussen die Emotionserkennung (beispielsweise Anastasi & Rhodes, 2005). Der „Own Age“-Bias wird hierfür häufig als Argument herangezogen. Bei Personen des eigenen Altersbereiches sollen demnach Emotionen besser erkannt werden als bei Personen aus einem anderen Altersbereich. Erklärt wird dies damit, dass man Gesichtern des eigenen Altersbereiches am häufigsten begegnet und deswegen mit der dort gezeigten Mimik am besten vertraut ist(Ebner, 2008). Ebner konnte in ihrer Studie einen Alterssstereotyp nachweisen: Ältere Gesichter wurden negativer bewertet als jüngere Gesichter. Vermutet wird, dass die veränderte Physiognomie (Faltenbildung, Runzeln et cetera) hierbei eine wichtige Rolle spielt. Zusätzlich ist davon auszugehen, dass im Alter einige perzeptuelle Fähigkeiten nachlassen, Verarbeitungsprozesse im Gehirn langsamer werden und auch die Reaktionsgeschwindigkeit abnimmt. Ebner und Johnson (2009) wiesen in ihrer Studie beispielsweise nach, dass die Wiederkennungsleistung bei jüngeren Probanden insgesamt besser war als bei älteren Probanden. Sullivan, Ruffman und Hutton (2007) diskutierten in ihrer Studie neurologische Veränderungen im Alter. Die von ihnen durchgeführten Studien zeigten, dass jüngere Probanden mehr die Augen der Stimulusperson detektieren und die emotionalen Ausdrücke besser zuordnen konnten als ältere Probanden. In der Entwicklung der Fähigkeit zur Emotionserkennung zeigt sich: Schon junge Kinder können Emotionen aus Gesichtsausdrücken herauslesen (Pollak, Messner, Kistler & Cohn, 2009). Beispielsweise durch die Rückmeldung der Eltern lernen Kinder dies mit der Zeit immer besser. Hier zeigt sich aber auch, dass frühe Erfahrungen die Erkennung emotionsspezifischer Informationen auch negativ beeinflussen. Beispielsweise können missbrauchte Kinder „Ärger“ viel früher erkennen als nicht missbrauchte Kinder (Pollak, Messner, Kistler & Cohn, 2009). Bei Kindern, die schon im jungen Alter relativ hohe Werte in der „Psychopathy Screening Device“ aufweisen, zeigen sich signifikante Defizite im Bereich der Trauer- und Angstwahrnehmung im Gegensatz zu unauffälligen Kindern (Blair, Colledge, Murray & Mitchell, 2001).

b) **Kultur**: Eine weitere Variable, die die Emotionserkennung zu beeinflussen scheint, ist der kulturelle Hintergrund (Matsumoto, 1992; Matsumoto, 2002). Zwar werden gemäß Universalitätshypothese keine kulturellen Unterschiede in der Emotionserkennung

angenommen, doch die empirische Befundlage ist heterogen. Mit der Entwicklung standardisierter Stimulussets war es möglich, kulturelle Unterschiede zu messen. Eine Studie von Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh und Ton (1997) zeigte, dass in verschiedenen Kulturen eine hohe Übereinstimmung bei der Bestimmung der Basisemotionen vorlag. Dies wurde anhand von Probanden aus Ungarn, Japan, Polen, Sumatra, den USA und Vietnam überprüft. Dennoch zeigten sich Unterschiede im exakten Grad der Übereinstimmung zwischen den Kulturen, für einen Vergleich herangezogen wurden die westlichen (USA, Polen, Ungarn) und die nicht-westlichen Ländern (Japan, Vietnam, Sumatra). Die westlichen Länder zeigten in dieser Studie insgesamt eine höhere Genauigkeit in der Erkennungsleistung. In einer Studie von Elfenbein, Beaupré, Lévesque und Hess (2007) wurden kulturelle Zusammenhänge differenzierter untersucht. In einer ersten Studie wurden Probanden aus Quebec (Kanada) oder Libreville (Afrika) gebeten, ohne Vorgaben zur Umsetzung bestimmte emotionale Ausdrücke zeigen (beispielsweise Freude, Ärger, Ekel). Diese natürlichen Emotionsausdrücke wurden fotografiert und Gruppen aus beiden Ländern, welche nicht an der Erstellung der Stimuli beteiligt waren, vorgegeben, um die Erkennungsleistung zu überprüfen. Zusätzlich wurden den Probanden in einer zweiten Studie auch standardisierte, das heißt nach EmFACS prototypisch angefertigte Bilder vorgegeben. Die Ergebnisse interpretierten die Autoren im Sinne es Dialektes bei der Emotionserkennung. Wurden prototypische Gesichtsausdrücke vorgegeben, so ergaben sich keine Unterschiede in der Erkennungsleistung zwischen den Kulturen. Wurde jedoch das Stimulusmaterial mit dem natürlichen Emotionsausdruck vorgegeben, so zeigte sich ein Eigengruppen-Vorteil.

c) **Status**: Als potentielle Einflussvariable auf die Genauigkeit bei der Emotionserkennung wurde auch der soziale Status näher untersucht. Pfiff und Kollegen beispielsweise nehmen an, dass für Personen mit einem niedrigeren sozialen Status eine höhere Notwendigkeit bestehe, sich an ihrem sozialen Umfeld zu orientieren und dies in einer höheren Emotionserkennungsleistung resultiert (Piff, Kraus, Côté, Cheng & Keltner, 2010). Die Autoren nehmen außerdem an, dass mit höherem sozialen Status eine größere Unabhängigkeit gegenüber der Gesellschaft bestehe, die es weniger notwendig mache, Emotionen präzise zu erkennen oder (adäquat) empathisch zu reagieren. Dies

wird untermauert durch den Befund, dass Menschen einer niedrigeren sozialen Schicht mehr interaktionsorientierte Verhaltensweisen (Kopfnicken, Lächeln) zeigen als Menschen höherer sozialer Schichten (Kraus & Keltner, 2009; Kraus, Pfiff & Keltner, 2009). In der Studie von Kraus, Côté und Keltner (2010) konnte für die Ausprägung der empathischen Genauigkeit, die Emotionserkennungsfähigkeit innerhalb einer Interaktion und die Emotionserkennungsfähigkeit aus statischen Bildern ein Vorteil für Menschen niedrigerer sozialer Schichten gefunden werden.

d) **Geschlecht**: Seit jeher ist die Erforschung von Unterschieden zwischen Männern und Frauen ein wichtiges Thema in der Wissenschaft. Für kognitive Prozesse (wie das abstrakt-logisches Denken) wird beispielsweise eine Überlegenheit des männlichen Geschlechts angenommen (Hyde & Mertz, 2009). Eine Überlegenheit des weiblichen Geschlechts wird seit jeher im emotionalen Bereich vermutet. Die Fähigkeit, mimische ausgedrückte Emotionen richtig zu deuten, scheint für Frauen von besonderer Bedeutung zu sein. Evolutionär betrachtet ist es denkbar, dass Frauen ihre körperliche Unterlegenheit durch das schnelle und effiziente Erkennen von Emotionen ausgleichen konnten und sich so einen wichtigen Überlebensvorteil sicherten (Bijlstra, Holland & Wigboldus, 2010). McClure (2000) konnte an einer Stichprobe bestehend aus Kleinkindern, Kindern und Jugendlichen nachweisen, dass weibliche Versuchsteilnehmer im Erkennen emotionaler Gesichtsausdrücke besser abschnitten. Diese Befunde werden durch eine Meta-Analyse von Hall, Carter und Horgan (2000) untermauert. In 80% der analysierten Studien war ein Vorteil zu Gunsten der Frauen erkennbar. In einer Studie von Hoffmann, Kessler, Eppel, Rukavina und Traue (2010) konnte nachgewiesen werden, dass Frauen emotionale Gesichtsausdrücke schon bei viel geringeren Intensitäten korrekt zuordnen können als Männer.

e) **Persönlichkeit**: In der Regel wird bei der Beschreibung der Persönlichkeit das Fünf-Faktoren-Modell herangezogen, welches „Extraversion“, „Neurotizismus“, „Gewissenhaftigkeit“, „Offenheit“ für „Erfahrungen“ und „Verträglichkeit“ als Faktoren umfasst. Gerade für die beiden erstgenannten Faktoren werden Einflüsse auf die Emotionserkennung vermutet. So konnten Chan, Goodwin und Harmer (2007) einen Zusammenhang zwischen „Neurotizismus“ und Emotionserkennung finden. Nicht nur, dass Personen mit hohen Neurotizismuswerten eher an Depressionen erkrankten, sie

fanden ebenfalls heraus, dass diese Personen schon im Vorfeld der Erkrankung durch eine schlechtere Erkennungsleistung von positiven Emotionen auffielen. Rubin, Munz und Bommer (2005) beschäftigten sich mit dem Zusammenhang von „Extraversion“ und Emotionserkennung bei Führungskräften. „Extraversion“ konnte als Moderator für den transformationalen Führungsstil detektiert werden, die Qualität des Führungsstils und die Performanz im Bereich der Emotionserkennung wiesen einen positiven Zusammenhang auf. Welling, DeBruite, Little und Jones (2009) wiesen für Frauen einen Zusammenhang zwischen „Extraversion“ und der Präferenz für bestimmte männliche Gesichter nach.

f) **Intensität und Natürlichkeit des Emotionsausdrucks und Attraktivität der fotografierten Person**: Bezüglich der Intensität konnten Studien (beispielsweise Hoffmann, Kessler, Eppel, Rukavina & Traue, 2010; Palermo & Coltheart, 2004) nachweisen, dass die Emotionserkennungsrate bei der Betrachung von emotionalen Gesichtsausdrücken vom Grad des Ausdrucks abhängt. Emotionen mit maximaler Intensität wurden besser erkannt als Emotionen, die lediglich subtil ausgedrückt wurden. Die Natürlichkeit des Gesichtsausdrucks ist für die ökologische Validität von Relevanz. Ekman und andere Autoren wiesen in der Vergangenheit darauf hin, dass neben den in EmFACS beschriebenen prototypischen Emotionsausdrücken noch alternative Möglichkeiten zum Ausdruck einer bestimmten Emotion bestehen (Ekman, 2007a; Carroll & Russell, 1997). In einer aktuellen Studie von Hoque und Picard (2011) konnten in den Antwortmustern kaum Übereinstimmungen hinsichtlich der Dynamik von Freude und Frustration beim Vergleich von natürlichen und prototypischen Gesichtsausdrücken finden. Trägt auf der einen Seite die Standardisierung zur Vergleichbarkeit von Untersuchungsergebnissen bei, so ist die Natürlichkeit auf der anderen Seite wichtig, um die Akzeptanz des Stimulusmaterials zu gewährleisten. Besonders relevant erscheint dies für die Emotion „Freude“ im Vergleich zu den anderen emotionalen Ausdrücken. In der Literatur wird das Duchenne-Lachen (natürlicher Freudeausdruck mit Beteiligung des Augenmuskels Zygomaticus) vom prototypischen und standardisierten Lachen unterschieden (Ekman, Davidson & Friesen, 1990). Auch für die Attraktivität der wahrgenommenen Person wird ein Einfluss auf die Emotionserkennung vermutet, auch wenn bisher systematische Studien fehlen. Hinweise ergeben sich aus folgenden

Befunden: Nach der Attraktivität einer Person werden beispielsweise Geschlechtspartner ausgewählt (O'Doherty, Winston, Critchley, Perrett, Burt & Dolan, 2003). Dies geht mit einem „Was schön ist, ist gut"-Stereotyp einher – beispielsweise werden attraktive Personen als charakterstärker und erfolgreicher eingeschätzt (Dion, Berscheid & Walster, 1972). Schon kleine Kinder sind an attraktiveren Gesichtern stärker interessiert (Samuels, Butterworth, Roberts, Graupner & Hole, 1994). Es ist zu vermuten, dass die Bevorzugung als attraktiv geltender Personen mit evolutionär erworbenen Konzepten einhergeht, beispielsweise dass attraktivere Partner gesünder und überlebensfähiger seien (Zaidel, Aarde & Baig, 2005).

2.3.2 Emotionserkennung und klinische Phänomene

Die bisherigen Ausführungen machen deutlich: Verschiedenste Variablen spielen bei der Emotionserkennung eine Rolle. Dies zeigt sich auch im klinischen Kontext. Beispielsweise weisen Patienten mit psychischen (z.B. Angststörung, Depression) und neurologischen (Apoplex, Läsionen) Störungen spezifische Defizite im Bereich der Emotionserkennung auf, die zu Problemen in der sozialen Interaktion führen können (zum Beispiel Bouhuys, Geerts Mersch & Jenner, 1997; Sachs, Steger-Wuchse, Kryspin-Exner, Gur & Katschnig, 2004; Braun, Traue, Frisch, Deighton & Kessler, 2005; Kessler, Schwarze, Filipic, Traue & von Wietersheim, 2006). Einige psychische Störungen werden mit emotionalen Exzessen, das heißt dem verstärkten Auftreten einer oder mehrer Emotionen, in Verbindung gebracht (Depression, Angst), andere wiederum mit Defiziten oder fehlender Kohärenz in diesem Bereich (zum Beispiel Psychopathie, Schizophrenie) verbunden. Ziel vieler therapeutischer Behandlungsansätze, beispielsweise bei Depression oder sozialer Phobie, ist unter anderem die Optimierung sozialer Kontakte. So weiß man inzwischen, dass das Erleben von Ärger intrapersonell zu einem Gefühl von Ungerechtigkeit führen kann. Für den Interaktionspartner führt die Wahrnehmung von Ärger häufig zu einer Angstreaktion. Keltner und Kring (1998) sehen diese fehlverlaufende emotionale Verarbeitung als wichtige Grundlage von Interaktionsstörungen beispielsweise bei Borderline-Patienten an. Der diagnostische und therapeutische Einsatz von emotionalen Gesichtausdrücken kann helfen, spezifische Defizite zu detektieren und zu behandeln.

In der Sektion Medizinische Psychologie der Universität Ulm wurden in Kooperation mit vielfältigen Forschergruppen einige Studien zur Emotionserkennungsfähigkeit an klinischen und neurologischen Stichproben durchgeführt. Grundlage bildet hierbei die Verwendung des FEEL-Tests, welcher in 3.2 ausführlich vorgestellt wird. Beispielsweise wurde in einer Studie von Braun, Traue, Frisch, Deighton und Kessler (2005) an Schlaganfallpatienten die Hypothese untersucht, ob eine rechtsseitige Verletzung im Gehirn zu größeren Beeinträchtigungen führt als eine linksseitige Verletzung. Diese Hypothese wurde aufgrund der Annahme postuliert, dass die Emotionsverarbeitung eher rechtshemisphärisch stattfindet. Beide Gruppen von Schlaganfallpatienten wurden außerdem mit einer gesunden Kontrollgruppe verglichen. Insgesamt wurden 24 Patienten untersucht, 13 mit linksseitige und 11 mit rechtsseitiger Läsion. Die Kontrollgruppe enthielt 29 für Alter, Geschlecht und Bildung gematchte Probanden. Im FEEL-Test konnten keine Unterschiede in den Leistungen zwischen den Patienten mit linksseitiger und rechtsseitiger Läsion gefunden werden, jedoch zeigten die Patienten allgemein eine schlechtere Emotionserkennungsleistung als die Probanden der Kontrollgruppe. Es wurde gemutmaßt, dass das Arbeitsgedächtnis und das Wortverständnis eventuell eine entscheidendere Rolle bei der Bewältigung dieser Aufgabe einnimmt als gedacht, sodass eine Kontrolle dieser Variablen in zukünftigen Studien notwendig erscheint.

In einer Studie von Gil, Ridout, Kessler, Neuffer, Schoechlin, Traue und Nickel (2009) wurden Unterschiede in der Emotionserkennung zwischen Personen mit somatoformer Störung (Experimentalgruppe mit 20 Probanden) und gesunden Probanden (Kontrollgruppe mit 20 Probanden) untersucht. Zusätzlich wurde die TAS-26 (kurz für: Toronto-Alexithymie-Skala 26) verwendet, um eine potentiell vorliegende Alexithymie erfassen zu können. Bei der Auswertung des FEEL-Tests zeigten sich Unterschiede zwischen Experimental- und Kontrollgruppe, wobei die Kontrollgruppe eine bessere Emotionserkennungsleistung aufwies. Diese Unterschiede zwischen den Gruppen verschwanden aber, wenn Alexithymie als Kontrollvariable eingeführt wurde. Laut den Autoren wäre es sogar denkbar, dass die Alexithymie Grundlage beziehungsweise Mitverursacher somatoformer Störungen sein könnte. Allerdings ist anzuführen, dass somatoforme Störungen grundsätzlich diagnostisch bislang nicht ausreichend exakt von anderen Störungsbildern abgegrenzt werden können und auch eine hohe Zahl an komorbiden Störungen aufweisen.

Ebenfalls wurde die Emotionserkennungsfähigkeit von essgestörten Frauen (Bulimia nervosa oder Anorexia nervosa) im Vergleich zu gesunden Kontrollprobandinnen (Kessler, Schwarze, Filipic, Traue & von Wietersheim, 2006) untersucht. Als Grundlage für diese Untersuchung wird angeführt, dass für Essgestörte häufig Defizite in sozialen Funktionen beschrieben werden. Zwar zeigte sich, dass die Experimentalgruppe alexithymer war als die Kontrollgruppe, Unterschiede bezüglich der Gesamtpunktzahl im FEEL-Test gab es jedoch nicht. Es wurde geschlussfolgert, dass die Emotionserkennung aus Gesichtern eine Basisfähigkeit darstellt, die nur einen Teilbereich des komplexen Systems für soziale Interaktionen und Erfahrungen darstellt.

Kessler, Roth, von Wietersheim, Deighton und Traue (2007) untersuchten den Einfluss von Panikstörungen auf die Emotionserkennungsfähigkeit. 37 Patienten mit Panikstörungen ohne Agoraphobie und ohne andere Komorbiditäten wurden in ihrer Emotionserkennungsfähigkeit mit 43 gesunden Kontrollprobanden verglichen. Die Performanz im FEEL-Test zeigte, dass die Emotionserkennung für Panikpatienten signifikant schlechter war als die der Kontrollgruppe. Dieser Effekt war besonders deutlich für die Emotionen Trauer und Ärger. Zusätzlich wurde eine Tendenz gezeigt, die Nicht-Ärger-Emotionen mit Ärger zu verwechseln.

Wie die aufgeführten Studien weisen auf einen starken Zusammenhang zwischen klinischen Phänomenen und spezifischen Defiziten in der Emotionserkennung hin. Einerseits unterstreicht dies die Annahme, dass die Emotionserkennung eine Basisfertigkeit darstellt. Andererseits lässt sich schlussfolgern, dass standardisiertes Bildmaterial zur Erfassung der Emotionserkennungsleistung im klinischen Kontext etwa zur Evaluation von Therapiefortschritten in Bezug auf soziale Interaktionen oder zu Trainingszwecken eingesetzt werden kann.

2.3.3 Entwicklung von Stimulusmaterial zur Überprüfung der Emotionserkennungsfähgkeit

Verschiedene Forschergruppen haben sich bereits mit der Erstellung von Bildmaterial zur Erfassung der Emotionserkennungsleistung auseinandergesetzt. In Tabelle 2 wird ein Überblick über die in der Literatur gefundenen Datenbanken gegeben (zur Übersicht siehe Gross, 2005). Auffällig ist, dass in aller Regel „nur“ Stimulusmaterial erstellt wird, eine

Einbettung in ein standardisiertes Testsystem jedoch nicht erfolgt. Neben dem FEEL-Test ist lediglich der VERT160/VIEMER (Hoheisel & Exner, 2005) als Testverfahren zu verstehen. Innerhalb der Bildersätze variieren die Art und Anzahl der zur Verfügung stehenden Emotionen, der kulturelle Hintergrund, die Verteilung der Geschlechter, Belichtung, Erstellung von Farb- oder Schwarz-Weiß-Fotografien sowie weitere Faktoren. Diese Unterschiede haben zur Folge, dass Untersuchungsergebnisse nur schwer vergleichbar sind, da die berichteten Erkennungsraten stark variieren. Für die existierenden Bildersets wurde bislang keine umfassende Überprüfung der psychometrischen Eigenschaften vorgenommen, welche die Qualität eines Testverfahrens beschreibt. Diese Lücke soll mit der Erstellung und Überprüfung der „Pictures of Facial Affect-Ulm“ geschlossen werden.

Tab. 2: Bildersets und Datenbanken anderer Forschergruppen

Name des Verfahrens (Abkürzung)	Kurzbeschreibung	Referenzliteratur
AR-Database	Etwa 3000 Bilder von 116 Personen in insgesamt 13 unterschiedlichen Bedingungen	Martinez & Benavente (1998)
BU-3DFE	3D-Ansichten der sechs Basisemotionen + neutral von 100 Personen im Alter von 18-70 Jahren. Insgesamt 2.500 Scans	Yin, Wei, Sun, Wang & Rosato (2006)
BU-4DFE	Dynamische 3D-Aufnahmen der sechs Basisemotionen von 101 Probanden	Yin, Chen, Sun, Worm & Reale (2008)
CAFE	80 Bilder in zwei Versionen, Basisemotionen + Zähne zeigend	Dailey, Cottrell & Reilly (2001), siehe auch: http://cseweb.ucsd.edu/users/gary/CAFE
CAS-PEAL	Fotos von 1040 Personen in sieben verschiedenen Bedingungen, keine Basisemotionen	Gao, Cao, Shan, Zhou, Zhang & Zhao, 2004
FACES	61 junge und 60 mittelalte und 58 alte Personen	Ebner, Riediger & Lindenberger (2010), siehe auch: http://faces.mpib-berlin.mpg.de
FEEL und Folgeversionen, u.a.: FEEL-M, areasFEEL, s-FEEL, i-FEEL	42 kaukasische und asiatische frontal fotografierte Personen	Kessler, Bayerl, Deighton & Traue, 2002; Kessler, Bachmayr, Walter, Hoffmann, Filipic & Traue, 2008; Hoffmann, Kessler, Eppel, Rukavina & Traue, 2010; Hoffmann, Traue, Bachmayr & Kessler, 2010
JACFEE/JACNeuF	56 farbige Fotografien von kaukasischen und japanischen Personen, pro Emotion 8 Personen	Von Matsumoto und Ekman, siehe Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh & Ton (1997)
KDEF	490 farbigen Bildern von 70 Menschen, 20 bis 30 Jahre, 7 Ausdrücke, unterschiedliche Winkel	Von Lundquist, Flykt & Öhmann (1998), siehe Goeleven, DeRaedt, Leyman & Verschuere (2008)
KFDB	Fotos von 1000 Koreanern, 7 Perspektiven, verschiedene Variationen	Hwang, Byun, Roh & Lee (2003)
Max Planck Institute for Biological Cybernetics Face Database	2D-Bilder von 200 Personen, pro Bilder 512 Referenzpunkte verfügbar. 2D-Bilder frei verfügbar. Inzwischen auch Videos verfügbar	Blanz & Vetter (1999), Cunningham, Kleiner, Bülthoff & Wallraven (2004), siehe auch: http://faces.kyb.tuebingen.mpg.de/ oder http://vdb.kyb.tuebingen.mpg.de
PoFA	110 schwarz-weiß Fotografien von sechs kaukasischen Männern und Frauen	Ekman & Friesen (1976)
RaFD	Farbige Fotografien von 72 Modellen, unterschiedliche Blickrichtungen und Winkel	Langer, Dotsch, Bijlstra, Wigboldus, Hawk & Knippenberg (2010)
UCDSEE	73 Bilder von 4 Personen (hell- und dunkelhäutig) zu Basisemotionen + Stolz, Scham und Peinlichkeit	Tracy, Robins & Schriber (2009)
THYMOS	Testverfahren zur Emotionserkennung in Gesichtern auf Grundlage des JACFEE/JACNeuF	Kessler, Retz-Junginger & Traue (1997)
VERT 160/VIEMER	VERT: Emotionserkennung bzgl. 4 Emotionen aus 160 Bildern. VIEMER: Messung des Erinnerns an Gesichtern	Hoheisel & Exner (2005)

2.4 Die Überprüfung der psychometrischen Qualität eines Testverfahrens

Die Testtheorie liefert Gütekriterien zur Überprüfung der Qualität psychometrischer Verfahren. Hauptanwendungsfeld ist hierbei die Testkonstruktion. Da diese Gütekriterien bei der Erstellung der „Pictures of Facial Affect-Ulm“ Berücksichtigung finden sollen, wird in der Folge ein kurzer Überblick gegeben. Als Hauptgütekriterien für ein Testverfahren gelten:

- Objektivität: Hier wird gefragt, ob Ergebnisse unabhängig von Einflüssen der Untersucher oder der Untersuchungssituation bei Durchführung, Auswertung und Interpretation zustande gekommen sind. Unterschieden werden die Durchführungs-, die Auswertungs- und die Interpretationsobjektivität.
- Reliabilität: Reliabilität bezeichnet den Grad der Genauigkeit, mit der ein bestimmtes Kriterium (in diesem Fall die Emotionserkennungsfähigkeit) durch ein Testverfahren gemessen wird. Die Reliabilität gibt also an, wie gut man sich auf ein Testergebnis „verlassen“ kann. Die Messgenauigkeit lässt sich mit unterschiedlichen Verfahren überprüfen. Ein mögliches Vorgehen ist die Überprüfung der Retest-Reliabilität. Hierbei wird ein Test mindestens zweimal an derselben Stichprobe eingesetzt und der Grad der Übereinstimmung berechnet. Angestrebt werden Zusammenhänge von r = .80 und höher. Dieses Verfahren wird sinnvollerweise dann eingesetzt, wenn Stabilität eines Instruments über die Zeit hinweg vorausgesetzt werden kann. Ein weiteres Maß zur Überprüfung der Reliabilität ist die interne Konsistenz (oftmals ermittelt durch Cronbach’s alpha). Die interne Konsistenz ist ein Maß dafür, wie stark die Items einer Skala miteinander zusammenhängen. Die Reliabilität eines Testverfahrens kann außerdem mittels Paralleltest-Reliabilität oder Split-Half-Reliabilität bestimmt werden.
- Validität: Die Validität bezieht sich auf die Güte der Operationalisierung und gibt Anhaltspunkte dafür, ob ein Testverfahren das Konstrukt misst, das es erfassen soll. Es werden die Konstrukt- und die Kriteriumsvalidität voneinander unterschieden. Die Konstruktvalidität bezieht sich auf die Zulässigkeit von Aussagen aufgrund der Operationalisierung über das gesamte dahinter liegende Konstrukt. Dies ist in der Regel dann der Fall, wenn der Bedeutungsumfang des Konstruktes vollständig, präzise und nachvollziehbar abgebildet ist. Als empirische Indikatoren der Konstruktvalidität gelten die konvergente und diskriminante Validität. Die Kriteriumsvalidität bezieht sich auf den Zusammenhang zwischen den Ergebnissen des Messinstruments und einem empirischen

Kriterium (Lienert & Raatz, 1994; Schnell, Hill & Esser, 2005). Als empirische Indikatoren gelten die Übereinstimmungsvalidität und die prognostische Validität.

- Normierung: Mit Normierung ist die Anpassung eines psychodiagnostischen Instruments an eine Personengruppe gemeint (Krohne & Hock, 2007). Ziel ist es, für unterschiedliche Zielgruppen gültige Referenzwerte zur Verfügung stellen zu können. Voraussetzung hierfür ist die endgültige Auswahl aller relevanten Items beziehungsweise Stimuli. Der Ablauf ist somit in der Regel folgendermaßen: In der Entstehungsphase des Testverfahrens wird ein Itempool erstellt, damit ist eine umfassende Sammlung eventuell passender Fragen oder Stimuli gemeint. Aus diesem Itempool wird in der Folge eine erste Auswahl getroffen. Daraufhin wird eine repräsentative Versuchsgruppe mit der vorläufigen Auswahl konfrontiert. Nun gilt es, eine Normalverteilung nachzuweisen, die für die Emotionserkennung als Grundlage angenommen wird. Hiermit wird ein bestimmter Typ der Wahrscheinlichkeitsverteilung beschrieben, der vielen Testverfahren (beispielsweise Intelligenztests) zugrunde liegt und von der angenommen wird, dass sie biologische Prozesse adäquat abbildet (Bortz, 2005). Wird eine Normalverteilung nicht erreicht, sondern sprechen die Ergebnisse beispielsweise eher für einen Deckeneffekt, so muss die Auswahl der Items beziehungsweise Stimuli überdacht werden.

Die Beachtung dieser psychometrischen Größen ermöglicht die Qualitätssicherung für psychodiagnostische Verfahren.

2.5 Ziele der vorliegenden Arbeit

Ziel der vorliegenden Arbeit ist es, ein neues valides und reliables Bilderset mit dem Namen PFA-U (kurz für „Pictures of Facial Affect-Ulm“) zu erstellen, das zur Erfassung der Emotionserkennungsleistung aus Gesichtern eingesetzt werden kann. Fokus soll hierbei einerseits auf die sorgfältige testtheoretische Überprüfung gelegt werden, andererseits auf die Maximierung der ökologischen Validität durch Bereitstellung von fünf unterschiedlichen Perspektiven des emotionalen Ausdrucks. Zusätzlich sollen Natürlichkeit und Intensität des Emotionsausdrucks sowie die Attraktivität des fotografierten Gesichtes Berücksichtigung finden.

Vor Erstellung Überprüfung des neuen Bildersets wurden folgende Anforderungen definiert:

- Minimierung ethnischer Unterschiede im Stimulusmaterial durch das Fotografieren von kaukasischen Personen.
- Definition der mimischen Muster der Basisemotionen nach EmFACS als ein objektives Maß des mimisch-emotionalen Ausdrucks.
- Gleichverteilung des Geschlechts (Hyde & Mertz, 2009; Hall, Carter & Horgan, 2000)
- Beschränkung des Altersbereiches auf 18 bis 35 (Anastasi & Rhodes, 2006; Ebner, 2008; Ebner & Johnson, 2009).
- Simultane Aufnahmen aus fünf Perspektiven, um den Alltagsbedingungen bei der Emotionserkennung näherzukommen und die der Emotionserkennung zugrundeliegenden (neuronalen) Prozesse noch besser zu verstehen.
- Auswahl der emotionalen Gesichtsstimuli mit dem Ziel, eine größtmögliche Übereinstimmung der durchschnittlichen Erkennungsraten sowie der Einschätzung von Intensität und Natürlichkeit des Gesichtsausdrucks sowie Attraktivität der fotografierten Person.
- Überprüfung der psychometrischen Qualität des Bildersets.
- Einbettung des neuen Stimulusmaterials in das bestehende Testsystem FEEL. Angestrebt werden eine Revision (FEEL-R) mit 48 Frontalbildern und eine Perspektivversion (FEEL-Rpersp) mit über die Perspektive variierenden emotionalen Gesichtsausdrücken von mindestens 30 Darstellern.

Eine derart ausführliche testtheoretische Untersuchung mimischen Stimulusmaterials ist bisher noch nicht vorgenommen worden. Darin gründet die Eigenständigkeit dieser Dissertation.

3. MATERIAL UND METHODIK

Dieser Abschnitt stellt die Grundlagen zur Entwicklung und Überprüfung des neuen Bildmaterials dar. So wird zunächst das Facial Action Coding System als objektive Grundlage zur Erfassung des mimischen Ausdrucks erläutert. Sodann wird der FEEL-Test als Ausgangspunkt der angestrebten Revision und Erweiterung des Testverfahrens ausführlich beschrieben. Im Anschluss daran wird die Erstellung des Bildersets PFA-U beschrieben.

3.1 Das Facial Action Coding System

Das menschliche Gesicht besteht aus einer etwa daumendicken Schicht aus Knorpeln, Adern, Nerven und 43 Muskeln über einem starren Knochengerüst (Eberle, 2004). Von diesem Knochengerüst, nämlich dem Schädel, vermag der Mensch lediglich den Unterkiefer frei zu bewegen. Dennoch sendet die dort vorhandene Muskulatur vielfältige Signale. So ist es dem Menschen möglich etwa die Freude über ein Geschenk genau wie die Überraschung bei einem unerwarteten Wiedersehen eines guten Freundes allein mit Hilfe der Mimik auszudrücken. Man weiß auch schon lange, dass ein Mensch noch meterweit entfernt die Angst aus einem Gesicht deuten und so schnellstmöglich adäquat auf die potentiell gefährliche Situation reagieren kann (Ekman, Brattesani, O'Sullivan & Friesen, 1979).

Das Facial Action Coding System (kurz: FACS) gilt als Basis für eine objektive Beschreibung der Mimik. Entwickelt wurde es von Ekman und Friesen (1978; Ekman, Friesen & Hager, 2002), welche in den 60er Jahren von Tomkins in die Grundlagen der Gesichteranalyse eingeführt wurden. Sie erkannten, welchen Gewinn es für die Emotionsforschung darstellen würde, wenn man emotionale Gesichtsausdrücke noch wesentlich feiner als zur damaligen Zeit möglich, analysieren könnte. Beim FACS liegt der Fokus auf dem, was beobachtbar ist. Stattdessen hätten Ekman und Friesen auch EMG-Daten (Elektromyogramm) oder die Durchblutung des Gesichts heranziehen können, was sicherlich in einigen Punkten vorteilhaft gewesen wäre. Beispielsweise ist eine (noch) objektivere Erfassung von Daten denkbar. Der große Nachteil wäre aber gewesen, dass Fotografien oder Filme nicht hätten ausgewertet werden können, weil die Person immer an die entsprechenden Gerätschaften hätte angeschlossen sein müssen (Ekman, Friesen & Hager, 2002). Somit kann mit dem FACS

flexibler mit vorliegenden Daten umgegangen werden und Codierer und zu codierende Person müssen sich nicht zwangsläufig treffen. Des Weiteren ist denkbar, dass physiologische Daten gar nicht so objektiv sind wie angenommen, sondern die Situation im Labor mit Elektroden im Gesicht für den Probanden oftmals artifiziell ist und dies einen Einfluss auf die Daten hat. Das FACS wurde explizit dafür entwickelt, visuell beobachtbare Unterschiede im Gesichtsausdruck einordnen zu können und eine gemeinsame Nomenklatur zu implementieren. Sowohl statische als auch dynamische Stimuli können bearbeitet werden.

Das Facial Action Coding System basiert auf der Einteilung in so genannte Action Units (kurz: AU für den Singular und AUs für den Plural). Der gesamten Gesichtsmuskulatur werden 46 solcher AUs ezugeteilt. Hierbei ist zu beachten, dass eine AU nicht zwingend nur einen einzelnen Muskel umfassen muss, sondern auch eine Muskelgruppierung umfassen kann, die häufig gemeinsam kontrahiert. Durch die Angabe unterschiedlicher AUs kann nahezu jeder Gesichtsausdruck beschrieben werden. Eine Übersicht über einige Muskeln des menschlichen Gesichts findet sich in Abbildung 4.

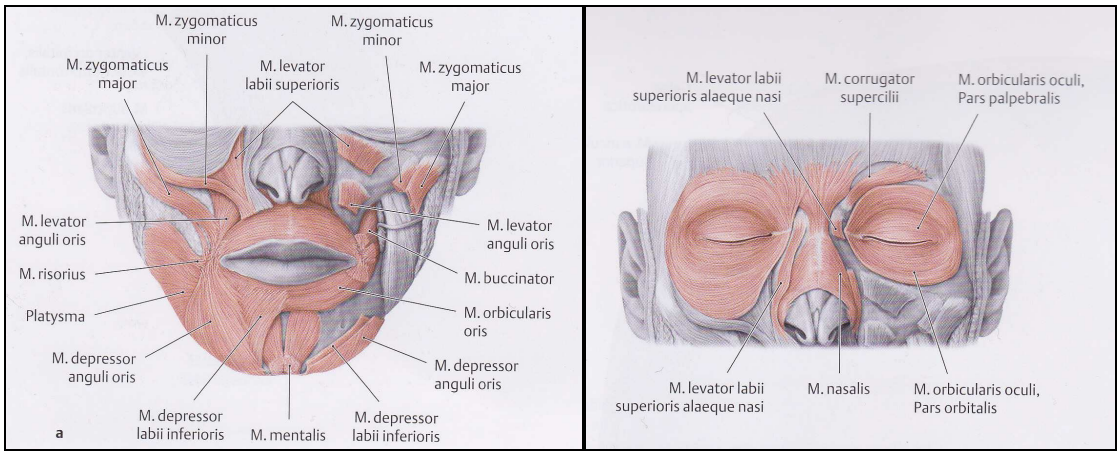


Abb. 4: Gesichtsmuskulatur (aus Prometheus, Lern-Atlas der Anatomie, S. 46)

Auch jede Basisemotion verursacht einen spezifischen Gesichtsausdruck (Ekman & Friesen, 1982). Im Umkehrschluss ist eine einwandfreie Zuordnung von Gesichtsausdrücken zu den Basisemotionen möglich. Ekman und Friesen führen an, dass unter anderem individuelle Unterschiede in der Gesichtsphysiognomie dazu führen, dass Unterschiede zwischen verschiedenen Personen entstehen (Ekman & Friesen, 2003; Ekman, 2003). Auch ist die Ausprägung mit der die betreffende Emotion gezeigt wird, von Relevanz. Bei einer

geringeren Intensität ist es möglich, nicht alle postulierten Muskelpartien des Gesichts anzuspannen (Ekman & Friesen, 2003). Bei fast allen Basisemotionen scheint es somit nicht nur „den einen“ adäquaten Gesichtsausdruck zu geben, sondern gleichwertige Alternativen, mit denen diese Basisemotion ausgedrückt werden kann. Ekel etwa kann mit geöffnetem oder geschlossenem Mund dargestellt werden. Eine Auflistung der protoypischen Gesichtsausdrücke der Basisemotionen auf Grundlage der Action Units findet sich im EmFACS-Manual (Ekman & Friesen, 1982). In der Regel wird für das Zuordnen zu den Basisemotionen folgende Einteilung verwendet (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & Knippenberg, 2010; Ekman, 2007a):

Ärger: 4CDE + 5CDE + 7CDE + 17 + 23 + 24

Ekel: 9 + [10 und/oder 16] + 19 + 26

Angst: 1 + 2 + 4 + 5ABCDE + 7 + 20ABCDE + 26

Freude: [6 und/oder 7] mit 12CDE

Trauer: 1 + 4 + [6 und/oder 7] + 15ABC + 64

Überraschung: 1CDE + 2CDE + 5AB + 26

Die Zahlen verweisen auf die AU, die Buchstaben auf die gezeigte Intensität. Beispielsweise entspricht die AU 12 einem Lächeln, die Kombination der AUs 1+2 einem Hochziehen der Augenbrauen. Es würde an dieser Stelle zu weit führen, die Entsprechungen aller AUs zur Mimik auszuführen. Hier sei deswegen auf den Artikel von Rosenberg in dem Buch „What the face reveals“ (Rosenberg, 1997) oder den „Investigator’s Guide“ zum FACS (Ekman, Friesen & Hager, 2002) verwiesen.

Das Codieren nach FACS kann potentiell jeder erlernen. Die Autoren raten an, dass zwei Personen FACS erlernen und so bei der Beurteilung eine Interrater-Reliabilität berechnet werden kann. Grundsätzlich ist die im „Investigator’s Guide“ mit 0,756 bis 0,822 angegebene Reliabilität (Ekman, Friesen & Hager, 2002) als hoch zu bewerten. Das sind Raten, die technische Geräte (beispielsweise eigens zur Mimikanalyse entwickelte Software) heute noch nicht erbringen können (Lien, Kanade, Cohn & Li, 2000), wobei hier schon enorme Fortschritte erzielt wurden (Cohn, 2010; Littlewort, Whitehill, Wu, Fasel, Frank, Movellan & Bartlett, 2011).

Das FACS wurde eigentlich nur zu dem Zwecke entwickelt, um statisches und dynamisches Material rückwirkend bewerten zu können. Bei ihrer Arbeit dazu fiel Ekman und Friesen dann jedoch auf, dass man die Personen auch einfach nur die bestimmten Muskelgruppen

anspannen lassen kann (z. B. für Trauer) und sich daraufhin bestimmte Gefühlsregungen einstellen (Ekman, 2007b). Das hierzu entwickelte und standardisierte Vorgehen wurde zusammenfassend als „Directed Facial Action Task“ beschrieben (Ekman, 2007b). Die in diesem Zusammenhang entwickelten Instruktionen bilden unter anderem auch die Grundlage für die Operationalisierung zur Erstellung der emotionalen Gesichtsausdrücke im PFA-U-Bildersatz.

3.2 Der FEEL-Test

Der FEEL-Test (kurz für Facially Expressed Emotion Labeling) baut auf den Ideen des THYMOS (kurz für Tachistoskopische Analyse emotionaler Sensitivität) auf (Kessler, Retz-Jungiger & Traue, 1997; Kessler, Bayerl, Deighton & Traue, 2002). Beide Verfahren wurden in der Sektion Medizinische Psychologie der Universität Ulm entwickelt. Der als PC-Programm (initial von Pierre Bayerl, nachfolgend re-programmiert von Holger Hoffmann) realisierte FEEL-Test misst die Fähigkeit einer Person, emotionale Gesichtsausdrücke zu erkennen (Emotionserkennungsleistung), indem sie sie ausgewiesenen Basisemotionen zuordnet. Grundlage der gezeigten Emotionen sind die Basisemotionen nach Ekman. Als Stimulusmaterial wurden Fotos aus dem JACFEE-Bildersatz (= Japanese and Caucasian Facial Expressions of Emotion) verwendet (Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh & Ton, 1997). Enthalten sind in diesem Bilderset Fotografien der sieben Basisemotionen nach Ekman sowie die korrespondierenden neutralen Gesichtsausdrücke (dieser Bildersatz wird „Japanese and Caucasian Neutral Faces“, kurz: JACNeuF genannt). Es handelt sich um farbige, frontal aufgenommene Portrait-Fotos. Pro Emotion gibt es je acht Darstellungen jeweils unterschiedlicher Personen. Die Fotografie eines neutralen Gesichtsausdrucks ist für jede Person vorhanden. Die Hälfte der abgebildeten Stimuluspersonen sind japanischer und die andere Hälfte kaukasischer Herkunft. Angenommen wurde hier, dass die unterschiedlichen Ethnien gemäß der Universalismushypothese (siehe 2.2.2) für die Erkennungsfähigkeit der dargestellten Emotion keine oder zumindest keine wesentliche Rolle spielen. Frauen und Männer sind gleich häufig vertreten.

Verschiedene Studien belegen die hohe Reliabilität der JACFEE-Bilder und eine große Übereinstimmungsrate zwischen unterschiedlichen Betrachtern (beispielsweise Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh & Ton, 1997; Matsumoto & Ekman, 1989). Eine zumindest ausreichende Validität der Bilder ist deswegen anzunehmen, da diese nach dem

bereits im ersten Abschnitt vorgestellten Facial Action Coding System codiert wurden. Somit ist davon auszugehen, dass die dargestellten Gesichter auch tatsächlich die gewünschten Emotionen widerspiegeln. Verwendet wurden für den Test lediglich sechs der Basisemotionen, die Emotion „Verachtung“ wurde nicht mit in den FEEL-Test aufgenommen (siehe 2.2.1). Die für den FEEL-Test aus dem JACFEE- beziehungsweise JACNeuF-Bildersatz ausgewählten Fotografien wurden in der Folge in Helligkeit und Größe aneinander angeglichen.

Der Ablauf des Tests ist folgendermaßen:

Vor dem eigentlichen Testbeginn steht ein Übungsdurchgang. Den Probanden werden sechs Bilder – eines je Basisemotion – gezeigt, welche dann von ihnen benannt werden sollen. Hierzu werden die Basisemotionen in Schriftform auf dem Bildschirm gezeigt, je eine davon kann durch Anklicken ausgewählt werden. Nach der Beurteilung der Übungsbilder erhalten die Probanden Rückmeldung über ihre Antwort (richtig oder falsch und die Antwortzeit). Dieser Vorlauf dient vor allem dazu, die Probanden mit dem Ablauf vertraut zu machen und so sicherzustellen, dass auftretende Fragen und Probleme noch vor dem Haupttest geklärt werden können. Dies hat zur Folge, dass Fehler im Haupttest nicht auf mangelnde Kenntnisse im Umgang mit dem Instrument zurückzuführen sind. Im Haupttest werden den Probanden dann 42 Fotos emotionaler Gesichter zur Bewertung vorgegeben. Die Präsentation des Stimulus und dessen Zuordnung zu einer der Basisemotionen verläuft während des gesamten Tests, bis auf die Rückmeldung, analog zu dem Vorgehen während der Übungsphase. Zunächst erscheint das Foto einer Person mit neutralem Gesichtsausdruck. Zeitgleich ertönt ein kurzes akustisches Signal zur Aufmerksamkeitsfokussierung des Probanden. Es folgt eine Pause von einer Sekunde Dauer. In dieser Zeit ist der Bildschirm grau, nur im oberen Drittel findet sich ein schwarzer Fixationspunkt. Nach der Pause ereignet sich die kurzzeitige Präsentation des Stimulus für 300 ms. Da im Sinne Ekmans bei der Emotionserkennung von einer Basisfunktion ausgegangen wird, würde eine zu lange Präsentationszeit zur Folge haben, dass quasi jeder die betreffende Emotion erkennen könnte. Bei dem emotionalen Gesichtsausdruck handelt es sich um dasselbe Gesicht, das vorher mit neutralem Gesichtsausdruck dargestellt wurde. Diese Zuordnung des emotionalen Ausdrucks zu einer der Basisemotionen stellt die eigentliche Anforderung an den Probanden bei der Durchführung dar. Als Antwortformat wurde ein Forced-Choiced-Format gewählt. Es erscheint eine Übersicht der sechs möglichen

Auswahlmöglichkeiten, von denen exakt eine durch Mausklick auf das entsprechende Feld gewählt werden soll. Der Proband hat maximal zehn Sekunden Zeit, sich für eine der Antwortmöglichkeiten zu entscheiden. Zwischen den Darstellungen der emotionalen Gesichtsausdrücke liegt eine zufällig variierte Pause zwischen vier und sechs Sekunden, um Gewöhnungseffekte zu vermeiden. In dieser Zeit ist der Bildschirm, wie bereits beschrieben, bis auf den schwarzen Fixaktionspunkt grau. Die Reihenfolge der Bilder wird für jeden Durchgang zufällig neu bestimmt, so dass Reihenfolgeeffekte vermieden werden können. Eine grafische Veranschaulichung des Testablaufs findet sich in Abbildung 5.

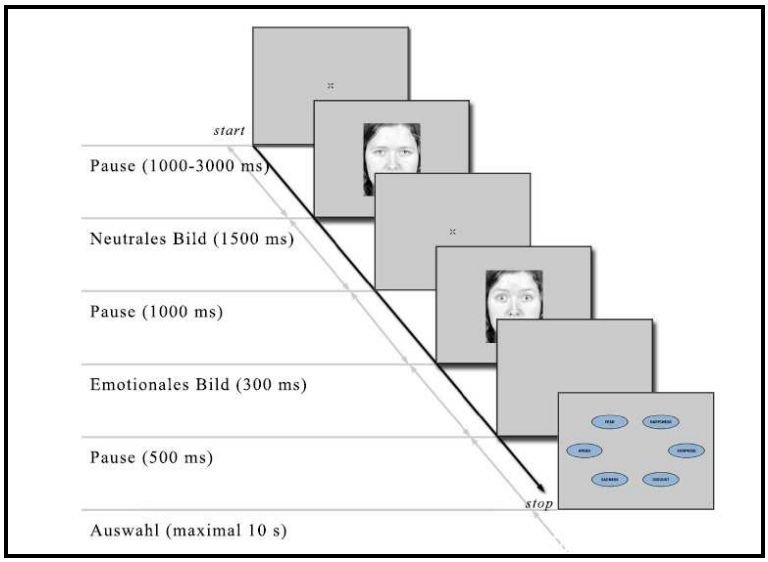


Abb. 5: Schematische Darstellung des Ablaufs des FEEL-Tests

Wurde der Test erfolgreich beendet, hat der Testleiter direkt Zugriff auf die Gesamtpunktzahl, auch FEEL-Score genannt, wie auch auf die Punktzahl pro Emotion, die durchschnittliche Beantwortungszeit und eine numerische und grafische Veranschaulichung der Verwechslungen bei der Benennung der Emotionen des Probanden. Da man pro richtig erkannter Emotion einen Punkt erhält, kann somit eine maximale Punktzahl von 42 als FEEL-Score erreicht werden. Die erzeugten Daten eines jeden Probanden sind kompatibel mit gängigen Statistikprogrammen wie beispielsweise SPSS und können ohne großen Aufwand in dieses importiert werden, um weitere Berechnungen vornehmen zu können. Dieser Datensatz umfasst dann beispielsweise auch die Antwortzeiten für jedes einzelne Bild, so dass tiefergehende Analysen möglich sind.

Ein Vorteil des FEEL-Tests ist, dass der Korpus dieses Verfahrens – die Software – leicht variiert werden kann. Zum Beispiel ist es möglich, die Darbietungsdauer der Stimuli oder die Sprache der Bildschirmausgabe zu ändern. Weiterhin ist es möglich, andere Bildersätze einzulesen und als Stimulusmaterial zu verwenden – so wie es für die Bilder der PFA-U angedacht ist. Kessler (Kessler, Bayerl, Deighton & Traue, 2002) gibt für den FEEL-Test an, dass im Rahmen einer ersten Studie mit N = 77 gesunden Probanden eine Reliabilität von r = .76 für Cronbach's alpha erzielt werden konnte. Die Erkennungsraten lagen bei: Trauer 70%, Angst 71%, Ekel 80%, Überraschung 84%, Freude 87% und Ärger 94%. Somit ist das Testverfahren hinreichend exakt, um sowohl bei der Messung von Gruppenunterschieden als auch interindividuellen Unterschieden eingesetzt werden zu können. Dies eröffnet ein breites Einsatzfeld für den FEEL-Test. Als valides und reliables Instrument ist er beispielsweise in neurobiologischen Studien einsetzbar, wie auch bei der Unterscheidung klinischer Gruppen hinsichtlich der Fähigkeit zur Erkennung von Emotionen aus Gesichtern.

Neben der herkömmlichen FEEL-Version wurden auch zahlreiche Weiterentwicklungen durchgeführt. Beispielsweise wurden Morphings erstellt (Kessler, Hoffmann, Bayerl, Neumann & Traue, 2005). Beim Morphing werden zwischen zwei Einzelbildern Zwischenübergänge berechnet, so dass in diesem Fall die Intensität des emotionalen Ausdrucks variiert werden konnte (Hoffmann, Kessler, Eppel, Rukavina & Traue, 2010). Die Aneinanderreihung dieser Bilder ermöglichte die Erstellung dynamischen Bildmaterials (Hoffmann, Traue, Bachmayr & Kessler, 2010). Auch Chimären wurden computerbasiert erstellt, bei denen jeweils eine Gesichtshälfte gespiegelt wurde (Kessler, Bachmayr, Walter, Hoffmann, Filipic & Traue, 2008).

Die Stimuli des PFA-U sollen im Rahmen dieser Arbeit in dieses standardisierte und gut untersuchte Instrument implementiert werden. Angestrebt wird einerseits eine Revision des bisher bestehenden FEEL-Tests. Hierbei werden die Frontalbilder durch die Frontalbilder der PFA-U ersetzt (FEEL-R). Zusätzlich soll eine Version entstehen, die die unterschiedlichen Perspektiven berücksichtigt (FEEL-Rpersp). Zwangsläufig ergibt sich aus der Verwendung neuer Stimuli die Notwendigkeit der Überprüfung testtheoretischer Aspekte, welche Fokus dieser Arbeit sind. Hierbei soll die Software des FEEL-Test in an die Studienfrage angepasster Form verwendet werden. Die Veränderungen im Ablauf werden jeweils bei den einzelnen Studien beschrieben.

3.3 Erstellung Bildermaterial PFA-U

Da Basisemotionen eindeutig an ihrer mimischen Aktivität erkennbar sind, können mit entsprechend hergestellten Gesichtsstimuli verschiedene Aspekte der Verarbeitung von wahrgenommenen Emotionen und der Emotionsregulation untersucht werden (Kessler, Bayerl, Deighton & Traue, 2002; Ebner & Johnson, 2009; Goeleven, De Raedt, Leyman & Verschuere, 2008). Grundlage hierfür ist das vorangehend beschriebene FACS (Ekman, Friesen & Hager, 2002) beziehungsweise EmFACS (Ekman & Friesen, 1982). Auf Grundlage der Zuordnung von AUs zu Basisemotionen können standardisierte Bildersets zur Erfassung der Erkennungsleistung von emotionalen Ausdrücken entwickelt werden. Methodisch und technisch müssen bei der Erstellung von mimischen Stimuli mehrere Faktoren beachtet werden, beispielsweise die Beleuchtung, die Schärfe des Bildes, die Identität der fotografierten Person, der kulturelle Hintergrund, die Qualität des emotionalen Ausdrucks, die Perspektive, die Hautfärbung, das Alter und potentielle Verdeckungen, sowie ablenkende Elemente im Gesicht (Haare, Bart, Brille, Piercings, Narben et cetera).

Im Folgenden soll beschrieben werden, wie das Bildmaterial, welches die Grundlage der im nächsten Kapitel beschriebenen Studien bildet, erstellt wurde. Die zu erfüllenden Anforderungen wurden in 2.4 ausführlich beschrieben. So wurde hier beispielsweise ein simultanes Fotografieren des emotionalen Ausdrucks aus fünf Perspektiven verwirklicht. Ziel dieses Vorgehens ist es, die ökologische Validität des statischen Bildmaterials zu erhöhen. Die Analyse der Erkennungsleistung aus unterschiedlichen Perspektiven kann wichtige zusätzliche Informationen zum Prozess der Emotionserkennung bringen.

3.3.1 Rekrutierung der Darsteller

Nach Ebner, Riediger und Lindenberger (2010) wurde die Idee umgesetzt, Schauspielschüler zu fotografieren, da diese – so die Überlegung – im Ausdruck von Emotionen trainierter seien. Im September 2009 wurde mit dem Direktor der Akademie für darstellende Kunst (kurz: adk) in Ulm, Herrn Reimann, Kontakt aufgenommen. Die adk ist ein staatlich anerkanntes Berufskolleg für die Ausbildungsgänge Schauspiel, Regie, Theaterpädagogik und Dramaturgie. Alle Studenten durchlaufen dieselbe einjährige (schauspielerische)

Grundausbildung, bevor sich eine dreijährige Fachausbildung anschließt. Somit haben alle Studenten nach diesem Jahr Grundkenntnisse im Bereich Schauspiel erworben.
Im Oktober 2009 fand ein Informationstermin für die Studenten statt, über eine Interessentenliste wurden die potentiellen Darsteller kontaktiert.
Einige Schauspielschüler – unter diese Benennung fallen im Folgenden alle Studenten der adk – konnten auf diesem Wege für die Aufnahmen gewonnen werden. Allerdings zeigte sich bald, dass auf diesem Wege nicht genügend Personen rekrutiert werden konnten, so dass zusätzlich Personen ohne Vorerfahrung im Schauspielfach herangezogen wurden. Wie Probeaufnahmen zeigten, erwiesen sich unerfahrene Personen als ebenso geeignet.

3.3.2 Training der Darsteller

Die Personen, mit denen ein Fototermin vereinbart werden konnte, erhielten ein Trainingsmanual (siehe Anhang A). Neben einigen theoretischen Grundlagen und Erläuterungen zu den Hintergründen des Forschungsbestrebens war Hauptanliegen dieses Manuals, den Probanden das Darstellen der Basisemotionen näherzubringen. Hierzu wurde folgendes Vorgehen angewandt: sowohl für den neutralen Gesichtsausdruck als auch für die sechs Basisemotionen Freude, Trauer, Ekel, Angst, Überraschung und Ärger gab es gesonderte Kapitel. In einer kurzen Einleitung wurde etwas über den Hintergrund und vermuteten Sinn dieses Gesichtsausdrucks erläutert, beispielsweise, dass die Mundbewegung bei Ekel auf eine Auswurfreaktion der Nahrung zurückzuführen sein könnte. Danach folgte eine detaillierte Beschreibung des jeweiligen Gesichtsausdrucks in Anlehnung an das Facial Action Coding System. Die Darsteller wurden gebeten anhand dieser Anleitung den jeweiligen Gesichtsausdruck vor dem Spiegel zu üben und mit dem zur Verfügung gestellten Bildmaterial abzugleichen. Hierzu war dem Manual von jedem Gesichtsausdruck ein prototypisches Bild in DIN A 4-Größe beigefügt worden – Bildmaterial, das dem JACFEE-Bildersatz entnommen wurde. Zur Vereinfachung der Anweisungen wurden diese jeweils in drei relevante Bereiche unterteilt: Augenpartie, Nasenpartie und Mundpartie.
In Abbildung 6 findet sich ein Beispiel für solch eine Anweisung.

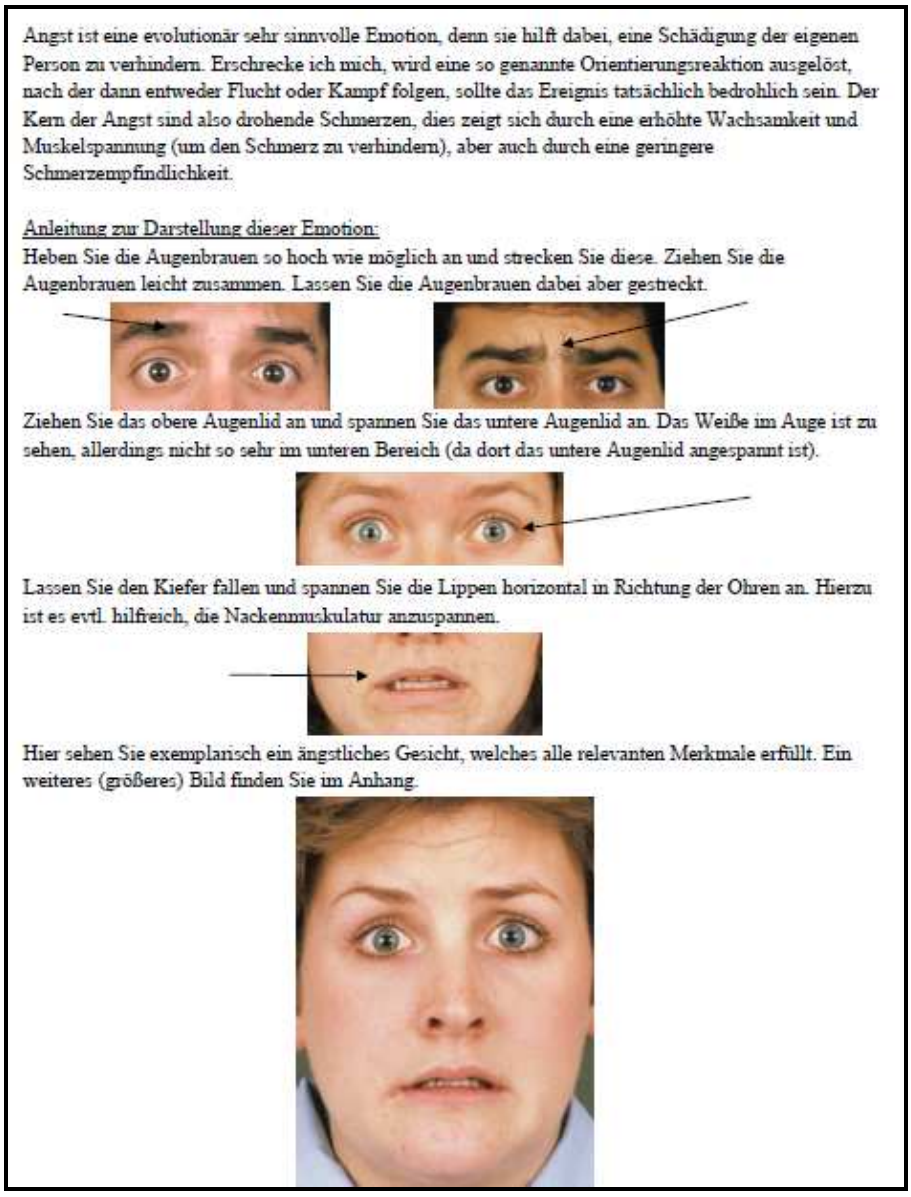


Abb. 6: Beispiel für die Anweisung nach EmFACS für einen spezifischen emotionalen Ausdruck (hierAngst)

Des Weiteren bekamen die Darsteller eine kurze Geschichte zum Hineinversetzen in die Emotion zur Verfügung gestellt. Diese wurden für die Emotionen Freude, Ekel, Ärger und Überraschung aus der Studie von Karin Pilz (2002) entnommen, die fehlenden Emotionen ergänzt. Der Hintergrund dieses Vorgehens ist, dass die Darsteller dazu angehalten werden sollten, ihr Gesicht möglichst differenziert nach Anweisung zu manipulieren, um einen Gesichtsausdruck entstehen zu lassen, der in erster Linie den Vorgaben des EmFACS entspricht, nicht einem natürlichen Emotionsausdruck. Es wurde angenommen, dass dies nicht für alle Personen gleichermaßen möglich sei, weswegen diesen das Darstellen dadurch erleichtert werden sollte, dass sie sich zunächst in einen natürlichen Emotionszustand

begeben und dann dem geforderten Gesichtsausdruck anzunähern hatten. Ein Beispiel für eine solche Geschichte findet sich in Abbildung 7. Bei diesem Vorgehen wurden unter anderem die Überlegungen von Carroll & Russell (1997) berücksichtigt, die bei der Analyse von Filmsequenzen herausfanden, dass die nach EmFACS postulierten prototypischen Gesichtsausdrücke im Alltag eher selten vorzukommen scheinen.

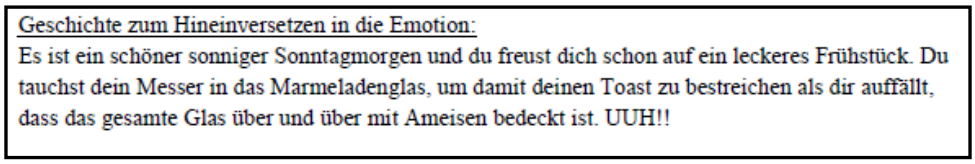

Geschichte zum Hineinversetzen in die Emotion:
Es ist ein schöner sonniger Sonntagmorgen und du freust dich schon auf ein leckeres Frühstück. Du tauchst dein Messer in das Marmeladenglas, um damit deinen Toast zu bestreichen als dir auffällt, dass das gesamte Glas über und über mit Ameisen bedeckt ist. UUH!!

Abb. 7: Beispiel für eine emotionsinduzierende Geschichte (hier Ekel)

Die Teilnehmer wurden aufgefordet, täglich etwa eine Stunde zu üben, um eine gewisse Vertrautheit mit dem eigenen Gesicht zu bekommen. Außerdem bot dies die Gelegenheit, etwaig auftretende Probleme im Vorfeld zu besprechen.
Um die Forschungsergebnisse bezüglich des Duchenne-Lächelns zu berücksichtigen (Ekman, Davidson & Friesen, 1990), sollten bei der Emotion Freude bevorzugt natürliche Gesichtsausdrücke fotografiert werden. Um eine Standardisierung des Manuals zu gewährleisten, sind auch für Freude beide Anleitungen enthalten.

3.3.3 Erstellung der Aufnahmen und Weiterverarbeitung des Bildmaterials

Mit der Erstellung des Stimulusmaterials wurde im November 2009 in Räumlichkeiten der Universität Ulm begonnen.
Die Darsteller saßen in einem Abstand von 1,20 m zu den fünf Kameras auf einem höhenverstellbaren Stuhl. Dies war notwendig, da die Kameras auf einer Höhe von ebenfalls 1,20 m auf Stativen fixiert worden waren, die Darsteller aber unterschiedliche Körpergrößen aufwiesen. Als Hintergrund wurde eine tragbare weiße Leinwand gewählt. Vorteil war, dass diese in eine gebogene Form gebracht werden konnte. Wie Abbildung 8 zu entnehmen ist, konnte somit ein gleichbleibender Abstand der zu fotografierenden Person zum Hintergrund gewährleistet werden. Die Kameras, fünf Spiegelreflexkameras der Marke Olympus (Typ E-420), waren in einem Viertelkreis um das Gesicht des Darstellers angeordnet, jeweils mit einem Abstand von 22,5 Grad. So konnten Aufnahmen von der Seite, aus der frontalen Ansicht und drei Zwischenschritte fotografiert werden. Der Aufbau erfolgte so, dass Bilder

der linken Gesichtshälfte aufgenommen werden konnten, da es Vermutungen darüber gibt, dass die linke Gesichtshälfte expressiver sei als die rechte (Sackheim, Gur & Saucy, 1978; Kessler, Bachmayr, Walter, Hoffmann, Filipic & Traue, 2008). Zur Beleuchtung des Raums wurde neben dem vorhandenen Kunstlicht aus Neonröhren ein professioneller Strahler verwendet (Multiblitz), welcher (siehe Abbildung 8 + 9) hinter Kamera 3 positioniert wurde. Durch einen angebrachten Schirm wurde die Schattenbildung minimiert.

Nach dem Eintreffen der Darsteller wurden diese nochmals über die Ziele der Studie aufgeklärt und unterschrieben eine Einwilligungserklärung, mit der sie ihre freiwillige Teilnahme versicherten und sich mit der Nutzung des Bildmaterials einverstanden erklärten (Ethikvotum 245/08_UBB/se). Danach wurde ihnen das Setup gezeigt und erklärt. Die Darsteller mussten abgepudert werden, um glänzenden Gesichtspartien bei den Aufnahmen entgegenzuwirken. Falls nötig, mussten die Darsteller sich abschminken und ihren Schmuck abnehmen. Es wurden bei allen Darstellern, soweit nötig, die Haare mittels Haarklammern aus dem Gesicht entfernt. Diese Maßnahmen waren notwendig, um eine möglichst hohe Standardisierung sicherzustellen. Alle Darsteller trugen einen weißen Umhang über ihrer normalen Kleidung und wurden im Vorfeld gebeten, für den Tag der Aufnahmen eine Oberbekleidung zu wählen, die den Halsbereich frei lässt.

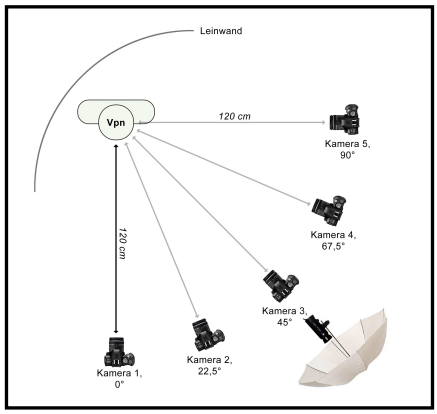


Abb. 8: Setup des Fotostudios in grafischer Darstellung

Nachdem die Vorbereitung abgeschlossen war, setzen sich die Darsteller auf den bereitgestellten Drehstuhl. Zur Überprüfung des eigenen Gesichtsausdrucks während des gesamten Shootings möglich war direkt unter Kamera 1 eine Videokamera angebracht worden, die das aufgenommene Bild über einen Beamer an eine große Leinwand übertrug. So konnten die Darsteller gleichzeitig Kamera 1 fixieren, wozu sie während des gesamten

Shootings angehalten wurden, um ihren Gesichtsausdruck zu überprüfen. Der Fotograf konnte zusätzlich während der Aufnahmen korrigierend eingreifen. Es handelt sich hierbei um die Autorin selbst, eine zertifizierte FACS-Coderin. Bei den Aufnahmen wurde nach dem Manual vorgegangen, mit dem die Darsteller im Vorfeld auch geübt hatten. Alle emotionalen Ausdrücke wurden mehrfach aufgenommen, um im Nachhinein auswählen zu können. Die Aufnahmen erfolgten für alle fünf Perspektiven simultan durch einen Funkauslöser (Typ Oly-1 der Firma Hama). Die Darsteller erhielten eine Aufwandsentschädigung von 50 Euro.

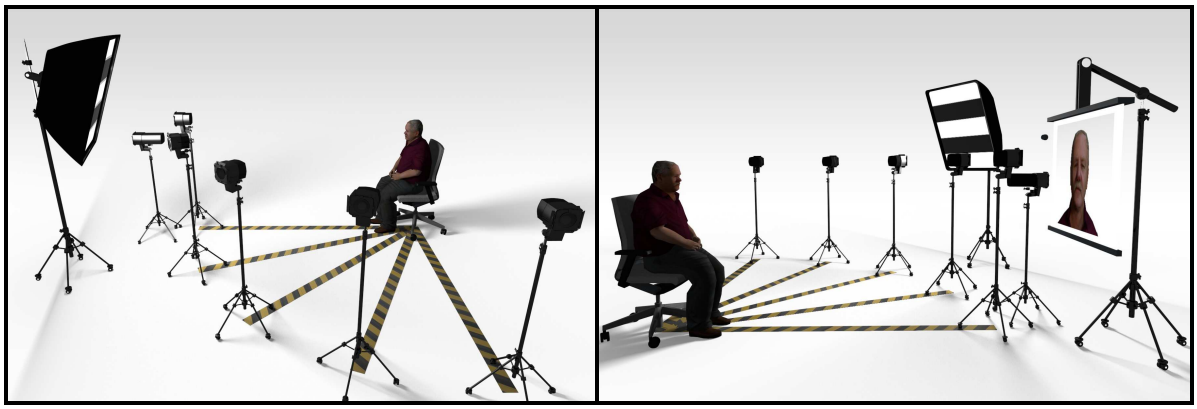


Abb. 9: Setup des Fotostudios aus unterschiedlichen Perspektiven

Die Aufnahmen dauerten mit Vorbereitung und Verabschiedung etwa eine Stunde. Die Anonymität aller verwendeten Daten wurde durch Vergabe eines Codes gewährleistet.

Insgesamt wurden 65 Personen (33 Frauen und 32 Männer) fotografiert. Hierbei wurden 2810*5 = 14050 Aufnahmen erstellt, berücksichtigt man die Aufnahmen aus den verschiedenen Perspektiven. Im Mittel entspricht dies 43 Bildern pro Person (bezogen auf eine Perspektive).

In einem ersten Schritt wurde die technische Qualität der Bilder überprüft: Wurde ein entsprechender mimischer Ausdruck fotografiert, waren die Augen offen, gab es Bewegungsartefakte und war die Bildschärfe befriedigend? Nach dieser Selektion verblieben 514 Bilder von 50 Personen (je 25 Männer und 25 Frauen; mittleres Alter M = 23,64; SA = 4,29) für die weiteren Untersuchungen. Diese wurden mittels Bildbearbeitung zugeschnitten, sowie einer Farbkorrektur unterzogen. Die Schattenbildung wurde reduziert, die Farbtiefe intensiviert.

3.3.4 Datenauswertung und statistische Analyse

Um standardisiertes und valides Stimulusmaterial zu erhalten, müssen Bilder ausgewählt werden, die der testtheoretischen Überprüfung standhalten. Bewertungen der Validität, Reliabilität und der Normierung werden anhand der Frontalbilder vorgenommen, da hier Vergleichsdaten aus anderen Datenbanken herangezogen werden können. Zusätzlich soll eine erste Überprüfung der Emotionswahrnehmung aus unterschiedlichen Perspektiven vorgenommen werden. Die zur Umsetzung dieser Ziele verwendeten technischen Programme werden direkt bei der jeweiligen Studie beschrieben. Auch die zusätzlich verwendeten Fragebögen werden nicht in diesem Kapitel vorgestellt, da diese für die durchgeführten Studien relativ unterschiedlich zusammengestellt wurden. Die verwendeten statistischen Verfahren werden für jede Studie einzeln vorgestellt, sie sind aber zumeist Verfahren für verbundene Stichproben. Diese wurde gewählt, da das mehrfache Bewerten von emotionalen Gesichtsausdrücken eine Unabhängigkeit in deren Bewertung nahezu ausschließt. Dies weist etwa ein Berliner Trainingsprogramm für Autisten (Dr. Isabel Dziobek: ww.languages-of-emotion.de/de/scott.html) nach. In diesem soll den Betroffenen ein besseren Zugang zu Emotionen anderer und Umgang mit diesen vermittelt werden. Hauptbestandteil des Traningsprogramms ist das systematische Betrachten und Analysieren mimischer Emotionsausdrücke. Die Autisten schaffen es tatsächlich nach dem Training, Emotionen besser wahrzunehmen, sowohl bei Darstellern, welche mehrmals gezeigt wurden als auch völlig unbekannten Personen. Daraus kann geschlussfolgert werden, dass die zugrundeliegenden Mechanismen zur Emotionserkennung übertragbar zu sein scheinen. Dies rechtfertigt dein Einsatz verbundener Stichproben.

In der Regel wurde alpha nicht adjustiert, da es sich um hypothesengenerierende und explorative Fragestellungen handelt. Es wurde angestrebt, für ähnliche Fragestellungen die gleichen statistischen Verfahren einzusetzen, um die Vergleichbarkeit der Ergebnisse zu erhöhen. Die Auswertung der in dieser Arbeit erhobenen Daten erfolgte mit der Statistik- und Analysesoftware SPSS 17 (www.spss.com/de/).

4. TESTTHEORETISCHE ÜBERPRÜFUNG DER FRONTALBILDER

In vielen Bereichen der Emotionsforschung werden Bilder mit mimisch-emotionalem Gesichtsausdruck zur Emotionsstimulation, zur Untersuchung von Emotionserkennung oder zur Deskription von Emotionskategorien benutzt. Da die Qualität der Forschungsergebnisse dabei unmittelbar von der Validität der Emotionsstimuli abhängt, ist es notwendig, qualitativ hochwertiges Stimulusmaterial zu erstellen und einzusetzen. Die Erfassung der psychometrischen Qualität wurde aufgrund der besseren Vergleichsmöglichkeiten mit bestehenden Bildersätzen anhand der Frontalbilder durchgeführt. In der Folge werden fünf Studien vorgestellt, die sich mit der testtheoretischen Überprüfung der Frontalbilder auseinandersetzen. Hierbei handelt es sich um:

- Überprüfung der Emotionserkennungsraten zur ersten Selektion
- Überprüfung der Kriteriumsvalidität
- Beurteilung der Intensität und Natürlichkeit des Emotionsausdrucks und der Attraktivität der fotografierten Person
- Auswahl des endgültigen Bildersatzes
- Normierung der Erkennungsleistung
- Vergleich der Erkennungsleistung über zwei Zeitpunkte (Retest-Reliabilität)

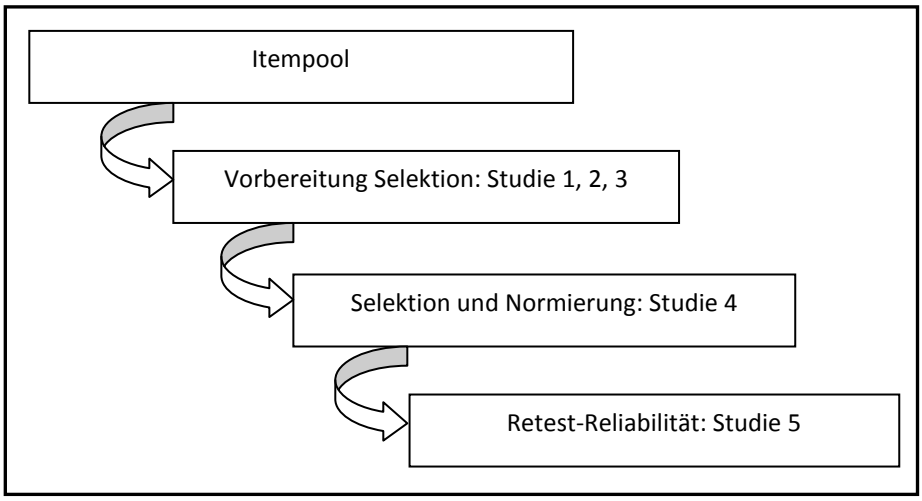


Abb. 10: Überblick über das Vorgehen bei der Untersuchung des Bildmaterials

4.1 Studie 1: Überprüfung der Emotionserkennungsraten zur ersten Selektion

Grundlage eines Testverfahrens zur Erfassung der Emotionserkennungsfähigkeit aus Gesichtern ist die Nutzung standardisierter Stimuli. Wichtigster Aspekt ist, dass die gezeigten emotionalen Stimuli sich hinlänglich in ihrer Ausdruckskraft ähneln, das heißt, potentiell gleiche Erkennungsraten erzeugen. Wäre dies nicht der Fall und einige Bilder leichter und andere schwerer zu erkennen, so ließen sich individuelle Unterschiede in der Erkennungsfähigkeit nicht mehr ausreichend abbilden. Somit ist für die Auswahl von Stimulusmaterial von Bedeutung, dass zumindest innerhalb jeder Emotionskategorie Bilder mit homogenen Erkennungsraten ausgewählt werden.

Hypothesen Studie 1

Hypothese 1: Es wird angenommen, dass nicht alle emotionalen Stimuli (innerhalb einer Emotionskategorie) gleich gut erkannt werden. Die Erkennungsraten ermöglichen die Selektion geeigneter Stimuli für ein standardisiertes Testverfahren.
Hypothese 2: Auch nach der Auswahl geeigneten Stimulusmaterials werden die sechs mimisch dargestellten Basisemotionen unterschiedlich gut erkannt. Es wird angenommen, dass „Freude“ am besten und „Angst“ am schlechtesten erkannt wird. Für die Rangfolge der Erkennungsleistung der anderen Emotionen wird keine Hypothese aufgestellt.

Erste Stimulusbewertung

Die nach der in 3.3.3 beschrieben Vorselektion im Bildersatz verbliebenen 514 Frontalbilder von 50 Personen (mittleres Alter = 23,64; SA = 4,29), wurden in ein in Delphi programmiertes Ratingprogramm eingelesen. Sie wurden auf eine Größe von 750 Pixeln Höhe und 578 Pixeln Breite mit einer Farbtiefe von 24 Bit gebracht und zufällig auf zwei Untersuchungsblöcke verteilt. Die Aufteilung auf zwei Blöcke war notwendig, da eine Bewertung aller Bilder in einem Durchgang aufgrund des zu großen Konzentrationsaufwandes nicht sinnvoll erschien. Nach einer zufälligen Aufteilung enthielt der erste Block 270 und der zweite Block 244 Bilder. Insgesamt nahmen 20 Personen (10 Frauen und 10 Männer) an dem Rating teil (mittleres Alter M = 27,15; SA = 3,39). Alle Beurteiler erklärten schriftlich ihre freiwillige Teilnahme, entsprechend der ethischen Vorgaben (Ethikvotum 245/08_UBB/se). Die Beurteilung der beiden Blöcke erfolgte in einem zeitlichen Abstand von durchschnittlich 3 Tagen (M = 3,15;

SD = 2,18). Der Versuchsablauf folgte diesem Schema: Zunächst wurden die Rater über die Intention des Versuches aufgeklärt, sie unterschrieben die Einwilligungserklärung und wurden mit der Ratingprozedur vertraut gemacht. Vor Beginn des Versuchs lasen die Probanden die in Abbildung 11 beschriebene Instruktion am Bildschirm.

Mit dem FEEL-Test wird die Emotionserkennung untersucht. Dafür werden auf dem Bildschirm farbige Fotos von Gesichtern gezeigt. Auf den Gesichtern ist jeweils ein anderes Gefühl erkennbar: und zwar Freude, Trauer, Überraschung, Angst, Ekel oder Ärger. Jede Darbietung folgt dem gleichen Schema: Sie sehen für eine Dauer von 4 Sekunden ein farbiges Bild mit einem Gesichtsausdruck auf dem Bildschirm.

Dann sehen Sie ein Antwortschema zu den Gefühlen. Wählen Sie per einfachem Mausklick dasjenige Gefühl aus, das vorher zu sehen war.

- **Freude** (auch wie Begeisterung, Glück, Zufriedenheit etc.)
- **Angst** (auch wie Furcht, Panik, Ängstlichkeit etc.)
- **Ekel** (auch wie Missfallen, Widerwärtigkeit, Verachtung etc.)
- **Ärger** (auch wie Wut, Zorn, Aggression etc.)
- **Überraschung** (auch wie Verblüffung, Erstaunen etc.)
- **Trauer** (auch wie Depression, Niedergeschlagenheit etc.)

Wenn Sie ein Gesicht gesehen haben, sollten Sie sich immer für ein Gefühl entscheiden, auch wenn das manchmal nicht einfach ist. Für die Antwort haben Sie 10 Sekunden Zeit. Wissen Sie nicht, welche Emotion zu sehen war, so wählen Sie „nicht erkannt“.

Abb. 11: Instruktion für die Bewertung des Bildmaterials in EmoIR

Danach begann die eigentliche Aufgabe der Emotionserkennung. Ein Fixationskreuz in der Mitte des Bildschirms wurde für eine Sekunde gezeigt, um die Aufmerksamkeit auf den nachfolgenden Stimulus zu lenken. Es folgte eine Pause von 2,25 bis 2,75 Sekunden Länge, bevor das erste emotionale Bild für vier Sekunden gezeigt wurde. Die Variation der Pausenzeit sollte Gewöhnungseffekten entgegenwirken. Nach Zeigen des Bildes folgte eine Pause von einer halben Sekunde, bevor für zehn Sekunden das Auswahlfeld erschien und der Proband sich für eine der Möglichkeiten entscheiden musste. Danach begann der Ablauf von vorne. Bei jeder Entscheidung wurde die Reaktionszeit erfasst. Die grafische Darstellung des Versuchsablaufs in Abbildung 12 verdeutlicht dies.

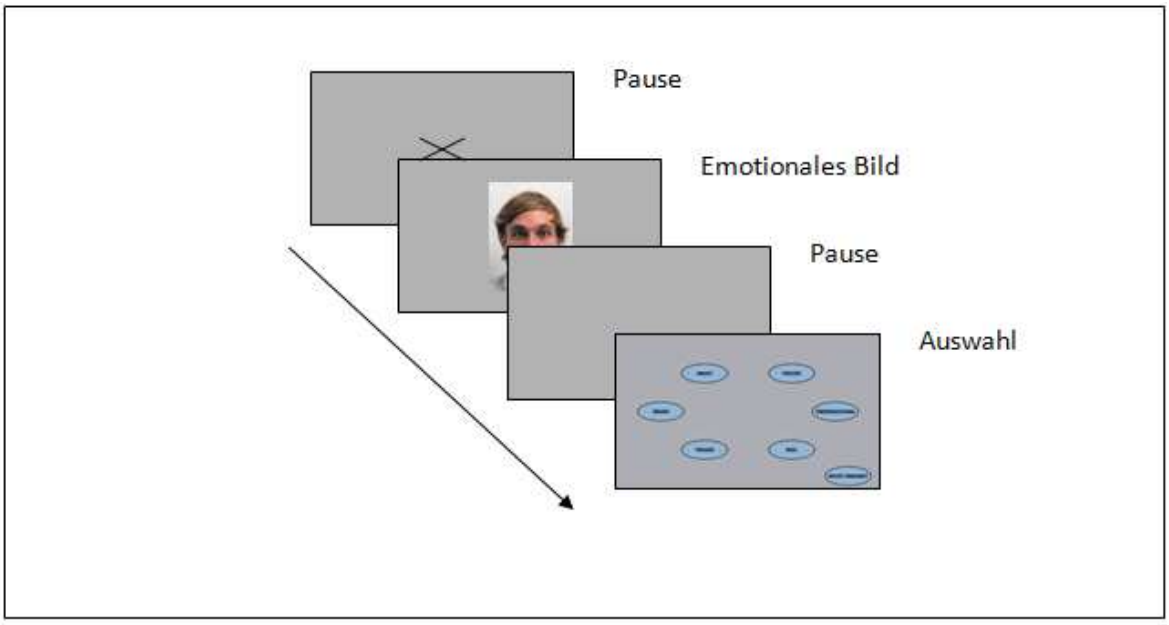


Abb. 12: Versuchsablauf Ratingprozedere

Erkennungsleistung und Verwechslungen

Insgesamt ergab sich folgende Erkennungsleistung: Freude (98,97%), Trauer (86,58%), Ekel (84,01%), Überraschung (82,33%), Ärger (81,53%) und Angst (61,4%). Die erzielten Resultate wiesen auf eine relative Heterogenität bezüglich der Erkennungsleistung hin. Dies wurde unterstützt durch die Betrachtung der Erkennungsraten der einzelnen Bilder, welche von 5% bis 100% variierte. Wie Tabelle 3 zu entnehmen ist, lagen für Freude kaum Verwechslungen mit den anderen Kategorien vor. Angst hingegen wurde zu fast 15% mit Überraschung verwechselt und zu weiteren 12% konnte der Stimulus nicht zugeordnet werden („nicht erkannt“).

Tab. 3: Zuordnung der Bewertung zu den einzelnen Emotionskategorien (514 Bilder, in Prozent)

		Angst	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Nicht erkannt**
Stimulus-emotion	**Angst**	**61,4**	0,07	15,07	4,6	7,07	0,26	11,53
	Freude	0	**98,97**	0,36	0	0,10	0	0,57
	Überraschung	13,64	1,19	**82,33**	0,06	0	0,11	2,67
	Ekel	1,63	0,06	0,12	**84,01**	5,41	5,41	5,47
	Trauer	2,81	0,14	0,48	0,68	**86,58**	1,58	7,74
	Ärger	1,16	0,05	0,21	1,42	8,16	**81,53**	7,47

Auswahl der emotionalen Stimuli

Ausgewählt wurden die 96 Bilder von 48 Darstellern, die von den meisten Beurteilern der korrekten Emotion zugeordnet werden konnten. Als Mindesterkennung wurde eine (willkürliche) Grenze von 55% für das Einzelbild festgelegt. Hierbei konnten pro fotografierter Person zwei Bilder mit unterschiedlichem emotionalen Ausdruck in die Auswahl eingehen. Bezüglich der Erkennungsleistung ergaben sich folgende Werte: Freude (98,79%), Überraschung (93,18%), Ärger (92,06%), Ekel (89,42%), Trauer (78,81%) und Angst (70,55%). Veränderungen in der Erkennungsleistung wurden mittels t-Test errechnet, da die Normalverteilungsvoraussetzungen für die zum Vergleich herangezogenen 514 Bilder mittels Shapiro-Wilks-Test errechnet wurden und als erfüllt angesehen werden konnten – einzige Ausnahme war Freude (Angst: $p = .379$, Freude: $p < .000$, Überraschung: .496, Ekel: .518, Trauer: $p = .063$, Ärger: $p = .149$). Es zeigte sich eine Verbesserung der Erkennungsraten für alle Emotionen außer Freude (Angst: $t(38) = -7{,}436$, $p < .000$; Ärger: $t(38) = -4{,}754$, $p < .000$; Ekel: $t(38) = -5{,}134$, $p < .000$; Trauer: $t(38) = -2{,}971$, $p < .000$; Überraschung: $t(38) = -4{,}104$, $p < .000$). Im Mittel betrug die Erkennungsleistung 93,64%. Ein Binomialtest ergab eine Erkennungsleistung, die signifikant oberhalb der Zufallswahrscheinlichkeit lag ($p < .000$). Das bedeutet, dass die Stimuli überzufällig häufig der korrekten Emotion zugeordnet wurden. Eine Übersicht über die Erkennungsraten für die einzelnen Emotionskategorien nach Auswahl der 96 Bilder findet sich in Tabelle 4.

Tab. 4: Zuordnung der Bewertung zu den einzelnen Emotionskategorien (96 Bilder, in Prozent)

		Angst	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Nicht erkannt**
Stimulus-emotion	**Angst**	**85,00**	0,31	4,69	3,44	2,18	0,00	4,38
	Freude	0,00	**99,36**	0,32	0,00	0,00	0,00	0,32
	Überraschung	3,13	0,62	**95,31**	0,00	0,00	0,00	0,94
	Ekel	0,32	0,00	0,00	**95,63**	0,31	1,87	1,87
	Trauer	2,80	0,00	0,00	0,00	**94,06**	0,30	2,84
	Ärger	0,63	0,00	0,00	1,25	3,12	**92,50**	2,50

Erkennungsleistung und Verwechslungen der ausgewählten emotionalen Stimuli

Im Folgenden wurden statistische Berechnungen zur Erfassung potentieller Unterschiede mit zwei verschiedenen Maßen durchgeführt: Dem Prozentsatz richtig erkannter

Emotionsausdrücke und der Reaktionszeit. Da aufgrund eines zuvor durchgeführten Shapiro-Wilks-Tests die Normalverteilung der aller Emotionsvariablen der 96 Stimuli außer Angst zu bezweifeln war (Angst: p = .192, Ärger: p = .007, Ekel: p < .000, Freude: p < .000, Trauer: p < .000, Überraschung: p <.000), wurde in der Folge mit einem non-parametrischen Verfahren für ein Design mit Messwiederholung (Friedman-Test) gerechnet.

Der Prozentsatz richtig erkannter Emotionsausdrücke ergab sich aus dem Anteil korrekt zugeordneter Emotionen an allen vorgenommenen Beurteilungen. Hier zeigte sich ein signifikanter Unterschied richtig erkannter Emotionsausdrücke bei den sechs verschiedenen Emotionen ($\chi^2(5) = 35{,}009$; p = .000). Eine detailliertere Analyse auf Einzelemotionsebene mit dem Wilcoxon-Test für verbundene Stichproben ergab: Angst wurde signifikant schlechter erkannt als alle anderen Emotionen (Z = -3,544, p < .000 für Freude; Z = -3,342, p = .001 für Trauer; Z = -3,089, p = .002 für Ekel; Z = -2,969, p = .003 für Überraschung; Z = -2,532, p = .011 für Ärger). Im Gegenzug wurde Freude signifikant besser erkannt als alle anderen Emotionen (Z = -3,544, p < .000 für Angst; Z = -3,400, p = .001 für Ärger; Z = -2,871, p = .004 für Trauer; Z = -2,267, p = .023 für Ekel; Z = -2,047, p = .041 für Überraschung). Ärger, Ekel, Trauer und Überraschung unterschieden sich nicht signifikant voneinander (p > .20 für alle verbleibenden möglichen Kombinationen). Somit ergab sich folgende Reihenfolge in der Erkennungsleistung der Fotos mit den emotionalen Gesichtsausdrücken: Angst < Ärger = Trauer = Überraschung = Ekel < Freude.

In der Folge wurden mögliche Geschlechtereffekte für die Darstellung der emotionalen Gesichtsausdrücke überprüft. Pro Geschlecht gingen 8 emotionale Stimuli in die Berechnungen ein. Lediglich für die Ärgerstimuli ergab sich ein signifikanter Unterschied (Z = -2,608; p = .009): Ärger wurde zu 96,25% richtig erkannt, wenn es von männlichen Darstellern präsentiert wurde, aber nur zu 88,13%, wenn die weiblichen Darsteller Ärger zeigten. Aufgrund der kleinen Stichprobe wurde auf die Berechnung von Geschlechterunterschieden in der Erkennungsleistung verzichtet.

Für einen Vergleich der Reaktionszeiten wurden Mittelwerte für die verschiedenen Emotionen gebildet (siehe Tabelle 5).

Tab. 5: Übersicht über die Reaktionszeiten für die einzelnen Emotionen

Reaktionszeit	Emotion	Mittelwert	Standardabweichung
	Angst	1323ms	593
	Ärger	1257ms	517
	Ekel	1034ms	350
	Freude	972ms	232
	Trauer	1161ms	335
	Überraschung	903ms	155
	Gesamt	1108ms	165

Hier zeigte sich ein signifikanter Unterschied bezüglich der Reaktionszeiten für die verschiedenen Emotionen ($\chi^2(5) = 28.857$; $p = .000$). Die Ergebnisse der Berechnungen auf Einzelemotionsebene (Wilcoxon-Test) sind Tabelle 6 zu entnehmen.

Tab. 6: Ergebnisse der Paarvergleiche für die Reaktionszeiten bei Auswahl der einzelnen Emotionen

	Ärger	Ekel	Freude	Trauer	Überraschung
Angst	p = .296	p = .000	p = .001	p = .279	p = .002
Ärger	-	p = .011	p = .006	p = .654	p = .003
Ekel		-	p = .391	p = .135	p = .117
Freude			-	p = .010	p = .093
Trauer				-	p = .002

Nachfolgend wurden potentielle Zusammenhänge zwischen dem Prozentsatz korrekt erkannter Emotionen und der Reaktionszeit errechnet. Die Ergebnisse sind Tabelle 7 zu entnehmen.

Tab. 7: Korrelation zwischen Reaktionszeit und Prozentsatz richtig erkannter Emotionen

Reaktionszeit x Erkennungsrate	
Angst	-.449*
Ärger	-.154
Ekel	.172
Freude	-.210
Trauer	-.242
Überraschung	-.641**

* = p < .05; ** = p < .001

<u>Diskussion der Erkennungsleistung und der Verwechslungen</u>

Mit dem ersten Rating und der Selektion aufgrund der besten Erkennbarkeit der Basisemotionen konnte eine erste Homogenisierung der Erkennungsraten erzielt werden. Freude wurde mit 99,36% am zuverlässigsten erkannt. Dies steht im Einklang mit den

Ergebnissen anderer Bildersets (siehe Tabelle 1). Beispielsweise beschrieben die Autoren der UCDSEE-Datenbank (Tracy, Robins & Schriber, 2009) eine Erkennungsleistung von ca. 94% für Freude, die Entwickler der RaFD-Bildersets (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & Knippenberg, 2010) konnten eine Erkennungsleistung von 98% für Freude nachweisen. In einer Validierungsstudie der KDEF (Goeleven, De Raedt, Leyman & Verschuere, 2008) wurde eine Erkennungsrate von 92,65% angegeben. Bei der Untersuchung des JACFEE-Bildersatzes wurden bei US-amerikanischen Bürgern für Freude Erkennungsleistungen für Freude von 98% erzielt (Matsumoto & Ekman, 1989).

Bezüglich der Präzision der Erkennung der einzelnen Emotionen konnte bislang kein einheitliches Ergebnis abgeleitet werden. Wurde im RaFD-Bilderset (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & Knippenberg, 2010) Ekel mit 79% am schlechtesten erkannt, so war es bei der UCDSEE-Datenbank (Tracy, Robins & Schriber, 2009) und dem KDEF-Bilderset (Goeleven, De Raedt, Leyman & Verschuere, 2008) Angst mit einer Erkennungsleistung von ca. 45% beziehungsweise 43,03%. Auch der hier vorliegenden Studie wurde Angst am schlechtesten erkannt, wobei nach der Selektion des Bildmaterials ein deutlicher Anstieg in der Erkennungsrate verzeichnet werden konnte. Es zeigt sich tendenziell, dass Überraschung häufig sehr gut erkannt wird (RaFD: 90%, UCDSEE: 90%). Dies stimmt mit den hier gewonnenen Ergebnissen überein. Unterschiede in der Reihenfolge der richtig erkannten Emotionen sind auf verschiedene Faktoren zurückzuführen: Antwortformat, Ethnie der fotografierten Personen und der bewertenden Personen, Stichprobengröße und Alter der Beurteiler sind nur einige Beispiele für Aspekte, die die Erkennungsleistung beeinflussen können. Insgesamt zeigten die Resultate jedoch eine gute Übereinstimmung mit bisherigen Forschungsergebnissen.

Es zeigten sich nur wenige Verwechslungen bei der Zuordnung der Emotion zum korrekten Emotionslabel, auch das zusätzliche Antwortfeld „nicht erkannt“ wurde nur selten verwandt. Angst wurde lediglich zu 85% korrekt erkannt und am häufigsten mit „Überraschung“ verwechselt (fast 5%). Dieses Ergebnis stimmt mit den Ergebnissen von Palermo und Coltheart (2004) überein, die in ihrer Studie die Erkennungsleistung von 336 emotionalen Bildern unterschiedlichster Bildersets überprüften. Denkbar ist, dass Angst und Überraschung aufgrund der großen Ähnlichkeiten in der mimischen Darstellung miteinander verwechselt wurden. Diese Vermutung wird dadurch unterstützt, dass in der hier

beschrieben Studie Überraschung trotz der hohen Erkennungsrate zu 3% mit Angst verwechselt wurde, also eine Verwechslungstendenz in beide Richtungen vorlag.
Zu vermuten ist außerdem, dass das Angebot der Auswahlmöglichkeit „nicht erkannt“ insgesamt zu einer Minimierung von Verwechslungen beigetragen hat. Die Probanden waren nicht gezwungen, sich innerhalb eines „Forced-Choice“-Formats für eine der sechs Emotionen zu entscheiden, sondern konnten auf diese siebte Kategorie ausweichen. Für eine vertiefende Diskussion potentieller Probleme des „Forced-Choice“-Formats siehe Studie 2. Wie Tabelle 4 zu entnehmen ist, wurde bei allen sechs Emotionen von der Antwortoption „nicht erkannt“ Gebrauch gemacht.
Ein Geschlechtereffekt ergab sich für die Darstellung von Ärger. Wurde Ärger von männlichen Darstellern präsentiert, war die Zuordnung präziser als bei der Darstellung durch Frauen. Möglicherweise sind Männer einerseits im Darstellen von Ärger trainierter als Frauen und andererseits dieser Emotionsausdruck bei Frauen kulturell weniger akzeptiert. Aufgrund der nur kleinen Stichprobe sollten die Ergebnisse als vorläufig bewertet werden.
Die Reaktionszeiten bei der Emotionserkennung geben einen Hinweis auf die Aufgabenschwierigkeit. Je einfacher eine bestimmte Emotion erkannt und einer Kategorie zuordnen lässt, desto schneller kann eine Zuordnung getroffen werden (Palermo & Coltheart, 2004; Batty & Taylor, 2003). In der hier vorgestellten Studie wurde Freude am schnellsten erkannt und zwar signifikant schneller als Ärger, Trauer und Angst. Für Freude zeigte sich außerdem die höchste Erkennungsleistung. Überraschung und Ekel, die Emotionen, die ebenfalls sehr gute Erkennungsleistungen aufwiesen, wurden fast ebenso schnell erkannt wie Freude. Im Gegensatz dazu wurde Angst nicht nur am schlechtesten, sondern auch am langsamsten erkannt, so dass sich hier ein Zusammenhang zwischen Erkennungsleistung und Reaktionszeit vermutet wurde. Die durchgeführte Korrelationsanalyse wies jedoch kaum signifikante Korrelationen auf. Nahezu alle Korrelationen wiesen aber negative Korrelationen auf. Es ist anzunehmen, dass aufgrund der geringen Power (N = 20) nur zwei signifikante Korrelationen gefunden werden konnten.

Eine Limitation bei der Interpretation der Ergebnisse ergibt sich potentiell aus der geringen Teilnehmerzahl. Jedoch war das Ziel dieser Studie lediglich eine erste Vorauswahl zu treffen. Der zeitliche Aufwand wäre an einer größeren Stichprobe nicht zu leisten gewesen.

4.2 Studie 2: Überprüfung der Kriteriumsvalidität

Ein gängiges Vorgehen zur Überprüfung der Emotionserkennungsfähigkeit ist in aller Regel, den Probanden emotionale Gesichtsausdrücke in Form von Fotos basierend auf den sechs Basisemotionen nach Ekman, zu präsentieren und sie anschließend anhand bestimmter Auswahlmöglichkeiten beurteilen zu lassen. Solch eine Vorgabe von Alternativen entspricht einem geschlossenen Antwortformat, auch „Forced Choice“ genannt. Vorteil dieses Vorgehens ist, dass eine schnelle und einfache Auswertung der Beurteilung erfolgen kann. Beispielsweise kann direkt im Anschluss an die Beurteilung ein Summenwert für die korrekt erkannten Emotionen und die Verwechslungen errechnet werden. Geschlossene Antwortformate werden daher auch häufig bei Tests und Fragebogen eingesetzt. Leider bietet die Verwendung dieses Antwortformates auch einige potentielle Nachteile. James A. Russell wies darauf hin (Russell, 1993), dass die Schlussfolgerungen, die aus der Verwendung eines geschlossenen Antwortformates gezogen werden können, nicht eindeutig seien. Durch das Vorhandensein mehrerer Optionen würde ein Relativurteil gefällt, das heißt, es würde die Antwort ausgewählt, die am wahrscheinlichsten erschiene. Die wahrscheinlichste Antwortalternative müsse aber nicht zwingend die vom Probanden intendierte Antwort widerspiegeln. Bei den Antworten handele es sich also nicht notwendigerweise um ein objektives Kriterium. Wagner (1997) wies des Weiteren darauf hin, dass Art und Anzahl der vorgegeben Kategorien immer Einfluss auf die Zuordnung nehmen könne. Wenn beispielsweise als positive Kategorie nur die Kategorie „freudig“ verfügbar wäre, so würden dieser alle positiven emotionalen Ausdrücke zugeordnet, auch wenn es sich vielleicht in Wirklichkeit um Zustände wie „amüsiert“ oder „friedlich“ handeln würde.

Werden Antwortalternativen vorgelegt, so ist denkbar, dass nicht nur die reine Erkennungsleistung, sondern auch einige „Glückstreffer“, also zufälliges Auswählen der korrekten Antwort, erfasst werden. Die Erkennungsrate würde dann nicht die tatsächlich erbrachte Leistung widerspiegeln. Möglich ist jedoch auch, dass die Erkennung der Basisemotionen einen nur so geringen fehleranfälligen elementar-neurologischen Prozess darstellt, dass der Einsatz eines geschlossenen Antwortformats beim „gesunden Durchschnittserwachsenen“ keine essentielle Beeinflussung bei der Bearbeitung bietet, sondern ohnehin vielmehr eine Art Deckeneffekt vorliegt. Hinweise hierauf ergeben sich beispielsweise auch aus der Sprachforschung. Nach Schwarz-Friesel (2007) ist Sprache ein

„humanspezifisches, genetisch determiniertes, neuronal repräsentiertes und von universalen Prinzipien bestimmtes Kenntnis- und Regelsystem“ (S. 18) und somit ebenso fest verankert wie spezifische Emotionen. Sprache ist neben der Mimik einer der Hauptkanäle zum Ausdruck von emotionalen Zuständen, mit ihr kann beispielsweise gelobt, gedroht, beleidigt oder geschmeichelt werden (Schwarz-Friesel, 2007, S. 22-23). Kommunikatives Handeln kann in diesem Zusammenhang als bewusst ausgeführte Tätigkeit zur Zielerreichung gesehen werden, wobei solche Ziele maßgeblich durch unser emotionales System gesteuert werden (Damasio, 1995). Sprache und Emotion stehen also in einem festen Zusammenhang (Wierzbicka, 2005). Es ist insofern davon auszugehen, dass spezifische emotionale Gesichtsausdrücke in einer gemeinsamen internen Repräsentation mit spezifischen semantischen Netzen (= Emotionskategorien) abgespeichert werden, die Frage die sich jedoch hierbei stellt, ist welche Stabilität eine solche Vernetzung aufweist.
In der vorliegenden Studie sollte nun ein objektives Kriterium gefunden werden. Es war anzunehmen, dass in diesem Fall nichts objektiver sei als die unabhängige Beurteilung selbst. Deswegen wurde nun überprüft, welchen Effekt ein freies Antwortformat auf die Emotionserkennung aus Gesichtern hat. Der Grad der Übereinstimmung weist die Güte der Kriteriumsvalidität aus.

Hypothesen Studie 2

Aufgrund der nur geringen empirischen Basis erscheint es hier nicht sinnvoll, eine gerichtete Hypothese aufzustellen. Vielmehr werden explorativ und hypothesengenerierend potentielle Zusammenhänge und Unterschiede zwischen einem geschlossenen und einem offenen Antwortformat untersucht.

Beurteilung der emotionalen Stimuli im offenen Antwortformat

Die Studie wurde an 33 Versuchspersonen durchgeführt (mittleres Alter = 26,39 Jahre; SA = 3,65). Alle Teilnehmer gaben ihr schriftliches Einverständnis (Ethikvotum 245/08_UBB/se). Der Anteil der Frauen betrug 69%. Rund 58% der VP waren Studenten, 36% Akademiker. Die verbleibenden 6,1% waren der Kategorie „Lehre, Zivildienst, Jobben“ zuzuordnen. Der Bildungsgrad war als hoch zu beurteilen, alle Probanden beendeten ihre Schullaufbahn mit dem Abitur. Des Weiteren wurden ausschließlich Muttersprachler in die Studie einbezogen.

Den Probanden wurden die 96 emotionalen Stimuli aus Studie 1 präsentiert, jeder Darsteller war in diesem Bilderset mit je zwei emotionalen Gesichtsausdrücken vertreten. Insgesamt wurden 16 Bilder pro Emotion gezeigt. Die Bilder wurden randomisiert in Papierdruck vorgegeben. Die Größe der gezeigten Bilder betrug 5,9 cm x 4,6 cm. Die Probanden wurden gebeten, diese zu beurteilen – ohne Vorgabe von Kategorien, das heißt im offenen Antwortformat. Des Weiteren wurde darauf hingewiesen, die subjektive Sicherheit auf einer fünfstufigen Likertskala von „1 = sehr unsicher (bei der Einschätzung)“ bis „5 = sehr sicher (bei der Einschätzung)“ anzugeben. Konnten die Probanden sich nicht für eine verbale Entsprechung des zu beurteilenden emotionalen Gesichtsausdrucks entscheiden, hatten sie die Möglichkeit, die alternativen verbalen Ausdrücke in einem weiteren Feld zu vermerken (für eine Übersicht siehe Abbildung 13).

Emotionserkennung aus Gesichtern – Bewertung der Gesichtsausdrücke:

Zu bewertendes Bild	Welche Emotion ist zu sehen?	Wie sicher bist du dir bei deiner Einschätzung? (1 = sehr unsicher, 2 = unsicher, 3 = weder/noch, 4 = sicher, 5 = sehr sicher → **bitte markiere die zutreffende Zahl fett!**	Andere mögliche Emotionen?
		1 – 2 – 3 – 4 – 5	

Abb. 13: Erfassung der Einschätzung emotionaler Gesichtsausdrücke im offenen Antwortformat

Alle verbalen Entsprechungen wurden anschließend zwei naiven und hypothesenblinden Personen unabhängig voneinander zum Rating vorgelegt, um eine Klassifikation der gewonnenen Daten in die sechs Kategorien der Basisemotionen vorzunehmen. Zusätzlich wurden zwei Kategorien eingeführt für Wörter, die den Basisemotionen nicht zuzuordnen sind. Hier sollten die Rater entscheiden, ob es sich um eine positive oder negative Emotion

handelte. Die Einschätzung erfolgte auf Grundlage eines eigens entwickelten Ratingmanuals (siehe Anhang B).

Beurteilung der emotionalen Stimuli im offenen und geschlossenen Antwortformat

Die Analyse der Emotionseinstufung ergab eine Nennung von insgesamt 245 verschiedenen Emotionswörtern, wobei Variationen eines Wortstammes mitgezählt wurden. 190 Emotionswörter wurden durch das erste Urteil erfasst, weitere 55 Wörter ergaben sich aus der zusätzlichen Beurteilungsmöglichkeit.

Die Rater erzielten bei der Zuordnung zu den Kategorien insgesamt eine Übereinstimmung von 72,25% (Kappa = 0,655; $p < .000$). Dies kann als beachtliche Übereinstimmung (Landis & Koch, 1977) angesehen werden. Die restlichen, unterschiedlich bewerteten Kategorien wurden in einem „Forced Consent" ein zweites Mal gemeinsam beurteilt. So konnten alle 245 verbalen Ausdrücke einer der acht möglichen Kategorien zugeordnet werden. Eine Aufteilung findet sich in Tabelle 8, eine Übersicht der genannten Emotionswörter in Anhang C.

Tab. 8: Zuordnung der 245 verbalen Ausdrücke zu den acht auswählbaren Kategorien

Kategorie	**Angst**	**Ärger**	**Ekel**	**Freude**	**Trauer**	**Überraschung**	**Pos. Emotion**	**Neg. Emotion**
Häufigkeit	13	20	6	26	40	34	26	80
Prozent	5,31%	8,16%	2,45%	10,61%	16,33%	13,88%	10,61%	32,65%
gesamt	245 (100%)							

In einem nächsten Schritt erfolgte eine Auszählung, wie oft welche verbale Entsprechung von den 33 Probanden genannt wurde. Die Analyse der Nennungen zeigt: Es gab einige wenige verbale Ausdrücke, die von vielen Probanden genannt wurden. Die größte Übereinstimmung fand sich für Ekel, dieser Begriff wurde insgesamt 458-mal genannt. Ebenfalls häufig genannt wurden die Begriffe Freude (428x), Trauer (304x), Überraschung (303x), Angst (297x) und Wut (227x). Ärger wurde 98-mal angegeben. Die Nennung dieser sieben Emotionswörter unterschied sich signifikant von der Nennung aller anderen Emotionswörter ($t(242) = -35,999$; $p < .000$).

Nach Durchführung der Ratingprozedur war eine Ermittlung der Erkennungsraten möglich: 94,27% für Freude, 88,13% für Ekel, 87,46% für Überraschung, 84,91% für Ärger, 81,04% für Trauer und 68,25% für Angst. Angst wurde signifikant schlechter erkannt als alle andere

Emotionen (Z = -2,934, p = .003 im Vergleich zu Ärger; Z = -3,698, p < .000 im Vergleich zu Ekel; Z = -4m551, p < .000 im Vergleich zu Freude; Z = -2,714, p = .007 im Vergleich zu Trauer; Z = -3,309, p = .001 im Vergleich zu Überraschung), Freude besser als alle anderen Emotionen (Z = -4,551, p < .000 im Vergleich zu Angst; Z = -2,440, p = .015 im Vergleich zu Ärger; Z = -2,336, p = .018 im Vergleich zu Ekel; Z = -4,115, p < .000 im Vergleich zu Trauer, Z = -2,570, p = .010 im Vergleich zu Überraschung). Trauer wurde außerdem signifikant schlechter erkannt als Ekel (Z = -2,460, p = .014). Für die Erkennungsleistung von Ärger, Ekel und Überraschung ergaben sich keine signifikanten Unterschiede (p > .05). Folgende Reihenfolge der Erkennungsleistung ließ sich daraus erstellen: Angst < Trauer = Ärger = Überraschung = Ekel < Freude. Zum Vergleich: In der vorherigen Studie mit „Forced Choice" (siehe Studie 1) wurden folgende Erkennungsraten angeführt: 99,36% für Freude, 95,63% für Ekel, 95,31% für Überraschung, 94,06% für Trauer, 92,5% für Ärger, und 85% für Angst. Mittels Mann-Whitney-U-Test wurden potentielle Unterschiede zwischen den Erkennungsraten mit offenem und geschlossenem Antwortformat berechnet. Für Angst (Z = -0,113; p = .910), Ärger (Z = -1,301; p = .193), Ekel (Z = -0,690; p = .490) und Freude (Z = -1,686; p = .092) ließen sich keine signifikanten Unterschiede aufzeigen. Trauerstimuli wurde in der Studie mit geschlossenem Antwortformat signifikant schlechter korrekt zugeordnet als in der Studie mit offenem Antwortformat (Z = -2,513; p = .012), für Fotos mit einem überraschten Gesichtsausdruck verhielt es sich genau andersherum (Z = -2,094; p = .036). Angst wies in dieser Studie eine Erkennungsrate von 68% auf und wurde am häufigsten mit Überraschung verwechselt (fast 10%). Außerdem konnte gezeigt werden, dass in der hier beschrieben Studie Überraschung zum größten Teil (ebenfalls etwa 10%) mit Angst verwechselt wurde, also eine Verwechslungstendenz in beide Richtungen vorlag. Zusätzlich wurde Angst aber auch relativ häufig mit Ekel (fast 9%) und Trauer (fast 5%) verwechselt. Ärger wurde kaum mit anderen Basisemotionen verwechselt, ebenso wie Freude – diese Emotion wurde fast zu 100% richtig erkannt. Ekel wurde zu 6% mit Ärger verwechselt, Trauer zu 4% mit Angst. Der Prozentsatz an Verwechslungen lag in keiner Kategorie über 10%. Zu vermuten ist, dass das Angebot der Auswahlmöglichkeit „positive bzw. negative Emotion" insgesamt zu einer Minimierung von Verwechslungen beigetragen hat. Wie Tabelle 9 zu entnehmen ist, wurde von diesen beiden zusätzlichen Kategorien bei allen sechs Emotionen als alternative Antwortmöglichkeit Gebrauch gemacht. Jedes einzelne Stimulusbild wurde zu mindestens 66% richtig erkannt.

Die Überprüfung von Geschlechtereffekten (Geschlecht der Darsteller beziehungsweise Beurteiler) ergab keine signifikanten Unterschiede in der korrekten Bezeichnung der gezeigten emotionalen Gesichtsausdrücke (alle Berechnungen $p > .05$).

Tab. 9: Erkennungsraten und Verwechslungen der emotionalen Stimuli nach Zuordnung durch die Rater (in Prozent)

Tatsächliche Emotion **Zuordnung im Rating**	**Angst**	**Ärger**	**Ekel**	**Freude**	**Trauer**	**Überraschung**
Angst	**68,25**	1,33	0,37	0	4,14	9,56
Ärger	0,36	**84,91**	6,22	0	2,4	0,38
Ekel	8,65	0,94	**88,13**	0	0,19	0
Freude	0	0	0	**94,27**	0	1,32
Trauer	4,67	2,83	0,19	0	**81,04**	0
Überraschung	9,79	0,75	0,19	0	0,74	**87,46**
Neg. Emotion	7,73	7,17	4,90	0,19	10,42	1,28
Pos. Emotion	0,55	2,07	0	5,54	1,11	0

Die 33 Probanden mussten neben der Beurteilung des Stimulusmaterials auch ihre Sicherheit bei der Beurteilung angeben. Sie betrug im Mittel 4,21 (auf der 5-stufigen Likert-Skala). Bezogen auf die einzelnen Emotionskategorien ließen sich nur geringe Unterschiede feststellen: Am geringsten wurde die Sicherheit bei der Beurteilung mit 3,79 bei Angst eingeschätzt, die Emotion mit den schlechtesten Erkennungsraten, und am höchsten mit 4,56 bei Freude, der Emotion mit der höchsten Erkennungsleistung. Die Ergebnisse einer für den Haupteffekt signifikant gewordenen einfaktoriellen ANOVA zeigte für die Scheffé-adjustierten post-hoc-Vergleiche lediglich signifikante Unterschiede zwischen der Einschätzung der subjektiven Sicherheit von Angst und Freude ($F(5, 172) = 4,097$; $p = .005$) und Angst und Trauer ($F(5, 172) = 4,097$; $p = .043$).

Es zeigte sich ein mittlerer positiver korrelativer Zusammenhang nach Pearson zwischen der angegebenen Sicherheit bezüglich des Urteils und der Korrektheit der genannten Emotion für Ärger ($r = .566$, $p = .001$), Ekel ($r = .500$, $p = .004$) und Überraschung ($r = .530$, $p = .003$). Das heißt eine höhere Bewertung der subjektiven Sicherheit ging bei diesen drei Emotionen mit einer höheren Korrektheit in der Beurteilung der gezeigten Emotionen einher. Für Angst ($r = .183$, $p = .352$), Freude ($r = .201$, $p = .277$) und Trauer ($r = .182$, $p = .336$) ergaben sich keine signifikanten Zusammenhänge, wenn auch ebenfalls Korrelationen im mittleren Bereich.

Diskussion der Verwendung offener Antwortformate bei der Beurteilung emotionaler Stimuli

Die Annahme von Russell (1993) und Wagner (1997), dass Basisemotionen ohne Vorlage eines geschlossenen Antwortformats nicht eindeutig erkannt würden, konnte in dieser Studie nicht bestätigt werden. Im Gegenteil: Es zeigte sich hier, dass die Basisemotionen auch ohne Vorgabe von Antwortalternativen zuverlässig der korrekten verbalen Entsprechung zugeordnet werden konnten. Dies spricht nicht nur für eine hohe Kriteriumsvalidität, sondern bietet auch Hinweise auf die universelle Gültigkeit und besonderen Relevanz von Basisemotionen. Die Probanden erkannten auch ohne Vorgabe von Alternativen die Emotionsausdrücke zuverlässig. Besonders deutlich konnte dies für Angst, Ärger, Ekel und Freude nachgewiesen werden. Die emotionalen Ausdrücke wurden ebenso präzise dem korrekten Emotionslabel zugeordnet wie in Studie 1. Trauer wurde in dieser Studie signifikant besser und Überraschung signifikant schlechter erkannt als in Studie 1, die Erkennungsraten lagen dennoch weiterhin deutlich über der Zufallswahrscheinlichkeit. Es zeigten sich keine Geschlechtereffekte, weder bei Betrachtung der Darsteller noch der Beurteiler.

Bei der Analyse der verbalen Entsprechungen fiel auf, dass bezogen auf die sechs Basisemotionen am meisten unterschiedliche Wörter zur Benennung von Überraschung verwendet wurden, am wenigsten für Ekel. Die Gründe hierfür sind unklar, in der Literatur finden sich nur wenige Hinweise. Möglicherweise lag die Divergenz an der Ursachenzuschreibung: Während der auslösende Reiz für das Gefühl Ekel relativ eindeutig scheint, kann Überraschung aufgrund einer unerwarteten, gefährlichen, positiven oder negativen Situation entstehen, so dass eventuell mehr Wörter zur Differenzierung benötigt werden. Schwarz-Friesel (2007) vermutete, dass positiv besetzte Wörter schneller und leichter verfügbar sind, der Abruf negativer emotionaler semantischer Netze jedoch einen erhöhten Schwellenwert besitzt. Des Weiteren zeigte sich, dass die Probanden eine Präferenz für die verbalen Ausdrücke, die in der Forschung für die sechs Basisemotionen verwendet werden, aufwiesen. Lediglich die Verwendung von Wut schien gebräuchlicher zu sein als die von Ärger und wurde 2,3x häufiger genannt, auch wenn dies im klinischen Kontext kritisch diskutiert wird. Wut wird demnach nur bei einem hohen Erregungsniveau und starker Intensität verwendet. Diese Unterscheidung scheint in der Normalbevölkerung jedoch nicht getroffen zu werden.

Die subjektive Sicherheit korrelierte teilweise mit der Erkennungsleistung. Das Referenzsystem des Menschen, nämlich die eigene Einschätzung, funktioniert demnach nicht fehlerfrei. Ob der Mensch dazu neigt, sich eher zu über- oder unterschätzen oder ob es Zusammenhänge mit bestimmten Persönlichkeitseigenschaften gibt, wurde in diesem Zusammenhang nicht systematisch untersucht. Dies wäre aber sicherlich eine interessante Fragestellung für Folgestudien.

Man kann also schlussfolgern: Die Verwendung geschlossener Antwortformate im Bereich der Überprüfung der Emotionserkennungsfähigkeit kann nach diesen Befunden als unproblematisch angesehen werden. Der Vorteil hierbei ist, dass eine einfache und schnelle Auswertung der Ergebnisse möglich ist. Im Vorfeld sollte jedoch genauestens untersucht werden, welche verbalen Ausdrücke die emotionalen Stimuli am besten charakterisieren. In dieser Studie zeigte sich beispielsweise, dass Wut weit häufiger verwendet wurde als das zumeist in Studien mit geschlossenem Antwortformat verwendete Label Ärger. Aufgrund der hohen Übereinstimmung der korrekten Zuordnungen der emotionalen Gesichtsausdrücke zu den Emotionslabeln in Studie 1 und 2 und Erkennungsleistungen, die deutlich über der Zufallswahrscheinlichkeit lagen, kann von einer guten Kriteriumsvalidität gesprochen werden.

Potentielle Limitationen ergeben sich hinsichtlich folgender Aspekte: Die Untersuchung wies lediglich eine Teilnehmerzahl von N = 33 Probanden auf. Aufgrund des hohen zeitlichen Aufwands bei der Bearbeitung von ein bis zwei Stunden war die Teilnahmebereitschaft eher gering ausgeprägt. Es wurden circa 70 Personen auf die Studie aufmerksam gemacht, so dass der Rücklauf nur knapp 50% betrug. Außerdem erscheint die Verwendung ein und derselben randomisierten Bildreihenfolge für alle Probanden problematisch. Potentielle Reihenfolgeeffekte konnten mit dem bestehenden Datensatz nicht analysiert werden. Schlussendlich wurden in dieser Studie lediglich Probanden der höheren Bildungsschicht untersucht. Dies führt zu einer eingeschränkten Generalisierbarkeit der vorliegenden Ergebnisse.

4.3 Studie 3: Beurteilung der Intensität und Natürlichkeit des Emotionsausdrucks und der Attraktivität der fotografierten Person

Neben der Erfassung der Erkennungsleistung für die verschiedenen Basisemotionen spielen auch noch andere Gesichtspunkte bei der Erstellung eines standardisierten Bildersets eine Rolle. Mit der Betrachtung eines menschlichen Gesichts werden automatisch soziale Informationen wie Alter, äußeres Erscheinungsbild und Geschlecht an den Betrachter übermittelt (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & Knippenberg, 2010). In dieser Studie wurden Intensität und Natürlichkeit der gezeigten Emotion und die Attraktivität der fotografierten Person beurteilt. Ziel dieser Studie war es, diese als (potentielle) Einflussfaktoren bei der Emotionserkennung berücksichtigen zu können. Die gewonnenen Daten sollten bei der Auswahl des endgültigen Bildersatzes Berücksichtigung finden.

Hypothesen Studie 3

Hypothese 1: In den bisherigen Studien (O'Doherty, Winston, Critchley, Perrett, Burt & Dolan, 2003; Dion, Berscheid & Walster, 1972; Carroll & Russell, 1997; Hoffmann, Kessler, Eppel, Rukavina & Traue, 2010; Palermo & Coltheart, 2004) zeigen sich Einflüsse Intensität, Natürlichkeit und Attraktivität auf die Beurteilung emotionaler Gesichtsausdrücke. Die Einschätzung dieser drei Faktoren für das vorliegende Bildmaterial ermöglicht die Auswahl geeigneter emotionaler Stimuli für ein standardisiertes Testverfahren.
Hypothese 2: Explorativ werden Zusammenhänge zum Geschlecht der fotografierten Personen und der Beurteiler sowie den verwendeten Fragebogenverfahren untersucht.

Beurteilung von Intensität, Natürlichkeit und Attraktivität

Die Bewertung der drei Faktoren wurde von 68 Medizinstudenten (mittleres Alter = 21,29 Jahre; SA = 2,58; 33 Frauen und 35 Männer) vorgenommen, welche allesamt freiwillig an der Studie teilnahmen (Ethikvotum 245/08_UBB/se). Der Versuchsablauf gestaltete sich wie folgt: Die Studenten sahen nacheinander die 96 emotionalen Bilder aus Studie 1 in einer zufälligen Reihenfolge. Jedes Bild wurde in einer Größe von 800 Pixeln zusammen mit dem korrekten Emotionslabel angezeigt, wobei sich die Präsentationsdauer nach der Dauer beim Ausfüllen der drei Skalen richtete, jedoch maximal 30 Sekunden betrug. Unter dem zu bewertenden Bild wurden die drei Skalen gleichzeitig eingeblendet, die Bewertung konnte

mittels Mausklick vorgenommen werden (siehe Abbildung 14). Jede Emotion war mit 16 Bildern vertreten, pro Person waren Fotos von zwei emotionalen Gesichtsausdrücken enthalten.

Die Intensität wurde mit einer zehnstufigen Skala (0%-100%) erfasst. Diese Skalierung wurde aufgrund guter Erfahrungen mit dieser Einteilung in anderen Studien (zum Beispiel Hoffmann, Kessler, Eppel, Rukavina & Traue, 2010) übernommen. Aufgabe der Probanden war es, einzuschätzen, wie intensiv die dargestellte Emotion gezeigt wurde. Die Attraktivität wurde in Anlehnung an die Skala zur Erfassung der Attraktivität aus einer Studie von Stirrat, Gumert & Perret (2011) auf einer siebenstufigen Skala („sehr unattraktiv“ bis „sehr attraktiv“) eingeschätzt. Die Beurteiler sollten angeben, wie attraktiv ihnen die gezeigte Person vorkam. Als drittes wurde die Natürlichkeit des fotografierten Emotionsausdrucks bewertet, ebenfalls auf einer siebenstufigen Skala („sehr unnatürlich“ bis „sehr natürlich“). Jeder Proband hatte jedes emotionale Bild hinsichtlich dieser drei Faktoren zu bewerten.

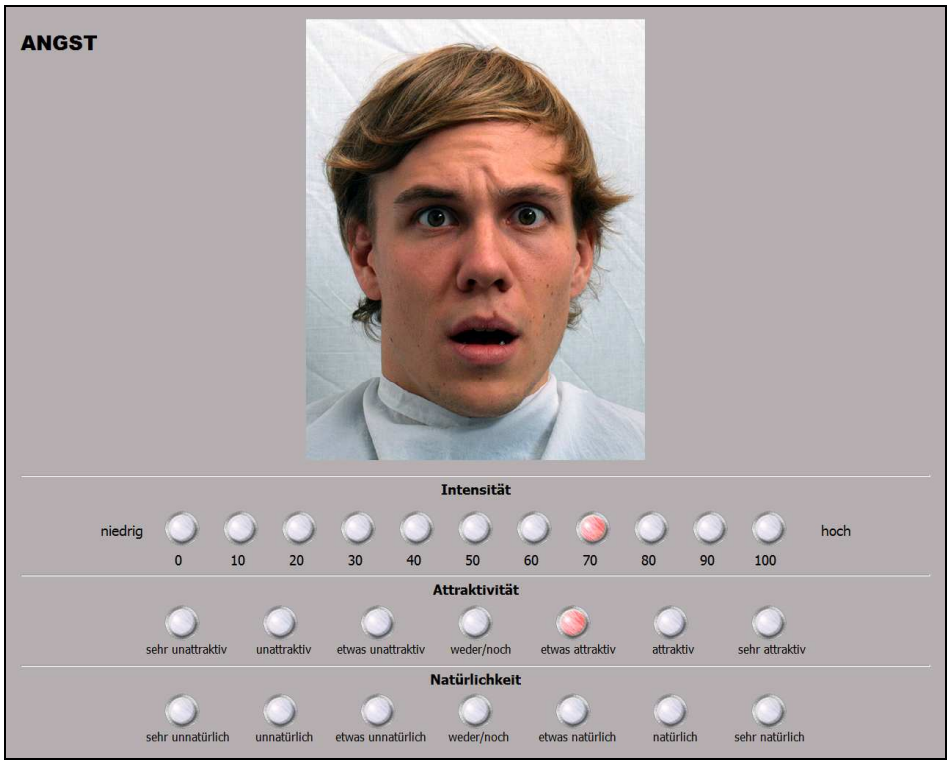


Abb. 14: Beispiel für die experimentelle Umsetzung der Bewertung von Natürlichkeit, Intensität, Attraktivität

Zusätzlich wurden der TeiQue (Cooper & Petrides, 2010; Petrides & Furnham, 2006) in seiner Kurzform, der ERQ (Abler & Kessler, 2009) und der NEO-FFI (Costa & McCrae, 2002; Borkenau & Ostendorf, 2008) als Fragebögen an die Probanden ausgegeben.

Beim TeiQue handelt es sich um ein Testverfahren zur Erfassung der Emotionalen Intelligenz als Trait. Unter Emotionaler Intelligenz versteht man hierbei die Fähigkeit, eigene und fremde Gefühle (korrekt) wahrzunehmen, zu verstehen und zu beeinflussen. Die Kurzform des TeiQue besteht aus 30 Items, aus denen ein globaler Wert errechnet werden kann (siehe Anhang D). Er basiert auf der Langform, welche 153 Items beinhaltet. Für die Kurzform wurden je zwei Items der 15 Subskalen ausgewählt, Auswahlkriterium war die Höhe der Korrelation mit dem Gesamtwert der jeweiligen Subskala. Antwortformat ist eine siebenstufige Likert-Skala. Grundsätzlich wird angenommen, dass ein korrelativer Zusammenhang zwischen der Emotionalen Intelligenz und der Erkennungsfähigkeit emotionaler Stimuli vorliegt. Vier Faktoren können aus den Werten des TeiQue gebildet werden: Wohlbefinden, Selbstkontrolle, Emotionalität und soziale Verträglichkeit sowie zusätzlich ein Gesamtwert.

Der ERQ (siehe Anhang E) in der deutschen Version (Abler & Kessler, 2009; basierend auf Gross & John, 2003) erlaubt die Untersuchung von Emotionsregulationsprozessen. Präferenzen für zwei häufig angewendete Strategien können erfasst werden: „suppression“ (Unterdrückung) und „reappraisal“ (Neubewertung). Zur Erfassung dieser beiden Größen werden zehn Items vorgegeben, welche von 1 („stimme überhaupt nicht zu“) bis 7 („stimmt vollkommen“) bewertet werden können. Die deutsche Version erreichte eine interne Konsistenz on r = .74 für „suppression“ und r = .76 für „reappraisal“, Ergebnisse welche mit denen der englischsprachigen Version vergleichbar erscheinen.

Der NEO-FFI (siehe Anhang F) stellt ein gut etabliertes Verfahren zur Messung fünf verschiedener Persönlichkeitszüge dar: Verträglichkeit, Offenheit, Extraversion, Neurotizismus und Gewissenhaftigkeit. Der NEO-FFI besteht aus 60 Fragen, 12 pro Domäne. Es basiert auf der populären Fünf-Faktoren-Theorie der Persönlichkeit (für mehr Informationen zu dieser Persönlichkeitstheorie siehe Amelang & Bartussek, 2001). Dieses Modell ist ein datenbasiertes, übergreifendes und empirisch überprüftes Modell. Für den NEO-FFI werden folgende interne Konsistenzen angegeben: Neurotizismus = .79, Extraversion = .79, Offenheit = .80, Verträglichkeit = .75, Gewissenhaftigkeit = .83. Antwortformat ist eine fünfstufige Likert-Skala. In Teilbereichen konnten Zusammenhänge zwischen der Ausprägung einzelner Persönlichkeitskonstrukte und der Verarbeitung von Gesichtern beziehungsweise emotionaler Stimuli nachgewiesen werden (Chan, Goodwin & Harmer, 2007; Rubin, Munz & Bommer, 2005; Welling, DeBruine, Little & Jones, 2009).

Ergebnisse der Beurteilung von Intensität, Natürlichkeit und Attraktivität

Zunächst wurden die Mittelwerte und Standardabweichungen für die drei Skalen auf Einzelemotionsebene und als Gesamtwert berechnet. Diese Werte sind Tabelle 10 zu entnehmen.

Tab. 10: Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit

Emotion	**Intensität (11-stufig)**		**Attraktivität (7-stufig)**		**Natürlichkeit (7-stufig)**	
	M	**SA**	**M**	**SA**	**M**	**SA**
Angst	6,77	1,13	4,05	0,63	4,23	0,72
Ärger	6,13	1,19	3,82	0,66	3,81	0,77
Ekel	7,14	1,12	3,25	0,69	4,17	0,94
Freude	7,35	1,19	4,74	0,66	5,77	0,53
Trauer	5,68	1,29	4,09	0,62	4,75	0,66
Überraschung	7,18	1,21	3,98	0,61	3,75	0,82
Gesamt	6,71	0,97	3,99	0,54	4,41	0,53

Aufgrund der bestätigten Normalverteilungsvoraussetzungen (Shapiro-Wilks-Test: $p > 0.639$ für Intensität; $p > 0.059$ für Attraktivität; $p > 0.356$ für Natürlichkeit) wurden im Folgenden parametrische Verfahren zur statistischen Analyse der Daten verwendet. Berechnungen mittels ANOVA (Emotion x Beurteilung) ergaben signifikante Unterschiede zwischen den verschiedenen Emotionen für die Intensität ($F(1, 5) = 21{,}147$; $p < .000$), die Natürlichkeit ($F(1, 5) = 86{,}530$; $p < .000$) und die Attraktivität ($F(1, 5) = 37{,}670$; $p < .000$. In der Folge wurden zur tiefergehenden Datenanalyse post-hoc-Tests durchgeführt. Eine detaillierte Übersicht findet sich in Tabelle 11. Die post-hoc-Analysen ergaben nur einige signifikante Unterschiede bei der Einschätzung der drei Faktoren für die sechs Basisemotionen ($p < .05$). Für die Skala Intensität ergab sich ein signifikanter Unterschied für die Beurteilung von Angst und Trauer ($p < .000$), sowie Ärger mit Ekel, Freude und Überraschung ($p < .000$ für alle aufgeführten Paarvergleiche). Außerdem konnten signifikante Unterschiede für die Einschätzung der Intensität von Trauer mit Ekel, Freude und Überraschung gefunden werden ($p < .000$ für alle aufgeführten Paarvergleiche). Für die Einschätzung der Attraktivität ergaben sich signifikante Unterschiede zwischen Ekel und allen anderen Emotionen sowie Freude und allen anderen Emotionen ($p < .000$ für alle aufgeführten Paarvergleiche). Die Beurteilung der Natürlichkeit ergab signifikante Unterschiede für Freude und Trauer mit allen anderen Emotionen ($p < .000$ für alle aufgeführten Paarvergleiche).

Tab. 11: Ergebnisse der Paarvergleiche für Intensität, Attraktivität und Natürlichkeit

	Ärger			**Ekel**			**Freude**			**Trauer**			**Überraschung**		
	In	**Att**	**Nat**	**In**	**Att**	**Nat**	**In**	**Att**	**Nat**	**In**	**Att**	**Nat**	**In**	**Att**	**Nat**
Angst	.088	.536	.068	.659	.000	.999	.153	.000	.000	.000	.999	.006	.531	.998	.019
Ärger				.000	.000	.179	.000	.000	.000	.432	.308	.000	.000	.823	.999
Ekel							.954	.000	.000	.000	.000	.001	1.00	.000	.062
Freude										.000	.000	.000	.984	.000	.000
Trauer													.000	.967	.000

In = Intensität, Att = Attraktivität, Nat = Natürlichkeit

Korrelationsanalysen ergaben etliche signifikante positive Korrelationen zwischen allen drei Variablen ($p < .05$) und alle sechs Basisemotionen im mittleren bis hohen Bereich (siehe Tabellen 12 bis 14). Dies bedeutet ganz grundsätzlich: Je höher die eine Variable ausgeprägt ist, desto höher ist auch die andere Variable ausgeprägt.

Tab. 12: Korrelationen der Faktoren Intensität und Attraktivität auf Emotionsebene

Attraktivität x Intensität	
Angst	.479**
Ärger	.406**
Ekel	.263*
Freude	.487**
Trauer	.520**
Überraschung	.555**

* = p < .05; ** = p < .001

Tab. 13: Korrelationen der Faktoren Intensität und Natürlichkeit auf Emotionsebene

Natürlichkeit x Intensität	
Angst	.322**
Ärger	.362**
Ekel	.331**
Freude	.295*
Trauer	.260*
Überraschung	.055

* = p < .05; ** = p < .001

Tab. 14: Korrelationen der Faktoren Natürlichkeit und Attraktivität auf Emotionsebene

Attraktivität x Natürlichkeit	
Angst	**.437****
Ärger	**.367****
Ekel	**.412****
Freude	**.330****
Trauer	**.355****
Überraschung	**.341****

* = p < .05; ** = p < .001

Analyse von Geschlechtereffekten

Da Frauen bereits bei geringen Intensitäten emotionale Gesichtsausdrücke korrekt zuordnen (Hoffmann, Kessler, Eppel, Rukavina & Traue, 2010), wurde das Geschlecht als möglicher differenzierender Faktor eingeführt. In einem zweiten Schritt wurden somit jeweils Bilder mit nur männlichen beziehungsweise nur weiblichen Darstellern untersucht (je acht pro Emotion). Eine Übersicht über die gefundenen Mittelwerte ist Tabelle 15 zu entnehmen.

Tab. 15: Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit getrennt nach dem Geschlecht der Darsteller

Mittelwert	**Intensität (11-stufig)**		**Attraktivität (7-stufig)**		**Natürlichkeit (7-stufig)**	
	Männliche Darsteller	**Weibliche Darsteller**	**Männliche Darsteller**	**Weibliche Darsteller**	**Männliche Darsteller**	**Weibliche Darsteller**
Angst	6,55	6,98	3,85	4,24	4,12	4,34
Ärger	6,21	6,06	3,68	3,96	3,77	3,86
Ekel	6,93	7,35	3,16	3,34	3,97	4,37
Freude	7,44	7,24	4,59	4,88	5,68	5,85
Trauer	5,37	5,99	3,64	4,61	4,57	4,94
Überraschung	7,19	7,17	3,73	4,24	3,58	3,91
Gesamt	6,62	6,79	3,78	4,21	4,28	4,55

In der Folge wurden potentielle Unterschiede zwischen den Gesamtmittelwerten und den Mittelwerten für männliche und weibliche Darsteller für die drei Faktoren untersucht. Hierzu wurde der t-Test eingesetzt. Für die Intensität zeigt sich lediglich ein signifikanter Unterschied zwischen männlichen und weiblichen Darstellern für die Emotion Angst (t(134) = -2,051; p = .042), Ekel (t(134) = -2,043; p = .043) und Trauer (t(134) = -2,696; p = .008). In allen drei Fällen wurde die Intensität für weibliche Darsteller signifikant höher eingeschätzt.
Für die Skala Attraktivität konnten signifikante Mittelwertsunterschiede für alle Emotionen außer Ekel gefunden werden (Angst: t(134) = -3,188; p = .002; Ärger: t(134) = -2,269; p = .025; Ekel: (t(134) = -1,385; p = .168; Freude: t(134) = -2,273; p = .025; Trauer: t(134) =

-8,046; p < .000; Überraschung: t(134) = -4,351; p < .000). Die weiblichen Darsteller wurden insgesamt attraktiver eingeschätzt.

Unterschiede bezüglich männlichen und weiblichen Darstellern hinsichtlich der Einschätzung der Natürlichkeit ergaben sich für die Emotionen Ekel (t(134) = -2,334; p = .021) und Trauer (t(134) = -2,831; p = .005). Die Emotionsausdruck der weiblichen Darsteller wurde als „natürlicher" bewertet.

Über alle Emotionen hinweg berechnet, zeigte sich für die Intensität kein signifikanter Unterschied zwischen männlichen und weiblichen Darstellern (t(134) = 1,170; p = .244), jedoch für die Natürlichkeit (t(134) = 2,777; p = .006) und die Attraktivität (t(134) = 4,392; p < .000).

In einem dritten Schritt wurden die Mittelwerte für die weiblichen und männlichen Beurteiler errechnet. Diese sind in Tabelle 16 zu finden.

Tab. 16: Mittelwerte der Skalen für Intensität, Attraktivität und Natürlichkeit getrennt nach Geschlecht der Beurteiler

Mittelwert	**Intensität (11-stufig)**		**Attraktivität (7-stufig)**		**Natürlichkeit (7-stufig)**	
	Männliche Beurteiler	**Weibliche Beurteiler**	**Männliche Beurteiler**	**Weibliche Beurteiler**	**Männliche Beurteiler**	**Weibliche Beurteiler**
Angst	6,74	6,79	4,05	4,04	4,23	4,23
Ärger	6,18	6,08	3,90	3,73	3,85	3,77
Ekel	7,12	7,15	3,29	3,21	4,25	4,08
Freude	7,36	7,35	4,71	4,77	5,75	5,78
Trauer	5,71	5,65	4,14	4,04	4,73	4,78
Überraschung	7,25	7,11	4,09	3,87	3,89	3,61
Gesamt	6,73	6,69	4,03	3,94	4,45	4,38

Zur Analyse möglicher Unterschieder wurden t-Tests eingesetzt. Für den Faktor Intensität konnte kein signifikanter Einfluss für das Geschlecht der Beurteiler gefunden werden (für alle Emotionen p > .05). Der Faktor Natürlichkeit lieferte ein signifikantes Ergebnis: Männliche Beurteiler schätzten die Natürlichkeit der überraschten Gesichtsausdrücke signifikant höher ein als weibliche Beurteiler (t(66) = -1,408; p = .034). Für Attraktivität zeigten sich zwei signifikante Unterschiede: Die Attraktivität von Personen, die Ärger und Überraschung ausdrückten wurde von männlichen Beurteilern signifikant höher eingeschätzt als von wieblichen Beurteilern (Ärger: t(66) = -1,064; p = 0.28, Überraschung: t(66) = -1,475; p = .044). Aufgrund der geringen Anzahl an signifikanten Unterschieden kann geschlussfolgert

werden, dass sich in dieser Studie aus dem Geschlecht der Beurteiler kein systematischer Einfluss ergab.

Auswertung der eingesetzten Fragebogenverfahren

Nachfolgend werden die Ergebnisse der eingesetzten Fragebögen vorgestellt. Eine Übersicht der Ergebnisse findet sich in Tabelle 17.

Tab. 17: Mittelwerte, Standardabweichung, Minimum und Maximum der Fragebögen ERQ, NEO-FFI und TeiQue

	Minimum	**Maximum**	**M**	**SA**
ERQ (Skala von 1-7)				
Reappraisal	1,67	6,67	4,54	1,04
Suppression	1,25	7,00	3,50	1,25
NEO-FFI (Skala von 0-4)				
Neurotizismus	0,67	3,25	1,75	0,61
Extraversion	1,42	3,50	2,46	0,49
Offenheit	0,67	3,67	2,49	0,60
Verträglichkeit	1,00	3,42	2,57	0,49
Gewissenhaftigkeit	1,17	4,00	2,75	0,61
TeiQue (Skala von 1-7)				
Gesamt	3,83	6,47	5,00	0,49
Wohlbefinden	3,67	7,00	5,56	0,72
Selbstkontrolle	3,00	6,67	4,78	0,70
Emotionalität	3,00	7,00	5,00	0,84
Soziale Verträglichkeit	3,17	7,00	5,01	0,74

Für die hier vorliegende Arbeit war es weniger von Belang, die erhobenen Werte aus den Fragebögen mit bestehenden Normen zu vergleichen. Bedeutsamer waren potentielle Zusammenhänge zwischen den erhobenen Fragebogen-Daten und der Beurteilung von Intensität, Attraktivität und Natürlichkeit. Hierzu wurden Korrelationen nach Pearson berechnet. Korreliert wurden die Mittelwerte der Fragebögen (siehe Tabelle 17) mit den Mittelwerten der Beurteilungen von Intensität, Attraktivität und Natürlichkeit. Es zeigten sich keine signifikanten Zusammenhänge.

Diskussion der Beurteilungen von Intensität, Natürlichkeit und Attraktivität

Im Unterschied zu den beiden anderen Faktoren wurde die Qualität der Intensität schon bei der Erstellung des Bildmaterials berücksichtigt. Alle Darsteller sollten die jeweilige Emotion in maximaler Intensität präsentieren. Aufgrund von Unterschieden in der Physiognomie und der Kompetenz bei der Darstellung der zugrunde liegenden muskulären Veränderungen war

jedoch trotzdem anzunehmen, dass sich bei der Bewertung Unterschiede zeigen würden. Insgesamt als am wenigsten intensiv wurden Trauer und Ärger eingeschätzt. Beides sind Emotionen, die den negativen Emotionen zuzuordnen sind. Es ist anzunehmen, dass gerade für die negativ konnotierten Emotionen besondere gesellschaftliche Rollen gelten (Ekman & Friesen, 2003). Beispielsweise beschreibt Ekman in seinem Buch (Ekman, 2007a), dass Trauer häufig mit Weinen und Schluchzen in Verbindung gebracht wird. Da die Bilder keinen natürlichen, sondern einen willentlich herbeigeführten Emotionsausdruck darstellten und alle Emotionsausdrücke im Sinne der Standardisierung lediglich über muskuläre Veränderungen dargestellt werden sollten, lagen keine Trauerbilder mit weinenden Personen vor. Schluchzen stellt einen dynamischen Vorgang dar, so dass dies ebenfalls nicht in die Erstellung des Bildmaterials einfließen konnte. Harwood, Hall und Shinkfield (1999) konnten jedoch in ihrer Studie nachweisen, dass gerade die Emotionen Trauer und Ärger besser erkannt wurden, wenn sie dynamisch präsentiert wurden. Es also nicht auszuschließen, dass Trauer intensiver eingeschätzt worden wäre, hätten diese beiden Aspekte berücksichtigt werden können. Ähnlich gestaltet sich die Diskussion bei Ärger. Ärger ist eine hochrelevante Emotion, die der Mensch in der Regel recht gut erkennen kann (beispielsweise Kessler, Bayerl, Deighton & Traue, 2002). Aus der Identifikation dieser Emotion leitet sich direkt ab, ob das eigene Leben durch eine andere Person bedroht scheint. Aufbauend auf den Ergebnissen von Harwood, Hall und Shinkfield (1999) ist denkbar, dass statisches Bildmaterial zwar aufgrund der Relevanz immer noch sehr gut erkannt, aber als weniger bedrohlich eingestuft wird. Eine geringere Bedrohlichkeit würde sich zugleich im Ausmaß der Intensität niederschlagen.

Natürlichkeit hat einen wichtigen Einfluss auf die ökologische Validität des Bildmaterials (Carroll & Russell, 1997). Freude wurde in diesem Fall als natürlichster Gesichtsausdruck bewertet. Dies erscheint insofern logisch, als dass Freude als einziger Gesichtsausdruck nur dann fotografiert wurde, wenn diese Emotion im Gespräch evoziert werden konnte, um ein Duchenne-Lächeln zur erhalten. Die Relevanz von Natürlichkeit wurde bisher nur in wenigen Studien untersucht. In einer Studie von Langner, Dotsch, Bijlstra, Wigboldus, Hawk & Knippenberg, 2010 wurde die Authentizität der gezeigten Emotion auf einer fünfstufigen Skala erfasst. Tiefergehende statistische Auswertungen wurden leider nicht vorgestellt, so dass keine Aussagen hinsichtlich der Vergleichbarkeit getroffen werden können. Elfenbein und Kollegen (2007) konnten nachweisen, dass die Natürlichkeit eine bedeutsame Rolle bei

der Emotionserkennung spielt. Während prototypische Emotionsausdrücke kulturübergreifend gut erkannt wurden, wurden natürlich erstellte Emotionsausdrücke nur im jeweiligen Kulturkreis gut erkannt. Ein höheres Maß an Natürlichkeit könnte durch die Verwendung dynamischer Stimuli erzielt werden oder wie nachfolgend in Kapitel 5 vorgestellt durch die Erstellung von Bildmaterial aus verschiedenen Perspektiven.
Attraktivität ist ein hochrelevanter Faktor in der sozialen Interaktion. Innerhalb von Sekunden realisieren wir, ob wir eine andere Person mögen oder nicht, beziehungsweise ob wir eine Person attraktiv finden oder nicht. Insgesamt betrachtet zeigt sich, dass Freude am attraktivsten und Ekel am unattraktivsten eingestuft wurde. Es wird vermutet, dass die die Bedeutung/der Gehalt einer jeden Emotion einen Einfluss auf die Einschätzung der Attraktivität hat. Beispielsweise kommuniziert Freude Zugänglichkeit und Ekel Abneigung (Becker, Kenrick, Neuberg, Blackwell & Smith, 2007). Es ist also nicht auszuschließen, dass positive Emotionen, welche Annäherung versprechen, mit höheren Attraktivitätsbeurteilungen einhergehen und negative Emotionen mit niedrigeren Attraktivitätswerten. Zusätzlich müssen sozio-kulturelle Geschlechterrollen beachtet werden.

Diskussion der Geschlechtereffekte

Männliche und weibliche Beurteiler wiesen nahezu keine signifikanten Unterschiede bei der Einschätzung von Intensität, Natürlichkeit und Attraktivität auf. Ein systematischer Geschlechtereffekt durch den Beurteiler ist auszuschließen. Jedoch spielte das Geschlecht der darstellenden Person in der vorliegenden Studie eine wichtige Rolle. Waren signifikante Unterschiede zu verzeichnen, so zeigte sich wiederkehrend die gleiche Tendenz: Die weiblichen Darsteller wurden attraktiver, im Ausdruck natürlicher und intensiver eingeschätzt als die männlichen Darsteller. Dies entspricht den im Vorfeld aufgestellten Vermutungen. Einige Studien (beispielsweise Kring & Gordon, 1998) wiesen nach, dass Frauen einen expressiveren Emotionsausdruck zeigen als Männer. Diskutiert wurden evolutionäre Ursachen, beispielsweise die Notwendigkeit Frauen aufgrund ihrer körperlichen Unterlegenheit vor körperlichen Angriffen zu schützen. Es ist denkbar, dass Frauen in der Folge gelernt haben, Emotionen stärker beziehungsweise facettenreicher ausdrücken zu können, um eine schnelle und effiziente Kommunikation mit möglichen „Beschützern“ gewährleisten zu können (Bijlstra, Holland & Wigboldus, 2010; Hess, Senécal, Kirouac,

Herera, Philippot & Kleck, 2000). Bradley, Codispot, Sabatinelli und Lang (2001) zeigten beispielsweise, dass Frauen eine ausgeprägtere physiologische Verteidigungsreaktion bei der Präsentation aversiven Bildmaterials aufwiesen als Männer. Sowohl biologische als auch strukturelle Aspekte können hier diskutiert werden. Für den Faktor Attraktivität entsprechen die Ergebnissen denen einer Studie von Tracy & Beall (2011), welche herausfanden, dass Frauengesichter in der Regel als liebenswürdiger und attraktiver eingeschätzt wurden als die Gesichter von Männern.

Diskussion der Ergebnisse aus den eingesetzten Fragebogenverfahren

Zwischen den eingesetzten Fragebögen und der hier geforderten Aufgabe konnte kein Zusammenhang gefunden werden. Es kann gemutmaßt werden, dass Zusammenhänge eher zwischen Emotionserkennung und den erfassten Fragebogenkonstrukten als bei der hier gestellten Bewertungsaufgabe gefunden werden.
Ein Vergleich mit Normstichproben zeigte: Für den TeiQue scheinen die hier gefundenen Werte mit denen bisheriger Normierungsstudien vergleichbar zu sein (Freudenthaler, Neubauer, Gabler & Scherl, 2008). Ein Einstichproben-t-Test ergab keine signifikanten Unterschiede für den Gesamtwert ($t(65) = 0{,}691$; $p = .492$). Die Ergebnisse des NEO-FFI wurden mit einer aktuellen britischen Normstichprobe verglichen (Egan, Deary & Austin, 2000). Die Werte für Neurotizismus ($t(69) = 1{,}678$; $p = .098$), Verträglichkeit ($t(69) = 1{,}668$; $p = .100$) und Gewissenhaftigkeit ($t(70) = 1{,}103$; $p = .274$) unterschieden sich nicht signifikant von der hier erhobenen Stichprobe, die Werte für Extraversion ($t(70) = 3{,}523$; $p = .001$) und Offenheit ($t(68) = 3{,}806$; $p = .000$) waren jedoch in der hier vorliegenden Stichprobe signifikant höher ausgeprägt. Die Werte des ERQ (Abler & Kessler, 2009) wurden jeweils für Männer und Frauen getrennt bewertet und sind mit den bestehenden Normen vergleichbar (Frauen: $t(35) = 1{,}083$; $p = .080$ für Unterdrückung und $t(35) = 0{,}485$; $p = .630$ für Neubewertung; Männer: $t(35) = -0{,}285$; $p = .777$ für Unterdrückung und $t(35) = -1{,}235$; $p = .225$ für Neubewertung).

Abschließende Bemerkungen

Insgesamt zeigte sich eine mittlere Konsistenz der drei Faktoren. Signifikante Unterschiede zwischen verschiedenen Emotionen waren für die Beurteilung von Intensität, Natürlichkeit

und Attraktivität beobachtbar. Insgesamt gilt: Je homogener der Bildersatz, desto eher können (potentielle) Einflüsse auf die Erkennungsleistung durch Intensität, Attraktivität und Natürlichkeit ausgeschlossen werden. Es lassen sich nur wenige Parameter heute technisch bearbeiten. Beispielsweise lässt sich die Intensität mittels Morphing aneinander angleichen. Dies ist allerdings nur schwer für die Natürlichkeit oder Attraktivität möglich. Möglichkeiten, die „reine“ Erkennungsleistung dennoch berechnen zu können, ergeben sich aus Verfahren wie der Kovarianzanalyse oder der Partialkorrelation. Bei diesen ist es ausreichend, wenn das Ausmaß des Einflusses bekannt ist. Wichtige Voraussetzung ist aber, dass die Varianz innerhalb einer Emotionskategorie nur gering ist, damit die generellen Unterschiede kontrollierbar bleiben. Dies scheint der Fall zu sein.

Es gilt festzuhalten, dass es bei der Erstellung eines Bildersets nicht ausreicht, lediglich die Erkennungsleistung zu berücksichtigen. Unterschiede in ihrer Ausprägung konnten für alle drei Faktoren – Natürlichkeit, Intensität und Attraktivität – nachgewiesen werden. Um eine systematische Einflussnahme dieser auszuschließen beziehungsweise zu verringern, sollte bei einer Bildauswahl ein Höchstmaß an Ähnlichkeit für die einzelnen Bilder angestrebt werden – und das für jede (bekannte) mögliche Einflussgröße.

Eine interessante Fragestellung, die hier leider nicht untersucht werden konnte, ist: Steht die Art der Emotion in einem direkten Zusammenhang mit der Einschätzung der drei Variablen Intensität, Natürlichkeit und Attraktivität? Hierzu müsste man von jeder Person Gesichtsausdrücke mit allen sechs Basisemotionen vorgeben, um Personeneffekte ausschließen zu können. Denkbar wäre beispielsweise, dass Fotos mit freudigem Gesichtsausdruck natürlicher und attraktiver eingeschätzt würden als ein Foto mit verärgertem Gesichtsausdruck.

Nicht untersucht werden konnten in dieser Studie außerdem der direkte Zusammenhang zwischen Emotionserkennung und der Beurteilung von Intensität, Natürlichkeit und Attraktivität. Eine Abfrage dieser beiden Parameter, Emotionserkennungsfähigkeit einerseits und die Beurteilung der drei Faktoren andererseits, würde Aussagen über Zusammenhänge ermöglichen. Beispielsweise könnten Korrelationen, Pfadanalysen oder multiple Regressionen berechnet werden.

4.4 Studie 4: Auswahl des endgültigen Bildersatzes und Normierung der Erkennungsleistung

Die hier durchgeführte Studie wurde aus zwei Gründen durchgeführt: Zum einen sollte eine Reduktion des Bildmaterials von 96 auf 48 Bilder, das heißt auf einen emotionaler Stimulus pro fotografierter Person erfolgen. Hierzu soll nicht nur die Erkennungsleistung herangezogen werden, sondern auch die Ergebnisse der Studien 2 und 3 Berücksichtigung finden. Zum anderen sollten erste Daten für eine Normierung erhoben werden. Zwei Ausgänge sind grundsätzlich denkbar: 1. es liegt eine Normalverteilung vor – das wäre der Idealfall, 2. es zeigen sich Deckeneffekte. Sollte diese zweite Option zutreffen, so muss über Maßnahmen nachgedacht werden, um eine Normalverteilung zu erreichen. Aus diesem Grund sollen Itemschwierigkeit und Trennschärfe berechnet werden, auf deren Grundlage weitere Schritte diskutiert werden können. Ebenfalls von Bedeutung ist in diesem Zusammenhang die interne Konsistenz (Cronbach's alpha). Die interne Konsistenz bezeichnet das Ausmaß, in dem die Items (also hier die Bildstimuli) miteinander in Beziehung stehen. Sie gibt einen Hinweis auf die Güte eines jeden Stimulus.

Hypothesen Studie 4

Hypothese 1: Die sechs mimisch gezeigten Basisemotionen werden unterschiedlich gut erkannt. Es wird angenommen, dass Freude am besten und Angst am schlechtesten erkannt wird. Für die Rangfolge der Erkennungsleistung der anderen Emotionen wird keine Hypothese aufgestellt.
Hypothese 2: Explorativ werden Zusammenhänge zwischen Erkennungsleistung und den Variablen Geschlecht, Alter, Persönlichkeit und Status untersucht.
Hypothese 3: Auf Grundlage der hohen Erkennungsleistungen für die einzelnen Emotionen in den vorherigen Studien (Studie 1+2) wird außerdem die Hypothese aufgestellt, dass keine Normalverteilung der Werte vorliegt und sich die Deckeneffekte negativ auf die Itemschwierigkeit auswirken.

Durchführung Onlinestudie zur Erfassung der Erkennungsleistung

Eine Normierung muss sich immer an einer bestimmten Zielgruppe orientieren, für die eine Aussage getroffen werden soll. Aus der Literatur ist bekannt (siehe 2.2.2), dass Alter, Status,

aber auch Geschlecht und Persönlichkeit die Emotionserkennung beeinflussen können. Zwar erschien es nicht möglich, für alle aufgeführten Faktoren hinreichend große Substichproben erheben zu können, dennoch sollten diese Variablen von jedem Teilnehmer erfasst werden, um potentielle Einflüsse errechnen zu können. Angestrebt wurde lediglich die systematische Erhebung einer älteren und einer jüngeren Stichprobe beiderlei Geschlechts, die nicht ausschließlich aus Probanden der Universität und somit der höheren Bildungsschicht besteht. Aus Gründen der Zugänglichkeit wurde eine internetbasierte Umsetzung für die Durchführung dieser Studie angestrebt. Es wurde angenommen, dass eine größere und facettenreichere Stichprobe rekrutiert werden könnte und somit eine möglichst hohe Repräsentativität gewährleistet werden kann. Für die Datenerhebung wurde Limesurvey verwendet (siehe http://www.limesurvey.org/de). Limesurvey ist eine in PHP geschriebene Software zur Datenerhebung mittels Online-Umfragen, die eine Vielzahl von Features umfasst. Für diese Studie wurde eine modifizierte Version entwickelt, welche es erlaubt, Bilder vorzuladen, damit es bei der Darstellung zu keinen Fehlern kommt. Hierzu geht ein JavaScript vor Beginn der Online-Studie alle Bilder durch, die während der Präsentation vorkommen und lädt diese. Das Programm läuft uni-intern auf einer Windows XP Maschine, ebenso wie die MySQL Datenbank, auf der die Ergebnisse gespeichert werden.

Insgesamt konnten über Mund-zu-Mund-Propaganda und Aushänge 189 Probanden rekrutiert werden. Bei 28 Probanden war eine Durchführung am heimischen PC leider aufgrund fehlender technischer Voraussetzungen (vermutlich aufgrund zu langsamer Internetverbindung und/oder zu wenig Arbeitsspeicher) nicht möglich. Es liegen für diese Personen keinerlei Daten vor. Alle Auswertungen wurden deswegen mit der verbleibenden Stichprobe von N = 161 durchgeführt. Alle Personen dieser Gruppe konnten die Online-Studie vollständig durchlaufen und die Ergebnisse am Ende abspeichern. Eine genauere Beschreibung der Stichprobe folgt im Ergebnisteil. Der Versuchsablauf war wie folgt: Über einen Link konnten die Probanden auf die Startseite der Studie gelangen. Sie wurden gebeten, die dort aufgeführten Informationen zur Studie (Ablauf, mögliche Risiken, Nutzung der Daten) genau durchzulesen. Es wurde darauf hingewiesen, dass, anstelle des schriftlichen Einverständnisses, das Klicken des „Weiter“-Buttons als Zustimmung zur Teilnahme an der Studie gelte (Ethikvotum 245/08_UBB/se). Außerdem wurde den Probanden erklärt, dass die Studie aus Zwecken der Vergleichbarkeit unter möglichst standardisierten Bedingungen durchgeführt werden müsse. Die Probanden wurden

deswegen darum gebeten, sich für die Durchführung der Studie etwa eine Stunde Zeit zu nehmen und dafür zu sorgen, in dieser Zeit nicht gestört zu werden. Außerdem wurde darum gebeten, den Bildschirm in einem Abstand von etwa 40 Zentimetern zum Gesicht auszurichten, da alle Bilder standardisiert in einheitlicher Größe dargestellt wurden (unabhängig von der Bildschirmgröße). Die Durchführung der Studie sollte des Weiteren alleine, ohne störende Geräusche und bei guter Beleuchtung stattfinden. In der Folge wurden einige demografische Daten erhoben: Alter, Geschlecht, monatliches Einkommen, Familienstand, aktuell ausgeübter Beruf und höchster Bildungsabschluss. Zusätzlich wurde zur Einschätzung der Persönlichkeit das „Ten Item Personality Inventory“ (Gosling, Rentfrow & Swann, 2003; deutsche Version von Hell & Muck) eingesetzt (siehe Anhang G). Hierbei handelt es sich um einen aus zehn Fragen für die Forschung entwickelten Fragebogen, der eine Einschätzung der Persönlickeitsdimensionen nach den „Big Five“ (Costa & McCrae, 2002) ermöglicht. Nach Geschlecht getrennte Normwerte stellen die Autoren zur Einordnung der Daten zur Verfügung. Nach diesem ersten Teil wurden die Probanden noch einmal gefragt, ob sie bereit seien für die Beurteilung der Stimuli. Nach Bestätigung per Mausklick wurden die 96 emotionalen Gesichtsausdrücke präsentiert. Wie in den vorherigen Studien handelt es sich um 16 Bilder pro Emotion, dargestellt von 48 Personen. Zunächst wurde das neutrale Bild einer Person für 1,5 Sekunden auf dem Computerbildschirm dargestellt, dann nach einer Pause von einer Sekunde der emotionale Ausdruck, ebenfalls für eine Sekunde. Dies stellt einen Unterschied zu den bisher beschriebenen Studien dar, bei denen kein neutrales Bild gezeigt wurde. Zugrundeliegende Idee war, dass die Aufgabenschwierigkeit erhöht ist, wenn kein neutrales Referenzbild angezeigt wird. Hier sollte jedoch der FEEL-Test in seiner bisherigen Form möglichst exakt nachgebildet werden, um die hier vorliegenden Erkennungsraten mit denen der Normierung des alten Bildmaterials vergleichen zu können. Die Studienteilnehmer hatten dann zehn Sekunden Zeit, diesen Ausdruck zu beurteilen. Hiefür standen ihnen acht Auswahloptionen zur Verfügung. Alternativ zu den sechs Basisemotionen konnten die Probanden auch die Option „nicht erkannt“ wählen, um der Möglichkeit potentieller „Zufallstreffer“ entgegenzuwirken. Da der technische Ablauf nicht immer reibungslos gewährleistet werden konnte, stand außerdem eine achte Option – „kein Bild angezeigt“ – zur Verfügung. So konnten später bei der Auswertung neue Grundgesamtheiten pro Versuchsperson auf Grundlage der angezeigten Bilder errechnet

werden. Nach Durchführung der Studie speicherten die Probanden die Daten ab und es wurde ihnen für die Teilnahme gedankt.

Ergebnisse der Onlinestudie: Allgemein

Zunächst erfolgt eine Beschreibung der Gesamtstichprobe. In der Folge werden Unterschiede zwischen möglichen Teilstichproben berichtet. Hierzu werden folgende Aspekte herangezogen: Alter, Status, Persönlichkeit und Geschlecht.
Die Gesamtstichprobe bestand zu 59% aus Frauen und 41% aus Männern. Das mittlere Alter lag bei M = 39,19 Jahren, der Median bei 35. Dieser wurde verwendet, um die Gruppe im Nachhinein in zwei Altersbereiche zu unterteilen. Eine detaillierte Erfassung von Schulabschluss, Einkommen und Kindern ist Tabelle 18 zu entnehmen.

Tab. 18: Verteilung von Schulabschluss, Einkommen und Kindern in der Gesamtstichprobe

Höchster Bildungsabschluss	**Einkommen**	**Anzahl Kinder**
0,6% gehen noch zur Schule	17,4% unter 500 Euro	52,8% keine Kinder
2,5% Hauptschule	11,8% 500 bis 1000 Euro	11,2% ein Kind
21,7% Realschule	29,8% 1000 bis 2000 Euro	26,1% zwei Kinder
5% Fachhochschulreife	14,9% 2000 bis 3000 Euro	5,6% drei Kinder
24,2% Abitur	8,1% 3000 bis 4000 Euro	3,7% mehr als drei Kinder
44,7% Studium	14,9% über 4000 Euro	0,6% sonstiges
1,2% anderes	3,1% anderes	

Die Erkennungsleistung für die 96 Bilder wurde zunächst für die Gesamtstichprobe berechnet. Wie bereits beschrieben, galt es zu beachten, dass aufgrund technischer Probleme bei einigen Probanden nicht alle Bilder angezeigt werden konnten. Dies konnte von Probanden durch die Möglichkeit zur Auswahl des Feldes „kein Bild angezeigt“ vermerkt werden, so dass für jeden einzelnen Probanden im Anschluss neue Grundgesamtheiten – bereinigt um die nicht angezeigten Bilder – errechnet werden konnten. Im Schnitt wurden 2,30% der Bilder nicht angezeigt. Eine Übersicht über die Zuordnung zu den verschiedenen Kategorien ist Tabelle 19 zu entnehmen.

Tab. 19: Zuordnung zu den einzelnen Emotionskategorien (96 Bilder, in Prozent)

		Angst	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Nicht erkannt**
Stimulus-emotion	**Angst**	**69,87**	0,07	10,79	13,71	2,29	1,65	1,61
	Freude	0,12	**97,92**	0,85	0,08	0,16	0,19	0,68
	Überraschung	6,39	1,88	**90,43**	0,16	0,12	0,21	0,81
	Ekel	0,58	0,21	1,73	**86,27**	0,59	9,31	1,31
	Trauer	10,65	0,41	3,09	3,20	**73,39**	5,55	3,69
	Ärger	1,78	0,43	1,48	2,55	1,77	**89,89**	2,11

Auf Grundlage der hier erzielten Erkennungsraten erfolgte die Auswahl des endgültigen Bildersatzes mit 48 emotionalen Stimuli – 8 pro Emotion – wobei jeder Darsteller jeweils nur mit einem emotionalen Ausdruck in das Bilderset einging (siehe Anhang H). Die Erkennungsraten sind Tabelle 20 zu entnehmen. Ziel war es, Stimulusmaterial mit möglichst hohen Erkennungsraten auszuwählen. Die Erkennungsleistung für die einzelnen Stimuli betrug mindestens 68%. Die durchschnittliche Erkennungsleistung über alle Emotionen hinweg 87,79%.

Tab. 20: Zuordnung zu den einzelnen Emotionskategorien (48 Bilder, in Prozent)

		Angst	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Nicht erkannt**
Stimulus-emotion	**Angst**	**75,38**	0,12	7,68	13,34	1,32	1,06	1,10
	Freude	0,16	**97,95**	0,85	0,08	0,24	0,16	0,56
	Überraschung	6,54	2,09	**90,28**	0,16	0	0,16	0,77
	Ekel	0,54	0,12	1,10	**89,28**	0,47	7,29	1,20
	Trauer	7,81	0,36	1,55	3,71	**82,24**	1,98	2,35
	Ärger	1,15	0,47	1,49	1,81	1,73	**91,66**	1,69

Für die nachfolgenden Berechnungen wurde zunächst mittels Shapiro-Wilks-Test die Normalverteilung überprüft. Dies kann weder für die sechs Basisemotionen als gegeben angesehen werden noch für die durchschnittliche Erkennungsleistung ($p < .000$ für alle durchgeführten Berechnungen). Zur Überprüfung grundlegender Unterschiede in der Erkennungsleistung einzelner Variablen wurde somit der non-parametrische Friedman-Test eingesetzt. Dieser ergab signifikante Unterschiede in der Erkennungsleistung der verschiedenen Emotionen ($\chi^2(5) = 210{,}307$; $p < .000$). Explorativ durchgeführte post-hoc Paarvergleiche ermöglichten eine umfassendere Betrachtung der Erkennungsgenauigkeit. Angst wurde signifikant schlechter erkannt als alle anderen Emotionen ($Z = -7{,}170$; $p < .000$ im Vergleich zu Ärger; $Z = -6{,}533$; $p < .000$ im Vergleich zu Ekel; $Z = -9{,}730$; $p < .000$ im

Vergleich zu Freude; Z = -3,518; p < .000 im Vergleich zu Trauer und Z = -6,416; p < .000 im Vergleich zu Überraschung). Freude wurde signifikant besser erkannt als alle anderen Emotionen (Z = -9,730; p < .000 im Vergleich zu Angst; Z = -5,727; p < .000 im Vergleich zu Ärger; Z = -6,753; p < .000 im Vergleich zu Ekel; Z = -9,017; p < .000 im Vergleich zu Trauer und Z = -5,994; p < .000 im Vergleich zu Überraschung). Auch Trauer unterschied sich signifikant von allen anderen Emotionen in der Erkennungsgenauigkeit (Z = -3,518; p < .000 im Vergleich zu Angst; Z = -5,603; p < .000 im Vergleich zu Ärger; Z = -4,262; p < .000 im Vergleich zu Ekel; Z = -9,017; p < .000 im Vergleich zu Freude und Z = -5,155; p < .000 im Vergleich zu Überraschung). Weitere signifikante Unterschiede in der Erkennungsleistung konnten nicht detektiert werden. Es ergab sich somit folgende Reihenfolge der Erkennungsleistung: Freude > Ärger = Überraschung = Ekel > Trauer > Angst.
Wie in der Einleitung beschrieben wurden für einige Faktoren Einflüsse auf die Erkennungsleistung vermutet. Betrachtet werden sollten: Geschlecht, Alter, Persönlichkeit und Status. Potentielle Einflüsse werden im Folgenden für jede Variable einzeln berichtet.

Ergebnisse der Onlinestudie: Einfluss Geschlecht

Zunächst wurden potentielle Geschlechtereffekte der Beurteiler betrachtet. Hierzu wurden Mann-Whitney U-Tests eingesetzt. Es konnten keine signifikanten Unterschiede in der Erkennungsleistung zwischen Männern und Frauen festgestellt werden (Z = -1,474; p = .140 für Ärger; Z = -1,388; p = .165 für Ekel; Z = -0,217; p = .828 für Freude; Z = -0,687; p = .492 für Trauer; Z = -0,457; p = .648 für Überraschung). Einzige Ausnahme bildete Angst (Z = -2,353; p = .019). Frauen erkannten Angstbilder mit einer Erkennungsrate von 78,43% signifikant besser als Männer, welche Angst nur zu 70,99% richtig erkannten.
Nachfolgend wurde getestet, ob sich Unterschiede in der Beurteilung weiblicher und männlicher Stimuli gefunden werden konnten. Dies war nicht der Fall (Z = -1,292; p = .196).

Ergebnisse der Onlinestudie: Einfluss Alter

Als nächstes wurden potentielle Alterseffekte betrachtet. Um die Gesamtstichprobe in eine jüngere und eine ältere Subpopulation zu unterteilen, wurde der errechnete Median von 35 Jahren verwendet. 35-jährige wurden der älteren Stichprobe zugeordnet. Es befanden sich in beiden Subgruppen 50% der untersuchten Personen. Das mittlere Alter der jüngeren

Stichprobe betrug M = 26,77 Jahre (SA = 3,960), das der älteren Stichprobe M = 50,58 Jahre (SA = 9,126). Der Mann-Whitney U-Test ergab keine signifikanten Unterschiede in der Erkennungsleistung zwischen den beiden Altersgruppen (Z = -0,043; p = .966 für Angst; Z = -0,418; p = .676 für Ärger; Z = -0,477; p = .633 für Ekel; Z = -1,253; p = .210 für Freude; Z = -0,619; p = .536 für Trauer; Z = -1,598; p = .110 für Überraschung). Zusätzlich wurden in der Folge Berechnungen mit den Extremgruppen (untere und obere 25% der Stichprobe) durchgeführt. Diese umfasste N = 90, wobei 47 Personen der Gruppe 27 Jahren und jünger und 43 Personen der 51 Jahre und älter zugeordnet werden konnten. Auch beim Vergleich dieser beiden Subgruppen konnten keine signifikanten Unterschiede in der Erkennungsleistung gefunden werden (Z = -0,986; p = .324 für Angst; Z = -1,461; p = .144 für Ärger; Z = -1,027; p = .307 für Freude; Z = -1,525; p = .127 für Trauer; Z = -1,463; p = .143 für Überraschung). Einzige Ausnahme war die Erkennungsleistung von Ekel (Z = -2,521; p = .012), hier schnitten die jüngeren Probanden mit einer Erkennungsleistung von 89,93% gegenüber einer Erkennungsleistung von lediglich 80,49% bei den älteren Studienteilnehmern deutlich besser ab.

Ergebnisse der Onlinestudie: Einfluss Persönlichkeit

In einem dritten Schritt wurde die Persönlichkeit berücksichtigt. Hierzu wurden die Ergebnisse des TIPI verwendet. Das Vorgehen war dabei Folgendes: Anhand der vorliegenden Normen (Gosling, Rentfrow & Swann, 2003) wurden die Teilnehmer für jede der drei Persönlichkeitseigenschaften Extraversion, Neurotizismus, Offenheit, Verträglichkeit und Gewissenhaftigkeit in eine von drei Gruppen eingeteilt: Unterdurchschnittlich ausgeprägt, durchschnittlich ausgeprägt, überdurchschnittlich ausgeprägt. Eine Übersicht über die Verteilung der Persönlichkeitskonstrukte findet sich in Tabelle 21. Es zeigte sich, dass eine Vielzahl der Probanden im Durchschnittsbereich lag, was für die Repräsentativität der Stichprobe spricht. Etwa 25-30% der Stichprobe war den Extrembereichen zuzuordnen.

Tab. 21: Verteilung der Persönlichkeitskonstrukte in der Normierungsstichprobe

	Unterdurchschnittlich	**Durchschnittlich**	**Überdurchschnittlich**
Extraversion	< 2,99: N = 22	2,99-5,89: N = 108	> 5,89: N = 31
Neurotizismus	< 3,41: N = 20	3,41-6,25: N = 115	> 6,25: N = 26
Offenheit	< 4,31: N = 28	4,31-6,45: N = 102	> 5,45: N = 31
Gewissenhaftigkeit	< 4,08: N = 27	4,08-6,72: N = 115	> 6,72: N = 19
Verträglichkeit	< 4,12: N = 34	4,12-6,34: N = 107	> 6,34: N = 20

Keines der Persönlichkeitskonstrukte führte eine grundsätzliche signifikante Veränderung in der Erkennungsleistung herbei ($p > .05$ für Angst, Ekel, Freude, Trauer und Überraschung bei Extraversion, $p > .05$ für Ärger, Ekel, Freude, Trauer und Überraschung bei Neurotizismus, $p > .05$ für Angst, Ärger, Ekel, Freude, Trauer und Überraschung bei Gewissenhaftigkeit, Verträglichkeit und Offenheit). Ausnahmen waren ein signifikanter Unterschied in der Erkennungsleistung für das Persönlichkeitskonstrukt Extraversion, wenn die Emotion Ärger betrachtet wurde ($\chi^2(2) = 6{,}029$; $p = .049$) und die Angsterkennung bei Neurotizismus ($\chi^2(2) = 7{,}547$; $p = .023$). Mittelwertsvergleiche zwischen den einzelnen Gruppen ergaben: Für die Ärgererkennung unterschied sich die Gruppe mit den durchschnittlich ausgeprägten Extraversions-Werten signifikant von der mit den hoch ausgeprägten Extraversions-Werten ($Z = -2{,}494$; $p = .13$). Während erstgenannte Gruppe Ärger lediglich zu 90,47% erkannte, erkannte die als zweites genannte Gruppe Ärger zu 96,37%. Angst wurde von der Gruppe mit durchschnittlichen Neurotizismus-Werten mit 88,66% deutlich besser erkannt als in der Gruppe mit unterdurchschnittlichen Neurotizismus-Werten ($Z = -2{,}795$; $p = .05$). Hier betrug die Erkennungsleistung lediglich 73,59%.

Ergebnisse der Onlinestudie: Einfluss Status

In einem letzten Schritt wurde der Status als potentieller Einflussfaktor überprüft. Eine Einteilung erfolgte in drei Gruppierungen: Niedrig, mittel, hoch. Hierzu wurde der Winkler-Index berechnet (Winkler & Stolzenberg, 2009; Siegrist, 2005). Dieser berücksichtigt Bildung/Ausbildung, Einkommen und die berufliche Stellung. Anhand vorgegebener Zuordnungsvorschläge können pro Kategorie 1 bis 7 Punkte vergeben werden. Der Index wird dann berechnet, indem die 1 bis 7 Punkte der Teilscores addiert werden. Der Winkler-Index kann somit Werte von 3 bis 21 annehmen. Der Wertebereich von 3 bis 8 definiert die Unterschicht, der Wertebereich von 9 bis 14 die Mittelschicht. Die Oberschicht wird durch den Wertebereich von 15 bis 21 definiert. In der hier vorliegenden Stichprobe waren 24,8% der Unterschicht, 38,5% der Mittelschicht und 36% der Oberschicht zuzuordnen. Für 0,6% konnte aufgrund fehlender Daten keine Zuordnung erfolgen. Es zeigte sich nur ein signifikanter Zusammenhang zwischen Status und Emotionserkennung und war für die Emotion Trauer ($\chi^2(2) = 6{,}658$; $p = .035$): Personen der Mittelschicht erkannten mit 77,57% Trauer signifikant schlechter als Personen aus der Oberschicht, welche Trauer zu 85,53% korrekt erkannt haben ($Z = -2{,}366$; $p = .018$).

Testtheoretische Berechnungen

Zur Überprüfung der testtheoretischen Qualität wurden die Trennschärfe, die Itemschwierigkeit und die interne Konsistenz berechnet. Die Überprüfung der Normalverteilung erfolgte bereits zu Beginn der Berechnungen mittels Shapiro-Wilks-Test, um eine Entscheidung für die zu verwendenden statistischen Verfahren treffen zu können. Es zeigte sich, dass diese weder für die einzelnen Emotionen noch für den Gesamtwert vorlag ($p < .000$ für alle aufgeführten Berechnungen).

Die Berechnung der internen Konsistenz erfolgte durch die Ermittlung von Cronbach's alpha. Die interne Konsistenz ist ein Maß dafür, wie die Items einer Skala miteinander zusammenhängen. Es wird quasi eine mittlere Korrelation berechnet. Die interne Konsistenz wurde für jede Emotion einzeln berechnet. Sie beträgt: $r = .657$ für Angst, $r = .644$ für Ärger, $r = .628$ für Ekel, $r = .711$ für Freude, $r = .513$ für Trauer und $r = .650$ für Überraschung. Hieraus ergibt sich ein Cronbach's alpha von $r = .634$ für die Gesamtskala – bezogen auf alle 48 Stimuli des Bildersets. Als Faustregel sollte ein beliebiges psychometrisches Instrument nur verwendet werden, wenn ein Wert für Cronbach's alpha von .65 oder mehr erreicht wird. Dieser Wert wurde in dieser Studie nur knapp verfehlt.

Die Itemschwierigkeit wird nach der Faustformel "richtige Antworten/alle Antworten" berechnet. Hier greift man somit auf die Erkennungsraten für die einzelnen Items zurück. Itemschwierigkeiten rangieren im Bereich zwischen 0 und 1. Die Interpretation der Werte erfolgt folgendermaßen: Ein Item mit einer Schwierigkeit nahe 0 gilt als sehr schwer, Kennwerte nahe 1 bedeuten, dass es sich um ein eher leichtes beziehungsweise leicht zu beantwortendes Item handelt. Deutlich wird dies, wenn man die erhaltenen Ergebnisse mit 100 multipliziert, um Prozente zu erhalten. Ein Wert von $p = 1$ würde dann 100% entsprechen, das heißt alle Probanden haben dieses Item richtig gelöst. Bei $p = 0$ hätten hingegen 0% der Probanden das Item lösen können. Ein Item, welches zu 100% beantwortet werden kann, stellt keine Herausforderung an den Studienteilnehmer und bringt somit auch keine zusätzlichen Informationen. Angestrebt werden sollten Itemschwierigkeiten zwischen 20% und 80%. In Tabelle 22 ist eine Übersicht über die mittlere Erkennungsleistung als Hinweis auf die Itemschwierigkeit und die Standardabweichung zu finden. Pro Emotion gingen 8 Stimuli in das Bilderset ein.

Tab. 22: Itemschwierigkeit (mittlere Erkennungsleistung in Prozent) und Standardabweichung aller 48 Stimuli

		Angst	**Ärger**	**Ekel**	**Freude**	**Trauer**	**Überraschung**
Stimulus 1	M	78,26	95,03	90,68	96,89	87,58	95,03
	SA	0,41	0,22	0,29	0,17	0,33	0,22
Stimulus 2	M	70,81	80,12	93,17	96,27	82,61	86,34
	SA	0,46	0,40	0,25	0,19	0,38	0,34
Stimulus 3	M	75,16	91,93	87,58	93,79	91,93	87,58
	SA	0,43	0,27	0,33	0,24	0,27	0,33
Stimulus 4	M	72,05	93,17	91,93	98,14	73,29	87,58
	SA	0,45	0,25	0,27	0,14	0,44	0,33
Stimulus 5	M	72,67	88,20	83,23	97,52	84,47	86,96
	SA	0,45	0,32	0,37	0,17	0,36	0,34
Stimulus 6	M	69,57	88,20	80,75	95,03	79,50	90,68
	SA	0,46	0,32	0,39	0,21	0,40	0,29
Stimulus 7	M	68,94	95,03	93,17	96,89	92,55	86,96
	SA	0,46	0,22	0,25	0,17	0,26	0,34
Stimulus 8	M	81,37	84,47	77,02	96,89	52,17	86,34
	SA	0,39	0,36	0,42	0,17	0,50	0,34

Die Trennschärfe eines Items wird definiert als die Korrelation des Items mit dem Gesamtergebnis eines Tests. Die Trennschärfe ermöglicht eine Einschätzung, wie gut ein Item zwischen Personen mit niedriger und hoher Merkmalsausprägung trennt. Hierzu wird der korrelative Zusammenhang eines einzelnen Items mit dem Gesamttest ermittelt. Der Wertebereich der Trennschärfe liegt zwischen r = –1 und r = 1. Eine hohe positive Trennschärfe bedeutet, dass das Item etwas Ähnliches wie der Gesamttest erfasst. Eine Trennschärfe nahe 0 weist darauf hin, dass ein Item mit dem restlichen Test wenig gemeinsam hat. Eine negative Trennschärfe hingegen kann bedeuten, dass ein Item umgekehrt als beabsichtigt verstanden wurde oder es bei der Auswertung umgepolt werden muss. Items mit einer Trennschärfe über r = .30 sind anzustreben. Die errechneten Trennschärfen für alle 48 Items sind Tabelle 23 zu entnehmen.

Tab. 23: Trennschärfen aller 48 Stimuli

	Gesamtwert					
	Angst	**Ärger**	**Ekel**	**Freude**	**Trauer**	**Überraschung**
Stimulus 1	.641	.536	.441	.611	.444	.325
Stimulus 2	.579	.615	.386	.512	.426	.615
Stimulus 3	.615	.442	.577	.608	.393	.556
Stimulus 4	.508	.642	.521	.655	.596	.569
Stimulus 5	.542	.581	.603	.649	.490	.491
Stimulus 6	.401	.551	.559	.460	.574	.562
Stimulus 7	.531	.492	.510	.528	.417	.532
Stimulus 8	.541	.491	.607	.696	.476	.615

Diskussion der Erkennungsleistung

Durch den Einsatz eines online-basierten Verfahrens konnte zur Selektion und Normierung des endgültigen Bildersets eine Stichprobe gewonnen werden, welche verschiedene Altersgruppen, Schichtzugehörigkeiten und Persönlichkeitsausprägungen repräsentierte. Auch wenn diese Stichprobe keinen Anspruch auf Repräsentativität erheben konnte, so hob sie sich dennoch deutlich von einer studentischen Stichprobe in Hinblick auf ihre Zusammensetzung ab. Die Erkennungsleistungen schwankten von 75,38% für Angst bis 96,95% für Freude. Im Vergleich wurden bei der Erhebung mit freiem Antwortformat Erkennungsleistungen von 68,25% für Angst bis 94,27% für Freude erzielt. Bei der Überprüfung der Erkennungsleistung der Stimuli des JACFEE-Bildersatzes im Rahmen des FEEL-Tests (Kessler, Bayerl, Deighton & Traue, 2002) wurden folgende Erkennungsleistungen erzielt: 70% Trauer, 71% Angst, 80% Ekel, 84% Überraschung, 87% Freude und 94% Ärger. Die Spanne der Erkennungsleistung scheint vergleichbar, die Reihenfolge jedoch nicht. Wie die Diskussion in Studie 1 jedoch bereits verdeutlicht hat, ist dies eher die Regel als die Ausnahme, vergleicht man verschiedene Bildersets hinsichtlich der erzielten Erkennungsleistungen.

Diskussion der Bedeutung von Geschlecht, Alter, Persönlichkeit und Status

In der hier durchgeführten Studie wurden potentielle Einflussfaktoren auf die Erkennungsleistung untersucht. Es zeigte sich hinsichtlich des Faktors Geschlecht, dass Frauen Angst signifikant besser erkannten. Die Untersuchung des Faktors Status erbrachte, dass Personen der Oberschicht Trauer signifikant besser erkennen als Personen der Mittelschicht. Auch der Faktor Persönlichkeit erbrachte Gruppenunterschiede: Personen mit hohen Extraversions-Werten scheinen Ärger besser erkennen zu können als Personen mit durchschnittlichen Extraversions-Werten. Angst hingegen wird von Personen mit durchschnittlich ausgeprägtem Neurotizismus signifikant besser erkannt als Personen mit unterdurchschnittlich ausgeprägtem Neurotizismus. Der Faktor Alter indes erbrachte lediglich beim Vergleich der Extremgruppen einen signifikanten Unterschied bei der Erkennung von Ekel. Es ist unklar, warum sich lediglich für diese Emotion ein Unterschied zwischen den Altersgruppen zeigt. In bisherigen Studien zur Untersuchung des Own-Age-Bias (beispielsweise Anastasi & Rhodes, 2006; Ebner & Johnson, 2009) zeigten sich eher

grundsätzliche Unterschiede in der Emotionserkennung zwischen den Altersgruppen. Möglicherweise wurden die Emotionen in dieser Untersuchung insgesamt zu gut erkannt (Deckeneffekte) oder die Unterschiede zwischen den Altersgruppen waren nicht groß genug. Sullivan, Ruffman und Hutton (2007) schlossen in ihrer Studie zur Untersuchung der Emotionserkennungsfähigkeit in Abhängigkeit vom Alter Probanden von 18-32 Jahren und älter 60 Jahren ein.

Bezüglich der Geschlechtsunterschiede bei der Emotionserkennung stellt sich die Frage: Warum sollten Frauen Angst grundsätzlich besser erkennen als Männer? Es kann gemutmaßt werden, dass evolutionäre und sozio-kulturelle Aspekte eine Rolle spielen. Aufgrund ihrer körperlichen Unterlegenheit, so vermuteten beispielsweise Kring und Gordon (1998), mussten Frauen lernen, Emotionen expressiver ausdrücken zu können, um auf sich aufmerksam zu machen. Darüber hinaus mussten Frauen lernen, Gefahr möglichst frühzeitig zu erkennen und zwar durch Beobachtung von Gesichtern. Zeigte ein Mitglied der Gruppe Angst, so bedeutete dies Gefahr. Buck, Miller und Caul (1974) konnten schon zu Beginn der systematischen Emotionsforschung zeigen, dass Frauen stärkere physiologische Reaktion bei der Konfrontation mit Angstreizen zeigen als Männer. Morris, Ohman und Dolan (1998) zum Beispiel argumentieren, dass die Antworten der Umwelt auf bestimmte (gefährliche) Stimuli überlebenswichtig seien und die Angsterkennung deswegen eine besondere Rolle spiele. Sanders, Sjodin & Chastelaine zeigten, dass Frauen in Phasen geringer Östrogenkonzentration „männerähnlichere“ Gehirnaktivitäten als in Zyklusphasen mit hoher Östrogenkonzentration aufwiesen. Damit einhergehend konnten Pearson & Lewis (2005) nachweisen, dass Frauen in Zyklusphasen mit hoher Östrogenkonzentration (also Phasen mit „männeruntypischen“ Gehirnaktivitäten) Angst besonders gut erkannten.

Bezüglich der Zusammenhänge zwischen Persönlichkeit und Emotionserkennung ist die Befundlage weniger eindeutig. Hohe Neurotizismus-Werte werden unter anderem mit emotionaler Labilität, Nervosität, schnellen und starken Stressreaktionen, Unsicherheit und einer Tendenz zu situationsbedingter Traurigkeit in Verbindung gebracht sowie mit zahlreichen körperlichen und psychischen Erkrankungen, beispielsweise Depressionserkrankungen. Der Zusammenhang zur Emotionserkennung und Neurotizismus wurde bislang kaum untersucht. In einer Studie von Chan, Goodwin & Harmer (2007) zeigte sich, dass Personen mit hohen Neurotizismus-Werten besser in der Angsterkennung abschnitten als eine Kontrollgruppe. Unterschiede zwischen Experimental- und

Kontrollgruppe zeigten sich aber unter anderem auch bei der Kategorisierung emotionaler Inhalte oder im EEG-Signal. Die Autoren vermuteten, dass Neurotizismus ein neuronales Korrelat vorweise, welches eine Art Prädisposition für psychische oder körperliche Erkrankungen darstelle. Die Befunde der hier vorgestellten Studie zeigten eine bessere Angsterkennung von Personen mit durchschnittlich ausgeprägtem Neurotizismus. Probanden mit hohen Neurotizismus-Werten unterschieden sich nicht signifikant. Es ist anzunehmen, dass die hier erfasste Population nicht ausreichend extrem in ihren Persönlichkeitsausprägungen ist, um Extremgruppen miteinander zu vergleichen wie es in oben angeführter Studie von Chan, Goodwin & Harmer (2007) der Fall war.

Extraversion zeichnet sich durch eine nach außen gewandte Haltung aus. Extraversion wird mit Eigenschaften wie gesprächig, bestimmt, aktiv, dominant, enthusiastisch oder abenteuerlustig in Verbindung gebracht. Die Rolle von Extraversion bei der Emotionserkennung wurde beispielsweise in einer Studie von Rubin, Munz & Bommer (2005) untersucht. Hierzu wurden die Auswirkungen der Emotionserkennung und der Persönlichkeit von 145 Führungskräften auf ihren Führungsstil herangezogen. Extraversion ist laut der Autoren ein wichtiges Eingangsmerkmal für den beruflichen Erfolg. Darüber hinaus konnten sie einen Zusammenhang zwischen Führung und emotionaler Intelligenz aufzeigen. Für die hier vorgestellte Studie zeigte sich kein Zusammenhang zwischen Schichtzugehörigkeit und Ausprägung der Extraversion. Hohe Extraversionswerte waren in allen Schichten gleichermaßen vertreten. Es fand sich lediglich ein Zusammenhang zwischen Extraversion und Ärgererkennung. Möglicherweise stellt die Ausprägung der Extraversion eine Prädisposition für den beruflichen Erfolg dar. Beruflicher Aufstieg geht mit Networking und dem Knüpfen wichtiger Kontakte einher. Es ist denkbar, dass das frühzeitige Erkennen von Ärger hierbei eine entscheidende Rolle spielt. Soweit bekannt, wurden diese Zusammenhänge allerdings noch nicht in einer Studie systematisch untersucht, um diese Hypothese empirisch untermauern zu können.

Hinsichtlich der Statuszugehörigkeit zeigte sich, dass Personen einer höheren sozialen Schicht signifikant besser in der Trauerkennung abschnitten als Personen der Mittelschicht. Dies widerspricht bisherigen Forschungsergebnissen, nach denen Personen der Unterschicht besonders einfühlend und sensibel beim Erkennen von Emotionen sein sollen (Kraus, Côté & Keltner, 2010). Die Autoren Kraus, Côté und Keltner argumentieren, dass Personen der Unterschicht mehr auf ihr soziales Umfeld angewiesen seien und deswegen Emotionen

sensibler wahrnehmen müssten. Möglicherweise bildet Trauer hier eine Ausnahme. Trösten kann nur der, der es sich leisten kann, nicht auf das eigene Wohlergehen zu achten, da eine Information transportiert wird, die im Gegensatz zu Emotionen wie Ärger oder Angst, keine Relevanz für das eigene Überleben aufweist. Doch auch hier müssen die Ergebnisse weiterer systematische Studien abgewartet werden, um dieses Ergebnis sinnvoll einordnen zu können.

Grundsätzlich lässt sich anführen, dass aufgrund der hohen Erkennungsleistungen Deckeneffekte zu vermuten sind. Diese erschwerten es, statistisch bedeutsame Unterschiede zwischen verschiedenen Subgruppen zu finden. Die hier gefundenen Ergebnisse sollten also dazu ermutigen, die hier berücksichtigten Variablen systematisch zu untersuchen – nach Möglichkeit in einem experimentellen Setting ohne Deckeneffekte bei der Zuordnung der emotionalen Stimuli zur dargestellten Emotion.

Erörterung der testtheoretischen Berechnungen

Die Werte für Cronbach's alpha variierten für die einzelnen Emotionen von r = .513 für Trauer bis r = .711 für Freude. Die mittlere interne Konsistenz lag nur knapp unter dem erwünschten Wert von r = .65. Zieht die interne Konsistenz als Maß für die Verlässlichkeit und Genauigkeit eines Messverfahrens heran, lässt sich sagen, dass diese hier ausreichend hoch war. Lediglich das Cronbach's alpha für Trauer fiel deutlich ab, vermutlich zurückzuführend auf Stimulus 8, welcher lediglich eine Erkennungsrate von 52% aufwies. Würde dieses emotionale Bild bei der Ermittlung von Cronbach's Alpha nicht berücksichtigt, so stiege die interne Konsistenz bereits auf r = .532. Eine Einschränkung der Varianz durch das Erstellen von emotionalen Stimuli mit noch homogeneren Erkennungswerten (bei Trauer variierte die Erkennungsleistung in dieser Studie ohne Stimulus 8 noch immer von 73% bis 93%), würde Cronbach's alpha weiter verbessern.

Die hier errechneten Trennschärfen erscheinen allesamt zufriedenstellend. Durch die relativ hohen Korrelationen der einzelnen Items mit dem Gesamtwert kann davon ausgegangen werden, dass „die Items das erfassen, was der Test intendiert". Sinnvoll wäre es nun in der Folge, die einzelnen Items mit einem Gesamttestwert zu korrelieren, der eine ähnliche aber nicht dieselbe Fragestellung erfasst. Würden hier nur niedrige Korrelationen entstehen, so könnte die Trennschärfe untermauert werden, da sich die Korrelation nur auf die zu

erfassende Skale beschränken würde. Sinnvoller erscheinen Vergleiche mit Verfahren zur Erfassung von Emotionsregulationsprozessen oder der Emotionalen Intelligenz.

Die errechnete Itemschwierigkeit ist als kritisch zu bewerten. Sie lag für alle Emotionen in einem Bereich nahe 1. Das ist bei dem bisherigen Vorgehen jedoch kaum verwunderlich. Ziel dieser Arbeit war es in Hinblick auf verschiedenste Faktoren ein möglichst einheitliches und standardisiertes Bilderset zu erstellen. Die bisherigen Studien zur Selektion waren allesamt darauf ausgelegt, die Stimuli auszuwählen, die am besten erkannt beziehungsweise am höchsten beurteilt wurden. Ein solches Vorgehen nimmt in Kauf, dass die Itemschwierigkeit abnimmt. Würde man willkürlich dem Bilderset (hier gleichzusetzen mit einem Test) Stimuli hinzufügen (gleichzusetzen mit Items), so wüsste man in unserem besonderen Fall nicht, ob der Emotionsausdruck nicht adäquat dargestellt wurde, die Intensität zu gering war oder eine andere Variable die Erkennungsleistung beeinflusst hat. Genau dies gilt es aber zu vermeiden. Deswegen muss zunächst eine möglichst hoch standardisierte Basis geschaffen werden, wo alle Stimuli möglichst gleich gut erkannt werden. Ist dies gewährleistet, kann technisch die Itemschwierigkeit manipuliert werden, beispielsweise durch Veränderung der Intensität mittels Morphing. Es ist definitiv, dass die Itemschwierigkeit dann eine höhere Variation aufweisen würde. Damit einhergehend würde dann auch eine Normalverteilung der Daten vorzufinden sein, welche im vorliegenden Fall (natürlicherweise) nicht gegeben ist. Die Normalverteilung kann in diesem Fall also durch die Manipulation der Items an sich hergestellt werden, nicht wie sonst üblich durch den Austausch von Items.

4.4.1 Überprüfung von Intensität, Natürlichkeit und Attraktivität des Stimulusmaterials am endgültigen Bilderset

Die getroffene Auswahl von 48 Bildern sollte eine Verbesserung hinsichtlich der Homogenität der Variablen Natürlichkeit und Intensität des Gesichtsausdrucks sowie Attraktivität der fotografierten Person erbringen. Aus diesem Grund wurden die in 4.3 ausgeführten Berechnungen nun noch einmal mit dem reduzierten Bildersatz vorgenommen. Zusätzlich wurde der statistische Zusammenhang zur Erkennungsleistung überprüft.

Zunächst wurden die Mittelwerte und Standardabweichungen für die drei Skalen auf Einzelemotionsebene und als Gesamtwert berechnet. Diese Werte sind Tabelle 24 zu entnehmen.

Tab. 24: Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit

Emotion	**Intensität (11-stufig)**		**Attraktivität (7-stufig)**		**Natürlichkeit (7-stufig)**	
	M	**SA**	**M**	**SA**	**M**	**SA**
Angst	6,01	1,14	3,98	0,65	4,29	0,83
Ärger	6,48	1,29	3,93	0,74	3,53	0,92
Ekel	7,63	1,09	3,14	0,72	4,25	1,02
Freude	7,09	1,24	4,80	0,61	5,66	0,53
Trauer	5,97	1,32	4,12	0,66	4,70	0,70
Überraschung	7,06	1,32	3,87	0,66	3,72	0,84
Gesamt	6,82	0,99	3,97	0,54	4,36	0,57

Bei der Überprüfung der Normalverteilung mittels Shapiro-Wilks-Test musste festgestellt werden, dass diese nicht mehr bei allen Variablen vorlag. Dennoch wurden weiterhin parametrische Verfahren eingesetzt. Dies hatte zwei Gründe: Erstens maximiert dies die Vergleichbarkeit der Ergebnisse zwischen dem doppelten und dem einfachen Stimulusset und zweitens sind die Verfahren zum Vergleich von Mittelwerten (t-Test, ANOVA) relativ unempfindlich gegenüber der Verletzung der Normalverteilungsannahme. Dennoch sollte dieses Vorgehen nur in besonderen Fällen angewendet werden.

Im Vergleich zu den Ergebnissen errechnet mit den 96 Bildern (siehe 4.3) ergaben sich für folgende Größen signifikante Veränderungen: bei der Intensität wurde eine signifikante Veränderung für die Emotionen Angst ($t(134) = -3{,}908$; $p < .000$) und Ekel ($t(134) = 2{,}576$; $p = .011$) festgestellt. Für die Variablen Natürlichkeit und Attraktivität ergaben sich keine signifikanten Veränderungen. Eine grafische Veranschaulichung der Einschätzung von

Natürlichkeit, Intensität und Attraktivität vor und nach der finalen Stimulusselektion findet sich in Abbildung 15.

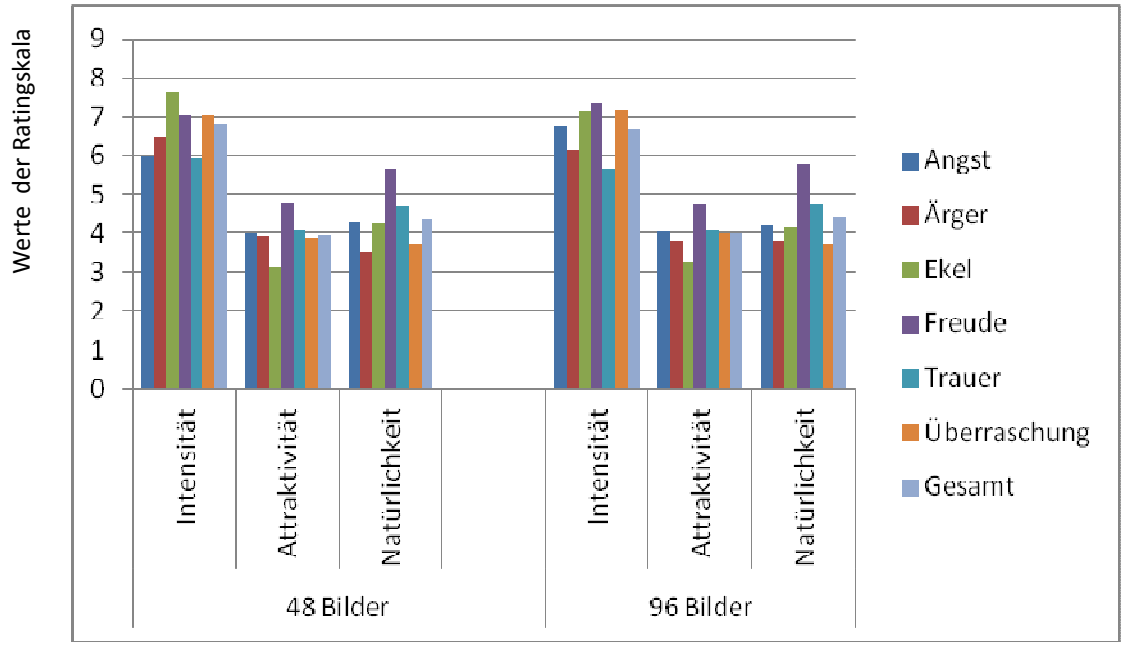


Abb. 15: Beurteilung von Intensität, Attraktivität und Natürlichkeit im Vergleich

Berechnungen mittels ANOVA ergaben signifikante Unterschiede zwischen den verschiedenen Emotionen für die Intensität ($F(1, 5) = 16{,}609$; $p < .000$), die Attraktivität ($F(5) = 42{,}841$; $p < .000$) und die Natürlichkeit ($F(1, 5) = 59{,}108$; $p < .000$). In der Folge wurden zur tiefergehenden Datenanalyse Paarvergleiche mittels post-hoc-Tests berechnet. Eine detaillierte Übersicht findet sich in Tabelle 25. Dieser ist zu entnehmen, dass tendenziell weniger Paarvergleiche signifikant wurden als in Studie 3 mit 96 emotionalen Stimuli.

Tab. 25: Ergebnisse der Paarvergleiche für Intensität, Attraktivität und Natürlichkeit

	Ärger			**Ekel**			**Freude**			**Trauer**			**Überraschung**		
	In	**Att**	**Nat**	**In**	**Att**	**Nat**	**In**	**Att**	**Nat**	**In**	**Att**	**Nat**	**In**	**Att**	**Nat**
Angst	.430	.999	.000	.082	.000	.999	.148	.000	.000	1.00	.917	.125	.423	.969	.007
Ärger				.000	.000	.102	.145	.000	.000	.327	.748	.000	.188	.998	.884
Ekel							.273	.000	.000	.000	.000	.072	.210	.000	.014
Freude										.000	.000	.000	1.00	.000	.000
Trauer													.000	.457	.000

In = Intensität, Att = Attraktivität, Nat = Natürlichkeit
Signifikanz wird zweiseitig angegeben

Korrelationsanalysen ergaben einige signifikante Korrelationen den drei Faktoren ($p < .05$) und alle sechs Basisemotionen im mittleren bis hohen Bereich (siehe Tabellen 26 bis 28).

Tab. 26: Korrelationen der Faktoren Intensität und Attraktivität auf Emotionsebene

Attraktivität x Intensität	
Angst	.525**
Ärger	.276*
Ekel	.201
Freude	.491**
Trauer	.416**
Überraschung	.559**

* = p < .05; ** = p < .001

Tab. 27: Korrelationen der Faktoren Intensität und Natürlichkeit auf Emotionsebene

Natürlichkeit x Intensität	
Angst	.295*
Ärger	.213
Ekel	.367**
Freude	.309*
Trauer	.276*
Überraschung	.051

* = p < .05; ** = p < .001

Tab. 28: Korrelationen der Faktoren Natürlichkeit und Attraktivität auf Emotionsebene

Attraktivität x Natürlichkeit	
Angst	**.363****
Ärger	**.412****
Ekel	**.335****
Freude	**.324****
Trauer	**.403****
Überraschung	**.359****

* = p < .05; ** = p < .001

In einem zweiten Schritt wurden zur Überprüfung potentieller Geschlechtereffekte jeweils die Bilder mit nur männlichen beziehungsweise nur weiblichen Darstellern untersucht. Eine Übersicht über die gefundenen Mittelwerte sind Tabelle 29 zu entnehmen.

Tab. 29: Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit getrennt nach dem Geschlecht der Darsteller

Mittelwert	**Intensität (11-stufig)**		**Attraktivität (7-stufig)**		**Natürlichkeit (7-stufig)**	
	Männliche Darsteller	**Weibliche Darsteller**	**Männliche Darsteller**	**Weibliche Darsteller**	**Männliche Darsteller**	**Weibliche Darsteller**
Angst	6,52	7,02	3,72	4,24	4,08	4,49
Ärger	6,32	6,63	3,78	4,07	3,50	3,56
Ekel	7,33	7,92	3,14	3,13	4,16	4,34
Freude	7,19	6,97	4,55	5,06	5,58	5,06
Trauer	5,99	5,93	3,43	4,80	4,61	5,01
Überraschung	7,11	7,00	3,88	3,86	3,44	3,99
Gesamt	6,73	6,91	3,75	4,19	4,05	4,49

In der Folge wurden potentielle Unterschiede zwischen den Gesamtmittelwerten und den Mittelwerten für männliche und weibliche Darsteller für die drei Skalen untersucht. Hierzu wurde der t-Test eingesetzt. Für die Intensität zeigte sich lediglich ein signifikanter Unterschied zwischen männlichen und weiblichen Darstellern und zwar für die Emotion Angst ($t(134) = -2{,}165$; $p = .032$) und Ekel ($t(134) = -2{,}723$; $p = .007$). In beiden Fällen wurde die Intensität für weibliche Darsteller signifikant höher eingeschätzt.

Für die Skala Attraktivität konnten signifikante Mittelwertsunterschiede für alle Emotionen außer Ekel und Überraschung gefunden werden (Angst: $t(134) = -2{,}165$; $p = .043$; Ärger: $t(134) = -2{,}269$; $p = .025$; Ekel: ($t(134) = -1{,}385$; $p = .168$; Freude: $t(134) = -2{,}273$; $p = .025$; Trauer: $t(134) = -8{,}046$; $p < .000$; Überraschung: $t(134) = -4{,}351$; $p < .000$). Die Attraktivität wurde jeweils höher für die weiblichen Darsteller eingeschätzt.

Unterschiede bezüglich männlichen und weiblichen Darstellern hinsichtlich der Einschätzung der Natürlichkeit ergaben sich für die Emotionen Angst ($t(134) = -2{,}568$; $p = .024$), Freude ($t(134) = -2{,}334$; $p = .021$), Trauer ($t(134) = -2{,}421$; $p = .009$) und Überraschung ($t(134) = -2{,}831$; $p = .005$). Die Natürlichkeit wurde jeweils für die weiblichen Darsteller höher bewertet.

Über alle Emotionen hinweg berechnet, zeigte sich für die Intensität kein signifikanter Unterschied zwischen männlichen und weiblichen Darstellern ($t(134) = -1{,}028$; $p = .306$), ebenso wie für die Natürlichkeit ($t(134) = -1{,}747$; $p = .083$), jedoch nicht für die Attraktivität ($t(134) = -4{,}372$; $p < .000$).

In einem dritten Schritt wurden die Mittelwerte für die weiblichen und männlichen Beurteiler errechnet. Diese sind in Tabelle 30 zu finden.

Tab. 30: Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit getrennt nach dem Geschlecht der Beurteiler

Mittelwert	**Intensität (11-stufig)**		**Attraktivität (7-stufig)**		**Natürlichkeit (7-stufig)**	
	Männliche Beurteiler	**Weibliche Beurteiler**	**Männliche Beurteiler**	**Weibliche Beurteiler**	**Männliche Beurteiler**	**Weibliche Beurteiler**
Angst	5,87	6,14	3,92	4,03	4,19	4,38
Ärger	6,45	6,50	3,87	3,98	3,51	3,55
Ekel	7,58	7,67	3,13	3,14	4,13	4,36
Freude	7,01	7,17	4,77	4,83	5,65	5,67
Trauer	5,89	6,03	4,04	4,19	4,64	4,76
Überraschung	6,97	7,14	3,80	3,93	3,58	3,84
Gesamt	6,75	6,88	3,92	4,02	4,28	4,43

Zur Analyse möglicher Unterschieder wurden t-Tests eingesetzt. Für den Faktor Intensität konnte kein signifikanter Einfluss für das Geschlecht der Beurteiler gefunden werden (für alle Emotionen $p > .05$). Der Faktor Natürlichkeit lieferte ein signifikantes Ergebnis: Männliche Beurteiler schätzten die Natürlichkeit der überraschten Gesichtsausdrücke signifikant höher ein als weibliche Beurteiler ($t(66) = -1{,}263$; $p = .014$). Für Attraktivität zeigte sich: Die Attraktivität von Personen, die Angst ausdrückten wurde von männlichen Beurteilern signifikant höher eingeschätzt als von wieblichen Beurteilern ($t(66) = -0{,}662$; $p = .012$). Aufgrund der geringen Anzahl an signifikanten Unterschieden kann geschlussfolgert werden, dass sich in dieser Studie aus dem Geschlecht der Beurteiler kein systematischer Einfluss ergab.

Als ein Maß für die Messung der Konsistenz der drei Skalen wurde die Abweichung vom Mittelwert herangezogen. Hierzu wurde die Standardabweichung herangezogen. Dies ist ein gängiges Verfahren beispielsweise im Bereich der Leistungsdiagnostik, wobei Werte, welche mehr als das Zweifache der Standardabweichung ausgeben als signifikante Differenz vom Durschnitt definiert werden. Werte unterhalb dieser Grenze sprechen für Einordnung in den Durchschnittsbereich. Wie Tabelle 31 zu entnehmen ist, schwankten die Werte von 1,1154-mal der Standardabweichung als geringste Abweichung (Attraktivitätseinschätzung bei Überraschung) bis 1,9015-mal der Standardabweichung als höchste Abweichung (Einschätzung der Natürlichkeit bei Angst).

Tab. 31: Konsistenzmaß für die drei Skalen

Emotion	Intensität	Attraktivität	Natürlichkeit
Angst	1,35495 x SA	1,5806 x SA	1,90155 x SA
Ärger	1,4962 x SA	1,5173 x SA	1,6165 x SA
Ekel	1,24355 x SA	1,4 x SA	1,253 x SA
Freude	1,3131 x SA	1,3702 x SA	1,65895 x SA
Trauer	1,42445 x SA	1,4641 x SA	1,11755 x SA
Überraschung	1,57135 x SA	1,1154 x SA	1,48995 x SA

Als Referenzwert für die Varianz wird der Faktor bestimmt, mit dem die Standardabweichung multipliziert werden muss, SA = Standardabweichung

Ziel der erneuten Berechnungen war, die Auswirkungen der Reduktion des Bildersatzes auf nunmehr 48 Stimuli zu überprüfen. Es zeigte sich, dass sich die Homogenität nur geringfügig verbessert hat. Betrachtete man allerdings die Übereinstimmung für jede Emotionskategorie getrennt, so konnten die gefundenen Ergebnisse als zufriedenstellend angesehen werden. Alle Abweichungen lagen innerhalb von zwei Standardabweichungen um den Mittelwert. Ein Effekt durch die Betrachter wurde ausgeschlossen. Hier zeigten sich kaum signifikante Unterschiede. Unterschiede zwischen den Darstellern blieben bestehen. Dieser Faktor kann bei zukünftigen Auswertungen berücksichtigt werden. Es kann beispielsweise von Interesse sein, bestimmten Stichproben nur männliche oder nur weibliche Stimuli vorzugeben. So gibt es Studien, die zeigen, dass Einfühlen beim gleichen Geschlecht einfacher geht und sich dies beispielsweise auf die Zeitwahrnehmung auswirken kann (Droit-Volet & Gil, 2009

4.4.2 Bestimmung der Kriteriumsvalidität am endgültigen Bilderset

Die Kriteriumsvalidität wurde zunächst am doppelten Bilderset mit 96 Stimuli überprüft (siehe Studie 2). Die Ergebnisse konnten als gut bis sehr gut eingeordnet werden. Allerdings gilt es diese Aussage für das finale Bilderset zu überprüfen.

Zunächst wurden die Erkennungsraten bestimmt. Hier ergaben sich für das reduzierte Bilderset folgende Prozentwerte: 71,39% für Angst, 86,15% für Ärger, 90,67% für Ekel, 92,15% für Freude, 85,79% für Trauer und 88,06% für Überraschung. Eine nachfolgend berechnete ANOVA aufgrund erfüllte Normalverteilungsvoraussetzungen ($p < .000$ für Ärger, Ekel, Freude, Trauer, Überraschung; $p = .003$ für Angst) ergab signifikante Unterschiede in der Erkennungsleistung für die verschiedenen Emotionen ($F(5, 192) = 4,954$; $p < .000$). Anschließend durchgeführte post-hoc-Tests zeigten, dass sich die Erkennungsleistung von Angst und Ekel ($p = .007$), Freude ($p = .002$) und Überraschung ($p = .034$) signifikant

unterschieden. Für alle anderen Paarvergleiche ergaben sich keine signifikanten Unterschiede ($p > .05$).

Darauffolgend wurde überprüft, ob ein signifikanter Unterschied in der Erkennungsleistung bei Betrachtung des Geschlecht der fotografierten Person zu beobachten. Dies war nicht der Fall ($t(64) = -0{,}124$; $p = 2.32$).

In einem nächsten Schritt wurde die Erkennungsleistung der 96 Stimuli mit der der 48 Stimuli verglichen. Hier zeigten sich keine signifikanten Unterschiede (Angst: $t(64) = -0{,}247$; $p = .806$; Ärger: $t(64) = -0{,}392$; $p = .696$; Ekel: $t(64) = -0{,}849$; $p = .399$; Freude: $t(64) = 0{,}462$; $p = .646$; Trauer: $t(64) = -0{,}980$; $p = .331$; Überraschung: $t(64) = -0{,}460$; $p = .647$). Die Reduktion des Bildersatz hat hier also weder zu Verbesserungen noch zu Verschlechterungen in der Erkennungsleistung beigetragen.

Auch der Vergleich der Ergebnisse aus der freien Beurteilung anhand der endgültigen 48 Bilder mit den Ergebnissen aus der Normierungsstichprobe erbrachte keine signifikanten Unterschiede (Angst: $t(192) = -1{,}347$; $p = .180$; Ärger: $t(192) = -1{,}391$; $p = .166$; Ekel: $t(192) = -1{,}382$; $p = .169$; Freude: $t(192) = -1{,}405$; $p = .162$; Trauer: $t(192) = -1{,}322$; $p = .188$; Überraschung: $t(192) = -1{,}396$; $p = .164$). Die Ergebnisse aus diesen beiden Studien sind somit vergleichbar.

4.4.3 Zusammenfassende Diskussion der Ergebnisse aus der Normierungsstudie

Betrachtet man die in 4.4 aufgeführten Berechnungen noch einmal zusammenfassend, so verdienen einige Aspekte besondere Beachtung:

Wie bereits weiter vorne beschrieben, war nicht mit einer Normalverteilung bezüglich der Erkennungsleistung zu rechnen. Ebenso war der Verlust der Normalverteilung bei der Reduktion von 96 auf 48 Stimuli auszugehen, betrachtet man die Einschätzungen von Intensität, Natürlichkeit und Attraktivität. Auch hier war die Zielvorgabe, das Bilderset hinsichtlich dieser Parameter einheitlicher zu machen. Eine hohe Homogenität verringert die Varianz, Trennschärfe und Normalverteilung leiden darunter. Dennoch wurde dieses Vorgehen bewusst gewählt. Sind die Stimuli alle maximal ähnlich – beispielsweise in ihrer Erkennungsrate – können diese mittels Morphing manipuliert werden, um eine Varianz zu erzeugen und die Trennschärfe zu erhöhen. Es kann dabei ausgeschlossen werden, dass die

Unterschiede in der Erkennungsleistung auf nicht zu kontrollierenden Variablen (beispielsweise uneindeutiger Gesichtsausdruck) beruhen.

Betrachtet man beispielsweise die Beurteilung der Intensität des emotionalen Ausdrucks, so zeigt sich, dass trotz der Vorgabe, jede Emotion mit maximaler Intensität darzustellen, Unterschiede in der Beurteilung zwischen den verschiedenen Emotionen auftraten. Spannend wäre es nun, wenn man in einer Folgestudie die Intensität technisch so manipulieren würde, dass die Intensität über alle Emotionen hinweg gleich eingeschätzt werden würde und man dann erneut die Erkennungsleistung erfassen würde. Möglicherweise würde sich diese dann noch weiter annähern.

Insgesamt sind die Erkennungsleistungen, die Trennschärfe und die interne Konsistenz mindestens zufriedenstellend, weitere relevante Parameter wie die Itemschwierigkeit oder die Umsetzung einer Normalverteilung können im Nachhinein bearbeitet werden. Es konnte eine hohe Übereinstimmung der Erkennungsraten aus der Normierungsstichprobe mit der Beurteilung mit offenem Antwortformat gefunden werden. Das spricht für eine hohe Kriteriumsvalidität. Ein Effekt für das Geschlecht der fotografierten Person konnte nicht für die Erkennungsleistung gefunden werden, jedoch bei der Einschätzung der Intensität, Natürlichkeit und Attraktivität das Stimulusmaterials scheint das Geschlecht der fotografierten Person eine Rolle zu spielen.

4.5 Studie 5: Vergleich der Erkennungsleistung über zwei Zeitpunkte (Retest-Reliabilität)

Zahlreiche psychische und neurologische Störungen gehen mit Defiziten in der Emotionserkennungsfähigkeit (und Expression) einher (beispielsweise Blair, 2000; Braun, Traue, Frisch, Deighton & Kessler, 2005). Ist es das Ziel, die Emotionserkennungsfähigkeit als Verlaufsdiagnostik im Sinne einer Symptomüberprüfung einzusetzen oder diese sogar zum Bestandteil der Therapie zu machen, so ist es notwendig, dieses Maß zu mindestens zwei Zeitpunkten zu erheben. Gewöhnungseffekte wären bei einer solchen Messung ungewollte Artefakte. Um zu überprüfen, inwieweit die „Pictures of Facial Affect-Ulm“ für ein solches Vorhaben einsetzbar sind, wurde angestrebt, die Retest-Reliabilität innerhalb eines kurzen Zeitrahmes von 1-2 Wochen zu erheben. Sollten sich hier keine oder nur geringe Gewöhnungseffekte abbilden lassen, so sollte dies auch innerhalb einer therapeutischen Behandlung beziehungsweise einem rehabilitativen Setting der Fall sein. Angestrebt wurden, in Anlehnung an gängige Maße bei der Konstruktion von psychologischen Testverfahren, Zusammenhänge von über 70%. Würde der Test diese Hürde nehmen können, so würden Veränderungen in der Emotionserkennungsfähigkeit auf einen therapeutischen Effekt schließen lassen.

Hypothesen Studie 5

Vermutet wird, dass die Emotionserkennung aus Gesichtern eine Basisfertigkeit darstellt. Diese ist zwar nicht unabhängig von Lerneffekten, es wird aber davon ausgegangen, dass solche Lerneffekte schon nach einigen Stimuli eine Sättigung erreichen und dann nur noch wenig Einfluss auf die Erkennungsleistung haben. Somit wird die Hypothese aufgestellt, dass die Befragung zu zwei Testzeitpunkten keine signifikanten Verbesserungen in der Erkennungsleistung erbringt.

Erfassung der Emotionserkennungsfähigkeit zu zwei Zeitpunkten

Die Überprüfung der Retest-Reliabilität erfolgte aufgrund der positiven Erfahrungen aus der vorherigen Studie erneut online-basiert (detaillierte Informationen sind Studie 4 zu entnehmen). Die untersuchte Stichprobe setzte sich aus Psychologie-Studenten des ersten und dritten Semesters zusammen. Alle Probanden wurden aufgefordert, ihre freiwillige Teilnahme dadurch kenntlich zu machen, dass nach Durchlesen der Einwilligungserklärung

der „Weiter“-Button angeklickt werden musste (Ethikvotum 245/08_UBB/se). An sie wurden die gleichen Bedingungen zur Konstanthaltung der Umgebungsfaktoren gestellt (siehe 4.4). Die Probanden wurden zusätzlich aufgefordert, beide Testungen unter gleichen Bedingungen durchzuführen. Die Probanden wurden mit einer Versuchspersonenstunde entlohnt. Der hohe Anteil von Frauen im Studiengang Psychologie spiegelt sich in der Zusammensetzung der Stichprobe wider, nur 10,45% der 67 Teilnehmer waren männlichen Geschlechts. Im ersten Durchgang konnten 63 der 67 Probanden erfolgreich an der Studie teilnehmen, im zweiten Durchgang nahmen noch 61 Probanden teil. Der Drop-Out ist mit 3,14% als sehr gering zu beurteilen. Das mittlere Alter der Stichprobe betrug M = 23,55 Jahre (SA = 4,38).

Der Ablauf sah folgendermaßen aus: Jeder Proband gab zu Beginn der Teilnahme einen sechsstelligen Code ein, der eine Zuordnung beider Erkennungsaufgaben ermöglichte. Dieser setzte sich zusammen aus dem zweiten Buchstaben des Vornamens und des Nachnamens, sowie dem Geburtsjahr (zweistellig). In der Folge beatwortete der Proband die zehn Items des TIPI (kurz für: Ten Item Personality Inventory von Gosling, Rentfrow & Swann, 2003; deutsche Version von Hell & Muck). Darauf folgte die eigentliche Aufgabe. Das finale Bilderset, bestehend aus 48 emotionalen Stimuli, wurde präsentiert, wobei der Ablauf exakt dem in 4.4 beschriebenen Vorgehen entsprach. Einige Tage später wurden die Probanden aufgefordert, erneut dieselben 48 Bilder zu bewerten. Der Abstand zwischen den Messungen im Mittel M = 7,05 Tage (SA = 2,91).

Die Probanden wurden nicht darüber aufgeklärt, ob sie die Bilder schon einmal gesehen hatten oder nicht. Gaben die Probanden ihre Einwilligung, die Anonymität aufzuheben, konnten Sie durch Nennung ihres Codes eine individuelle Auswertung anfordern. Für 59 Probanden (M = 23,93 Jahre; SA = 4,41) konnte der Code, welcher zu Messzeitpunkt 1 angegeben wurde, dem zu Messzeitpunkt 2 angegebenen Code zugeordnet werden. Mit dieser Stichprobe wurden die weitergehenden Analysen durchgeführt. Etwa 5% der vorgegebenen Stimuli wurden aufgrund technischer Probleme nicht angezeigt. Durch die Einführung eines zusätzlichen Auswahlfeldes („nicht angezeigt“) konnten die entsprechenden Grundgesamtheiten berechnet werden.

Vergleich der erhobenen Erkennungsleistungen

Gängiges Vorgehen zur Überprüfung der Retest-Reliabilität ist die Berechnung einer Korrelation nach Pearson. Dieses Vorgehen wurde deswegen zunächst auch hier angewandt. Korreliert wurden die prozentualen Erkennungsraten (siehe Tabelle 32). Dabei fiel auf, dass die Aussagekraft dieser Berechnung im vorliegenden Fall nicht ausreichend ist. Aufgrund der variierenden Grundgesamtheiten durch nicht angezeigte Bilder ergaben sich unterschiedliche Erkennungsraten. Beispielsweise führte eine Erkennungsleistung von sieben Bildern bei einer Grundgesamtheit von acht Bildern zu einer Erkennungsleistung von 87,5%, die von sechs Bildern bei einer Grundgesamtheit von sieben zu einer Erkennungsleistung von 85,7%. Die tatsächliche Übereinstimmung wurde somit unterschätzt. Des Weiteren erkannte SPSS bei Freude, eine Emotion, die größtenteils Erkennungsraten von 100% zu beiden Testzeitpunkten aufwies, die Variable nicht als (ausreichend) stetig an. Die Korrelation betrug lediglich r = -.055, obwohl ein Zusammenhang nahe 1 per Augenmaß abgeschätzt werden konnte. Außerdem berücksichtigt die Bildung von prozentualen Erkennungsraten nicht, ob ein und dasselbe Bild in beiden Durchgängen angezeigt wurde oder aufgrund technischer Probleme nur einmal. Es sollte aber angestrebt werden, nur tatsächliche Paare (Bilder die zu Messzeitpunkt 1 und 2 angezeigt wurden) miteinander zu vergleichen.

Tab. 32: Prozentuale Erkennungsraten für die sechs Emotionen zu zwei Messzeitpunkten

	Messzeitpunkt 1	**Messzeitpunkt 2**
Angst	79,79%	82,47%
Ärger	94,54%	96,34%
Ekel	89,44%	90,79%
Freude	99,57%	98,35%
Trauer	84,31%	86,34%
Überraschung	93,21%	95,81%
Gesamt	90,14%	91,68%

Aus diesen Gründen wurde ein alternatives Übereinstimmungsmaß gebildet. Hierbei wurde folgendermaßen vorgegangen: Für jeden Teilnehmer wurde für jede einzelne Emotion geschaut, welche Bilder zu Messzeitpunkt 1 und 2 angezeigt wurden und ob diese gleich beurteilt wurden. Die Anzahl korrekter Übereinstimmungen wurde pro Emotion aufsummiert und die Anzahl der angezeigten Paare geteilt. Hieraus ergab sich die prozentuale Übereinstimmung zwischen den zwei Messzeitpunkten pro Emotion und Person. Diese konnte über alle Studienteilnehmer aufsummiert werden. Die gefundenen

Überstimmungen pro Emotionen waren: 73,24% für Angst, 91,33% für Ärger, 84,98% für Ekel, 98,13% für Freude, 82,09% für Trauer und 92,01% für Überraschung. Geht man davon aus, dass Korrelationen von r = .7 (übersetzt Zusammenhänge von 70%) gute Retest-Reliabilitäten widerspiegeln, können die hier gewonnenen Ergebnisse als höchst zufriedenstellend angesehen werden (Amelang & Zielinski, 1994).

Zusätzlich wurde berechnet, ob sich signifikante Unterschiede in der Erkennungsleistung zwischen Messzeitpunkt 1 und 2 ergeben, da die prozentualen Zusammenhänge zu Messzeitpunkt 2 tendenziell höher erschienen (siehe Tabelle 32). Hierzu wurde zunächst mittels Shapiro-Wilks-Test die Normalverteilung überprüft. Die Nullhypothese wurde in den berechneten Fällen verworfen ($p < .05$), eine Normalverteilung konnte nicht angenommen werden. Es wurde in der Folge der non-parametrische Wilcoxon-Test für verbundene Stichproben eingesetzt. Dieser zeigte keine signifikanten Unterschiede zwischen den beiden Testzeitpunkten für Angst ($Z = -1,159$; $p = .246$), Ärger ($Z = -1,249$; $p = .212$), Ekel ($Z = -0,300$; $p = .764$), Freude ($Z = -1,869$; $p = .062$), Trauer ($Z = -0,461$; $p = .645$) und Überraschung ($Z = -1,1550$; $p = .121$) sowie die Gesamterkennungsrate ($Z = -1,415$; $p = .157$) an.

Beurteilung der Retest-Reliabilität

Die Reliabilität gibt Hinweise darauf, wie genau ein Testverfahren die intendierte Fragestellung erfasst. Die Retest-Reliabilität vergleicht hierzu die erbrachten Ergebnisse zu zwei verschiedenen Zeitpunkten. Große Übereinstimmungen bedeuten ein ähnliches Antwortmuster über die Zeit hinweg. Daraus lässt sich schlussfolgern, dass andere Faktoren wie Lerneffekte, Kontextvariablen et cetera eine untergeordnete Rolle spielen und das Messinstrument zur wiederholten Messung eingesetzt werden kann. Die hier gewonnen Übereinstimmungsraten von mindestens 70% erfüllen den geforderten Standard. Diese sind vor dem Hintergrund der nur geringen Abstände zwischen den beiden Messzeitpunkten noch einmal beeindruckender. Es ist anzunehmen, dass im Rahmen einer Verlaufsdiagnostik mehr als eine Woche zwischen den beiden Erhebungszeitpunkten liegen würde, so dass der Einfluss potentieller Lerneffekte noch unwahrscheinlicher wird. Die Höhe der Retest-Reliabilität spricht dafür, die „Pictures of Facial Affect-Ulm“ ohne Probleme im therapeutischen oder rehabilitativen Kontext einsetzen zu können, um die Emotionserkennung im Verlauf überprüfen zu können. Erinnerungseffekte konnten in dieser Studie nicht gefunden werden, die Veränderung in den Erkennungsleistungen zwischen den

Zeitpunkten wurde nicht signifikant. In dieser Studie wurden nur zwei Messzeitpunkte realisiert, angelehnt an Prä-Post-Messungen in klinischen Setting. Es ist durchaus denkbar, dass sich Erinnerungseffekte abbilden ließen, wenn die Bilder noch häufiger gezeigt würden. Dies zu untersuchen, wäre eine spannende Fragestellung für eine Folgestudie.

5. EMOTIONDERKENNUNG AUS UNTERSCHIEDLICHEN PERSPEKTIVEN

Ein Alleinstellungsmerkmal dieses Bildersets ist, dass es neben Bildern der Frontalperspektive auch Bilder anderer Perspektiven zur Verfügung stellt. Es wurden simultan Bilder aus fünf Perspektiven angefertigt, von 0° bis 90° in 22,5°-Abständen (siehe Abbildung 16).

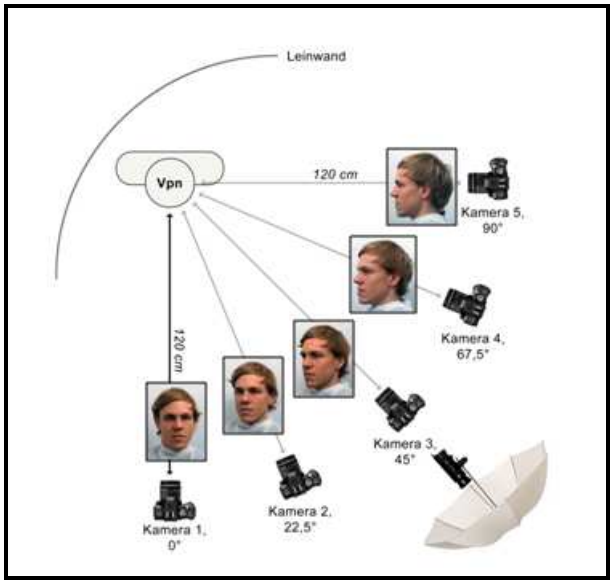


Abb. 16: Übersicht über die fünf fotografierten Perspektiven am Beispiel eines neutralen Gesichtsausdrucks

Diese Erweiterung wurde eingeführt, um eine differenziertere Erfassung der Emotionserkennung zu ermöglichen und den Realitätsbezug statischer Bilder zu maximieren. Im Alltag muss der Mensch in der Lage sein, Emotionen auch dann zu erkennen, wenn das Gegenüber nicht frontal zu einem steht. Vermutet wird, dass diese Kompetenz davon abhängt, wie gut relevante Gesichtsmerkmale noch erkannt werden können, beispielsweise das Zusammenziehen der Augenbrauen oder das Zusammenpressen der Lippen. Die Merkmale, an denen eine Emotion erkannt wird, unterscheiden sich von Emotion zu Emotion, weswegen anzunehmen ist, dass sie die Erkennungsleistung über die verschiedenen Perspektiven hinweg für die sechs Basisemotionen unterschiedlich verhält. Zusätzlich spielen vermutlich gesellschaftliche Konventionen eine Rolle. Während Freude eine gesellschaftlich akzeptierte Emotion darstellt, wird Trauer häufig nicht offen im Alltag

gezeigt. Daraus lässt sich ableiten, dass Trauer, wenn diese auftrifft, eher zu einem Abwenden des Gesichts vom Interaktionspartner führt. Da diese Emotion dennoch hochrelevant ist, ist es denkbar, dass der Mensch hier besondere Fähigkeiten entwickelt hat, diese Emotion auch noch dann zu erkennen, wenn der Interaktionspartner sich abgewendet hat.

Bisher scheint die Emotionserkennung aus verschiedenen Perspektiven kaum untersucht worden zu sein. Studien, welche Perspektivbilder beinhalten, weisen in aller Regel folgendes Paradigma auf: Probanden lernen Gesichter aus unterschiedlichen Blickwinkeln in einem ersten Schritt kennen. In einem zweiten Schritt werden diese Bilder zusammen mit unbekannten Bildern vorlegt. Aufgabe der Probanden war, anzugeben, ob sie das betreffende Bild schon einmal gesehen haben (bekannt) oder nicht (unbekannt). Die Ergebnisse sind uneinheitlich. Während in einigen Studien ein Vorteil für die 45°-Perspektive gefunden wurde (beispielsweise Bruce, Valentine & Baddeley, 1987), konnte dieses Ergebnis in anderen Studien nicht repliziert werden (beispielsweise Liu & Chaudhuri, 2002; Hill, Schyns & Akamatsu, 1997). Gerade neuere Literatur weist darauf hin, dass die in vorherigen Studien gefundenen Effekte möglicherweise zu einem großen Teil auf Defizite in der Operationalisierung der Studien zurückzuführen sind. So wurden Probanden beispielsweise nicht unter verschiedenen Perspektiven trainiert und abgefragt. Auch neurologisch ist eher ein Vorteil für frontale beziehungsweise seitlich dargestellte Gesichter anzunehmen, Spiegelneurone sind bei diesen Perspektiven besonders aktiv (O'Toole, Edelmann & Bülthoff, 1998).

Auch die Blickrichtung ist eine bedeutsame Information in der sozialen Interaktion (Haxby, Hoffman & Gobbini, 2002; Mogg, Garner & Bradley, 2007). Die Blickrichtung gibt ebenso wie die Orientierung des Kopfes in eine bestimmte Richtung Auskunft darüber, wohin die Aufmerksamkeit des Interaktionspartners gerichtet ist. Zeigt ein Interaktionspartner Ärger, sein Blick ist jedoch auf eine andere Person gerichtet, so ist die soziale Relevanz deutlich geringer, als wenn der verärgerte Blick auf der eigenen Person ruht. Beim hier vorgestellten Bildmaterial wurde die Blickrichtung nicht variiert. Die Darsteller schauten frontal in die Kamera (siehe Abbildung 9).

Zusammenfassen ist ein Abfall der Erkennungsleistung mit Drehung des Gesichtes zur seitlichen Perspektive hin anzunehmen. Vermutet wird dies aus zwei Gründen: Einige der relevanten Analysepunkte im Gesicht (Mund, Augen, Augenbrauen, Nase) sind seitlich

weniger deutlich zu erkennen. Da diese Hinweisreize bei den verschiedenen Emotionen eine unterschiedliche Relevanz aufweisen, wird ein Abfall der Erkennungsleistung nicht für alle Emotionen gleichermaßen angenommen. Zusätzlich ist die soziale Relevanz unterschiedlich ausgeprägt. Während bei Trauer anzunehmen ist, dass diese aufgrund kultureller Normen oftmals nur versteckt ausgedrückt wird und eventuell eine höhere Sensitivität bei der Erkennung aus unterschiedlichen Winkeln vorliegen muss, ist Ärger, der auf eine andere Person gerichtet ist, weniger bedeutsam und wird somit möglicherweise in den nicht-frontalen Perspektiven schlechter erkannt. Zusätzlich könnte das Geschlecht des Beurteilers eine wichtige Rolle spielen: Wie bereits in Studie 3 diskutiert, gibt es Hinweise darauf, dass Frauen aufgrund ihrer körperlichen Unterlegenheit gelernt haben, bedrohliche Emotionen schnell und präzise wahrzunehmen. Beispielsweise Angst und Ärger auch aus seitlichen Perspektiven wahrnehmen zu können, stellt einen Wettbewerbsvorteil dar.

Ein Hinweis zu den hier vorgestellten Studien: Betrachtet man Emotion und Perspektive gleichermaßen, so sind pro Kategorie immer nur wenige Stimuli in die Studie eingegangen. Bei Studie 1 wurden beispielsweise lediglich 3 Stimuli pro Emotion und Perspektive bereitgestellt, in Studie 2 5 Stimuli. Die auf diesen geringen Stimuli basierenden Berechnungen müssen somit als vorläufig angesehen werden.

5.1 Studie 6: Variation der Stimuli über Emotion und Perspektive

In dieser Studie wurde die Emotionserkennung über unterschiedliche Perspektiven an einem Paradigma getestet, welches potentielle Effekte der Wiedererkennung ausschließen sollte. Grund hierfür war die Überlegung, dass die Erkennungsleistung möglicherweise davon beeinflusst werden könnte, dass eine Person mit einem emotionalen Ausdruck über verschiedene Perspektiven hinweg gezeigt würde. Angenommen wurde, dass Person und emotionaler Ausdruck wiedererinnert werden würden und dann keine tatsächliche Messung der Erkennungsleistung mehr möglich wäre. Ziel war es zu schauen, wie sich die Emotionserkennungsleistung über die verschiedenen Perspektiven hinweg entwickelt, wobei angenommen wurde, dass eine grundsätzlich abnehmende Tendenz zu beobachten sein müsste, welche aber für die verschiedenen Emotionen unterschiedlich stark ausgeprägt sei.

Hypothesen Studie 6

Hypothese 1: Wie in der Einleitung beschrieben, wird ein grundsätzlicher Abfall er Erkennungsleistung mit zunehmender Annäherung an eine Profilansicht der Stimuli angenommen.
Hypothese 2: Es wird aufgrund der Unterschiede bezüglich der relevanten Gesichtsmerkmale und der sozialen Relevanz angenommen, dass die Veränderungen für die verschiedenen Emotionen unterschiedlich stark ausgeprägt sein werden.

Bewertung der perspektivisch variierenden Stimuli

Die Bewertung der Perspektivbilder wurde von 81 Medizinstudenten (mittleres Alter = 21,55 Jahre; SA = 2,41; 43% Männer, 54% Frauen, 3% ohne Angabe) vorgenommen, welche allesamt freiwillig an der Studie teilnahmen (Ethikvotum 245/08_UBB/se). Der Versuchsablauf gestaltete sich wie folgt: Zunächst wurde ein Probedurchgang durchgeführt. Alle sechs Emotionen und fünf Perspektiven wurden präsentiert, die Probanden erhielten nach ihrer Beurteilung eine Rückmeldung. Danach begann der eigentliche Versuch. Die Studenten sahen nacheinander die 90 emotionalen Bilder aus unterschiedlichen Perspektiven in einer zufälligen Reihenfolge. Jedes Bild wurde in einer Höhe von 800 Pixeln angezeigt, die Breite variierte je nach Perspektive. Die Präsentationsdauer betrug 1 Sekunde. Nach Darbietung des emotionalen Bildes wurden die Probanden gebeten anzugeben, welche

Emotion präsentiert wurde. Hierfür stand ein abgestuftes Auswahlformat zur Verfügung, welches einerseits die sechs Basisemotionen, andererseits die subjektive Sicherheit in vier Abstufungen erfasst (siehe Abbildung 17). Dieses neue Antwortformat wurde entwickelt, um die Antwort der Probanden differenzierter auswerten zu können. Es ist angedacht, dieses Format in Zukunft in den FEEL-Test zu übernehmen. Die Teilnehmer hatten maximal 10 Sekunden Zeit für eine Auswahl. Die Zusammensetzung der Stimuli erfolgte dergestalt: 45 Darsteller (23 Männer und 22 Frauen) gingen jeweils mit zwei emotionalen Ausdrücken in den Itempool ein. Die Perspektive wurde für jeden Stimulus zufällig ausgewählt. Somit standen 90 Stimuli zur Verfügung – je 18 pro Emotion beziehungsweise 15 pro Perspektive. Jeder Darsteller war nur einmal pro Emotionsausdruck vertreten.

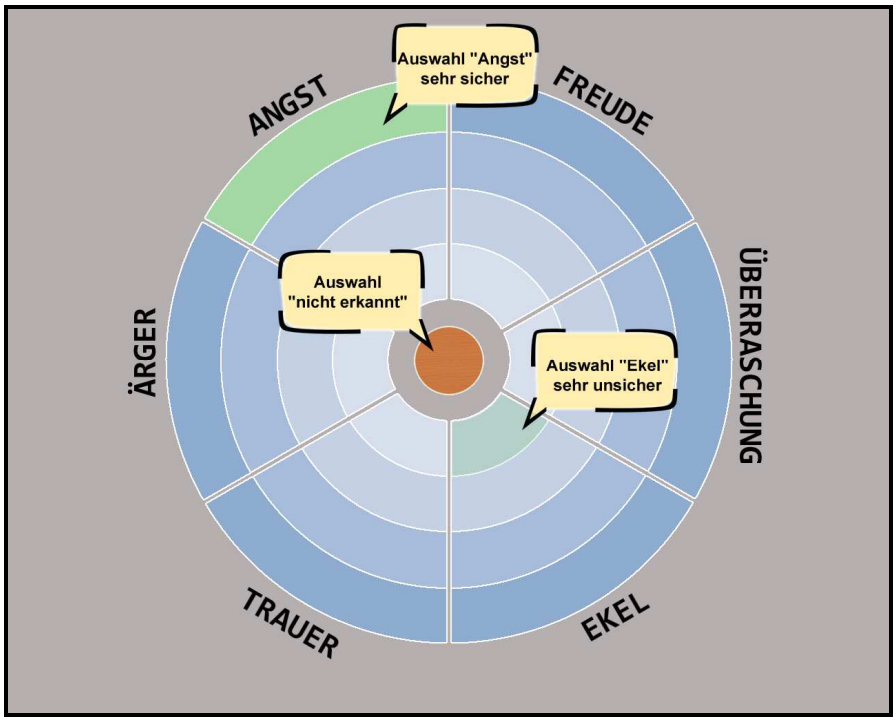


Abb. 17: Veranschaulichung des neuen Antwortformates

Den Probanden wurden verschiedene Fragebögen vorgelegt. Neben den bereits beschriebenen Instrumenten NEO-FFI (Costa & McCrae, 2002; Borkenau & Ostendorf, 2008) und TeiQue-SF (Cooper & Petrides, 2010; Petrides & Furnham, 2006) wurde der „Relationship-Questionnaire“ (Bartholomew & Horowitz, 1991) eingesetzt. Steele, Steele und Croft (2008) konnten in ihrer Studie mit Kindern nachweisen, dass der Bindungsstil die Emotionserkennung im Schulkindalter vorhersagen kann. Es wird von den Autoren angenommen, dass die frühen Erfahrungen mit Bindungspersonen einen maßgeblichen

Einfluss darauf haben, ob ein Kind sehr empfindlich auf emotionale Signale reagieren „muss“ (beispielsweise bei einem unsicheren Bindungsstil, bei der die Mutter wenig verlässlich in sozialen Interaktionen auftritt) oder nicht. Der „Relationship-Questionnaire“ erfasst den Bindungsstil Erwachsener anhand kurzer Beschreibungen der vier Bindungsstile (sicher, unsicher-vermeidend, unsicher-ambivalent, desorganisiert). Die Teilnehmer lesen sich diese durch und können eine Bewertung dazu abgeben, welche der vier Beschreibungen am ehesten auf sie zutrifft. Zusätzlich wird angegeben, wie stark die vier Stile auf einer Skala von 1 (gar nicht zutreffend) bis 7 (sehr stark zutreffend) auf die betreffende Person zutreffen. Hieraus können zwei Skalenwerte errechnet werden: die Ausprägung von „Avoidance“ (Vermeidung) und „Anxiety“ (Ängstlichkeit).

- Vermeidung: (unsicher-ambivalent + sicher) – (unsicher desorganisiert + unsicher-ängstlich)
- Ängstlichkeit: (unsicher-ambivalent + unsicher desorganisiert) – (sicher + unsicher-ängstlich)

Beide Skalen können somit Ausprägungen im negativen und positiven Bereich annehmen.
Der genaue Inhalt der Beschreibungen ist Anhang I zu entnehmen. Explorativ wurden Korrelationen zur Emotionserkennung untersucht.

<u>Auswertung der Erkennungsleistung der perspektivisch variierenden Stimuli</u>

Zunächst wurden die durchschnittlichen Erkennungsleistungen errechnet. Diese sind Tabelle 33 zu entnehmen.

Tab. 33: Erkennungsleistung aufgeschlüsselt nach Emotion und Perspektive Studie 6

	Emotion						
Perspektive	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Angst**	**Gesamt**
1	99,59	91,36	74,9	74,9	76,95	63,79	80,24
2	98,35	89,3	86,42	85,6	93,42	72,43	87,58
3	98,35	92,18	93,42	85,6	88,07	65,43	87,17
4	99,59	87,24	87,24	78,6	73,25	79,42	84,22
5	99,18	92,18	90,95	68,72	92,18	71,6	85,80
Gesamt	99,01	90,45	86,58	78,68	84,77	70,53	-

Nachfolgend wurden potentielle Unterschiede zwischen den verschiedenen Perspektiven berechnet. Im Vorfeld wurde für jede Emotion innerhalb der verschiedenen Perspektiven die Normalverteilung mittels Shapiro-Wilks-Test überprüft. Diese konnte für alle 30 Berechnungen nicht nachgewiesen werden ($p < .05$), weswegen in der Folge non-parametrische Verfahren eingesetzt wurden. Der Friedman-Test ergab signifikante Unterschiede in der Erkennungsleistung zwischen den verschiedenen Perspektiven ($\chi^2(4) = 43{,}299$; $p < .000$). Zur besseren Einordnung dieses Befundes wurden anschließend Paarvergleiche mittels Wilcoxon-Test berechnet. Die Erkennungsleistung für Perspektive 1 unterschied sich signifikant von Perspektive 2 ($Z = -5{,}737$; $p < .000$), Perspektive 3 ($Z = -4{,}922$; $p < .000$), Perspektive 4 ($Z = -3{,}394$; $p = .001$) und Perspektive 5 ($Z = -4{,}268$; $p < .000$). Darüber hinaus unterschied sich Perspektive 2 signifikant von Perspektive 4 ($Z = -2{,}498$; $p = .013$) und Perspektive 5 ($Z = -1{,}886$; $p = .049$). Es ergaben sich keine weitere signifikante Unterschiede der Erkennungsraten zwischen den einzelnen Perspektiven ($p > .05$). Betrachtete man zusätzlich die einzelnen Emotionen, so ergab sich folgendes Bild: Für die Emotion Angst konnten signifikante Unterschiede in der Erkennungsleistung zwischen den einzelnen Perspektiven gefunden werden. Perspektive 1 unterschied sich signifikant von Perspektive 2 ($Z = -2{,}419$; $p = .016$), Perspektive 4 ($Z = -4{,}400$; $p < .000$) und Perspektive 5 ($Z = -1{,}963$; $p = .050$). Darüber hinaus unterschied sich Perspektive 2 in der Erkennungsleistung signifikant von Perspektive 3 ($Z = -1{,}999$; $p = .046$) und Perspektive 4 ($Z = -2{,}075$; $p = .038$). Die Erkennungsleistung fiel außerdem signifikant für Perspektive 4 von Perspektive 3 ($Z = -3{,}673$; $p < .000$) und Perspektive 5 ($Z = -2{,}133$; $p = .033$). Eine ähnliche Ergebnisstruktur ließ sich für Ärger feststellen. Die Erkennungsleistung unter Perspektive 1 unterschied sich signifikant von Perspektive 2 ($Z = -5{,}246$; $p < .000$), Perspektive 3 ($Z = -3{,}635$; $p < .000$) und Perspektive 5 ($Z = -4{,}983$; $p < .000$). Darüber hinaus zeigte sich ein signifikanter Unterschied in der Erkennungsleistung von Perspektive 2 zu Perspektive 3 ($Z = -2{,}343$; $p = .019$) und Perspektive 4 ($Z = -5{,}212$; $p < .000$). Des Weiteren unterschieden sich Perspektive 3 und 4 ($Z = -4{,}053$; $p < .000$) und Perspektive 4 und 5 ($Z = 4{,}940$; $p < .000$) beim Vergleich der Erkennungsleistungen. Die Ekelerkennung unterschied sich signifikant zwischen der ersten und allen anderen Perspektiven (1 zu 2: $Z = -3{,}300$; $p < .000$; 1 zu 3: $Z = -4{,}668$; $p < .000$; 1 zu 4: $Z = -4{,}033$; $p < .000$; 1 zu 5: $Z = -4{,}215$; $p < .000$). Ekel wurde des Weiteren aus Perspektive 2 signifikant unterschiedlich erkannt als aus Perspektive 3 ($Z = -2{,}270$; $p = .023$) und aus Perspektive 4 ($Z = -2{,}040$; $p = .041$). Für Freude ließen keine signifikanten Unterschiede in der

Erkennungsleistung über die verschiedenen Perspektiven hinweg berichten, alle Paarvergleiche lagen über p = .05. Die Trauererkennung fiel signifikant unterschiedlich aus zwischen Perspektive 1 und 2 (Z = -3,457; p = .001), Perspektive 1 und 3 (Z = -3,96; p = .002), Perspektive 2 und 4 (Z = -2,160; p = .031), Perspektive 2 und 5 (Z = -4,055; p < .000), Perspektive 3 und 4 (Z = -2,217; p = .027), Perspektive 3 und 5 (Z = -4,087; p < .000) sowie Perspektive 4 und 5 (Z = -2,427; p = .015). Für Überraschung wiederum ließen keine signifikanten Unterschiede in der Erkennungsleistung über die verschiedenen Perspektiven hinweg berichten, alle Paarvergleiche lagen über p = .05. Die Richtung der Unterschiede sind Tabelle 33 zu entnehmen.

Zur besseren Übersicht sind die Ergebnisse noch einmal zusammenfassend in Tabelle 34 dargestellt.

Tab. 34: Übersicht signifikante Paarvergleiche Studie 6

	Emotion					
Signifikante Paarvergleiche der verschiedenen Perspektiven	**Angst**	**Ärger**	**Ekel**	**Freude**	**Trauer**	**Überraschung**
1 vs. 2	1 < 2*	1 < 2**	1 < 2**	-	1 < 2**	-
1 vs. 3	-	1 < 3**	1 < 3**	-	1 < 3**	-
1 vs. 4	1 < 4**	-	1 < 4**	-	-	-
1 vs. 5	1 < 5*	1 < 5**	1 < 5**	-	-	-
2 vs. 3	2 > 3*	2 > 3*	2 < 3*	-	-	-
2 vs. 4	2 < 4*	2 > 4**	2 < 4*	-	2 > 4*	-
2 vs. 5	-	-	-	-	2 > 5**	-
3 vs. 4	3 < 4*	3 > 4**	-	-	3 > 4*	-
3 vs. 5	3 < 5*	-	-	-	3 > 5*	-
4 vs. 5	-	4 < 5**	-	-	4 > 5*	-

< = kleiner, > = größer; * = p < .05; ** = p < .001

In einem nächsten Schritt wurde mittels Mann-Whitney U-Test überprüft, ob das Geschlecht der Beurteiler einen Einfluss auf die perspektivische Erkennungsleistung hat. Dies ist lediglich für Ärger der Fall. Es lassen sich signifikante Unterschiede in der Erkennungsleistung zwischen Männern und Frauen in der perspektivischen Emotionserkennung finden, vergleicht man die Ärgererkennung für Perspektive 2 (Z = -2,524; p = .012), 3 (Z = -2,426; p = .015) und 4 (Z = -2,330; p = .020) miteinander. In jedem dieser drei Fälle erkannten Frauen Ärger signifikant besser als die Männer.

Bei der Betrachtung potentieller Unterschiede in der Erkennungsleistung der verschiedenen Emotionen wurde die über die Perspektiven gemittelte Gesamterkennungsleistung

herangezogen. Der Friedman-Test zeigte signifikante Unterschiede an ($\chi^2(5) = 139{,}907$; $p < .000$), nachfolgend durchgeführte Wilcoxon-Tests ergaben signifikante Unterschiede zwischen allen sechs Emotionen ($p < .05$), so dass folgende Rangreihe der allgemeinen Erkennungsleistung erstellt wurde: Freude > Überraschung > Ekel > Ärger > Trauer > Angst.
Zusätzlich wurde die subjektive Sicherheit (4 = sehr sicher bis 1 = sehr unsicher) ausgewertet. Eine Übersicht findet sich in Tabelle 35.

Tab. 35: Subjektive Sicherheit aufgeschlüsselt nach Emotion und Perspektive Studie 6

	Emotion						
Perspektive	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Angst**	**Gesamt**
1	3,90	3,67	3,44	3,32	3,30	3,23	3,48
2	3,85	3,64	3,58	3,35	3,57	3,30	3,55
3	3,84	3,61	3,67	3,26	3,44	3,52	3,56
4	3,88	3,41	3,59	3,19	3,06	3,41	3,42
5	3,84	3,67	3,55	2,98	3,42	3,36	3,47
Gesamt	3,86	3,60	3,57	3,22	3,36	3,36	-

Bei der Einschätzung der subjektiven Sicherheit wurde die Bandbreite der Skala nicht ausgenutzt. Da in erster Linie Zusammenhänge zwischen der subjektiven Sicherheit und der Erkennungsleistung interessierten, wurden in der Folge Korrelationen berechnet. Diese sind Tabelle 36 zu entnehmen.

Tab. 36: Korrelation Erkennungsleistung mit subjektiver Sicherheit Studie 6

	Erkennungs-leistung allgemein					
	Freude	**Über-raschung**	**Ekel**	**Trauer**	**Ärger**	**Angst**
Subjektive Sicherheit (über alle Perspektiven)	.449**	.608**	.085	.372**	.497**	247*

* = $p < .05$; ** = $p < .001$

<u>Auswertung der eingesetzten Fragebogenverfahren</u>

Abschließend fand eine Auswertung der Fragebögen statt. Für die hier vorliegende Arbeit war es weniger von Belang, die erhobenen Werte mit bestehenden Normen zu vergleichen. Vielmehr bedeutsam erschienen Zusammenhänge zwischen diesen Maßen und der Emotionserkennungsfähigkeit. Hierzu wurden Korrelationen nach Pearson berechnet. Korreliert wurden die Mittelwerte der Fragebögen mit den mittleren Erkennungsleistungen pro Emotion. Eine Übersicht über die Ergebnisse des TeiQue und NEO-FFI sowie die

Ergebnisse der Dimensionen „Anxiety“ (Ängstlichkeit) und „Avoidance“ (Vermeidung) aus dem „Relationship-Questionnaire“ findet sich in Tabelle 37, eine Auflistung der Bindungsstile in Tabelle 38. Da nicht alle Probanden Angaben gemacht haben, mussten die Analysen teilweise mit geringeren Fallzahlen getätigt werden.

Tab. 37: Mittelwerte, Standardabweichung, Minimum und Maximum von NEO-FFI und TeiQue Studie 6

	Minimum	**Maximum**	**M**	**SA**
NEO-FFI (Skala von 0-4)				
Neurotizismus	0,50	2,92	1,67	0,61
Extraversion	1,42	3,67	2,58	0,56
Offenheit	1,17	3,58	2,75	0,52
Verträglichkeit	1,50	3,58	2,69	0,48
Gewissenhaftigkeit	1,25	4,00	2,65	0,61
TeiQue (Skala von 1-7)				
Gesamt	3,57	6,33	5,23	0,54
Wohlbefinden	3,67	7,00	5,75	0,84
Selbstkontrolle	3,00	6,33	4,64	0,76
Emotionalität	2,88	6,63	5,33	0,69
Soziale Verträglichkeit	2,50	6,50	5,10	0,80
Relationship-Questionnaire				
Anxiety	-6,00	9,00	1,57	2,87
Avoidance	-8,00	10,00	3,06	4,04

Tab. 38: Übersicht Bindungsstile Studie 6

	Häufigkeit	**Prozent**
Sicher	4	5,1
Unsicher-vermeidend	10	12,8
Unsicher-ambivalent	36	46,2
Unsicher-desorganisiert	7	9,0

Die Auswertung ergab keine signifikanten korrelativen Zusammenhänge. Die Überprüfung mittels Kruskal-Wallis-Test, ob der Bindungsstil (siehe Tabelle 38) die Erkennungsleistung emotionaler Bilder beeinflusst, blieb ebenfalls ohne signifikantes Ergebnis (Angst: $\chi^2(3) = 0{,}601$; $p = .968$; Ärger: $\chi^2(3) = 2{,}151$; $p = .542$; Ekel: $\chi^2(3) = 4{,}452$; $p = .217$; Freude: $\chi^2(3) = 3{,}167$; $p = .367$; Trauer: $\chi^2(3) = 5{,}105$; $p = .164$; Überraschung: $\chi^2(3) = 1{,}200$; $p = .753$).

Diskussion der perspektivischen Erkennungsleistung

Diese erste Studie zur Überprüfung der Emotionserkennungsfähigkeit mit variierender Perspektive erbrachte einige interessante Ergebnisse. Es zeigten sich signifikante Unterschiede in der Erkennungsleistung mit Veränderung der Ansicht des emotionalen Reizes. Bei Betrachtung der Erkennungsraten (siehe Tabelle 33) wurde deutlich, dass die

Hypothese einer abfallenden Erkennungsleistung nicht für alle Emotionen haltbar ist. Für Angst, Ärger, Ekel und Trauer zeigte sich eine signifikante Verbesserung der Erkennungsleistung von Perspektive 1 zu Perspektive 2. Ekel wurde beispielsweise bei Vorgabe der Perspektive 3 am besten erkannt. Freude und Überraschung wurden über alle Perspektiven hinweg gleich gut erkannt. Dieser Befund regt zu der Annahme an, dass wie in Hypothese 2 formuliert, die Erkennung der sechs Basisemotionen durch verschiedene Verarbeitungsprozesse bedient. Angst, Ärger und Ekel sind sozial hochrelevante Emotionen, die bei rechtzeitigem Erkennen, das Überleben sichern können. Es erscheint somit durchaus sinnvoll, dass diese Emotionen auch aus anderen Perspektiven gut erkannt werden. Die Ergebnisse lassen aber darauf schließen, dass eine seitliche Ansicht die Erkennungsleistung sogar noch steigert. Möglicherweise konnte hier also ein Nachweis gefunden werden, dass die eingangs berichtete Überlegenheit der 45°-Perspektive zutrifft (Bruce, Valentine & Baddeley, 1987). Möglicherweise haben sich neuronale Verarbeitungsprozesse im Laufe der Evolution auf die Tatsache eingestellt, dass beispielsweise Angst besonders dann eine große Aussagekraft besitzen, wenn ein anderes Individuum einen angsteinflößenden Reiz entdeckt, der nicht die eigene Person betrifft. Man stelle sich beispielsweise vor, dass zwei Steinzeit-Frauen auf ihren Nachwuchs aufpassen. Beide sitzen am Lagerfeuer. Die eine der beiden Frauen entdeckt am Horizont einen Säbelzahntiger, ihr Gesicht spiegelt die aufkommende Angst wider. Es ist überlebenswichtig, dass die andere Frau diesen Gesichtsausdruck sofort erkennt, die Kinder zusammenruft und in Sicherheit bringt, während die angsterstarrte Frau eventuell erst nach einigen Sekunden wieder aus der „Betäubung“ erwacht. Für die anderen Emotionen ist eine Erklärung weniger einfach abzuleiten, da die perspektivische Emotionserkennung bisher kaum Gegenstand wissenschaftlicher Studien war. Es wird vermutet, dass die soziale Relevanz in allen Fällen eine Rolle spielt. Deutlich wird dies auch bei der Betrachtung der Geschlechterunterschiede. In drei von fünf Perspektiven fiel die Ärgererkennung der Frauen signifikant höher aus als die der Männer. Vermutlich ist sie für Frauen aufgrund ihrer körperlichen Unterlegenheit wichtiger (Bijlstra, Holland & Wigboldus, 2010). Dieser Befund sollte in nachfolgenden Studien systematisch untersucht werden.

In der hier durchgeführten Studie konnte ein Zusammenhang zwischen der über die Perspektiven gemittelten subjektiven Sicherheit und der allgemeinen Erkennungsleistung für alle Emotionen außer Ekel nachgewiesen werden: Je besser eine Emotion erkannt wurde, desto höher war die subjektive Sicherheit (bezüglich eines korrekt gefällten Urteils).

Trotz dieser ermutigenden Ergebnisse weist die Studie leider einige Einschränkungen auf. Um Lerneffekten entgegenzuwirken, wurde keine Person mehrmals mit gleichem Emotionsausdruck in unterschiedlichen Perspektiven gezeigt. Die Kehrseite dieses Vorgehens ist jedoch, dass nicht ausgeschlossen werden kann, dass systematische Personeneffekte vorliegen. So ist es denkbar, dass Unterschiede zwischen den Perspektiven lediglich darauf beruhen, dass zufällig einer Zelle (beispielsweise Angst, Perspektive 2) Bilder zugeordnet wurden, die von den Probanden besonders gut erkannt wurden. Außerdem ist die Stimulusanzahl mit lediglich drei Bildern pro Zelle als sehr gering einzuschätzen. Um diese beiden Defizite auszugleichen, wurde eine zweite Studie durchgeführt, bei der der Emotionsausdruck einer Person über alle Perspektiven variierte und eine Zellbesetzung von zumindest von fünf Stimuli umgesetzt werden konnte.

Diskussion der Ergebnisse aus den Fragebögen

Ein Zusammenhang zwischen den eingesetzten Fragebögen und der perspektivischen Erkennungsleistung konnte nicht gefunden werden. Gründe hierfür sind unklar. Für den TeiQue beispielsweise ergab sich ein signifikant höherer Mittelwert als in der Normierungsstichprobe (Freudenthaler, Neubauer, Gabler & Scherl, 2008), so dass die hier erhobene Stichprobe nicht als repräsentativ angesehen werden kann ($t(59) = 4{,}608$; $p < .000$). Die Ergebnisse des NEO-FFI im Vergleich mit einer aktuellen Vergleichsstichprobe (Egan, Deary & Austin, 2000) erbrachte, dass lediglich die Werte für Neurotizismus ($t(63) = 0{,}485$; $p = .629$) und Gewissenhaftigkeit ($t(63) = -0{,}058$; $p = .954$) repräsentativ verteilt sind. Die Werte für Extraversion ($t(63) = 5{,}533$; $p < .000$), Offenheit ($t(63) = 8{,}842$; $p < .000$) und Verträglichkeit ($t(63) = 3{,}604$; $p = .001$) unterschieden sich aber signifikant von der Norm. Für den „Relationship-Questionnaire“ liegen bislang keine Vergleichsnormen vor. Es wurde deswegen der Fragebogen „Experience in Close Relationships – Revised“ herangezogen (Ehrenthal, Dinger, Lamla, Funken & Schauenburg, 2009), welcher ebenfalls Ängstlichkeit und Vermeidung erfasst. Der „Relationship-Questionnaire“ wurde im Vorfeld ausgewählt, da er sowohl die Messung der vier Bindungskategorien als auch der zwei Dimensionen ermöglicht. Die Ergebnisse für die beiden Dimensionen für beide Fragebögen vergleichbar, da nahezu identisch erfasst. Während die Erfassung der Vermeidungs-Dimension vergleichbar erscheint ($t(57) = 1{,}396$; $p = .168$), ist dies für die Ängstlichkeits-Dimension augenscheinlich nicht der Fall ($t(57) = -3{,}290$; $p = .002$). Die Zuordnung der Bindungsstile in

die vier verschiedenen Kategorien erscheint ebenfalls nicht repräsentativ. Im Regelfall weisen etwa 2/3 der Personen einen sicheren Bindungsstil auf, hier wurden 46,2% der Stichprobe der unsicher-ambivalenten Bindungskategorie zugeordnet. Somit konnten Zusammenhänge in diesem Kontext zwar nicht aufgezeigt werden, sind aber nicht grundsätzlich auszuschließen. Eine Untersuchung mit einer für die erfassten Konstrukte repräsentativen Stichprobe mit größeren Teilstichproben erscheint hier sinnvoll und notwendig.

5.2 Studie 7: Konstanthaltung der Stimuli über die Perspektive

Wie oben beschrieben, ergab sich aus der Verwendung unterschiedlicher Personenstimuli für jede Emotion und Perspektive ein Problem: Die Darsteller wurden unterschiedlich gut erkannt, so dass eine Konfundierung der individuellen Erkennungsleistung (für die verschiedenen Darsteller) mit der Perspektivwahrnehmung nicht auszuschließen ist. Deswegen wurde eine weitere Studie durchgeführt, in der jeder der 30 eingehenden Darsteller nur eine Emotion, die Fotos dieser Emotion aber über alle fünf Perspektiven variierten. Insgesamt stellten jeweils fünf Darsteller einen Emotionsausdruck dar, so dass insgesamt 150 Stimuli zur Bewertung vorlagen (6 Emotionen x 5 Perspektiven x 5 Darsteller). Das Geschlecht wurde gleichmäßig verteilt, je 15 Darsteller waren männlich respektive weiblich.

Hypothesen Studie 7

Es werden dieselben Hypothesen wie in Studie 6 aufgestellt.

Bewertung der perspektivisch variierenden Stimuli

Der Versuchsablauf entsprach dem in Studie 6. Neben dem Austausch der Stimuli wurden keine weiteren Veränderungen vorgenommen. Es wurden die gleichen Fragebogenverfahren eingesetzt wie in der ersten Studie zur Beurteilung der perspektivisch variierenden Stimuli. Bewertet wurde das Stimulusmaterial von 74 Medizinstudenten, welche alle ihr schriftliches Einverständnis zur freiwilligen Teilnahme gaben (Ethikvotum 245/08_UBB/se). Der Anteil an Frauen betrug 59%, der der Männer 38%, 3% machten keine Angaben zu ihrem Geschlecht. Das mittlere Alter betrug 20,93 Jahre (SA = 2,32).

Auswertung der Erkennungsleistung der perspektivisch variierenden Stimuli

Zunächst wurden die durchschnittlichen Erkennungsleistungen errechnet. Diese sind Tabelle 39 zu entnehmen.

Tab. 39: Erkennungsleistung aufgeschlüsselt nach Emotion und Perspektive Studie 7

	Emotion						
Perspektive	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Angst**	**Gesamt**
1	98,11	90,81	95,68	89,73	94,32	84,05	92,12
2	98,11	89,73	95,95	90,54	91,89	75,96	90,36
3	97,57	90,54	95,41	92,97	92,43	75,68	90,77
4	97,29	90,54	95,13	87,84	92,16	75,14	89,68
5	98,11	86,22	94,32	87,03	93,51	72,16	88,56
Gesamt	97,84	89,57	95,30	89,62	92,86	76,60	-

Nachfolgend wurden potentielle Unterschiede zwischen den verschiedenen Perspektiven berechnet. Im Vorfeld wurde für jede Emotion innerhalb der verschiedenen Perspektiven die Normalverteilung mittels Shapiro-Wilks-Test überprüft. Diese konnte für alle Berechnungen nicht nachgewiesen werden ($p < .05$), weswegen in der Folge non-parametrische Verfahren eingesetzt wurden. Berechnungen mit dem Friedman-Test ergaben signifikante Unterschiede in der Erkennungsleistung zwischen den verschiedenen Perspektiven ($\chi^2(4) = 16,916$, $p = .002$). Anschließend wurden Paarvegleiche mittels Wilcoxon-Test durchgeführt, um die Ergebnisse detaillierter betrachten zu können. Über alle Emotionen hinweg betrachtet, unterschied sich die Erkennungsleistung aus Perspektive 1 signifikant von Perspektive 4 ($Z = -2,497$; $p = .013$) und Perspektive 5 ($Z = -3,274$; $p = .001$). Perspektive 2 unterschied sich in der Erkennungsleistung signifikant von Perspektive 5 ($Z = -2,052$; $p = .040$). Darüber hinaus unterschieden sich die Erkennungsraten der Perspektive 3 von Perspektive 4 ($Z = -1,986$; $p = .047$) und 5 ($Z = -2,900$; $p = .004$) signifikant voneinander. Alle anderen Vergleiche wurden nicht signifikant ($p > .05$).

Betrachtet man die perspektivische Erkennungsleistung für die einzelnen Emotionen, so ergab sich folgendes Bild: Für die Emotion Angst konnte ein signifikanter Unterschied zwischen Perspektive 1 und allen anderen Perspektiven bezüglich der Erkennungsleistung gefunden werden (1 zu 2: $Z = -2,086$; $p = .037$; 1 zu 3: $Z = -2,220$; $p = .026$; 1 zu 4: $Z = -2,634$; $p = .008$; 1 zu 5: $Z = -3,377$; $p = .001$). Für Ärger, Ekel und Freude ließen sich keine signifikanten Paarvergleiche nachweisen ($p > .05$). Die Trauererkennung fiel signifikant unterschiedlich aus zwischen Perspektive 3 und 4 ($Z = -2,547$; $p = .011$) sowie Perspektive 3 und 5 ($Z = -2,975$; $p = .003$). Für Überraschung wiederum fanden sich signifikanten Unterschiede in der Erkennungsleistung für den Vergleich von Perspektive 1 mit 5 ($Z = -2,086$; $p = .037$), Perspektive 3 mit 5 ($Z = -2,271$; $p = .023$) und Perspektive 4 mit 5 ($Z =$

-2,309; p = .020). Eine Übersicht der Ergebnisse ist Tabelle 40 zu entnehmen. Für die Richtung der signifikanten Paarvergleiche sei auf Tabelle 39 verwiesen.

Tab. 40: Übersicht signifikante Paarvergleiche Studie 7

	Emotion					
Signifikante Paarvergleiche der verschiedenen Perspektiven	**Angst**	**Ärger**	**Ekel**	**Freude**	**Trauer**	**Überraschung**
1 vs. 2	1 > 2*	-	-	-	-	1 > 2*
1 vs. 3	1 > 3*	-	-	-	-	-
1 vs. 4	1 > 4*	-	-	-	-	-
1 vs. 5	1 > 5*	-	-	-	-	-
2 vs. 3	-	-	-	-	-	-
2 vs. 4	-	-	-	-	-	-
2 vs. 5	-	-	-	-	-	-
3 vs. 4	-	-	-	-	-	-
3 vs. 5	-	-	-	-	3 > 5**	3 > 5*
4 vs. 5	-	-	-	-	-	4 > 5*

< = kleiner, > = größer; * = p < .05; ** = p < .001

In einem nächsten Schritt wurde mittels Mann-Whitney U-Test überprüft, ob das Geschlecht der Beurteiler einen Einfluss auf die perspektivische Erkennungsleistung hat. Es ließen sich signifikante Unterschiede in der Erkennungsleistung zwischen Männern und Frauen in der perspektivischen Emotionserkennung finden, verglich man die Ärgererkennung für Perspektive 1 (Z = -2.068; p = .039), sowie der Angsterkennung für Perspektive 2 (Z = -2,341; p = .019), 3 (Z = -2,659; p = .008) und 5 (Z = -2,167; p = .030) miteinander. In jedem dieser Fälle erkannten Frauen Ärger beziehungsweise Angst signifikant besser als die Männer.

Bei der Betrachtung potentieller Unterschiede in der Erkennungsleistung der verschiedenen Emotionen wurde die über die Perspektiven gemittelte Gesamterkennungsleistung herangezogen. Der Friedman-Test zeigte signifikante Unterschiede an ($\chi^2(5) = 156,080$; p < .000), nachfolgend durchgeführte Wilcoxon-Tests ergaben signifikante Unterschiede zwischen allen sechs Emotionen (p < .05) außer Überraschung und Trauer (Z = -1,107; p = .915), so dass folgende Rangreihe der allgemeinen Erkennungsleistung erstellt wurde: Freude > Ekel > Ärger > Trauer = Überraschung > Angst.

Zusätzlich wurde die subjektive Sicherheit (4 = sehr sicher bis 1 = sehr unsicher) ausgewertet. Eine Übersicht findet sich in Tabelle 41.

Tab. 41: Subjektive Sicherheit aufgeschlüsselt nach Emotion und Perspektive Studie 7

	Emotion						
Perspektive	**Freude**	**Überraschung**	**Ekel**	**Trauer**	**Ärger**	**Angst**	**Gesamt**
1	3,95	3,72	3,79	3,63	3,81	3,61	3,75
2	3,94	3,71	3,79	3,72	3,66	3,53	3,73
3	3,91	3,69	3,79	3,64	3,72	3,46	3,70
4	3,92	3,70	3,74	3,52	3,68	3,44	3,67
5	3,94	3,59	3,75	3,32	3,70	3,42	3,62
Gesamt	3,93	3,68	3,77	3,57	3,71	3,49	-

Die Einschätzung der subjektiven Sicherheit war insgesamt recht hoch. Da in erster Linie Zusammenhänge zwischen der subjektiven Sicherheit und der Erkennungsleistung interessierten, wurden in der Folge Korrelationen berechnet. Diese sind Tabelle 42 zu entnehmen.

Tab. 42: Korrelation Erkennungsleistung mit subjektiver Sicherheit Studie 7

	Erkennungs-leistung allgemein					
	Freude	**Über-raschung**	**Ekel**	**Trauer**	**Ärger**	**Angst**
Subjektive Sicherheit (über alle Perspektiven)	.388**	.439**	.330**	.531**	.465**	.300*

* = p < .05; ** = p < .001

<u>Vergleich der Erkennungsleistungen aus Studie 6 und Studie 7</u>

Verglich man mittels Mann-Whitney U-Test die Erkennungsleistungen, die in Studie 6 und Studie 7 erzielt wurden, so zeigten sich einige signifikante Unterschiede. Für Perspektive 1 unterschied sich die Erkennungsleistung für Angst (Z = -4,562; p < .000), Ärger (Z = -5,954; p < .000), Ekel (Z = -5,258; p < .000) und Trauer (Z = -3,882; p < .000). Für Perspektive 2 fand sich ein signifikanter Unterscheid für die Ekelerkennung zwischen den beiden Studien (Z = -3,111; p = .002). Für Perspektive 3 ließ sich dies für die Erkennung von Angst (Z = -2,002; p = .045) berichten. Für Perspektive 4 wurden Angst (Z = -2,098; p = .036), Ärger (Z = -4,763; p < .000) und Trauer (Z = -2,192; p = 0.12) in Studie 6 und 7 signifikant unterschiedlich erkannt. Für Perspektive 5 war ein signifikanter Unterschied bei der Trauererkennung (Z = -4,374; p < .000) und der Erkennung von Überraschung (Z = -2,341; p = .019) festzustellen. Die Emotionserkennung war in allen signifikanten Fällen in Studie 7 höher als in Studie 6.

Auswertung der eingesetzten Fragebogenverfahren

Abschließend fand eine Auswertung der Fragebögen statt. Wie bereits in der vorangegangenen Studie beschrieben, war es für die hier vorliegende Arbeit weniger von Belang, die erhobenen Werte mit bestehenden Normen zu vergleichen. Vielmehr bedeutsam erschienen Zusammenhänge zwischen den erhobenen Maßen und der Emotionserkennungsfähigkeit. Hierzu wurden Korrelationen nach Pearson berechnet. Korreliert wurden die Mittelwerte der Fragebögen mit den mittleren Erkennungsleistungen pro Emotion. Eine Übersicht über die Ergebnisse des TeiQue und NEO-FFI sowie die Ergebnisse der Dimensionen „Anxiety“ (Ängstlichkeit) und „Avoidance“ (Vermeidung) aus dem „Relationship-Questionnaire“ findet sich in Tabelle 43, eine Auflistung der Bindungsstile in Tabelle 44. Da nicht alle Probanden Angaben gemacht haben, mussten die Analysen teilweise mit geringeren Fallzahlen getätigt werden.

Tab. 43: Mittelwerte, Standardabweichung, Minimum und Maximum von NEO-FFI und TeiQue Studie 7

	Minimum	**Maximum**	**M**	**SA**
NEO-FFI (Skala von 0-4)				
Neurotizismus	0,58	3,42	1,65	0,65
Extraversion	1,17	3,58	2,59	0,49
Offenheit	1,00	3,58	2,59	0,49
Verträglichkeit	1,17	3,75	2,74	0,51
Gewissenhaftigkeit	1,17	4,08	2,79	0,68
TeiQue (Skala von 1-7)				
Gesamt	3,60	6,47	5,33	0,63
Wohlbefinden	3,17	7,00	5,73	0,97
Selbstkontrolle	2,83	6,50	4,96	0,85
Emotionalität	3,00	6,75	5,38	0,76
Soziale Verträglichkeit	3,50	6,67	5,13	0,85
Relationship-Questionnaire				
Anxiety	-3,00	8,00	1,66	2,37
Avoidance	-8,00	8,00	0,44	4,28

Tab. 44: Übersicht Bindungsstile Studie 7

	Häufigkeit	**Prozent**
Sicher	2	2,8
Unsicher-vermeidend	7	9,9
Unsicher-ambivalent	34	47,9
Unsicher-desorganisiert	15	21,1

Die Auswertung ergab keine signifikanten korrelativen Zusammenhänge. Die Überprüfung mittels Kruskal-Wallis-Test, ob der Bindungsstil (siehe Tabelle 44) die Erkennungsleistung

emotionaler Bilder beeinflusst, blieb ebenfalls ohne signifikantes Ergebnis (Angst: $\chi^2(3) = 0{,}673$; $p = 0{,}880$; Ärger: $\chi^2(3) = 1{,}176$; $p = .759$; Ekel: $\chi^2(3) = 1{,}730$; $p = .630$; Freude: $\chi^2(3) = 2{,}661$; $p = .447$; Trauer: $\chi^2(3) = 0{,}751$; $p = .861$; Überraschung: $\chi^2(3) = 5{,}719$; $p = .126$).

Diskussion der perspektivischen Erkennungsleistung

In dieser Studie zur Überprüfung der perspektivischen Emotionserkennung wurde der emotionale Stimulus einer Person über alle fünf Perspektiven variiert, um Personeneffekte auszuschließen. Es zeigten sich für einige Emotionen signifikante Unterschiede in der Erkennungsleistung mit Veränderung der Ansicht des emotionalen Reizes. Für Ärger, Ekel und Freude konnten keine signifikanten Unterschiede für die perspektivische Erkennungsleistung gefunden werden. Für Angst jedoch zeigt sich eine signifikant bessere Erkennungsleistung aus Perspektive 1. Trauer wird hingegen bei Vorgabe der Perspektive 3 am besten erkannt. Für Überraschung zeigte sich eine homogene Erkennungsleistung über alle Perspektiven, die Erkennungsrate für Perspektive 5 fiel im Vergleich jedoch deutlich ab. Diese Ergebnisse unterschieden sich teilweise von denen der ersten Studie. Freude wurde in beiden Studien über alle Perspektiven gleich gut erkannt. Angst wurde aus der zweiten Perspektive am besten erkannt. Die in Studie 6 gefundenen Unterschiede in der perspektivischen Erkennungsleistung für Ärger und Ekel konnten in Studie 7 nicht bestätigt werden. Die Ergebnisse für Überraschung waren bis auf Perspektive 5 vergleichbar. Auch wenn die Erkennungsraten in Studie 7 auf einen massiven Deckeneffekt schließen lassen, scheint dieses Paradigma zur Überprüfung der perspektivischen Emotionserkennung besser geeignet zu sein, da Zusammenhänge zu den gezeigten Personen ausgeschlossen werden können. Möglicherweise lagen in Studie 7 Lerneffekte vor. Potentielle Maßnahmen zur Einschränkung dieser werden in 5.3 diskutiert. Die Möglichkeit, Lernkurven zu erstellen, um potentielle Lern- und Saturierungseffekte nachweisen zu können, war in diesem Fall leider nicht gegeben, da die Reihenfolge der Randomisierung für jeden einzelnen Probanden nicht erfasst wurde. Dies sollte in Folgestudien in jedem Fall geschehen. Zeigt sich in Folgestudien anfangs ein deutlich Anstieg in der Erkennungsleistung, welcher sich dann stabilisiert, so ist zu vermuten, dass die Personen mit ihren (gleichbleibenden) emotionalen Ausdrücken wiedererkannt werden und Lerneffekte vorliegen, die die Messung der eigentlichen Erkennungsleistung stören.

In der hier durchgeführten Studie konnte ein Zusammenhang zwischen der subjektiven Sicherheit und der Erkennungsleistung gefunden werden. Je sicherer der Proband in der Einschätzung war, desto besser war die Erkennungsleistung. Möglicherweise spielte in dieser Studie das Wiedererkennen der Stimuli aus unterschiedlichen Perspektiven eine Rolle. Die Studienteilnehmer konnten ihr vorläufig getroffenes Ergebnis für einen Stimulus wiederholt bestätigen und in der Folge eventuell auch eine stabilere Einschätzung des vorgenommenen Urteils gewinnen. Da die Stimuli hier randomisiert vorgegeben wurden, wäre eine Untersuchung, welche diesen potentiellen Effekt systematisch untersucht von Bedeutung.
Es konnte wie in Studie 6 ein Geschlechtereffekt gefunden werden. Neben der besseren Ärgererkennung für Perspektive 2 konnte eine Überlegenheit der Frauen für die Erkennung von Angstreizen aus Perspektive 2, 3 und 5 nachgewiesen werden. Es wird, wie in der vorherigen Studie diskutiert, angenommen, dass Frauen aufgrund ihrer körperlichen Unterlegenheit darauf angewiesen waren, potentielle Gefahrenreize möglichst schnell zu erkennen, um so ihr Überleben zu sichern. Andere Personen der Gruppe als Referenz heranzuziehen kann als strategischer Überlebensvorteil gedeutet werden. Hierbei sind gerade die Angst- und Ärgerausdrücke von Bedeutung.

Diskussion der Ergebnisse aus den Fragebögen

Ein Zusammenhang zwischen den eingesetzten Fragebögen und der perspektivischen Erkennungsleistung konnte wie auch in Studie 6 nicht gefunden werden. Es werden hierfür ähnliche Gründe angenommen: Nicht ausreichende Repräsentativität der Stichprobe und zu kleine Gruppengrößen. Für den TeiQue beispielsweise ergab sich ein signifikant höherer Mittelwert als in der Normierungsstichprobe (Freudenthaler, Neubauer, Gabler & Scherl, 2008), so dass die hier erhobene Stichprobe nicht als repräsentativ angesehen werden kann ($t(62) = 4{,}617$; $p < .000$). Die Ergebnisse des NEO-FFI wurden wiederum mit einer aktuellen Vergleichsstichprobe (Egan, Deary & Austin, 2000) verglichen. Der Vergleich erbrachte, dass lediglich die Werte für Neurotizismus ($t(68) = 0{,}377$; $p = .707$) und Gewissenhaftigkeit ($t(68) = 1{,}402$; $p = .165$) repräsentativ verteilt waren. Die Werte für Extraversion ($t(68) = 5{,}487$; $p < .000$), Offenheit ($t(68) = 6{,}393$; $p < .000$) und Verträglichkeit ($t(68) = 4{,}285$; $p = .001$) unterscheiden sich aber signifikant von der Norm. Für den „Relationship-Questionnaire“ wurde wiederum der Fragebogen „Experience in Close Relationships – Revised“ herangezogen. (Ehrenthal, Dinger, Lamla, Funken & Schauenburg, 2009). Die Ergebnisse

beider Dimensionen waren nicht mit der Norm vergleichbar (Vermeidungs-Dimension: $t(66) = -4{,}428$; $p < .000$; Ängstlichkeits-Dimension: $t(66) = -3{,}833$; $p < .000$).

5.3 Diskussion der Emotionserkennung über verschiedene Perspektiven

Unterschiede in der Erkennungsleistung zwischen den verschiedenen Basisemotionen scheinen sich mithilfe von perspektivischem Stimulusmaterial noch differenzierter untersuchen zu lassen. Während beispielsweise die Erkennungsleistung für Freude in beiden Studien konstant blieb, nahm die Erkennungsleistung für Trauer insgesamt signifikant ab. Dies weist darauf hin, dass die Hinweisreize unterschiedlich verarbeitet werden, möglicherweise in verschiedenen neuronalen Verarbeitungsschleifen (siehe 2.2.2). Ursachen hierfür wurden in 5.1 und 5.2 diskutiert, es ist denkbar dass einerseits die soziale Relevanz der Emotionen (beispielsweise für das Überleben) einen Einfluss hat, andererseits sind einige Gesichtsmerkmale bei zunehmend seitlicher Perspektive nur noch schwer zu detektieren. Inwieweit diese beiden Aspekte miteinander zusammenhängen, kann an dieser Stelle nicht eingeschätzt werden. Ein grundsätzlicher Vorteil für die 45°-Perspektive kann nicht berichtet werden. Beide Studien wiesen sehr hohe Erkennungsraten auf, insbesondere Studie 7. Grundsätzlich ist die Art der Operationalisierung aus Studie 7 zu bevorzugen, da Effekte auf Grundlage verschiedener Identitäten ausgeschlossen werden können. Es erscheint notwendig, die Aufgabenschwierigkeit zu erhöhen, um Lerneffekte zu minimieren. Hierzu stehen eine Reihe von Möglichkeiten zur Verfügung, beispielsweise die Verkleinerung der Stimuli oder die Verkürzung der Präsentationsdauer. Eine Studie mit einer Präsentationsdauer von lediglich 300 ms (hier war es eine Sekunde) wurde bereits geplant und durchgeführt. Eine Auswertung der Ergebnisse folgt. Von großem Interesse wäre die Durchführung einer Eyetracking-Studie zur Überwachung der Blickbewegungspfade. Hieraus würden sich Hinweise ergeben, welche Gesichtsmerkmale bei der perspektivischen Emotionserkennung eine besondere Rolle spielen. In einer Studie von Barton, Radcliffe, Cherkasova, Edelman und Intriligator (2006) wurden in einer Eyetracking-Studie Blickbewegungspfade untersucht, welche aufrecht-frontal oder invertiert dargestellt wurden. Es zeigte sich, dass bei invertierten Gesichtern der Blick eher auf dem Mund und unteren Gesicht lag, ansonsten Augen und oberes Gesicht fixiert wurden.

Ein Faktor, der bei dem hier vorliegenden Stimulusmaterial nicht berücksichtigt wurde ist die Blickrichtung. Sie spielt eventuell eine Rolle als Kovariate, beispielsweise dass die soziale Relevanz höher eingeschätzt wird, wenn der Blick auf einen gerichtet ist – unabhängig von der Orientierung des Kopfes. Hinweise hieraus ergeben sich beispielsweise aus einer Studie

von Doi und Shinohara (2009), die die subjektiv wahrgenommene Zeit für ärgerliche und fröhliche Gesichter unter Variation der Blickrichtung untersucht haben. Eine Veränderung der Zeitwahrnehmung trat nur für ärgerliche Gesichter auf und nur dann, wenn der Blick auf den Betrachter ausgerichtet war. Die Autoren vermuten ein höheres Arousal, was zu einer Verstellung der inneren Uhr führen könnte. Mit den hier vorgestellten Studien konnte ein Anfang bezüglich der Untersuchung und Deutung der perspektivischen Emotionserkennung getätigt werden. Studien mit ähnlichem Experimentalaufbau sind zum jetzigen Zeitpunkt nicht bekannt. Es konnte jedoch aufgezeigt werden, dass zahlreiche systematische Folgestudien notwendig sind, um die gefundenen Ergebnisse sinnvoll einbetten zu können.

6. ALLGEMEINE DISKUSSION

Ziel dieser Arbeit war es, ein Bilderset zu erstellen, das zur wissenschaftlichen Untersuchung der Emotionserkennung aus Gesichtern eingesetzt werden kann. Im Vorfeld wurde die bisherige Literatur gesichtet, in der Absicht alle relevanten Aspekte zu berücksichtigen, die notwendig sind, um ein Höchstmaß an testtheoretischer Güte zu erzielen.

Demnach sollte das Bilderset folgende Aspekte einbeziehen:

- Minimierung ethnischer Unterschiede im Stimulusmaterial. Für diesen Bildersatz wurden deshalb nur kaukasische Personen fotografiert.
- Definition der mimischen Muster der Basisemotionen nach EmFACS als objektives Maß des mimisch-emotionalen Ausdrucks.
- Gleiche Verteilung beiderlei Geschlechts.
- Beschränkung des Altersbereiches der Darsteller auf 18 bis 35, um dem „Own-Age-Bias“ entgegenzuwirken.
- Simultane Aufnahmen aus fünf Perspektiven, um den realen Bedingungen bei der Emotionserkennung im Alltag näherzukommen und die der Emotionserkennung zugrundeliegenden (neuronalen) Prozesse noch besser zu verstehen.
- Kontrolle der Variablen Intensität und Natürlichkeit des Emotionsausdrucks sowie die Attraktivität der fotografierten Personen.
- Auswahl der emotionalen Gesichtsstimuli mit dem Ziel, eine größtmögliche Übereinstimmung der durchschnittlichen Erkennungsraten der Einschätzung von Intensität und Natürlichkeit des Gesichtsausdrucks sowie Attraktivität der fotografierten Person.
- Überprüfung der psychometrischen Qualität des Bildersets.
- Nachträgliche Einbettung des neuen Stimulusmaterials in das bestehende Testsystem FEEL als Revision (FEEL-R) mit 48 Frontalbildern und als Perspektivversion (FEEL-Rpersp) mit über die Perspektive variierenden emotionalen Gesichtsausdrücken von mindestens 30 Darstellern.

Zur Umsetzung dieser Ziele wurden 65 Personen unter standardisierten Bedingungen fotografiert. Die verschiedenen Auswahlstufen durchlaufend, blieben schlussendlich 48

Personen, die mit jeweils einem emotionalen Ausdruck in das Bilderset eingingen. Für die Perspektivversion wurden Bilder von maximal 45 Personen eingesetzt

Die durchgeführten Studien ergaben:

a) **Allgemeine Erkennungsleistung**: Die allgemeine Erkennungsleistung lag in der Normierungsstichprobe bei 75,38% für Angst, 82,24% für Trauer, 89,28% für Ekel, 90,28% für Überraschung, 91,66% für Ärger und 97,95% für Freude. Die hier gefundenen Ergebnisse zeigten keine signifikanten Unterschiede zu denen aus der Befragung mit dem freien Antwortformat auf.

b) **Normierung**: In einer Studie mit N = 161 Personen konnten verschiedene Altersgruppen, Schichtzugehörigkeiten und Persönlichkeitsstile berücksichtigt werden. Insgesamt zeigten sich kaum Unterschiede zwischen den verschiedenen Gruppen in ihrer Erkennungsleistung. Dennoch konnten einige wichtige Unterschiede gefunden werden: Frauen konnten Angst besser erkennen als Männer, der Status beeinflusse die Trauererkennung und Neurotizismus und Extraversion wirkten sich auf die Angst- beziehungsweise Ärgererkennung aus. Aufgabe von zukünftigen Studien ist es, diese Effekte systematisch zu untersuchen.

   Eine Normalverteilung der Erkennungsleistung lagnicht vor, die Gründe hierfür wurden beispielsweise in 4.4 ausführlich erläutert.

c) **Objektivität**: Inwieweit ein Test objektiv ist, lässt sich auf drei Ebenen beurteilen: Es stehen einem die Durchführungs-, Auswertungs- und Interpretationsobjektivität zur Verfügung. Die Durchführungsobjektivität beschreibt den Grad der Unabhängigkeit der Testergebnisse von Verhalten des Untersuchers während der Erhebung. Die Durchführungsobjektivität lässt sich durch Standardisierung maximieren (Lienert & Raatz, 1994). Dies wurde bei allen Studien mit den „Pictures of Facial Affect-Ulm“ konsequent umgesetzt. Die Probanden bekamen alle Instruktionen auf einem Bildschirm präsentiert, es gab keine zusätzlichen Erläuterungen durch den Versuchsleiter. Einflüsse durch Anwesenheit anderer Personen, Temperatur, Lärmpegel oder Beleuchtung konnten nicht vollständig kontrolliert werden, da teilweise ein Online-Verfahren herangezogen wurde. Die Auswertungsobjektivität ist hingegen vollständig gegeben. Für jeden Stimulus gibt es nur eine richtige Antwort. Somit lassen sich Berechnungen zur Leistung einzelner Probanden schnell und einfach durchführen. Durch die Möglichkeit des Datenimports in gängige Berechnungsprogramme wie SPSS oder Excel lassen sich Berechnungsfehler

vollständig ausschließen. Durch den Versuch einer Normierung kann auch die Interpretationsobjektivität als gegeben angesehen werden. Dem Nutzer werden Normen bereitgestellt, um die erhobenen Ergebnisse interpretieren zu können. Anhand dieser Normen können Ergebnisse aus anderen Studien sinnvoll eingeordnet werden.

d) **Validität**: Eine Überprüfung der Kriteriumsvalidität erfolgte über den Einsatz eines freien Urteilsformates. Hier zeigte sich, dass die Verwendung der bestehenden Emotionskategorien sinnvoll ist und überzufällig häufig eine korrekte Benennung der Stimuli erfolgte. Es konnte hinreichend eingegrenzt werden, dass Probanden nicht nur aus Mangel an Alternativen, sich im geschlossenen Antwortformat für eine der Antworten entschieden, was darauf schließen lässt, dass die intendierten Emotionen korrekt erkannt wurden. Dies lässt annehmen, dass mit den „Pictures of Facial Affect-Ulm“ tatsächlich die Emotionserkennungsleistung erfasst wird.

   Die ökologische Validität konnte durch zwei Maßnahmen maximiert werden: Einerseits wurde die Natürlichkeit des Gesichtsausdrucks bewertet, so dass Bilder mit geringer Natürlichkeit im Ausdruck aussortiert werden konnten. Andererseits wurden Bilder aus fünf verschiedenen Perspektiven angefertigt, um den Alltagsbezug zu erhöhen.

   Hinweise zur Konstruktvalidität würden sich aus dem Vergleich der hier gewonnenen Erkennungsleistungen mit denen des FEEL-Tests oder anderer Bildersets ergeben – sicherlich ein interessanter Vergleich für zukünftige Studien.

e) **Reliabilität**: Die Retest-Reliabilität kann mit Übereinstimmungen von 73,24% für Angst, 82,09% für Trauer, 84,98% für Ekel, 91,33% für Ärger, 92,01% für Überraschung und 89,13% für Freude als zufriedenstellend bis gut angesehen werden. Es sei an dieser Stelle noch einmal darauf hingewiesen, dass nicht wie normalerweise eine Korrelation zwischen beiden Testzeitpunkten berechnet werden konnte. Eine genaue Erläuterung der Vorgehensweise findet sich in 4.5.

   Auch Cronbach’s Alpha ist mit r = .634 als zufriedenstellend zu bezeichnen. Zwar zeigt gerade die Trauerskala mit r = .511 leichte Defizite, ist damit aber die einzige Emotion mit geringerer interner Konsistenz. Grund hierfür ist die größere Varianz der einzelnen Stimuli bezüglich der Erkennungsleistung.

Die Erkennungsleistung aus verschiedenen Perspektiven wurde in dieser Form zum augenblicklichen Kenntnisstand erstmalig untersucht. Die Ergebnisse sind interessant,

beispielsweise ist ein Abfall der Erkennungsleistung in Tendenz zu beobachten. Gerade in der zweiten Studie zeigten sich jedoch Deckeneffekte, die Erkennungsleistungen waren insgesamt sehr hoch. Folgestudien sind hier also notwendig. Eine Studie wird derzeit durchgeführt, die dem Experimentalaufbau von Studie 2 entspricht, die Darbietungsdauer der Stimuli wurde aber von einer Sekunde auf 300 ms verkürzt. Ergebnisse liegen noch nicht vor, es ist aber anzunehmen, dass sich Unterschiede in der korrekten Zuordnung der gezeigten Emotion über die Perspektiven aufgrund der erschwerten Bedingungen deutlicher zeigen. Eine Vielzahl von Fragestellungen kann mit diesem neuen Stimulusmaterial bearbeitet werden, um die Prozesse bei der Verarbeitung emotionaler Stimuli noch detaillierter zu erfassen.

Auffällig ist, dass Freude in vielen Studien zur Überprüfung der Emotionserkennung am besten erkannt wird. Es ist nicht auszuschließen, dass dies damit im Zusammenhang steht, dass Freude bei den meisten Fotosets natürlich dargestellt wurde, während die anderen Emotionen zumeist geschauspielert wurden (beispielsweise Langner, Dotsch, Bijlstra, Wigboldus, Hawk, & Knippenberg, 2010) wurden. Auch in den hier beschriebenen Studien wurde Freude signifikant besser erkannt als die anderen Emotionen. Daraus lässt sich die Überlegung ableiten, natürliche Emotionen zu fotografieren und bewerten zu lassen, um so die Emotionserkennungsfähigkeit einer Person zu messen. Neben großen kulturellen und individuellen Unterschieden bezüglich des Emotionsausdrucks (beispielsweise Elfenbein, Beaupré, Lévesque & Hess, 2007), besteht aber bei solch einem Vorgehen ein großes Problem darin, Emotionen überhaupt adäquat zu induzieren (Merten, 2003). Möglichkeiten sind die Imagination, das Lesen einer emotionalen Vignette, das Erinnern eines emotionalen Ereignisses oder das Ansehen emotionaler Videoausschnitte. Studien zeigen (Merten, 2003), dass sich negative Emotionen im Allgemeinen besser induzieren lassen als positive (d = 0,526 zu d = 0,407). Am besten schneiden Filme und Geschichten ab (d = 0,743 für negative Emotionen und 0,726 für positive Emotionen), wenn die Personen explizit aufgefordert wurden, sich in die Emotion zu versetzen. Die Kehrseite eines solchen Vorgehens wäre der Verlust der Standardisierung, welche durch die EmFACS-Anweisungen (Ekman & Friesen, 1982) gewährleistet werden. Somit ist ein solches Vorgehen im Vorfeld genau abzuwägen.

Bezüglich der Homogenität der Stichprobe lässt sich Folgendes anführen: In einer Vielzahl der Studien wurden Studierende der Humanmedizin als Versuchspersonen rekrutiert, was

eine relative Homogenität bedingt. Die hier gewonnenen Daten können somit nicht als grundsätzlich repräsentativ angesehen werden. Die Interpretierbarkeit ist eingeschränkt, auch wenn die hier gewonnen Daten jeweils relativ gut mit bestehenden Normen übereinstimmen. Aus diesem Grund wurde bei der Normierungsstichprobe explizit darauf geachtet, Probanden unterschiedlichen Alters und Bildung für die Studienteilnahme gewinnen zu können. Interessanterweise scheinen die Ergebnisse bezüglich der Erkennungsleistung relativ vergleichbar zu sein, was darauf hinweist, dass die Emotionserkennungsfähigkeit eine basale Fähigkeit des Menschen darstellten könnte.

Abschließend lässt sich anführen, dass eine derart ausführliche und durch testtheoretische Überlegungen geprägte Untersuchung von mimischem Stimulusmaterial bisher einzigartig ist. Dabei erscheint gerade diese „Basisarbeit“ notwendig, um den Aussagegehalt späterer Studien zu maximieren. Wie den obigen Ausführungen zu entnehmen ist, sind viele der psychometrischen Maße zufriedenstellend bis gut, einige Aspekte konnten noch nicht zufriedenstellend umgesetzt werden und zeigen die Notwendigkeit für weiterführende Studien auf. So ist es sinnvoll und notwendig, auf dem vorliegenden Stimulusmaterial aufbauend, die Itemschwierigkeit zu optimieren und Deckeneffekte zu reduzieren. Hierzu würde sich beispielsweise eine Variation der Intensität – technisch mittels Morphing herbeigeführt – optimal eignen. Es ist anzunehmen, dass sich dann auch eine Normalverteilung der Erkennungsleistung einstellt, die eine wichtige Voraussetzung für psychometrisch gutes Testverfahren darstellt. Sinnvoll wäre es, die hier vorgestellten Berechnungen dann noch einmal mit dem angepassten Stimulusmaterial durchzuführen.

7. AUSBLICK

„Abenteuerliche Anwendungen werden erträumt: Flughäfen könnten die Gesichter von Fluglotsen nach Zeichen von gefährlicher Erschöpfung abtasten, Marketingfirmen garantiert unverbrauchte Verbraucherreaktionen messen. Geldautomaten könnten sich selber sperren, sobald die Kamera eine verdächtige Miene einfängt und Flughafencomputer Alarm auslösen, wenn es bei der Frage: ‚Irgendetwas Gefährliches im Koffer?', im Gesicht des Passagiers gefährlich zuckt. Die Technologie dazu wird in 2-3 Jahren vorhanden sein, so Jeffrey Cohn, Psychologe an der Universität Pittsburgh, der solche Computerprogramme mitentwickelt hat. Aber die Maschinen können allenfalls Mikroausdrücke einfangen, interpretieren können sie diese aber nicht. Das kann nur der Mensch, am besten einer mit viel Lebenserfahrung und Intuition. Denn flackernde Aggression um den Mund verrät noch nicht, ob der Fluggast einen Anschlag plant oder nur das Geschnatter seiner Frau nicht mehr hören kann. Und die Angst eines Schuldigen gefasst zu werden, sieht genau so aus, wie die Angst eines Unschuldigen, dass man ihm nicht glauben könnte." (Eberle, 2004).

Emotionen sind für das soziale Zusammenleben von enormer Bedeutung – sowohl hinsichtlich des Informationsgehalts für den Interaktionspartner als bei der Regulation interner Prozesse. In neuerer Zeit konnten vor allem neurologische und psychiatrische Studien wichtige Erkenntnisse hinsichtlich der Verarbeitung emotionaler Stimuli erbringen. **Emotionserkennung** scheint eine basale Fähigkeit zu sein, die bereits von frühester Kindheit an effektiv funktioniert. Psychische Störungen sind in aller Regel mit Defiziten oder Exzessen in der Emotionsverarbeitung verknüpft. Allerdings wird dieser Aspekt in bisherigen Therapien nur selten genutzt. Mit dem hier vorgestellten Bilderset wird für Forscher und Kliniker ein Feld für die standardisierte Diagnostik der **Emotionserkennungsfähigkeit** (Screening der Defizite und weiterführende Untersuchungen) eröffnet. Darüber hinaus sind Verlaufskontrollen des therapeutischen Erfolgs denkbar. Auch direkte therapeutische Interventionen sind vorstellbar, beispielsweise ein explizites Training im Bereich der Emotionserkennung zur Steigerung positiver sozialer Kontakte.

Doch wie die obigen Überlegungen zeigen, birgt die Notwendigkeit des Erlebens und Erkennens von Emotionen in unserem Alltag auch gewisse Risiken. Schon zum jetzigen Zeitpunkt stehen uns hochentwickelte Technologien zur Verfügung, die den Anspruch erheben, menschliche Verhaltensweisen analysieren zu können. Eingesetzt werden sollen diese zum Schutz derer, die sie analysieren, der Menschheit. Es bleibt die Frage bestehen, inwieweit der Mensch ersetzbar ist. Kann eine Maschine so fein differenzieren wie ein Mensch? Ist es möglich, in einer Technologie alle Kontextvariablen zu berücksichtigen und situationsspezifisch zu gewichten, so wie der Mensch es tut? Auch wenn diese Fragen in der näheren Zukunft nicht beantwortet werden können, so zeigen sie doch auf, dass ein achtsamer Umgang mit solchen Technologien oberste Priorität einnehmen sollte.
Zusätzlich stellt sich die Frage, ob eine reine Gesichtsanalyse ausreichen wird, um Emotionen zuverlässig zu erkennen. Es gibt Studien, die darauf hinweisen, dass der Zusammenhang zwischen Gesichtsausdruck und Emotion fraglich ist (für einen Überblick zur Facial-Feedback-Hypothese siehe Meta-Analyse von McIntosh, 1996). Unabhängig davon ist die technische Umsetzung noch weit davon entfernt, zuverlässige Einteilungen vornehmen zu können. Klassifikationen einer aktuellen Studie von Baltrušaitis und Kollegen (2011) liegen unter Hinzunahme von Gestik bei etwa 40% korrekter Zuordnung zur gezeigten Emotion in Echtzeit-Paradigmen. Eine zuverlässige Klassifikation scheint die Hinzunahme anderer Parameter, beispielsweise Hautleitwert, Herzschlag, EEG oder EMG, notwendig zu machen (beispielsweise Hernandez, Morris & Picard, 2011; Walter, Scherer, Schels, Glodek, Hrabal, Schmidt, Böck, Limbrecht, Traue & Schwenker, 2011; Böck, Limbrecht, Walterm Glüge, Hrabal, Wendemuth & Traue, 2012).

Nichts desto trotz ist für die Erfassung der menschlichen Fähigkeit zur Emotionserkennung ein standardisiertes Bilderset ein großer Gewinn. Mit den „Pictures of Facial Affect-Ulm“ ist es gelungen, ein Bilderset zu erstellen, welches allgemeinen testtheoretischen Überprüfungen standhalten konnte. Basierend auf diesem Bilderset lassen sich zahlreiche Ideen für Weiterentwicklungen ableiten, um beispielsweise (wie im vorherigen Kapitel bereits angesprochen) die Itemschwierigkeit zu optimieren. Mögliche Erweiterungen sind:

- Variation der Intensität durch die Bildung von gemorphten Zwischenbildern zwischen dem neutralen Stimulus und dem Stimulusbild mit voller Expressivität.

- Erstellung dynamischer Stimuli durch die Aneinanderreihung von in der Intensität an- und absteigender Bilder.
- Diese beiden Schritte lassen sich für das Stimulusmaterial aus fünf unterschiedlichen Perspektiven umsetzen.

Durch die Umsetzung dieser Folgeschritte könnte eine umfassende Testbatterie zusammengestellt werden, wobei der FEEL-Test die notwendige Hard- und Software bietet, um das Stimulusmaterial standardisiert einzusetzen.

Denkbar wäre auch eine Erweiterung der Emotionskategorien. Es gibt weiterhin Zweifel, ob die von Ekman postulierten Basisemotionen tatsächlich die einzigen Basisemotionen darstellen (Frijda & Parrott, 2011). Zusätzlich gilt es zu bedenken, ob die Basisemotionen die Emotionen abdecken, welche für das Zusammenleben und soziale Interaktionen (ausschließlich) bedeutsam sind. In der heutigen Zeit sind Emotionen wie Scham, Stolz und so weiter von nicht zu unterschätzender Relevanz. Es ist nicht auszuschließen, dass die Erfassung der Erkennungsleistung von weiteren Emotionen wichtige Hinweise bei der Erforschung psychischer oder neurologischer Störungen und ihrer Heilung bietet.

8. ZUSAMMENFASSUNG

Die Fähigkeit, Emotionen auszudrücken, zu erkennen und zu interpretieren ist für die soziale Interaktion bedeutsam und ein wichtiges Element in den Konstrukten wie Empathie, Emotionsregulation und Emotionaler Intelligenz. Hierbei fällt der Mimik eine besondere Rolle zu. Jede Emotion hat einen korrespondierenden Gesichtsausdruck beziehungsweise eine mimische Entsprechung. Im Rahmen der zwischenmenschlichen Interaktion ist es wichtig, diese mimischen Entsprechungen korrekt der zugrundeliegenden Emotion zuzuordnen. Zur Untersuchung der menschlichen Emotionserkennungsfähigkeit werden Stimuli benötigt, welche emotionale Gesichtsausdrücke valide anzeigen. Kernstück der vorliegenden Arbeit stellte die Konzeption und testtheoretische Überprüfung des neuen Bildersets „Pictures of Facial Affect-Ulm“ (kurz: PFA-U) dar. Im Vordergrund stand einerseits der Einbezug der ökologischen Validität: Bildmaterial aus fünf unterschiedlichen Perspektiven sollte der Wahrnehmung von mimischen Ausdrücken im Alltag stärker entsprechen. Außerdem wurde die Natürlichkeit und Intensität des Gesichtsausdrucks und die Attraktivität des abgebildeten Gesichts bei der Auswahl berücksichtigt. Andererseits sollte die psychometrische Qualität des Bildmaterials überprüft werden, welche für die Verwendung psychologischer Testverfahren definiert wurden. Insgesamt wurden hierfür sieben Studien durchgeführt, die in der Folge zusammenfassend dargestellt werden sollen. Eine ausführliche Auflistung der Ergebnisse ist Punkt 6 (Allgemeine Diskussion) dieser Arbeit zu entnehmen. Im Vorfeld der Studien wurde das Bildmaterial der mimisch expressiven Gesichter erstellt. Insgesamt wurden 65 Probanden fotografiert. Die emotionalen Gesichtsausdrücke wurden im Vorfeld anhand eines Trainingsmanuals eingeübt. Alle Probanden wurden simultan aus fünf Perspektiven fotografiert, die Bedingungen bei der Erstellung waren standardisiert. Die Qualität des Bildes sowie die Qualität des Emotionsausdrucks waren bei einer ersten Auswahl entscheidend. 514 Frontalbilder von 50 Darstellern wurden für Studie 1 ausgesucht. In dieser erfolgte die Überprüfung der Erkennungsraten der 514 Bilder zur ersten Selektion. Bilder mit einer Erkennungsrate über 55% wurden für die weiteren Studien ausgewählt – dies entsprach 96 Bildern von 48 Darstellern. Jeder Darsteller ging mit zwei Fotos unterschiedlicher emotionaler

Gesichtsausdrücke in die weiteren Studien ein. In Studie 2 wurde die Kriteriumsvalidität überprüft, indem die 96 Frontalbilder ohne Vorgabe von Alternativen bewertet wurden. Die Ergebnisse zeigten, dass die emotionalen Gesichter auch in einem offenen Antwotformat zuverlässig erkannt wurden. Dies weist darauf hin, dass es gelungen ist, die intendierten emotionalen Ausdrücke fotografisch festzuhalten. Da vermutet wurde, dass verschiedene Faktoren die Emotionserkennung beeinflussen, wurde in Studie 3 die Beurteilung von Natürlichkeit und Intensität des Emotionsausdrucks sowie die Attraktivität der fotografierten Person beurteilt. Es zeigte sich, dass diese drei Faktoren zwar innerhalb einer emotionalen Kategorie relativ ähnlich beurteilt wurden, es jedoch große Unterschiede zwischen den verschiedenen Basisemotionen gab. In Studie 4 wurde das Bildmaterial einer großen Stichprobe von 161 Probanden zur Beurteilung vorgelegt. Die dort gefundenen Erkennungsraten sowie die Ergebnisse aus den Studien 2 und 3 wurden verwendet, um das endgültige Bilderset aus 48 Bildern auszuwählen. In der Folge wurden Itemschwierigkeit, Trennschärfe und Normalverteilung überprüft. Aufgrund der möglichst homogenen Erkennungsraten als Auswahlkriterium, lag keine Normalverteilung vor, die Itemschwierigkeit war nicht zufriedenstellend. Es wurde angedacht, beispielsweise die Intensität in Folgestudien mittels Morphing zu variieren, um dieses Problem zu lösen. Folgende Erkennungsraten konnten als erste Norm ermittelt werden: 75,38% für Angst, 82,24% für Trauer, 89,28% für Ekel, 90,28% für Überraschung, 91,66% für Ärger und 97,95% für Freude. In Studie 5 wurde die Retest-Reliabilität des Bildmaterials überprüft. Sie kann mit Übereinstimmungen von 73,24% für Angst, 82,09% für Trauer, 84,98% für Ekel, 91,33% für Ärger, 92,01% für Überraschung und 89,13% für Freude als zufriedenstellend bis gut angesehen werden. Dies eröffnet die Möglichkeit, das Bilderset auch als Verlaufsdiagnostik im klinischen Kontext einzusetzen, da enge Zusammenhänge zwischen psychischen und neurologischen Störungen und der Emotionserkennungsfähigkeit beschrieben werden. In den Studien 6 und 7 wurde die perspektivische Emotionserkennungsfähigkeit überprüft. Trotz deutlicher Deckeneffekte in den Erkennungsraten zeigten sich Unterschiede in der perspektivischen Emotionserkennung für die verschiedenen Basisemotionen. Zusätzlich zeigte sich ein leichter Vorteil in der Erkennung von Angst und Ärger aus den verschiedenen Perspektiven für weibliche Probanden. Die systematische und kriterienorientierte Erstellung von Stimulusmaterial zur Überprüfung der humanen Emotionserkennung hat aufgezeigt, dass die „Pictures of Facial Affect-Ulm“ gute psychometrische Qualitäten aufweisen.

LITERATURVERZEICHNIS

Abler B, Kessler H: Emotion Regulation Questionnaire – Eine deutsche Version des ERQ von Gross & John. Diagnostica 55: 144-152 (2009)

Amelang M, Bartussek D: Differentielle Psychologie und Persönlichkeitsforschung. Kohlhammer, Stuttgart (2001)

Amelang M, Zielinski W: Psychologische Diagnostik und Intervention. Springer, Berlin Heidelberg (1994)

Anastasi JS, Rhodes MG: Evidence for an own-age bias in face recognition. N Am J Psychol 8: 237-252 (2006)

Baltrušaitis T, McDuff D, Ntombikayise B, Mahmoud M, El Kaliouby R, Robinson P, Picard R: Real-time inference of mental states from facial expressions and upper body gestures. Automatic Face & Gesture Recognition and Workshops, IEEE International Conference (2011)

Bartholomew K, Horowitz LM: Attachment styles among young adults: A test of a four-category model. J Pers Soc Psychol 61: 226-244 (1991)

Barton JJS, Radcliffe N, Cherkasova MV, Edelman J, Intriligator JM: Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. Perception 35: 1089-1105 (2006).

Batty M, Taylor MJ: Early processing of the six basic facial emotional expressions. Cogn Brain Res 17: 613-620 (2003)

Becker DV, Kenrick DT, Neuberg SL, Blackwell KC, Smith DM: The confounded nature of angry men and happy women. J Pers Soc Psychol 92: 179-190 (2007)

Biehl M, Matsumoto D, Ekman P, Hearn V, Heider K, Kudoh T, Ton V: Matsumoto and Ekman's Japanese and Caucasian facial expression of emotion (JACFEE): Reliability data and cross-national differences. J Nonverbal Behav 21: 3-21 (1997)

Bijlstra G, Holland RW, Wigboldus DHJ: The social face of emotion recognition: Evaluations versus stereotypes. J Exp Soc Psychol 46: 657-663 (2010)

Blair RJR: Expression recognition and behavioural problems in early adolescence. Cogn Dev 15: 421-434 (2000)

Blair RJR, Colledge E, Murray L, Mitchell D: A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. J Abnorm Child Psychol 29: 491-498 (2001)

Blanz V, Vetter T: (1999). A morphable model for the synthesis of 3D faces. Proceedings of the 26th annual conference on Computer graphics and interactive techniques: 187-194 (1999)

Böck R, Limbrecht K, Walter S, Glüge S, Hrabal D, Wendemuth A, Traue HC: Intraindividual and interindividual multimodal emotion analyses in human-machine-interaction. In Proceedings of 2012 IEEE Conference on Cognitive Methods in Situation Awareness and Decision Support (zur Publikation angenommen)

Borkenau P, Ostendorf F: NEO-Fünf-Faktoren-Inventar nach Costa & McCrae: 2. Neu normierte und vollständig überarbeitete Auflage. Hogrefe, Göttingen (2008)

Bortz J: Statistik für Human- und Sozialwissenschaftler. Springer, Heidelberg (2005)

Bouhuys AL, Geerts E, Mersch PP: Relationship between perception of facial emotions and anxiety in clinical depression: does anxiety-related perception predict persistence of depression? J Affect Disord 43: 213-223 (1997)

Bradley MM; Codpispoti M; Sabatinelli D, Lang PJ: Emotion and Motivation II: Sex differences in picture processing. Emotion 1: 300-319 (2001)

Bradley MM, Lang PJ: Measuring emotion: The Self-Assessment Manikin and the semantic differential. J Behav Ther Exp Psychiatry 25: 49-59 (1994)

Braun M, Traue HC, Frisch S, Deighton RM, Kessler H: Emotion recognition in stroke patients with left and right hemispheric lesion: Results with a new instrument – the FEEL Test. Brain Cogn 58: 193-201 (2005)

Bruce V, Valentine T, Baddeley AD: The basis of 3/4 view advantage in face recognition. Appl Cogn Psychol 1: 109-120 (1987)

Buck R: The Communication of Emotion. Guildford Press, New York (1984)

Buck R, Miller RE, Caul WF: Sex, personality, and physiological variables in the communication of affect via facial expression. J Pers Soc Psychol 30: 587-596 (1974)

Cacioppo J: Feelings and emotions: Roles for electrophysiological markers. Biol Psychol 67: 235–243 (2004)

Calder AJ, Young J, Keane J, Dean, M: Configural information in facial expression perception. J Exp Psychol Hum Percept Perform 26: 527-551 (2000)

Carroll JM, Russell JA: Facial expressions in Hollywood's portrayal of emotion. J Pers Soc Psychol 72: 164-176 (1997)

Chan SW, Goodwin GM, Harmer CJ: Highly neurotic never-depressed students have negative biases in information processing. Psychol Med: 1-11 (2007)

Cohn JF: Advances in behavioural science using automated facial image analysis and synthesis. IEEE Transactions on Signal Processing Magazine 6: 128-133 (2010)

Cooper A, Petrides KV: A psychometric analysis of the Trait Emotional Intelligence Questionnaire-Short Form (TEIQue-SF) using Item Response Theory. J Pers Assess 92: 449-457 (2010)

Costa PT, McCrae RR: Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI). Psychological Assessment Resources, Odessa (2002)

Cunningham DW, Kleiner M, Bülthoff HH. Wallraven C: The components of conversational facial expressions. Proceedings of the 1st Symposium on Applied perception in graphics and visualization: 143-150 (2004)

Daily MN, Gottrell GW, Reilly J: California facial expressions, CAFE. Computer Science and Engineering Department, San Diego. Abrufbar: http://www.cs.ucsd.edu/users/gary/CAFE/ [24.08.2009]

Damasio AR: Descartes' Irrtum, Fühlen, Denken und das menschliche Gehirn. München, List (1995)

Dion K, Berscheid E, Walster E: What is beautiful is good. J Pers Soc Psychol 24: 285-290 (1972)

Doi H, Shinohara K: The perceived duration of emotional faces is influenced by the gaze direction. Neurosci Lett 457: 97-100 (2009)

Droit-Volet S, Gil S : The time-emotion paradox. Philos Trans R Soc London 364: 1943-1953 (2009)

Eberle U: Die Weltsprache der 43 Muskeln. Bild der Wissenschaft online, 1: 24 (2004)

Ebner NC: Age of face matters: Age-group differences in ratings of young and old faces. Behav Res Methods 40: 130-136 (2008)

Ebner NC, Johnson MK: Young and older emotional faces: Are there age group differences in the expression identification and memory? Emotion 9: 329-339 (2009)

Ebner NC, Riediger M, Lindenberger U: FACES – A database of facial expressions in young, middle-aged and older women and men: Development and validation. Behav Res Methods 42: 351-362 (2010)

Egan V, Deary I, Austin E: The NEO-FFI: Emerging British norms and an item-level analysis suggest N, A and C are more reliable than O and E. Pers Individ Dif 29: 907-920 (2000)

Ehrenthal JC, Dinger U, Lamla A, Funken B, Schauenburg H: Evaluation der deutschsprachigen Version des Bindungsfragebogens „Experience in Close Relationships – Revised (ECR-RD). Psychother Psychosom Med Psychol 59: 215-223 (2009)

Ekman P: Felt, false, and miserable smiles. J Nonverbal Behav 6: 238-252 (1982)

Ekman P: An argument for basic emotions. Cogn Emot 6: 169-200 (1992)

Ekman P: Facial expression and emotion. Am Psychol 48: 384-392 (1993)

Ekman P: Emotions revealed. Times Books, New York (2003)

Ekman P: Gefühle lesen. Spektrum, München (2007a)

Ekman P: The directed facial action task. In: Coan JA (Hrsg) Handbook of emotion elicitation and assessment. Oxford: Oxford University Press, Oxford, S. 47-53. (2007b)

Ekman P, Brattesani KA, O'Sullivan M, Friesen WV: Does image size affect judgments of the face? J Nonverbal Behav 4: 57-61 (1979)

Ekman P, Cordano D: What is meant by calling emotions basic. Emot Rev 3: 364-370 (2011)

Ekman P, Davidson RJ, Friesen WV: The Duchenne smile: Emotional expression and brain physiology II. J Pers Soc Psychol 58: 342-353 (1990)

Ekman P, Friesen, WV: Pictures of facial affect. Consulting Psychologists Press, Palo Alto (1976)

Ekman P, Friesen WV: Facial Action Coding System. Consulting Psychologists Press, Palo Alto (1978)

Ekman P, Friesen WV: EmFACS. San Francisco, Californien (1982 unveröffentlichtes Dokument)

Ekman P, Friesen WV: Unmasking the face: A guide to recognizing emotions from facial expressions. Malor Books, Cambridge (2003)

Ekman P, Friesen WV, Hager JC: Facial Action Coding System. Manual and Investigator's Guide. Salt Lake City, UT (2002)

Ekman P, Sorenson ER, Friesen WV: Pancultural elements in facial displays of emotions. Sci 164*:* 86-88 (1969)

Elfenbein HA, Beaupré M, Lévesque M, Hess U: Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. Emotion 7: 131-146 (2007)

Freudenthaler HH, Neubauer AC, Gabler P, Scherl WG, Rindermann H: Testing and validating the Trait Emotional Intelligence Questionnaire (TEIQue) in a German-speaking sample. Pers Individ Dif 45: 673-678 (2008)

Frijda NH: The emotions. Cambridge University Press, London (1986)

Frijda NH: The laws of emotion. Am Psychol 43: 349-358 (1988)

Frijda NH, Parrott WG: Basic emotions or Ur-emotions? Emotion Review 3: 406-415 (2011)

Gao W, Cao B, Shan S, Zhou D, Zhang X, Zhao D: The CAS-PEAL Large-Scale Chinese Face Database and evaluation protocols, Technical Report No. JDL_TR_04_FR_001, Joint Research & Development Laboratory (2004)

Gil FP, Ridout N, Kessler H, Neuffer M, Schoechlin C, Traue HC, Nickel M: Facial recognition and alexithymia in adults with somatoform disorders. Depress Anxiety 26: 26-33 (2009)

Goeleven E, De Raedt R, Leyman L, Verschuere B: The Karolinska Directed Emotional Faces: A validation study. Cogn Emot 22: 1094-1118 (2008)

Gosling SD, Rentfrow PJ, Swann WB Jr.: A very brief measure of the Big-Five personality domains. J Res Pers 37: 504-528 (2003)

Gross R: Face databases. In: Li S, Jain A (Hrsg) Handbook of Face Recognition. New York: Springer, New York, S. 301-328 (2005)

Gross JJ, John OP: Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. J Pers Soc Psychol 85: 348-362 (2003)

Gur RC, Schroeder L, Turner T, McGrath C, Chan RM, Turtsky BI, Alsop D, Maldjian J, Gur RE: Brain activation during facial emotion processing. Neuroimage 16: 651-662 (2002)

Hall JA, Carter J, Horgan T: Gender differences in the nonverbal communication of emotion. In: Fisher A (Hrsg) Gender and emotion: Social psychological perspectives. Cambridge University Press, Paris, S. 97-117 (2000)

Harwood NK, Hall LJ, Shinkfield AJ: Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. Am J Ment Retard 104: 270-278 (1999)

Haxby JV, Hoffman EA, Gobbini MI: Human neural systems for face recognition and social communication. Biol Psychiatry 51: 59-67 (2002)

Hernandez J, Morris RR, Picard RW: Call center stress recognition with person-specific models. In: D'Mello S et al. (Hrsg) ACII 2011, Part I, LNCS 6974. Berlin, Heidelberg: Springer, Berlin Heidelberg, S. 125-134 (2011)

Hess U, Senécal S, Kirouac G, Herrera P, Philippot P, Kleck RE:. Emotional expressivity in men and women: Stereotypes and self-perception. Cogn Emot 14*:* 609-642 (2000)

Hill H, Schyns PG, Akamatsu S: Information and viewpoint dependence in face recognition. Cognition 62 201-222. (1997)

Hoffmann H, Kessler H, Eppel T, Rukavina S, Traue HC: Expression intensity, gender and facial emotion recognition: Women recognize only subtle emotions better than men. Acta Psychol 135: 278-283 (2010)

Hoffmann H, Traue HC, Bachmayr F, Kessler H: Perceived realism of dynamic facial expressions of emotion: Optimal durations for the presentation of emotional onsets and offsets. Cogn Emot 24: 1369-1378 (2010)

Hoheisel B, Kryspin-Exner I: Emotionserkennung in Gesichtern und emotionales Gesichtergedächtnis. Zeitschrift für Neuropsychologie 16: 77-87 (2005).

Hoque ME, Picard RW: Acted vs. natural frustration and delight: Many people smile in natural frustration. 9th IEEE International Conference on Automatic Face and Gesture Recognition (2011)

Hyde JS, Mertz JE: Gender, culture, and mathematics performance. Proceedings of the National Academy of Sciences USA 106: 8801-8807 (2009)

Hwang B-W, Byun H, Roh M-C, Lee S-W: Performance evaluation of face recognition algorithms on the asian face database, KFDB. Proceedings of the 4th international conference on Audio- and video-based biometric person authentication (AVBPA'03): 557-565 (2003)

Izard CE: The face of emotion*.* Appleton-Century-Crofts, New York (1971)

Kappas A: What facial activity can and cannot tell us about emotions. In: Katsikitis M (Hrsg) The human face: Measurement and meaning. Kluwer Academic Publishers, Dordrecht, S. 215-234 (2003)

Keltner D, Kring AM: Emotion, social function, and psychopathology. Rev Gen Psychol 2: 320-342 (1998)

Kessler H, Bachmayr F, Walter S, Hoffmann H, Filipic S, Traue HC: Symmetries in a standardized set of emotional facial expressions (JACFEE). Psychosoc Med 5: 1-4 (2008)

Kessler H, Bayerl P, Deighton RM, Traue HC: Facially expressed emotion labeling (FEEL): PC-gestützer Test zur Emotionserkennung. Verhaltenstherapie und Verhaltensmedizin 23: 297-306 (2002)

Kessler H, Hoffmann H, Bayerl P, Neumann H, Basic A, Deighton RM, Traue HC: Die Messung von Emotionserkennung mittels Computer-Morphing: Neue Methoden für Forschung und Klinik. Nervenheilkunde 24: 611-614 (2005)

Kessler H, Roth J, von Wietersheim J, Deighton RM, Traue HC: Emotion recognition patterns in patiens with panic disorder. Depr Anxiety 24: 223-226 (2007)

Kessler H, Schwarze M, Filipic S, Traue HC, von Wietersheim J: Alexithymia and facial emotion recognition in patients with eating disorders. Int J Eat Disord 39: 245-251 (2006)

Kessler M, Retz-Junginger P, Traue HC: Tachistoskopische Analyse emotionaler Sensitivität (THYMOS). Praxis Klinische Verhaltensmedizin und Rehabilitation 10: 73-80 (1997)

Kraus MW, Côté S, Keltner D: Social class, contextualism, and empathic accuracy. Psychol Sci 21: 1716-1723 (2010)

Kraus MW, Keltner D: Signs of Socioeconomic Status: A thin-slicing approach. Psychol Sci 20: 99-106 (2009)

Kraus MW, Pfiff PK, Keltner D: Social class, sense of control, and social explanation. J Pers Soc Psychol 97: 902-1004 (2009)

Kring AM, Gordon AH: Sex differences in emotion: Expression, experience, and physiology. J Pers Soc Psychol 74: 686-703 (1998)

Krohne HW, Hock M: Psychologische Diagnostik. Grundlagen und Anwendungsfelder. Kohlhammer, Stuttgart (2007)

Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics 33: 159-174 (1977)

Lang PJ, Greenwald MK, Bradley MM, Hamm AO: Looking at pictures: Affective, facial, visceral and behavioral reactions. Psychophysiology 30: 261–273 (1993)

Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, Knippenberg A: Presentation and validation of the Radboud Faces Database. Cogn Emot 24: 1277-1388 (2010)

LeDoux JE: Emotion: Clues from the brain. Annu Rev Psychol 46: 209-235 (1995)

LeDoux JE: Das Netz der Gefühle. dtv, München (2003)

Lien JJ-J, Kanade T, Cohn JF, Li C-C: Detection, tracking, and classification of Action Units in facial expression. Rob Auton Syst 31: 131-146 (2000)

Liu CH, Chaudhuri A: Reassessing the ¾ view effect in face recognition. Cognition 83: 31-48 (2002)

Lienert GA, Raatz U: Testaufbau und Testanalyse. Beltz, Weinheim (1994)

Littlewort G, Whitehill J, Wu T, Fasel I, Frank M, Movellan J, Bartlett M: The Computer Expression Recognition Toolbox (CERT). Proceedings IEEE International Conference on Automatic Face and Gesture Recognition: 298-305 (2011)

Martinez AM, Benavente R: The AR-face database. Computer Vision Center, Technical Report 24 (1998)

Matsumoto D: American-Japanese cultural differences in the recognition of universal facial expressions. J Cross Cult Psychol 23: 72-84 (1992)

Matsumoto D (2002). Methodological requirements to test a possible in-group advantage in judging emotions across cultures: Comments on Elfenbein and Ambady (2002) and evidence. Psychol Bull 128: 236-242 (2002)

Matsumoto D, Ekman P: Japanese and Caucasian facial expressions of emotion (JACFEE) and neutral faces (JACNeuF) (1988) Dia-Abbildungen (erhältlich bei Dr. Paul Ekman, Department of Psychiatry, University of California, San Francisco, 401 Parnassus, CA 94143-0984)

Matsumoto D, Ekman P: American-Japanese cultural differences in judgements of facial expressions of emotion. Motiv Emot 13: 143-157 (1989)

McClure EB: A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. Psychol Bull 126: 424-453 (2000)

McIntosh DN; Facial Feedback Hypotheses: Evidence, implications, and directions. Motiv Emot 20: 121-147 (1996)

Merten J: Einführung in die Emotionspsychologie. Kohlhammer, Stuttgart (2003)

Meyer W-U, Reisenzein R, Schützwohl A: Einführung in die Emotionspsychologie, Band I: Die Emotionstheorien von Watson, James und Schachter. Hans Huber, Bern u a (2003)

Meyer W-U, Schützwohl A, Reisenzein R: Einführung in die Emotionspsychologie, Band II: Evolutionspsychologische Emotionstheorien. Hans Huber, Bern u a (2003)

Mogg K, Garner M, Bradley BP: Anxiety and orienting of gaze to angry and fearful faces. Biol Psychol 76: 163-169 (2007)

Morris JS, Öhman A, Dolan RJ: Conscious and unconscious learning in the human amygdala. Nature 393: 467-470 (1998)

Murphy FC, Nimmo-Smith I, Lawrence AD: Functional neuroanatomy of emotions: A meta-analysis. Affect Behav Neurosci 3: 207-233 (2003)

Oatley K, Johnson-Laird PN: Towards a cognitive theory of emotions. Cogn Emot 1: 37-41 (1987)

O'Doherty J, Winston J, Critchley HD, Perrett D, Burt DM, Dolan RJ: Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. Neuropsychologia 41: 147-55 (2003)

Ortony A, Turner TJ: What's basic about basic emotions? Psychol Rev 97: 315-331 (1990)

Otto JH, Euler HA, Mandl H: Emotionspsychologie – ein Handbuch. Beltz, Weinheim (2000)

O'Toole AJ, Edelman S, Bülthoff HH: Stimulus-specific effects in face recognition over changes in viewpoint. Vision Res 38*:* 2351-2363 (1998)

Palermo R, Coltheart M: Photographs of facial expression: accuracy, response times, and ratings of intensity. Behav Res Methods Instrum Comput 36: 634-638 (2004)

Panksepp J: Toward a general psychobiological theory of emotions. Behav Brain Sci 5: 407-467 (1982)

Pearson R, Lewis MB: Fear recognition across the menstrual cycle. Horm Behav 47: 267-271 (2005)

Petrides KV, Furnham A: Trait emotional intelligence: Behavioural validation in two studies of emotion recognition and reactivity to mood induction. Eur J Pers 17: 39-57 (2003)

Pilz K, Breidt M, Thornton IM, Bülthoff HH: Video capture of non-rigidly moving faces out of five different viewpoints (2002)

Plutchik R: Emotion: A psychoevolutionary synthesis. Harper & Row, New York (1980)

Pollak SD, Messner M, Kistler DJ, Cohn JF: Development of perceptual expertise in emotion recognition. Cognition 110: 242-247 (2009)

Rhodes G, Brake S, Atkinson AP: What's lost in inverted faces? Cognition 47: 25-57 (1993)

Roseman IJ: Emotional behaviors, emotivational goals, emotion strategies: Multiple levels of organization integrate variable and consistent responses. Emot Rev 3: 434-443 (2011)

Rosenberg EL: Emotions as unified responses. In: Ekman P, Rosenberg EL (Hrsg) What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press, Oxford, S. 86-88 (1997)

Ross MD, Owren MJ, Zimmermann E: Reconstructing the evolution of laughter in Great Apes and Humans. Curr Biol 19: 1-6.

Rubin RS, Munz DC, Bommer WH: Leading from within: The effects of emotion recognition and personality on transformational leadership behavior. Acad Manage J 48: 845-858 (2005)

Russell JA: Culture and the categorization of emotions. Psychol Bull 110: 426-450 (1991)

Russell JA: Forced-choice response format in the study of facial expression. Mot Emot 17: 41–51 (1993)

Russell JA: Is there universal recognition of emotion from facial expression? A review of cross-cultural studies. Psychol Bull 115: 102-141 (1994)

Sachs G, Steger-Wuchse D, Kryspin-Exner I, Gur RC, Katschnig H: Facial recognition deficits and cognition in schizophrenia. Schizophr Res 68: 27-35 (2004)

Sackheim H, Gur RC, Saucy M: Emotions are expressed more intensely on the left side of the face. Science 202: 434-435 (1978)

Samuels CA, Butterworth G, Roberts T, Graupner L, Hole G: Facial aesthetics: Babies prefer attractiveness to symmetry. Perception 23: 823 – 831 (1994)

Sanders G, Sjodin M, Chastelaine M: On the elusive nature of sex differences in cognition: Hormonal influences contributing to within sex variation. Arch Sex Behav 31: 145-152 (2002)

Scherer K: What are emotions? And how can they be measured? Soc Sci Inf 44: 693-727 (2005)

Schnell R, Hill PB, Esser E: Methoden der empirischen Sozialforschung. Oldenbourg Wissenschaftsverlag, München (2005)

Schünke M, Schulte E, Schumacher U, Voll M, Wesker K: Prometheus, Lern-Atlas der Anatomie. Kopf und Neuroanatomie. Thieme, Stuttgart (2006)

Schwarz-Friesel M: Sprache und Emotion. A. Francke, Tübingen Basel (2007)

Siegrist J: Medizinische Soziologie. Urban & Fischer, München Jena (2005)

Sprengelmeyer R, Rausch M, Eysel UT, Przuntek H: Neural structures associated with recognition of facial expressions of basic emotions. Proc. R. Soc. Lond. B 265: 1927-1931 (1998)

Steele H, Steele M, Croft C: Early attachment predicts emotion recognition at 6 and 11 years old. Attach Hum Dev 10: 379-93 (2008)

Stirrat M, Gumert M, Perrett D: The effect of attractiveness on food sharing preferences in human mating markets. Evol Psychol 9: 79-91 (2011)

Sullivan S, Ruffman T, Hutton SB: Age differences in emotion recognition skills and the visual scanning of emotion faces. J Ger 62: 53-60 (2007)

Tomkins SS: Affect Theory. In: Scherer KR, Ekman P (Hrsg) Approaches to emotion. Erlbaum, Hillsdale NJ, S. 163-195 (1984)

Tracy JL, Beall AT: Happy guys finish last: The impact of emotion expressions on sexual attraction. Emotion 11: 1379-1387 (2011)

Tracy JL, Robins RW, Schriber RA: Development of a FACS-verified set of basic and self-conscious emotion expressions. Emotion 9: 554-559 (2009)

Traue HC: Cerebrale Laterlaität, emotionale Prozesse und Krankheit. In: Ehlers U, Traue HC, Czogalik D (hrsg) Bio-psycho-soziale Grundlagen für die Medizin. Berlin u.a: Springer, S. 67-91 (1988)

Traue HC: Emotion und Gesundheit – die psychobiologische Regulation durch Hemmungen. Spektrum, Heidelberg Berlin (1998)

Traue HC, Kessler H: Psychologische Emotionskonzepte. In: Walter SA (Hrsg) Natur und Theorie der Emotion. Paderborn: Mentis, Paderborn, S. 20-33 (2003)

Wagner HL: Methods for the study of facial behaviour. In: Russell JA, Fernandez-Dols JM (Hrsg) The Psychology of Facial Expression. Cambridge University Press, Cambridge, S. 41-54 (1997)

Wagner HL: Methods for the study of facial behaviour. In: Russell JA, Fernandez-Dols JM (Hrsg) The Psychology of Facial Expression. Cambridge University Press, Cambridge, S. 271-313 (1997)

Wagner HL: The accessibility of the term „contempt“ and the meaning of the unilateral lip curl. Cogn Emot 63: 271-313 (2000)

Walter S, Scherer S, Schels M, Glodek M, Hrabal D, Schmidt M, Böck R, Limbrecht K, Traue HC, Schwenker F: (2011). Multimodal emotion classification in naturalistic user behavior. In Proceedings of 14th International Conference on Human-Computer Interaction: 3, 603 – 611 (2011)

Welling LLM, DeBruine LM Little AC, Jones BC: Extraversion predicts individual differences in women’s face preferences. Pers Individ Dif 47: 996-998 (2009)

Wierzbicka A: Emotions across languages and cultures – diversity and universals. Cambridge University Press, Paris (2005)

Winkler J, Stolzenberg H: Adjustierung des Sozialen-Schicht-Index für die Anwendung im Kinder- und Jugendgesundheitssurvey (KiGGS) 2003/2006. Wismarer Diskussionspapiere, 7 (2009)

Wundt W: Grundriss der Psychologie. Wilhelm Engelmann, Leipzig (1905)

Yin L, Wei X, Sun Y, Wang J, Rosato MJ: A 3D Facial Expression Database For Facial Behavior Research. IEEE International Conference on Automatic Face and Gesture Recognition: 211-216 (2006)

Yin L, Chen X, Sun Y, Worm T, Reale M: A high-resolution 3D dynamic facial expression database. IEEE International Conference on Automatic Face and Gesture Recognition: 1-6 (2008)

Zaidel DW, Aarde SM, Baig K: Appearance of symmetry, beauty, and health in human faces. Brain Cogn 57: 261-263 (2005)

TABELLENVERZEICHNIS

		Seite
Tabelle 1	Auszug der von verschiedenen Forschern postulierten Basisemotionen (nach Ortony & Turner, 1990)	8
Tabelle 2	Bildersets und Datenbanken anderer Forschergruppen	22
Tabelle 3	Zuordnung der Bewertung zu den einzelnen Emotionskategorien (514 Bilder, in Prozent)	44
Tabelle 4	Zuordnung der Bewertung zu den einzelnen Emotionskategorien (96 Bilder, in Prozent)	45
Tabelle 5	Übersicht über die Reaktionszeiten für die einzelnen Emotionen	47
Tabelle 6	Ergebnisse der Paarvergleiche für die Reaktionszeiten bei der Auswahl der einzelnen Emotionen	47
Tabelle 7	Korrelation zwischen Reaktionszeit und Prozentsatz richtig erkannter Emotionen	47
Tabelle 8	Zuordnung der 245 verbalen Ausdrücke zu den acht auswählbaren Kategorien	53
Tabelle 9	Erkennungsraten und Verwechslungen der emotionalen Stimuli nach Zuordnung durch die Rater (in Prozent)	55
Tabelle 10	Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit	61
Tabelle 11	Ergebnisse der Paarvergleiche für Intensität, Attraktivität und Natürlichkeit	62
Tabelle 12	Korrelationen der Faktoren Intensität und Attraktivität auf Emotionsebene	62
Tabelle 13	Korrelationen der Faktoren Intensität und Natürlichkeit auf Emotionsebene	62
Tabelle 14	Korrelationen der Faktoren Natürlichkeit und Attraktivität auf Emotionsebene	63
Tabelle 15	Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit getrennt nach dem Geschlecht der Darsteller	63
Tabelle 16	Mittelwerte der Faktoren für Intensität, Attraktivität und Natürlichkeit getrennt nach dem Geschlecht der Beurteiler	64
Tabelle 17	Mittelwerte, Standardabweichung, Minimum und Maximum der Fragebögen ERQ, NEO-FFI und TeiQue	65
Tabelle 18	Verteilung von Schulabschluss, Einkommen und Kindern in der Gesamtstichprobe	73
Tabelle 19	Zuordnung zu den einzelnen Emotionskategorien (96 Bilder, in Prozent)	74
Tabelle 20	Zuordnung zu den einzelnen Emotionskategorien (48 Bilder, in Prozent)	74
Tabelle 21	Verteilung der Persönlichkeitskonstrukte in der Normierungsstichprobe	76
Tabelle 22	Itemschwierigkeit (mittlere Erkennungsleistung in Prozent) und Standardabweichung aller 48 Stimuli	79

Tabelle 23	Trennschärfen aller 48 Stimuli	79
Tabelle 24	Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit	85
Tabelle 25	Ergebnisse der Paarvergleiche für Intensität, Attraktivität und Natürlichkeit	86
Tabelle 26	Korrelationen der Faktoren Intensität und Attraktivität auf Emotionsebene	87
Tabelle 27	Korrelationen der Faktoren Intensität und Natürlichkeit auf Emotionsebene	87
Tabelle 28	Korrelationen der Faktoren Natürlichkeit und Attraktivität auf Emotionsebene	87
Tabelle 29	Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit getrennt nach dem Geschlecht der Darsteller	88
Tabelle 30	Mittelwerte der Faktoren Intensität, Attraktivität und Natürlichkeit getrennt nach dem Geschlecht der Beurteiler	89
Tabelle 31	Konsistenzmaß für die drei Skalen	90
Tabelle 32	Prozentuale Erkennungsrate für die sechs Emotionen zu zwei Messzeitpunkten	95
Tabelle 33	Erkennungsleistung aufgeschlüsselt nach Emotion und Perspektive Studie 6	103
Tabelle 34	Übersicht signifikante Paarvergleiche Studie 6	105
Tabelle 35	Subjektive Sicherheit aufgeschlüsselt nach Emotion und Perspektive Studie 6	106
Tabelle 36	Korrelation Erkennungsleistung mit subjektiver Sicherheit Studie 6	106
Tabelle 37	Mittelwerte, Standardabweichung, Minimum und Maximum von NEO-FFI und TeiQue Studie 6	107
Tabelle 38	Übersicht Bindungsstile Studie 6	107
Tabelle 39	Erkennungsleistung aufgeschlüsselt nach Emotion und Perspektive Studie 7	112
Tabelle 40	Übersicht signifikante Paarvergleiche Studie 7	113
Tabelle 41	Subjektive Sicherheit aufgeschlüsselt nach Emotion und Perspektive Studie 7	114
Tabelle 42	Korrelation Erkennungsleistung mit subjektiver Sicherheit Studie 7	114
Tabelle 43	Mittelwerte, Standardabweichung, Minimum und Maximum von NEO-FFI und TeiQue Studie 7	115
Tabelle 44	Übersicht Bindungsstile Studie 7	115

ABBILDUNGSVERZEICHNIS

		Seite
Abbildung 1	Die sechs Basisemotionen nach Ekman (aus dem JACFEE-Bildersatz: Matsumoto & Ekman, 1988)	7
Abbildung 2	Übersicht Hybrid-Modell von Roseman (2011, S. 438)	10
Abbildung 3	Übersicht über die für die Emotionsverarbeitung relevanten Hirnareale (Copyright: www.julius-ecke.de)	12
Abbildung 4	Gesichtsmuskulatur (aus Prometheus, Lern-Atlas der Anatomie, Kopf und Neuroanatomie, S. 46)	28
Abbildung 5	Schematische Darstellung des Ablaufs des FEEL-Tests	32
Abbildung 6	Beispiel für die Anweisung nach EmFACS für einen spezifischen emotionalen Ausdruck (hier Angst)	36
Abbildung 7	Beispiel für eine emotionsinduzierende Geschichte (hier Ekel)	37
Abbildung 8	Setups des Fotostudios in grafischer Darstellung	38
Abbildung 9	Setup des Fotostudios aus unterschiedlichen Perspektiven	39
Abbildung 10	Überblick über das Vorgehen bei der Untersuchung des Bildmaterials	41
Abbildung 11	Instruktion für die Bewertung des Bildmaterials in EmoIR	43
Abbildung 12	Versuchsablauf Ratingprozedere	44
Abbildung 13	Erfassung der Einschätzung emotionaler Gesichtsausdrücke im offenen Antwortformat	52
Abbildung 14	Beispiel für die experimentelle Umsetzung der Bewertung von Natürlichkeit, Intensität und Attraktivität	59
Abbildung 15	Beurteilung von Intensität, Attraktivität und Natürlichkeit im Vergleich	86
Abbildung 16	Übersicht über die fünf fotografierten Perspektiven am Beispiel eines neutralen Gesichtsausdrucks	98
Abbildung 17	Veranschaulichung des neuen Antwortformats	102

ANHANG **CD-ROM**

A Trainingsmanual

B Ratingmanual

C Übersicht Emotionswörter aus der freien Beurteilung

D TeiQue-SF

E ERQ

F NEO-FFI

G Ten Item Personality Inventory

H Übersicht Auswahl endgültiger Bildersatz

I Relationship-Questionnaire

DANKSAGUNG

„Dies war für mich ein denkwürdiger Tag, da er gewaltige Veränderungen in mir bewirkte. Doch das gibt es in jedem Leben. Man stelle sich vor, ein ganz bestimmter Tag würde gelöscht, und überlege dann, wie anders dieses Leben verlaufen wäre. Du, der dies liest, halt ein und denke für einen Augenblick an die lange Kette aus Eisen und Gold, aus Dornen und Blumen, die dich niemals gefesselt hätte, wäre nicht an einem denkwürdigen Tage ihr erstes Glied geschmiedet worden." (Charles Dickens, Große Erwartungen)

Emotionen zu empfinden und auszudrücken sind Grundpfeiler des menschlichen Lebens. Nichts desto trotz stellen sie uns vor viele ungelöste Geheimnisse. Was sind Emotionen ganz genau? Wie entstehen sie? Wie nimmt man sie wahr? Bis all diese Fragen befriedigend beantwortet werden können, ist noch ein weiter Weg zurückzulegen. Die hier vorliegende Arbeit soll hierzu einen Beitrag leisten.
Ich danke für die Begleitung und Unterstützung: Hauke, meinen Eltern, meinem Bruder und Kristin, sowie der ganzen Familie. Außerdem danke ich insbesondere Harald, Lucia und allen aus der Sektion Medizinische Psychologie, Herrn Muche und Kornelia Blohm sowie all jenen, die mich auf diesem Weg begleitet haben. Die fachliche und emotionale Unterstützung lassen mich dankbar sein.

Ich habe in dieser Dissertation stets die männliche Form der relevanten Substantive verwendet. Dies dient der besseren Lesbarkeit, keineswegs der Abwertung der Leserinnen.

Die Dissertation wurde im Rahmen des Sonderforschungsbereichs/Transregio 62 „Eine *Companion*-Technologie für kognitive technische Systeme" angefertigt, welcher von der Deutschen Forschungsgesellschaft (DFG) finanziert wird. Bei diesem Forschungsvorhaben handelt es sich um eine Zusammenarbeit der Universität Ulm, der Otto-von-Guericke-Universität Magdeburg und dem Leibnitz-Institut für Neurobiologie in Magdeburg. Die vorliegende Arbeit ist dem Teilprojekt C4 „Psychobiologische Emotionserkennung" zuzuordnen.

Lebenslauf wurde aus Gründen aus des Datenschutzes entfernt.