



If he's happy, do you know it?

Analyse des neuen Test of Perceptual Competence of Facial Affective Expression (PCFAE)

Mark Peter Richter

Klinisches Institut für Psychosomatische Medizin und Psychotherapie,

Universitätsklinikum der Heinrich-Heine-Universität, Düsseldorf.

Anschlussarbeit des Studiengangs Psychologie (MSc.)

Wintersemester 2018/19

Gutachter: Dr. Ralf Schäfer und Prof. Dr. Jochen Musch

Zusammenfassung

Die Erkennung affektiver Gesichtsausdrücke ist von grundlegender Bedeutung für das menschliche Zusammenleben. Einen besonderen Stellenwert nehmen die Basisaffekte Angst, Ekel, Freude, Trauer, Überraschung und Wut ein, welche nicht erlernt werden müssen, über alle Kulturen hinweg gleichermaßen verstanden werden und somit eine universelle nonverbale Kommunikation affektiver Zustände ermöglichen. Im diagnostischen Kontext ist das Messen der Wahrnehmungsleistung für diese Basisaffekte in der qualitativen und zeitlichen Domäne von zunehmendem Interesse. Die vorliegende Studie analysierte das computerbasierte Instrument zur Erfassung der Affektwahrnehmungskompetenz, den Test of Perceptual Competence of Facial Affective Expression (PCFAE), mittels einer Stichprobe von $N = 154$ Versuchspersonen hinsichtlich seiner Itemgüte, Reliabilität, zeitlichen Stabilität nach erneuter Testung (4 Wochen) und explorierte dessen Dimensionalität. Das leichte bis mittelschwere Testverfahren erweist sich trotz aufgetretener systematischer Verwechslungen der Affekte unter Verwendung der Reaktionszeit als Leistungsmaß trennscharf, hinreichend reliabel und zeitlich stabil. Ein Korrelierte Faktoren Modell und ein Bi-Faktor Modell werden als mögliche Teststruktur vorgeschlagen. Item- und durchführungsbezogene Optimierungsmöglichkeiten sowie Beiträge zur Konstruktvalidität werden diskutiert.

Schlüsselwörter: PCFAE, Affekterkennung, Reliabilität, Retest, Itemanalyse

Einleitung

Die Auseinandersetzung des Menschen mit dem Wesen seiner Emotionen begleitet die menschliche Geschichte. Bereits vor 2000 Jahren befasste sich der Philosoph Epikur mit dem affektiven Erleben und postulierte die menschliche Motivation, hedonische Affekte zu suchen und aversive zu vermeiden (Rudolph, 2013). Während die altertümlichen Stoiker Epikurs Lustprinzip und die damit verbundenen Affekte noch als „vernunftlose, naturwidrige Gemütsbewegungen“ kritisierten (Apel & Ludz, 1976, S.7), erkannte Darwin (1872/1965) deren evolutionäre Bedeutung: Affekte gewährleisten eine automatisierte, schnelle erste Beurteilung situativer Kontexte und erhöhen so die Überlebens- und Fortpflanzungswahrscheinlichkeit des Individuums. Die Fähigkeit zur mimischen Kommunikation des Affekts überträgt diesen Vorteil auf die Gruppe.

Die moderne Psychologie widmet sich ebenfalls der Erklärung der menschlichen Emotionen und liefert eine Vielzahl verschiedener theoretischer Ansätze und Klassifikationsversuche. Diese können beispielsweise behavioristisch (Watson, 1919), evolutionspsychologisch (Darwin 1872/1965; Plutchik, 1980), kognitiv (Arnold, 1960; James, 1884; Lazarus, 1966; Mandler, 1984; Schachter, 1964) oder attributional (Weiner, 1986) begründet sein. Einen weiteren möglichen Zugang stellt die inhaltliche Annäherung über die Klassifikation sogenannter Basisaffekte (Ekman & Friesen, 1970; hier als Basisemotionen bezeichnet) dar. Während die frühe Forschung vor allem darauf ausgerichtet war, das Wesen und die Funktion der Emotionen bzw. Affekte zu untersuchen, zielt die aktuelle psychologische Forschung auch auf die Erfassung der Leistungskomponente der Affektwahrnehmung sowie inter- und intraindividuellen Unterschiede.

Das im Rahmen dieser Arbeit analysierte Instrument zur Erfassung der Affektwahrnehmungskompetenz stellt den Versuch dar, mittels eines computerbasierten Tests die Leistung der menschlichen Affekterkennungsfähigkeit für Basisaffekte zu messen, diese

anhand entsprechender Parameter zu operationalisieren und in der Folge für diagnostische Zwecke vergleichbar machen zu können. Konstituierend für das zugrundeliegende Paradigma ist die angenommene Fähigkeit zum Rückschluss vom mimischen Ausdruck auf den zugrundeliegenden Basisaffekt sowie dessen Klassifizierung nach dem Verständnis der Forschergruppe um Paul Ekman. Dies soll in den folgenden Kapiteln dargestellt werden.

Affekte und Affektwahrnehmung

Arbeitsbegriffsdefinition und Abgrenzung. Trotz der langen Wissenschaftshistorie existiert bis zum heutigen Tag kein ganzheitliches Modell oder eine einheitliche Definition der Emotion. Der Begriff Emotion wird in der Regel als Oberbegriff für verschiedene kognitive, neurophysiologische, motivationale, motorische und gefühlserlebende Prozesse verwendet, die im Zusammenhang mit dem Phänomen Emotion auftreten (Scherer, 1990). Emotionen lassen sich als psychophysische Reaktionsmuster beschreiben, die auf der Bewertung einer Reizsituation beruhen, einem Lebewesen Ziele und Bedürfnisse signalisieren, eine physiologische Zustandsänderung und Verhaltenstendenz aufweisen und mit einem subjektiven Gefühlserleben einhergehen (Goschke & Dreisbach, 2010).

Hinsichtlich der zeitlichen Ausdehnung können Emotionen in Stimmung und Affekt differenziert werden. Die Stimmung wird als Gefühlskonsequenz von Ereignissen oder Lebensumständen verstanden, welche mit vergleichsweise geringer Intensität und weniger bewusst erlebt wird (Wirth, 2013), dabei nicht in direktem Bezug zu konkreten Ursachen steht und über einen längeren Zeitraum aufrechterhalten wird (Scherer, 2005). Es fehlt eine starke physiologische Aktivierung, sodass eher von einer emotionalen „Grundtönung“ (Sachse & Langens, 2014, S. 17) gesprochen werden kann. Demgegenüber steht der Affekt für ein kurzes aber intensives emotionales Erleben mit deutlicher Verhaltenstendenz, vor allem hinsichtlich des mimischen Ausdrucks (Reischies, 2007). Affekte können folglich als

phasische Zustände verstanden werden, die eine Verlaufsform mit Einsetzen, endlicher Dauer und Aussetzen beinhalten (Adolphs, 2002).

Basisaffekte. Obwohl bereits Darwin (1872/1965) von einer genetischen Verankerung emotionalen Ausdrucksverhaltens ausging, nahm zu Beginn des 20. Jahrhunderts ein wesentlicher Teil der Psychologen an, affektive Gesichtsausdrücke wären durch soziale Einflüsse erlernt und deren Ausprägung unterläge kulturellen Unterschieden. Weiterhin bestünde keine feste Beziehung zwischen dem fazial-motorischen Ausdruck selbst und dem Affekt, der ausgedrückt werden soll (Ekman, 1992). Ekman, Sorenson & Friesen (1969) sowie Izard (1971) konnten jedoch zeigen, dass Fotos affektiver Gesichtsausdrücke über eine Vielzahl von Kulturen hinweg konsistent den gleichen Affekten zugeordnet werden¹. Dies traf ebenfalls für vollständig isoliert lebende Kulturen zu: Obwohl für sie keine Möglichkeit bestanden hatte, affektive Gesichtsausdrücke anderer Kulturen zu erlernen, erkannten sie diese umgehend korrekt (Ekman, 1972, 1988; Ekman et al., 1987; Ekman & Friesen, 1971).

Ekman (1982; Ekman & Cordaro, 2011) schlussfolgerte daraus, dass eine begrenzte Anzahl an Affekten, die sich durch ein spezifisches Gefühlserleben, eine spezifische physiologische Reaktion und ein spezifisch mimisches Ausdrucksverhalten charakterisieren, bereits angeboren² und damit evolutionär bedingt vorliegen müssen. Diese sogenannten Basisaffekte können als lern- und kulturunabhängige affektive Grundausrüstung des Menschen verstanden werden. Sie unterscheiden sich gemäß Ekman & Davidson (1994) durch eine diskrete Ausprägung in mimischen, sprachlichen und autonom physiologischen Abläufen grundlegend voneinander. Ein umfassender Klassifizierungsversuch nach 11 bzw. 13 Kriterien, die für einen Basisaffekt überwiegend erfüllt sein müssen, findet sich bei

¹ Weitere Evidenz: Eine kulturell universelle Erkennung von Basisaffekten ist ebenfalls durch Musik vermittelbar (Fritz et al., 2009).

² Experimenteller Beleg: Babys zeigen den Affekt Freude durch Lächeln, wenn sie Zuckerlösung verabreicht bekommen und Ekel bei faulen Eiern (Steiner, 1979).

Ekman (2003) sowie Ekman & Cordaro (2011). Diese Kriterien sind durch die 6 Affekte *Angst, Ekel, Freude, Trauer, Überraschung* und *Wut* erfüllt (siehe ebenfalls Ekman, Friesen & Ellsworth, 1982; im weiteren Verlauf als die 6 Basisaffekte bezeichnet). Die experimentell psychologischen Befunde zur Diskriminierbarkeit der Basisaffekte werden durch spezifische psychophysiologische Korrelate flankiert (Kragel & Bar, 2013; Lench, Flore & Bench, 2011; Stephens, Christie & Friedman, 2010; Vytal & Hamann, 2010). Ortony & Turner (1990)¹ erfassen darüber hinaus in ihrer Metastudie wesentliche Gemeinsamkeiten der 6 Basisaffekte über die unterschiedlichen Emotionstheorien hinweg.

Die von Ekman & Davidson (1994) postulierte vollständige Unabhängigkeit der Basisaffekte – vor allem auf mimischer und interpretatorischer Ebene – ist jedoch nicht abschließend belegt (Wegrzyn et al., 2017). *Verachtung* wurde von Ekman & Cordaro (2011) als weiterer Kriterien erfüllender Basisaffekt vorgeschlagen, dies wurde jedoch aufgrund geringer universeller Erkennungsraten (Elfenbein & Ambady, 2002) kritisiert. Verschiedene neuere Publikationen führen Verachtung jedoch als siebten Basisaffekt mit an (Adolphs, 2002). Die Autoren des PCFAE Tests orientierten sich bei der Erstellung an der empirischen Befundlage Ekmans, sodass die im Test gezeigten Stimuli sich auf die 6 genannten Basisaffekte *Angst, Ekel, Freude, Trauer, Überraschung* und *Wut* beschränken (Ingenhag, Schäfer & Franz, 2007; unveröffentlicht).

Funktion der Basisaffekte. Darwin (1872/1965) beurteilte die Funktion der Basisaffekte vor allem als eine Steigerung der Überlebenswahrscheinlichkeit, vermittelt durch eine organismische und eine sozial kommunikative Komponente (ebenfalls Buss, 1999; Ekman, 1972). Beispielsweise erlauben weit geöffnete Augen bei Angst eine effektive visuelle Wahrnehmung der potenziell bedrohlichen Situation, während ein geöffneter Mund leises Atmen und somit eine verringerte Entdeckungswahrscheinlichkeit ermöglicht (Seidler,

¹ Verschiedene Theorien diskutieren das Vorliegen von 3 (Gray, 1982) bis 11 (Arnold, 1960) Basisaffekten. Die genannte Metaanalyse bietet einen Überblick über die verschiedenen Ansätze und betont vor allem die Gemeinsamkeiten.

2008). Diese Signale werden von anderen Personen ohne Notwendigkeit verbaler Kommunikation wahrgenommen und ermöglichen ein ebenfalls angepasstes Verhalten. Das Affektsystem ermöglicht folglich eine schnell verfügbare erste Beurteilung der Situation und weitet diese durch universell verständliche mimische Kommunikation des Affekts auf die Gruppe aus (Parkinson, 1996). Dies ermöglicht, Ereignisse zügig in ihrer Bedeutsamkeit zu bewerten und zu kommunizieren, eine angemessene Reaktion vorzubereiten (Sachse & Langens, 2014) und angepasste Verhaltensweisen für zukünftige gleichartige Situationen zu erlernen (Ekman & Cordaro, 2011). Die sozial kommunikative Komponente tritt in Abwesenheit Anderer in den Hintergrund. Der Affekt wird in der Regel nur dann mimisch ausgedrückt, wenn sich Artgenossen in der Nähe befinden (Fridlund, 2014).

Affektwahrnehmung. Die Bedeutung des Gesichtsausdrucks für die Wahrnehmung des zugrundeliegenden Affekts stellt den Konsens¹ in der Affektforschung dar. Gesichtsausdruck und Affekt werden als zusammengehörig (Ekman & Friesen, 1972; Lazarus, 1991), gemeinsam auftretend (Ekman, 1973), sich wechselseitig beeinflussend (James, 1884; experimenteller Beleg: Strack, Martin & Stepper, 1988) und als entscheidend zur Erkennung und Differenzierung von Affekten (Darwin, 1872/1965; Ekman & Friesen, 1975; Tomkins, 1962) beschrieben. Neuere experimentelle Befunde betonen bei der Affekterkennung und -differenzierung vor allem den Fokus auf Augen und Mund (Adolphs et al., 2005; Guo, 2012; Hager, Ekman & Friesen, 2002; Smith et al., 2005). Jeder Basisaffekt zeichnet sich dabei durch eine spezifische Aktivierung verschiedener Gesichtspartien aus (Wegrzyn et al., 2017), eine vollständige Unabhängigkeit wird jedoch nicht erreicht. Systematische Verwechslungen treten insbesondere zwischen Wut und Ekel (Pochedly, Widen & Russell, 2012) sowie Angst und Überraschung (Palermo, Bonnar & Gosselin, 2002) auf. Demgegenüber wird Freude am besten erkannt und am wenigsten verwechselt (Hahn,

¹ McDougall (1908/1960) sieht hingegen kaum eine Bedeutung des mimischen Gesichtsausdrucks für den Rückschluss auf den zugrundeliegenden Affekt.

2016; Leppänen & Hietanen, 2004; Seidler, 2008). Diese Befunde stehen im Einklang mit der evolutionspsychologischen Perspektive, denn die bedeutsamsten Affekte mit potenziell überlebenswichtiger Signalwirkung sind gut zu diskriminieren (Wut/Ekel signalisieren Bedrohung, Reaktion: Abwehr; Angst/Überraschung signalisieren mögliche Gefahr, Reaktion: Flucht; Freude signalisiert keine Bedrohung, Reaktion: Entwarnung; Wegryn et al., 2017).

Neurophysiologische Grundlagen der Affektwahrnehmung. Übersichtsarbeiten zu neurofunktionalen Grundlagen der Affekterkennung liefern Adolphs (2002), Fusar-Poli et al. (2009) und Phan et al. (2002). Im thematischen Zusammenhang mit dieser Arbeit steht die Vorhersage der Affektwahrnehmungskompetenz mittels der N170 Komponente im Elektroenzephalogramm (EEG) bei Csukly et al. (2014) im Sinne eines alternativen Messansatzes. Starke Hinweise auf die tiefe neuronale Verankerung und grundlegende biologische Bedeutsamkeit der Affektwahrnehmung verdeutlichen de Gelder et al. (1999) sowie Morris et al. (2001) und zeigen, dass selbst Cortexblinde trotz des Fehlens bewusster visueller Wahrnehmung noch in der Lage sind, über die amygdale Verarbeitungsrouten affektive Gesichtsausdrücke zu unterscheiden. Für eine Studie unter kombiniertem Einsatz von EEG und Elektromyografie (EMG) zur Affektwahrnehmung unter Verwendung des PCFAE Paradigmas siehe Jansen (2018a; unveröffentlicht).

Einschränkungen der Affektwahrnehmungskompetenz. Eine Vielzahl psychischer Erkrankungen geht mit einer Einschränkung der Affektwahrnehmungskompetenz einher. Einschränkungen wurden experimentell belegt für die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (Beyer von Morgenstern, Becker & Sinzig, 2013; Jusyte, Gulewitsch & Schönenberg, 2017: Ausschließlich eingeschränkte Genauigkeit der Affekterkennung), Alexithymie (Parker, Taylor & Bagby, 1993), Anorexia Nervosa (Sfärlea et al., 2018), Antisoziales Verhalten (Cigna, Guay & Renaud, 2017), Autismus Spektrum Störung (Lozier,

Vanmeter & Marsh, 2014), Bipolare Persönlichkeitsstörung (Demirel et al., 2014), Borderline Persönlichkeitsstörung (Aebi, 2012: Keine Einschränkung der Affekterkennungsleistung, jedoch verlangsamte Reaktionszeit), Depression (Suslow, Junghans & Arolt, 2001; Yalcin-Siedentopf et al., 2014), Gewalttäterschaft (Chapman, Gillespie & Mitchell, 2018), Körperschemastörung (Toh, Castle & Rossell, 2015), Multiple Sklerose (Donell, 2015; Prochnow et al., 2011), Psychose (Addington & Barbato, 2015), Psychopathie (Boll & Gamer, 2016; Dawel et al., 2012), Posttraumatische Belastungsstörung (Javdani et al., 2017), Schizophrenie (Csukly et al., 2014; Drusch et al., 2014: Fähigkeitsverbesserung durch Training; ebenfalls bei Popova et al., 2014), Schizotypische Persönlichkeitsstörung (Canli, Özdemir & Kocak, 2015) und bei traumatischer Hirnverletzung (Genova et al., 2017; Rigon et al., 2017). Hinweise auf eine mögliche Verstärkung bereits bestehender Einschränkungen sind bei Kokainmissbrauch gegeben (Bedi et al., 2016).

Zusammenfassend beinhalten Affekte eine situativ evaluierende und eine sozial kommunikative Komponente. Über eine kurze Dauer findet ein intensives emotionales Erleben, begleitet durch eine physiologische Reaktion mit starker (mimischer) Verhaltenstendenz statt. Der affektive Gesichtsausdruck spielt eine entscheidende Rolle bei der Kommunikation und Erkennung des Affekts. Das spezifische Zusammenspiel der Gesichtsmuskulatur differenziert zumindest 6 empirisch gestützte Basisaffekte (Angst, Ekel, Freude, Trauer, Überraschung, Wut), deren vollständige Unabhängigkeit jedoch nicht abschließend geklärt ist. Eine Vielzahl psychischer Störungsbilder geht mit einer Einschränkung der Affekterkennung in der qualitativen und/oder zeitlichen Domäne einher. Dies begründet den Bedarf nach validen, sensitiven, spezifischen und im diagnostischen Alltag einfach anzuwendenden Messinstrumenten, welche sowohl die Genauigkeit der Affekterkennung, als auch die Reaktionszeit erfassen und in ein Maß integrieren.

Messung der Affektwahrnehmungskompetenz

Das Verständnis menschlicher Affektwahrnehmung hat eine hohe Relevanz für die medizinische und psychologische Forschung. PubMed.gov listet über 13.000 Publikationen unter dem Suchbegriff affect recognition. Daraus leitet sich ein ebenfalls großer Bedarf nach geeigneten Messverfahren und Stimulusinventaren ab, welcher sich auch im Umfang der verschiedenen publizierten Verfahren und Inventare niederschlägt. Eine Übersichtsarbeit liefern Paiva-Silva et al. (2016). Im Folgenden werden die für die Erstellung und Modifikation des PCFAE entscheidenden Beiträge dargestellt.

Ekman 60-Faces Test (EK-60F). Duchenne (1862) schlussfolgerte aus seinen elektrophysiologischen Experimenten, dass die gemeinsamen Kontraktionen bestimmter Gesichtsmuskeln spezifische, affektiv mimische Gesichtsausdrücke evozieren. Auf Grundlage dieses Befundes entwickelten Ekman & Friesen (1971) zunächst die Facial Affect Scoring Technique (FAST) zur kategorialen Klassifizierung diskreter Affekte. Dies ermöglichte die Erstellung eines ersten, gut validierten Stimulussets affektiver Gesichtsausdrücke (Pictures of Facial Affects, PFA; Ekman & Friesen, 1976). Die PFA stellen das in wissenschaftlichen Studien bis heute am häufigsten verwendete Stimulusmaterial zur Affektwahrnehmung dar (Calvo & Lundqvist, 2008). Ekman & Friesen (1978) entwickelten weiterhin ein Kodierungssystem, welches unter anderem eine Beurteilung der Darstellungsgüte affektiver Gesichtsausdrücke - beispielsweise des Stimulusmaterials innerhalb eines Inventars affektiver Gesichtsausdrücke - ermöglicht (Facial Action Coding System, FACS). Diese Arbeiten bildeten die Grundlage bei der Erstellung des Ekman 60-Faces Tests (Young et al, 2002), dem im klinischen Kontext bisher am häufigsten eingesetzten Test zur Erfassung von psychisch oder neurologisch bedingten Einschränkungen der Affektwahrnehmungskompetenz (Dodich, 2013).

Der Test beinhaltet 60 Einzelbilder (schwarz/weiß Aufnahmen) affektiver Gesichtsausdrücke mit maximaler Affektausprägung. Es werden die 6 Basisaffekte nach Ekman, Friesen & Ellsworth (1982) sowie neutrale Gesichtsausdrücke mit einer Präsentationsdauer von 5 Sekunden gezeigt. Anschließend soll der Affekt qualitativ klassifiziert werden. Es findet keine Erfassung der Reaktionszeit statt. Die Split-Half Reliabilität wird mit $\rho_t = 0.62$ angegeben (Young et al, 2002). Der Test ist somit als wenig messgenau zu bewerten (Bühner, 2011).

Karolinska Directed Emotional Faces (KDEF). Die schwedische Forschergruppe um Lundqvist bewerteten insbesondere die geringe Anzahl an Fotomodellen (5 Männer, 6 Frauen) bei der Erstellung des EK-60F als unzureichend und entwickelte in der Folge eines der bisher umfangreichsten Stimulussets unter Verwendung von 35 männlichen und weiblichen Modellen. Es entstanden insgesamt 4900 Fotografien menschlich affektiver Gesichtsausdrücke. Diese unterteilen sich in die 6 Basisaffekte und einen neutralen Gesichtsausdruck. Die Gesichter wurden aus 5 verschiedenen Blickwinkeln fotografiert (Lundqvist, Flykt & Öhman, 1998; Calvo & Lundqvist, 2008). Das Bilderset ist validiert (Goeleven et al., 2008). Es steht darüber hinaus ein aus den KDEF Stimuli gemitteltes Set (Averaged Karolinska Directed Emotional Faces, AKDEF; Lundqvist & Litton, 1998) zur Verfügung. Das KDEF Bilderset bildete die Grundlage für die Erstellung der Erwachsenenstimuli des PCFAE.

Weitere Inventare. Neben Art und Umfang wurde die mangelnde ethnische Varianz und Altersdiversität bestehender Stimulusinventare sowie die statische Stimuluspräsentation kritisiert (Phillips et al., 1997). Das Max Planck Institut lieferte mit FACES (Ebner, Riediger & Lindenberger, 2010) ein Bilderset verschiedener Altersgruppen (min 19 Jahre). Das National Institute of Mental Health Child Emotional Faces Picture Set (Egger et al., 2011) bildet Jugendliche im Alter von 10 – 17 Jahren ab, weitere Verfahren die

Altersgruppen von 6 bis 16 Jahren (Dartmouth Database of Children's Faces; Dalrymple, Gomez & Duchaine, 2013) sowie von 2 bis 8 Jahren (Child Affective Facial Expression – CAFE; LoBue & Thrasher, 2014). Das NimStim Inventar (Tottenham et al., 2009) enthält Stimulusmaterial von Darstellern asiatischer, afrikanischer, latein-amerikanischer und europäischer Abstammung. Dynamisches Stimulusmaterial wurde von Jusyte, Gulewitsch & Schönenberg (Radboud Faces Database, 2017) unter der Verwendung von FantaMorph (Abrosoft Co., 2002) implementiert, während Phillips et al. (1997) die Intensität der maximalen Stimulusausprägung abstuften.

Kritik und Implikationen. Wie vorab dargestellt, decken verfügbare Simulussets und Testverfahren häufig nur spezifische Bereiche ab. Zudem wird oftmals nur eine qualitative Einschätzung des gezeigten Affekts gefordert und somit nur die Antwortgenauigkeit erfasst wird. Nur wenige Verfahren erfassen zusätzlich die Antwortlatenz oder Fehlerraten. Nahezu alle Verfahren sind im leichten bis mittleren Schwierigkeitsgrad anzusiedeln (70 – 88% korrekte Antworten im Mittel über alle Sets; Egger et al., 2011), sodass eine Differenzierung im Bereich hoher Fähigkeitsausprägung nicht gewährleistet ist. Statische Stimuli weisen nur geringe ökologische Validität auf (Barbeau, Joubert & Felician 2009; Cigna, Guay & Renaud, 2017; Kosson et al., 2002).

Demgegenüber verbessert die Verwendung dynamischen Bildmaterials die Erkennungsleistung (Wehrle et al., 2000), selbst bei geringer Intensität des affektiven Ausdrucks (Ambadar, Schooler & Cohn, 2005). Dynamische Stimuli werden als natürlicher empfunden (Sato, Kochiyama & Uono, 2015), führen zu stärkerer Aktivität in affektverarbeitenden Hirnarealen (Kilts et al., 2003; LaBar et al., 2003; Trautmann, Fehr & Herrmann, 2009) und stärkerer Nachahmung des Affekts (Rymarcczyk et al., 2011; Sato, Fujimua & Suzuki, 2008; Weyers et al., 2006). Die Verwendung neutraler Gesichtsausdrücke kann eine Fehlerquelle darstellen: Personen mit Borderline Persönlichkeitsstörung und

Kinder nehmen diese, möglicherweise aufgrund unsicherer Interpretation, eher als bedrohlich wahr (Donegan et al., 2013; Thomas et al., 2001). Egger et al. (2011) legen weiterhin dar, dass neutrale Stimuli, obwohl sie emotionale Intensität vermeiden sollen, hinsichtlich ihrer Eindrucksstärke höher bewertet werden, als traurige Stimuli. Aus diesem Grund wurde anstelle eines neutralen Gesichtsausdrucks die Verwendung gelassener Gesichter vorgeschlagen (Tottenham et al., 2009).

Es ist zu berücksichtigen, dass existierende Verfahren oft für einen bestimmten Verwendungszweck erstellt wurden. Folglich muss zwischen verfahrensspezifischer Kritik und allgemeingültigen Implikationen unterschieden werden, bzw. die Zweckausrichtung des Verfahrens bei der Umsetzung dieser Punkte berücksichtigt werden. Die bei der Erstellung des PCFAE berücksichtigten Aspekte werden in der Folge dargestellt.

Test of Perceptual Competence of Facial Affective Expression (PCFAE)

Der PCFAE Test wurde von Ingenhag, Schäfer & Franz (2007) in seiner ersten Version als computerbasiertes, schnell und einfach einzusetzendes Diagnostikverfahren entwickelt. Der Test soll die Affektwahrnehmungskompetenz bei Gesunden bzw. den Grad der Einschränkung bei Krankheitsbildern, die mit einer reduzierten Affekterkennungsfähigkeit einhergehen, messen. Mögliche Einsatzbereiche sind die psychologische Forschung und Leistungsdiagnostik sowie die Anwendung im klinischen (Prochnow et al., 2011) und therapeutischen Kontext. Eine zugehörige Trainings-App für elektronische Geräte (Smartphone, Tablet) unter Verwendung des Stimulusmaterials des PCFAE ist in Erstellung (Jansen, 2018b; unveröffentlicht). Der Test wurde mehrfach modifiziert, die Analyse der aktuellen Version des PCFAE stellt den Schwerpunkt dieser Arbeit dar.

Konzeptionelle Vorüberlegungen. Der PCFAE sollte die Kritikpunkte und Implikationen, die sich aus den Erfahrungen vorangegangener Testverfahren und

Stimulussets zur Affektwahrnehmung ergeben hatten so umsetzen, dass das entstehende Verfahren die vielschichtigen Anforderungen seiner geplanten Anwendungsbereiche valide erfüllen kann. Die Verbesserungen gegenüber anderen Verfahren umfassten folgende Punkte:

- Verbesserung der Bildqualität des Stimulusmaterials,
- Verbesserung der ökologischen Validität durch Verwendung einer (quasi)dynamischen Darstellung,
- Verbesserung der Fähigkeitseinschätzung durch Implementierung neuer Leistungsmaße,
- kulturelle Unabhängigkeit durch Verwendung gemittelter Bilder,
- Altersdiversität durch Verwendung von Erwachsenen- und Kinderstimuli,
- Berücksichtigung des Geschwindigkeits-Genauigkeits-Ausgleichs bei der Affektwahrnehmung.

Testbeschreibung und Durchführung. Der PCFAE ist ein computerbasierter Leistungstest zur Bestimmung der Fähigkeit von Personen, die 6 Basisaffekte korrekt erkennen bzw. diskriminieren zu können. Der Test umfasst insgesamt 24 Items. Je 4 Items bilden einen Basisaffekt ab und werden als entsprechende Affektskala zusammengefasst. Die skaleninternen Items untergliedern sich in je 2 Items weiblichen und männlichen Geschlechts. Die geschlechterspezifischen Items werden einmal als Kindergesicht und einmal als Erwachsenengesicht präsentiert¹. Die Präsentation der Items erfolgt in vollständig randomisierter Reihenfolge auf einem Monitor, wobei jedes Item im Testverlauf 2 Mal gezeigt wird. Zur Auswertung erfolgt eine Mittelung der Leistungsmaße zwischen beiden Präsentationen der Items.

Die in der 2007 erstellten Erstversion des PCFAE verwendeten Items wurden dem AKDEF Stimulusset (Lundqvist & Litton, 1998) entnommen und entsprachen stilistisch einer

¹ Die Items *Angst weiblich Kind*, *Angst weiblich Erwachsener*, *Angst männlich Kind* und *Angst männlich Erwachsener* bilden folglich die Skala Angst.

gemittelten Form des von Ekman & Friesen (1976) erstellten PFA Sets. Diese ursprünglichen Stimuli erwiesen sich qualitativ (schwarz/weiß Darstellung, unscharfe Abbildung) als unzureichend, sodass mittels moderner Bildbearbeitungssoftware (FantaMorph; Abrosoft Co., 2002) und in Absprache mit den Autoren aus dem qualitativ gut geeigneten KDEF Set (scharfe Farbdarstellung; Lundqvist, Flykt & Öhman, 1998) neu gemittelte Stimuli hoher Darstellungsgüte erstellt wurden. Als Mittelung wird der softwareseitige Morphingprozess bezeichnet, bei dem aus einer Vielzahl von Originalaufnahmen affektiver Gesichtsausdrücke ein personenunabhängiges Bild des Affektausdrucks als Mittelwert der individuellen Gesichtsmerkmale erstellt wird. Abbildung 1 zeigt Beispiele der Stimuli des EK 60-F, des AKDEF Sets und eines aus dem KDEF Set neu gemittelten Stimulus des PCFAE.

Der PCFAE wurde weiterhin um die Darstellung gemittelter, FACS kodierter Kinderbilder ergänzt (Picture-Set of Young Children Affective Facial Expression – PSYCAFE, Altersspanne 4-6 Jahre; Franz et al., 2018; unveröffentlicht; Rampoldt, 2014). Anhang A enthält eine vollständige Auflistung der Originalbilder als Ausgangsmaterial und die daraus gemittelten Stimuli des PCFAE. Die Validierung des modifizierten Stimulusmaterials wird im Abschnitt Bisherige Testanalyse beschrieben. Alle Bilder des

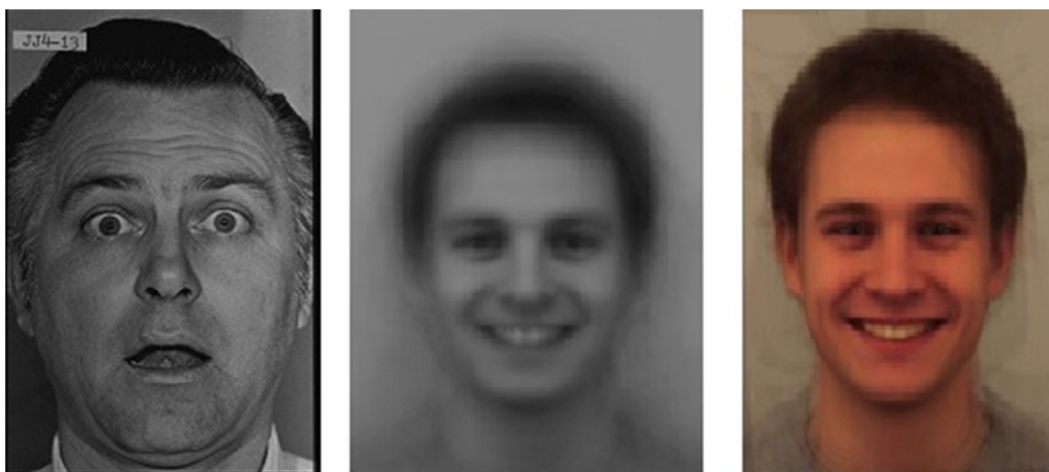


Abbildung 1. Vergleich eines Items des EK-60F (links), eines in der Urversion des PCFAE verwendeten Stimulus des AKDEF Sets (Mitte) und eines aus dem KDEF Set neu gemittelten Stimulus (Itemgrundlage des neuen PCFAE, rechts).

Stimulusmaterials wurden zusätzlich mit einem schwarzen Rahmen versehen, sodass der Fokus des Betrachters ausschließlich auf der affektexpressiven Mimik liegt und die Peripherie des Gesichts (Haare, Ohren, etc.) ausgeblendet wird.

Zur Erstellung der Testitems wurden mittels FantaMorph aus je einem Bild des Stimulusmaterials mit neutralem Gesichtsausdruck und dem zugehörigen Stimulus eines spezifischen Affekts maximaler Ausdrucksintensität 28 weitere Einzelbilder ansteigender Intensität erstellt. Es ergibt sich eine Sequenz von 30 Einzelbildern mit stetig ansteigendem Affektausdruck. Diese werden beginnend mit dem neutralen Gesichtsausdruck (Bild 1) schrittweise bis zum Affektapex mit maximaler Affektausdrucksintensität präsentiert (Bild 30). Abbildung 2 zeigt beispielhaft den Beginn einer Itemsequenz mit neutralem Ausdruck, 50% Affektausdrucksintensität und Affektapex. Die Präsentationsdauer von 700 Millisekunden pro Einzelbild erzeugt einen langsam fließenden, quasidynamischen

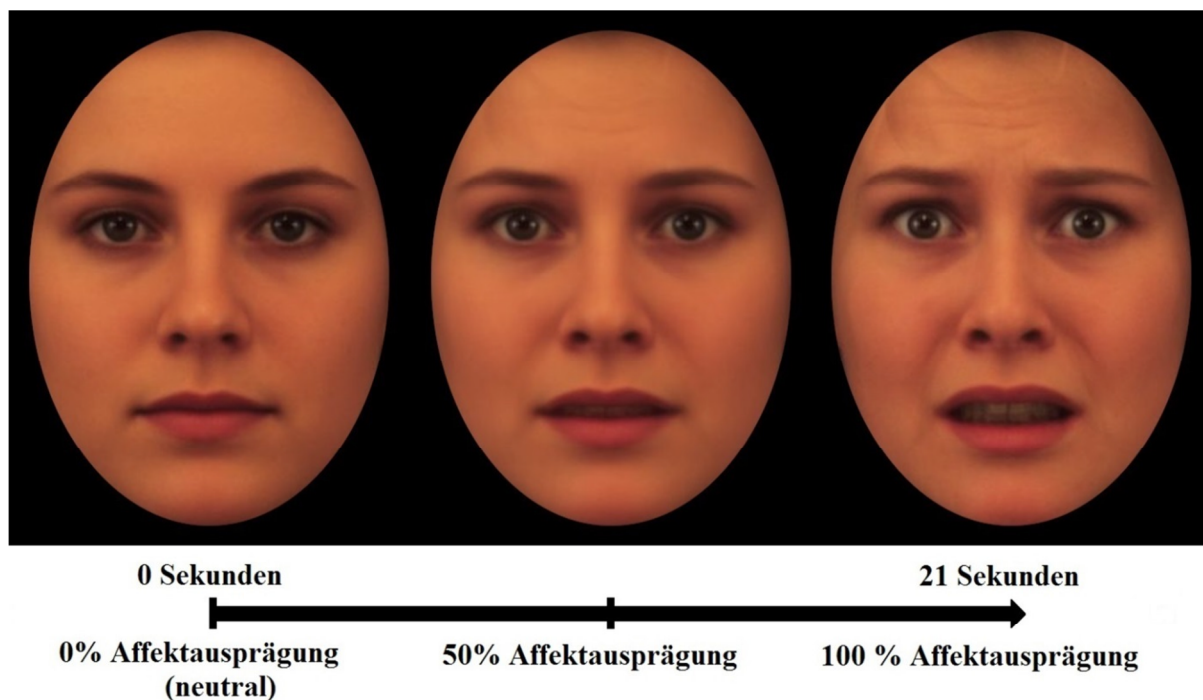


Abbildung 2. Stimulusverlauf des Items *Angst weiblich Erwachsener* des neuen PCFAE von neutralem Gesichtsausdruck zu maximaler Affektausprägung.

Affekteindruck als Kompromiss zwischen einer natürlicheren, vollständig dynamischen Videodarstellung und den Anforderungen an eine vereinfachte Auswertung. Die maximale Präsentationsdauer bis zum Erreichen des Affektapex beträgt folglich 21 Sekunden pro Item. Die Versuchsperson hat somit für jede Stufe der Affektausprägung eine kurze Entscheidungszeit und ist angehalten, mittels Druck der Leertaste anzugeben, ob der dargestellte Affekt erkannt wurde. Die Itempräsentation wird dann unterbrochen, das Item ist für die folgende Affektauswahl nicht mehr einsehbar. Mittels Radiobutton¹ ist aus einem Auswahlménü der entsprechende Basisaffekt auszuwählen (forced choice Aufgabe). Die Auswahlreihenfolge (Freude, Angst, Ekel, Wut, Überraschung, Trauer) wird, unter Inkaufnahme von Übungseffekten, für jedes Item konstant gehalten, um eine ständige Neuorientierung innerhalb des Durchgangs zu vermeiden.

LeistungsmaÙe. Der PCFAE erfasst die Antwortgenauigkeit (*ACC*; Anzahl korrekt beantworteter Items, Wertebereich 0 – 1, 1 entspricht 100% korrekt beantworteter Items; die Fehlerrate kann hieraus abgeleitet werden), die Reaktionszeit (*RT*; Zeit von Beginn der Itempräsentation bis der Proband durch Tastendruck anzeigt, den Affekt erkannt zu haben) sowie die Reaktionssicherheit (*RSI*; Entscheidungszeit vom Zeitpunkt der Anzeige der Affekterkennung bis zur Auswahl des erkannten Affekts aus den 6 Basisaffekten).

Die zusätzliche Erfassung der RT orientiert sich an den Befunden von Newcombe (1979): Personen mit Störungsbildern, die mit einer eingeschränkten Fähigkeit der Gesichtswahrnehmung einhergehen, nutzen Kompensationsstrategien, die auf Kosten der RT eine mit Gesunden vergleichbare ACC erzielen (Geschwindigkeits-Genauigkeits-Ausgleich). Dies gilt ebenso für die Wahrnehmung affektiver Gesichtsausdrücke (Aebi, 2012; Young et al, 1993). Für die Auswertung des PCFAE wird im Rahmen dieser Arbeit ein Leistungsquotient vorgeschlagen, der ACC und Fehlerrate unter Einbezug der RT integriert.

¹ Ein Radiobutton bezeichnet einen anklickbaren Auswahlbutton, welcher bis zur Eingabebestätigung Änderungen der Auswahl zulässt.

Hohe Werte zeigen eine hohe Personenfähigkeit an und werden durch hohe Erkennungsraten, geringe Fehlerraten und kurze Reaktionszeiten erreicht.

Die Eignung der RSI als Leistungsmaß wurde im Rahmen dieser Arbeit überprüft. Probanden könnten bereits frühzeitig die Stimulusdarbietung unterbrechen, ohne den Affekt zweifelsfrei erkannt zu haben. Die RSI würde dann ebenfalls Informationen zur Fähigkeit der Affekterkennung liefern: Eine kurze Entscheidungsphase spräche für eine klare Affekterkennung. Für die RSI wurde ebenfalls ein Leistungsquotient vorgeschlagen und auf Eignung geprüft. Die Auswertung beider Reaktionszeitmaße als Leistungsmaß erfolgte für korrekt beantwortete Items. Eine schnelle Reaktionszeit allein, ohne das Item richtig erkannt zu haben, ließe keine sinnvolle Aussage über die Fähigkeitsausprägung zu.

Bisherige Testanalyse. Eine Analyse und Validierung des PCFAE in seiner ersten Version wurde von Seidler (2008) durchgeführt. Die Validierung des PCFAE mit dem neuen Stimulusmaterial erfolgte durch Hahn (2016).

Itemanalyse. Seidler (2008) gibt eine Trennschärfe der 6 Affektskalen des PCFAE von $r_{it} = 0.27 - 0.41$ an. Somit erfüllte die PCFAE Urversion die Anforderungen an die Trennschärfe von $r_{it} = 0.30$ weitestgehend (Fisseni, 2004), jedoch auf Mindest- bis mittlerem Niveau. Die durchschnittliche zufallskorrigierte Itemschwierigkeit nach Affektskala wies mit $P_{ZK} = 56.87 - 96.25$ auf einen leichten bis mittelschweren Test hin. Seidler (2008) schlug nach durchgeführter Hauptkomponentenanalyse (Varimax Rotation) ein eindimensionales Konstrukt des PCFAE vor, auf dessen Faktor alle Items luden (Faktorladungen: $\lambda = 0.46 - 0.76$). Es wurden keine weiteren Modelle exploriert. Für eine detaillierte Darstellung der skalen- und itemspezifischen Werte sowie der Hauptkomponentenanalyse siehe Seidler (2008).

Reliabilitätsanalyse. Die interne Konsistenz des PCFAE wurde mit $\alpha = 0.75$ (Cronbach, 1951) angegeben, die Split Half Reliabilität mit $\lambda_4 = 0.82$ (Guttman, 1945) und

der Spearman-Brown-Koeffizient mit $\rho_{tt} = 0.82$ (Brown, 1910; Spearman, 1910). Seidler (2008) schlug für eine weitere Einschätzung der Reliabilität des PCFAE die Berechnung der Test-Retest-Reliabilität vor.

Konstruktvalidität. Seidler (2008) gab die konvergente Validität zum EK-60F (Ekman & Friesen, 1976) mit $r_{tc} = 0.35$ auf Grundlage der ACC an und bewertete diese als „mäßig“ (Seidler, 2008, S.96). Eine Berücksichtigung der RT bzw. des Geschwindigkeits-Genauigkeits-Ausgleich konnte nicht umgesetzt werden, da die RT beim EK-60F nicht erfasst wird. Die 20-Item Toronto Alexithymia Scale (Taylor, Bagby & Parker, 2003) und die Skalen zum Erleben von Emotionen (Behr & Becker, 2004) als inhaltlich ähnliche Verfahren wurden mit erfasst, es wurden jedoch keine Korrelationen mit dem PCFAE berechnet.

Neues Stimulusmaterial. Die Analyse des modifizierten Stimulusmaterials des PCFAE von Hahn (2016) zeigte, dass Wut und Angst in Erwachsenengesichtern besser und schneller erkannt wurde als in Kindergesichtern. Im Gegenzug wurde Freude bei Kindern als intensiver nachempfunden bewertet. Auf Gesamttestebene wurde kein Alterseffekt der Stimuli für die ACC nachgewiesen, während für die RT korrekt beantworteter Items ein Alterseffekt mittlerer Effektstärke auftrat (Cohen, 1988). Kindergesichter wurden im Mittel nach 10.88 Sekunden korrekt erkannt, während Affekte in Erwachsenengesichtern im Mittel nach 9.33 Sekunden korrekt erkannt wurden.

Zusammenfassend ist der PCFAE als ein einfach durchzuführendes Diagnoseinstrument zur Beurteilung der Affektwahrnehmungskompetenz konzipiert. Er beurteilt die Fähigkeit der Affekterkennung für die 6 Basisaffekte Angst, Ekel, Freude, Trauer, Überraschung und Wut hinsichtlich Antwortgenauigkeit und Reaktionszeit und kann so Kompensationsstrategien bei Einschränkungen der Affektwahrnehmungskompetenz (Geschwindigkeits-Genauigkeits-Ausgleich) erfassen. Eine Analyse der Urversion des

PCFAE lieferte erste Hinweise zur Test- und Itemstruktur. Der Test sollte inhaltlich ähnlichen Verfahren konzeptionell überlegen sein. Insbesondere das Stimulusmaterial erwies sich jedoch qualitativ als noch unzureichend, sodass in der Folge kontinuierliche Modifizierungen stattfanden.

Fragestellung

Die umfangreichen Modifizierungen des PCFAE erforderten eine erneute Analyse des Tests. Diese Arbeit beabsichtigt einen Überblick über das aktuelle Testverfahren zu schaffen und prüft folgende Fragestellungen:

- Welche Güte weisen die Items des PCFAE hinsichtlich Trennschärfe, Itemschwierigkeit und Itemhomogenität auf?
- Welche Leistungsmaße des PCFAE eignen sich zur Erfassung der Affektwahrnehmungskompetenz?
- Welche inhaltlich sinnvoll interpretierbaren Strukturmodelle lassen sich explorieren?
- Wie messgenau ist der PCFAE in seiner neuen Form und weist er zeitliche Stabilität auf?
- Weist der PCFAE konvergente oder diskriminante Validität zu inhaltlich ähnlichen Verfahren auf?

Kein Bestandteil dieser Arbeit ist die Überprüfung intra- bzw. interindividueller Unterschiede der Affektwahrnehmungskompetenz der Versuchspersonen sowie deren Performanz innerhalb der Affektskalen.

Methode

Versuchspersonen

Die Stichprobe wurde an der Heinrich Heine Universität Düsseldorf und am Universitätsklinikum der Heinrich-Heine-Universität Düsseldorf mittels Aushängen rekrutiert. Zugelassen wurden Personen mit ausreichender deutscher Sprachkenntnis und einem Alter zwischen 18 – 65 Jahren. Die Gesamtstichprobengröße betrug $N = 168$ Versuchspersonen (VP). Insgesamt wurden 14 Personen ausgeschlossen. Die endgültige Stichprobe umfasst $N = 154$ Teilnehmer (91.67% der ursprünglich erhobenen Stichprobe), davon 88 (57.14%) weiblichen und 66 (42.86%) männlichen Geschlechts. Ausgeschlossen wurden Personen, die nur einen Erhebungstermin absolvierten (6 VP), die nachweislich den Test nicht ernsthaft bearbeitet hatten (2 VP) oder die Versuchsanweisungen nicht befolgt hatten (6 VP)¹. Das Alter der Teilnehmer betrug $M = 23.87$ ($SD = 6.77$, $min = 18$, $max = 58$) Jahre. Die Stichprobe war hoch gebildet. 126 (81.82%) Teilnehmer gaben das Abitur als Schulabschluss an, 28 (18.18%) Teilnehmer hatten bereits einen Hochschulabschluss erworben.

Weitere Ausschlusskriterien für die Teilnahme waren eine bestehende schwerwiegende psychische Vorerkrankungen bzw. eine starke psychische Belastung, schwere Stoffwechselstörungen, unkorrigierte Visusminderung, akute Einnahme psychoaktiver Substanzen oder zentralnervös wirksamer Medikamente sowie Verdacht auf vorliegende Prosopagnosie (Traitartige Unfähigkeit, bekannte Personen anhand ihres

¹ Die nicht ernsthafte Bearbeitung wurde erkennbar durch eine mittlere Reaktionszeit von 1 Sekunde im PCFAE Paradigma bei einer Antwortgenauigkeit auf Rateniveau ($P = 0.17$ bei 6 Antwortmöglichkeiten). Da sich der Affekt innerhalb der Präsentation eines Items in 700 Millisekunden Schritten aufbaut, ist nach einer Sekunde lediglich eine Affektausprägung von 6,66% (nahezu neutraler Gesichtsausdruck) erreicht, eine valide Antwort ist zu diesem Zeitpunkt unwahrscheinlich. Dies wird anhand der Antwortgenauigkeit deutlich. Das Nichtbefolgen der Versuchsanweisung wurde erkennbar durch eine mittlere Reaktionszeit von 21 Sekunden, welche gleichzeitig der maximalen Präsentationsdauer der Items entsprach. Die entsprechenden Probanden haben sich folglich so schnell wie möglich durch den Test gearbeitet bzw. jedes Item bis zur maximalen Affektausprägung von 100% betrachtet, ohne die Itempräsentation bei Erkennen des Affekts zu unterbrechen. Eine Beurteilung der zeitlichen Komponente der Affekterkennung war somit nicht möglich.

Gesichts zu erkennen; Bodamer, 1947). Die Abklärung des Vorliegens ausreichender Sprachkenntnisse und Prüfung auf Ausschlusskriterien erfolgte durch ein Anamnesegespräch. Dieses wurde durch ein computergestütztes Screening mittels Brief Symptom Inventory (BSI, Derogatis, 1993) sowie dem Famous Faces Test (FFT, Duchaine, 2008; Adaption durch Jansen, Schäfer & Franz, 2018; unveröffentlicht) unterstützt. Eine starke psychische Belastung wurde gemäß einem Globalen BSI Kennwert (GSI, Global Severity Index) von $T_{GSI} \geq 63$ definiert, Verdacht auf Vorliegen einer Prosopagnosie bei einem FFT Score $\leq 50\%$ (Rateniveau). Es lagen posthoc jedoch keine Befunde vor, die zum nachträglichen Ausschluss weiterer Teilnehmer geführt hätten.

Durchführung

Die Durchführung des Experiments (Registrierungs-ID: 2015023246, Ethik Votum Nr.: 6051R, siehe Anhang D) fand in einem reizarmen Gruppenraum am Klinischen Institut für Psychosomatische Medizin und Psychotherapie am Universitätsklinikum der Heinrich-Heine-Universität Düsseldorf unter konstanten Bedingungen (Beleuchtung, Lärmpegel, Temperatur) statt. Es wurden maximal 3 Personen gleichzeitig erhoben. VP wurden mittels standardisierter Probandeninformationen über Thema, Ziel, Sinn und Zweck der Studie informiert. Anschließend wurden die Probanden aufgefordert, die Kenntnisnahme der Belehrung zum Datenschutz und die Freiwilligkeit der Teilnahme durch eine Unterschrift zu bestätigen. Es folgte die Erhebung demografischer Daten mittels Fragebogen und die Generierung des Versuchspersonencodes. Anhang B enthält Muster dieser Unterlagen. Anschließend wurde den Probanden ein Laptop zugewiesen und das Experiment gestartet. Der Retest erfolgte mit einem Abstand von 4 Wochen. Nach vollständigem Abschluss des Experiments wurden die VP für ihren zeitlichen Aufwand mit 20 € entschädigt.

Material und experimentelle Aufgabe

Das Experiment wurde vollständig mithilfe des Computerprogramms PsychoPy2

(Version 1.85.1; Pierce, 2007) auf Laptops präsentiert. In der ersten Testung waren 5 Einzelabschnitte zu bearbeiten: Beginnend mit dem Famous Faces Test (Duchaine, 2008; Adaption durch Jansen, Schäfer & Franz, 2018; unveröffentlicht) folgten die psychometrischen Fragebögen 20-Item Toronto Alexythymia Scale (Bagby et al., 1994), Skalen zum Erleben von Emotionen (Behr & Becker, 2004), das Brief Symptom Inventory (Derogatis, 1993; deutschen Version Franke, 1999) und abschließend der PCFAE. Bei der Durchführung des Retests entfiel die Bearbeitung des Famous Faces Test, da eine mögliche Prosopagnosie bereits zu Messzeitpunkt T1 abgeklärt wurde. Die verwendeten psychometrischen Inventare und Leistungstests werden nachfolgend beschrieben und sind in Anhang C abgebildet. Das Stimulusmaterial des PCFAE kann Anhang A entnommen werden. Die Reihenfolge der Inventare und Leistungstests wurde für alle Versuchspersonen konstant gehalten, die Reihung der nachfolgenden Beschreibung der Verfahren entspricht der Reihung innerhalb des Experiments.

Rechnerausstattung. Für die Präsentation des experimentellen Paradigmas wurden 3 unterschiedliche Laptops verwendet. Die Zuweisung der Rechner an die VP erfolgte zu beiden Messzeitpunkten zufällig. Die Rechner besaßen folgende Systemspezifikationen:

Rechner 1: Acer Extensa 7630EZ mit Intel Pentium Dual Core Prozessor 2GHz, 3GB RAM, Betriebssystem Windows XP Professional, Bildschirmauflösung 1440*900 Pixel auf 17 Zoll Monitor.

Rechner 2: Lenovo ThinkPad L530 mit Intel Core I3 Prozessor 2.4GHz, 4GB RAM, Betriebssystem Windows 7 Professional 64 Bit, Bildschirmauflösung 1366*768 Pixel auf 15.6 Zoll Monitor.

Rechner 3: Acer TravelMate P276M56WM mit Intel Core I5 Prozessor 2.4GHz, 4GB RAM, Betriebssystem Windows 7 Professional 64 Bit, Bildschirmauflösung 1600*900 Pixel auf 17.3 Zoll Monitor.

Der mögliche Einfluss der Verwendung nicht einheitlicher Laptops wurde überprüft.

Famous Faces Test (FFT). Der Leistungstest Famous Faces Test stellt, in Anlehnung an ein vergleichbares Instrument der Universitäten Harvard und London (Duchaine, 2008; Adaption durch Jansen, Schäfer & Franz, 2018; unveröffentlicht), ein einfaches Verfahren zur Verdachtsdiagnose auf Vorliegen einer Prosopagnosie dar. Die experimentelle Aufgabe sieht ein Identifizieren und Wiedererkennen insgesamt 30 bekannter Personen des öffentlichen Lebens vor. Als Verdacht auf Prosopagnosie wurde eine Erkennungs- und Wiedererkennungseistung auf oder unterhalb des Rateniveaus von 50% gewertet. Dies hätte im Rahmen dieser Arbeit zum Ausschluss des betreffenden Probanden geführt (exakte Beschreibung der Testauswertung bei Jansen, 2018a; unveröffentlicht). Das Verfahren befindet sich zurzeit in der Validierung.

20-Item Toronto Alexithymia Scale (TAS-20). Der Selbstauskunftsfragebogen TAS-20 dient der Erfassung alexithymer Persönlichkeitsmerkmale. Alexithymie bezeichnet neben weiteren Symptomen die Unfähigkeit, eigene Gefühle angemessen wahrzunehmen und verbal zu beschreiben (Nemiah, Freyberger & Sifneos, 1976). Das Inventar umfasst 20 Items in 5-stufiger Likert-Skalierung, welchen faktorenanalytisch begründet 3 Subskalen zugeordnet werden (Taylor, Bagby & Parker, 2003). Das Verfahren wird als reliabel mit $\alpha = 0.81$ (Cronbach, 1951) und zeitlich stabil mit einer Retest Reliabilität (3 Wochen) von $\rho_{tt} = 0.71$ beschrieben (Bagby, Parker & Taylor, 1994). Im Rahmen dieser Arbeit wurde der Fragebogen zur Überprüfung der Konstruktvalidität des PCFAE herangezogen.

Skalen zum Erleben von Emotionen (SEE). Der Selbstauskunftsfragebogen SEE dient, auf Grundlage der Person-zentrierten Persönlichkeitstheorie Rogers (1951) bzw. der emotionszentrierten Psychotherapie (Greenberg, Rice & Elliott, 1996), der Erfassung des individuellen Gefühlserlebens und der Gefühlsbewertung. Das Inventar umfasst 42 Items in 5-stufiger Likert-Skalierung, welchen faktorenanalytisch begründet 7 Subskalen zugeordnet

werden. Das Verfahren ist anhand mehrerer Stichproben normiert ($N = 3197$). Die interne Konsistenz der einzelnen Subskalen variiert zwischen $\alpha = 0.70$ und $\alpha = 0.86$ (Cronbach, 1951), die Retest Reliabilität (2 – 14 Wochen) zwischen $\hat{\rho}_t = 0.70$ und $\hat{\rho}_t = 0.90$. Das Inventar kann somit als zeitlich stabil und reliabel beurteilt werden (Behr & Becker, 2004). Für die vorliegende Untersuchung bedeutsam sind die Skalen „Erleben von Emotionsüberflutung“ (EEU) sowie „Erleben von Emotionsmangel“ (EEM). Diese beschreiben ein Erleben von subjektiv zu viel bzw. zu wenig Emotion – folglich eine dysfunktionale Wahrnehmung der eigenen Emotionen. Beide Skalen weisen konvergente Validität mit der Skala „Mangelnde Identifikation von Gefühlen“ des TAS-26 Fragebogen auf¹ (Kupfer, Brosig & Braehler, 2001). Im Rahmen dieser Arbeit wurde die Konstruktvalidität des PCFAE mithilfe des SEE (Subskalen EEU und EEM) untersucht.

Brief Symptom Inventory (BSI). Das BSI ist ein Verfahren zur Beurteilung und Klassifizierung der psychischen Belastung. Es umfasst 53 Items in 5-stufiger Likert-Skalierung, die 9 klinisch-erfahrungsbezogen konstruierten und faktorenanalytisch überprüften Symptomskalen zugeordnet wurden. Diese Subskalen können in 3 globale Kennwerte überführt werden, die eine schnelle Beurteilung der psychischen Belastung erlauben. Das Verfahren ist anhand zweier Stichproben ($N = 1189$) sowie weiterer klinischer Stichproben normiert. Die interne Konsistenz der Skalen variiert zwischen $\alpha = 0.70$ und $\alpha = 0.89$ (Cronbach, 1951), die Retest Reliabilität (1 Woche) zwischen $\hat{\rho}_t = 0.68$ und $\hat{\rho}_t = 0.93$, abhängig von der jeweils herangezogenen Stichprobe (Franke, 1999). Das BSI kann somit als reliabel beurteilt werden. Im Rahmen dieser Untersuchung diente das BSI dem Probandenscreening und eventuellem -ausschluss bei vorliegender starker psychischer

¹ Das TAS-20 Inventar stellt die Weiterentwicklung des TAS-26 dar. Die Korrelationen der TAS-26 Skala „Mangelnde Identifikation von Gefühlen“ zu den SEE Skalen EEM und EEU wird mit $r = 0.41$ sowie $r = 0.55$ angegeben. Nach Kenntnis des Autors liegen derzeit keine Kennwerte für die konvergente Validität der aktuellen TAS-20 Version mit dem SEE Inventar vor, sodass die genannten Werte als Hinweis zu verstehen sind.

Belastung.

Test of Perceptual Competence of Facial Affective Expression (PCFAE). Testbeschreibung und -durchführung siehe Kapitel Einleitung.

Versuchsplan und statistische Auswertung

Die Untersuchung erfolgte mit einer angestrebten Stichprobengröße von $N = 150$ VP in einem 2×3 Messwiederholungsdesign (within subjects). Die Innersubjektfaktoren sind Messzeitpunkt (UV1, Faktorstufen *T1*, *T2*) und Affektskala (UV2, Faktorstufen *Angst*, *Ekel*, *Freude*, *Trauer*, *Überraschung*, *Wut*). Die abhängigen Variablen sind Reaktionszeit (RT, AV1), Reaktionssicherheit (RSI, AV2) und Antwortgenauigkeit (ACC, AV3). Außerhalb der Test-Retest Reliabilitätsanalyse (4 Wochen) wurden die Messzeitpunkte T1 und T2 getrennt voneinander betrachtet. Multivariate Analysen zur Erfassung möglicher Beziehungen der AVn untereinander sind nicht Bestandteil dieser Arbeit.

Die statistische Auswertung umfasste eine Stichprobenbeschreibung sowie eine Voraussetzungsprüfung für nachfolgende statistische Verfahren. Mögliche konfundierende Einflüsse der uneinheitlichen Rechnerausstattung wurden varianzanalytisch überprüft, ebenso wie das Vorliegen von Testwiederholungs- und Übungseffekten. Schwerpunkt der Arbeit bildete die Analyse des PCFAE. Berichtet werden zur Itemanalyse die Trennschärfe, Itemhomogenität, zufallskorrigierte Itemschwierigkeit und explorierte Dimensionalität sowie Reliabilitätsmaße und Beiträge zur Konstruktvalidität.

Aufgrund der Vielzahl an statistischen Operationen enthält die Ergebnisdarstellung methodische Anmerkungen, um eine bessere Nachvollziehbarkeit und Verständlichkeit zu erreichen. Diese beziehen sich auf die Beschreibung der einzelnen Verfahren und Berechnungsgrundlagen sowie notwendige Korrekturverfahren bei Voraussetzungsverletzungen. Mit Ausnahme der Darstellung der deskriptiven Statistiken und der

Überprüfung des Einflusses der Rechnerausstattung erfolgte die Analyse des PCFAE auf Grundlage der RT und RSI für korrekt beantwortete Items.

Die statistische Auswertung erfolgte mithilfe des Statistikprogramms R, Version 3.5.0 (R Core Team, 2018) und Microsoft Excel 2010. Effektstärkenberechnungen wurden mit dem Programm Psychometrica (Lenhard & Lenhard, 2016) durchgeführt. Folgende R Pakete fanden Anwendung:

- *ez* (Lawrence, 2016): Berechnung der Varianzanalysen,
- *ggplot2* (Wickham, 2016): Erstellen von Plots und Darstellungen,
- *nFactors* (Raiche, 2010): Berechnung von Parallelanalyse und Scree Test,
- *psych* (Revelle, 2018): Berechnung der explorativen Faktorenanalyse, Item-, Reliabilitäts-, und Validitätsanalyse,
- *GPArotation* (Bernaards & Jenrich, 2014): Rotation innerhalb der Faktorenanalyse,
- *reshape2* (Wickham, 2008): Datenumstrukturierung,
- *stats* (R Core Team, 2018): Statistische Testung auf Normalverteilung, Varianzhomogenität, grundlegende statistische Operationen,
- *xlsx* (Dragulescu, 2014): Einlesen der Rohdaten.

Innerhalb der Pakete verwendete Skripte werden gesondert im Rahmen der Ergebnisdarstellung genannt.

Ergebnisse

Einfluss der Rechnerausstattung

Varianzanalytisch betrachtet wurde der Einfluss des verwendeten Rechners (UV, between subjects, Faktorstufen Rechner 1 – 3) auf die Reaktionszeit (RT in Sekunden, AV1), die Reaktionssicherheit (RSI in Sekunden, AV2) sowie die Antwortgenauigkeit (ACC Anteil richtig gelöster Items, AV3) zu Messzeitpunkt T1 und T2. Aufgrund der angenommenen Varianzhomogenität (Levene Test für alle ANOVA n.s.) wurden die Unterschiede der Zellbesetzungen der UV als unkritisch bewertet, ebenso das Nichtvorliegen der Normalverteilung (Robustheit der ANOVA gegen diese Voraussetzungsverletzung, Nachtigall & Wirtz, 2013). Die Zellbesetzungen und deskriptiven Statistiken sind Tabelle 1 zu entnehmen. Die Betrachtung der AVn erfolgte univariat. Der verwendete Rechner hatte zu keinem Zeitpunkt einen signifikanten Einfluss auf die RT (Messzeitpunkt T1: $F(2,151) =$

Tabelle 1

Deskriptive Statistiken der Leistungsmaße RT, RSI und ACC nach Versuchsrechner.

Platz	Messzeitpunkt T1							Messzeitpunkt T2						
	<i>N</i>	RT		RSI		ACC		<i>N</i>	RT		RSI		ACC	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Rechner 1	33	8.73	3.08	1.94	2.91	0.87	0.07	30	8.06	1.58	1.79	1.04	0.86	0.09
Rechner 2	65	9.16	3.60	1.88	2.94	0.85	0.09	54	7.74	1.13	1.66	1.53	0.85	0.11
Rechner 3	56	9.10	3.31	1.96	3.39	0.86	0.07	70	7.92	1.34	1.76	1.43	0.87	0.09

Hinweis. N = Zellbesetzung der VP je Rechner. RT = Reaktionszeit in Sekunden. RSI = Reaktionssicherheit in Sekunden. ACC = Antwortgenauigkeit, Wertebereich von 0 bis 1, ein Wert von 1 entspricht 100% korrekter Antworten. M = Mittelwert. SD = Standardabweichung.

0.58, $p = 0.56$, $\eta^2 = 0.01$; Messzeitpunkt T2: $F(2,151) = 0.22$, $p = 0.81$, $\eta^2 = 0.00$), die RSI (Messzeitpunkt T1: $F(2,151) = 0.51$, $p = 0.60$, $\eta^2 = 0.01$; Messzeitpunkt T2: $F(2,151) = 0.84$, $p = 0.44$, $\eta^2 = 0.01$) oder die ACC (Messzeitpunkt T1: $F(2,151) = 1.47$, $p = 0.23$, $\eta^2 = 0.02$; Messzeitpunkt T2: $F(2,151) = 1.00$, $p = 0.37$, $\eta^2 = 0.01$).

Stichprobenbeschreibung und Voraussetzungsprüfung

Die VP erkannten die Affekte korrekt nach einer mittleren Reaktionszeit von $M =$

9,44 Sekunden ($SD = 3,48$) zu Messzeitpunkt T1 bzw. $M = 8,07$ Sekunden ($SD = 3,25$) zu T2. Dies entsprach einer mittleren Affektausprägung von 45% zu T1 bzw. 38% zu T2. Die Auswahl des korrekten Affekts erfolgte mit einer mittleren Reaktionssicherheit von $M = 1,84$ Sekunden ($SD = 1,26$) zu T1 und $M = 1,65$ Sekunden ($SD = 0,92$) zu T2 bei einer mittleren Antwortgenauigkeit von $M = 86\%$ korrekter Antworten ($SD = 0.27/0.28$ zu T1/T2) zu beiden Messzeitpunkten. Der Affekt Freude wurde am schnellsten, sichersten und mit der geringsten Fehlerrate erkannt. Die Affekte Ekel und Überraschung werden mittel schnell und sicher erkannt. Am langsamsten und mit der geringsten Reaktionssicherheit wurden die Affekte Angst, Trauer und Wut erkannt. Während Trauer, Überraschung und Wut überwiegend korrekt erkannt wurden, wurden die Affekte Angst und Ekel häufiger falsch erkannt. Von Messzeitpunkt T1 zu T2 nahmen die RT und RSI ab, während die ACC sich nicht substantiell veränderte. Abbildung 3 stellt die Antwortgenauigkeit nach Affektskala grafisch dar.

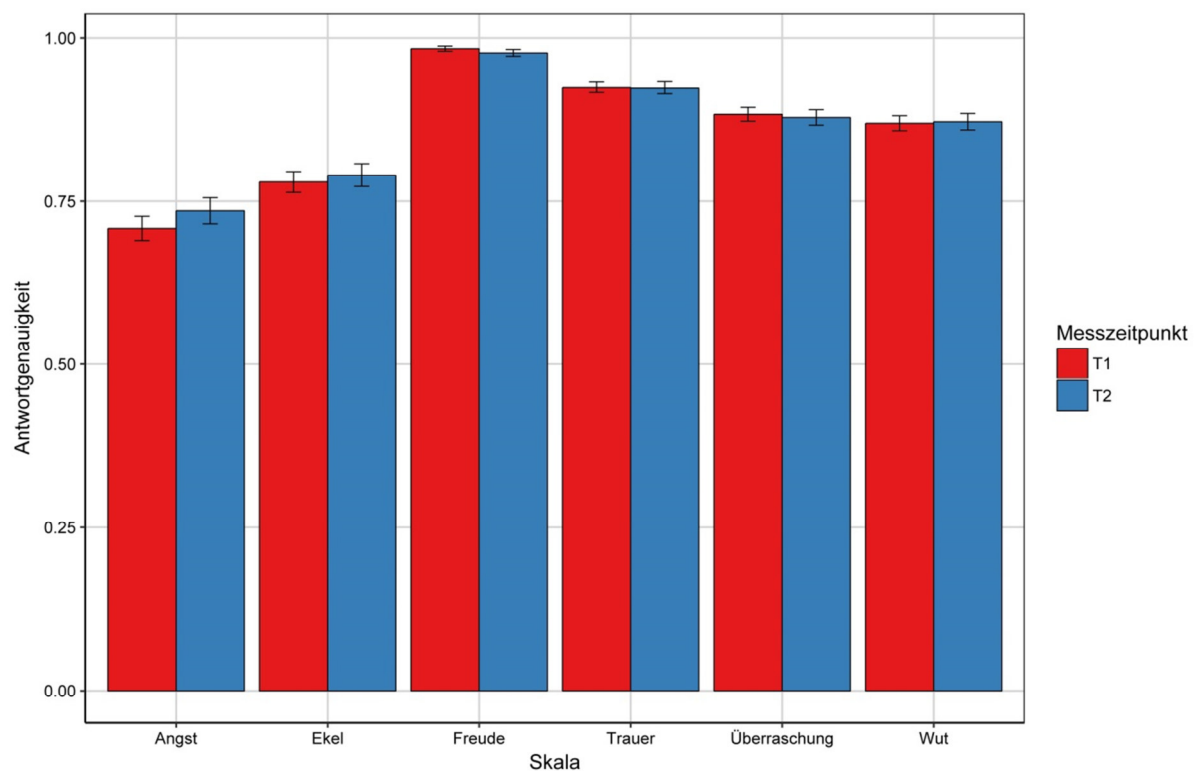


Abbildung 3. Mittelwerte und Standardfehler des Mittelwerts der Antwortgenauigkeit der Versuchspersonen des PCFAE nach Affektskala und Messzeitpunkt. Ein Wert von 1 entspricht 100% korrekter Affekterkennung.

Verwechslungen lagen für alle Items vor. Besonders häufig wurden jedoch die Items *Angst männlich Kind*, *Angst männlich Erwachsener* sowie *Ekel männlich Kind* falsch beantwortet. Tabelle 2 berichtet den Anteil der Verwechslungen je Item im prozentualen Verhältnis gegenüber den Gesamtantworten des jeweiligen Items, zusammengefasst für beide Messzeitpunkte. Bei der Interpretation ist zu berücksichtigen, dass in der Auswertung nicht zwischen tatsächlichen Verwechslungen aufgrund Falscherkennung des dargestellten Affekts und Fehlern aufgrund versehentlich falscher Affektauswahl unterschieden werden kann.

Tabelle 2
Gesamte Verwechslungen der Items für beide Messzeitpunkte.

Item				Item			
Skala	Geschlecht	Alter	Verwechslung %	Skala	Geschlecht	Alter	Verwechslung %
Angst	weiblich	Erwachsener	26.30	Trauer	weiblich	Erwachsener	7.47
		Kind	17.86			Kind	13.31
	männlich	Erwachsener	32.47		männlich	Erwachsener	4.87
		Kind	45.78			Kind	4.87
Ekel	weiblich	Erwachsener	18.18	Überr.	weiblich	Erwachsener	7.47
		Kind	15.91			Kind	14.61
	männlich	Erwachsener	21.75		männlich	Erwachsener	17.21
		Kind	39.61			Kind	7.79
Freude	weiblich	Erwachsener	0.65	Wut	weiblich	Erwachsener	10.06
		Kind	2.27			Kind	23.05
	männlich	Erwachsener	1.62		männlich	Erwachsener	4.22
		Kind	1.30			Kind	12.01

Hinweis. Die Tabelle berichtet den prozentualen Anteil der Verwechslungen im Verhältnis zu den Gesamtantworten je Item. Überr. = Überraschung

Tabelle 3 berichtet die deskriptiven Statistiken für beide Messzeitpunkte T1/T2 für die 6 Affektskalen des PCFAE sowie global für den Gesamttest. Berichtet werden die Werte für die AVn RT, RSI sowie ACC. Zusätzlich berichtet werden die Prüfung auf Vorliegen der Normalverteilung mittels Kolmogorov-Smirnov-Anpassungstest (KSA, nur *p*-Wert) und die Prüfung auf Varianzhomogenität mittels des gegen Verletzung der Normalverteilung robusteren Brown-Forsythe-Test (BF, nur *p*-Wert, zusätzlich zwischen Messzeitpunkt T1 und T2).

Tabelle 3

Deskriptive Statistiken, Normalverteilungs- und Varianzhomogenitätstests der Skalen des PCFAE für die Leistungsmaße RT, RSI und ACC.

Maß	Skala	Messzeitpunkt T1				Messzeitpunkt T2				T1 vs. T2
		<i>M</i>	<i>SD</i>	KSA (<i>p</i>)	BF (<i>p</i>)	<i>M</i>	<i>SD</i>	KSA (<i>p</i>)	BF (<i>p</i>)	BF (<i>p</i>)
RT	Angst	9.96	2.94	<.001	0.21	8.44	2.78	<.001	0.66	0.08
	Ekel	9.27	3.46	<.001	<.001	7.61	3.20	<.001	<.05	<.01
	Freude	7.15	2.71	<.001	0.09	6.44	2.80	<.001	0.05	0.32
	Trauer	10.46	3.38	<.001	<.001	8.83	3.30	<.001	<.001	0.44
	Ueberr.	8.20	2.77	<.001	0.13	7.01	2.82	<.001	0.53	0.36
	Wut	10.94	3.42	<.001	<.001	8.95	3.12	<.001	<.05	<.05
	global	9.33	3.39	<.001	<.001	7.88	3.15	<.001	<.001	<.001
RT cor	Angst	10.56	2.91	<.001	0.16	8.98	2.73	<.001	0.31	0.27
	Ekel	9.44	3.50	<.001	<.05	7.84	3.24	<.001	0.07	<.05
	Freude	7.16	2.72	<.001	0.07	6.43	2.79	<.001	<.05	0.34
	Trauer	10.49	3.45	<.001	<.001	8.90	3.30	<.001	<.001	0.21
	Ueberr.	8.14	2.93	<.001	0.09	6.94	2.77	<.001	0.74	0.10
	Wut	11.05	3.49	<.001	<.001	9.45	3.36	<.001	0.05	0.18
	global	9.44	3.48	<.001	<.001	8.07	3.25	<.001	<.001	<.001
RSI	Angst	2.08	1.78	<.001	0.64	1.82	0.96	<.001	0.34	<.01
	Ekel	1.95	1.51	<.001	0.31	1.70	0.74	<.001	0.23	<.01
	Freude	1.35	0.81	<.001	0.98	1.23	0.59	<.001	0.24	0.05
	Trauer	2.21	1.25	<.001	<.05	1.96	1.38	<.001	0.17	0.20
	Ueberr.	1.86	1.08	<.001	0.58	1.81	2.66	<.001	0.51	0.65
	Wut	2.10	1.01	<.001	<.05	1.88	0.82	<.001	0.05	<.01
	global	1.92	1.31	<.001	<.001	1.73	1.40	<.001	<.01	<.001
RSI cor	Angst	1.97	1.84	<.001	0.58	1.73	0.99	<.001	0.85	<.01
	Ekel	1.79	1.03	<.001	0.08	1.65	0.69	<.001	0.06	<.05
	Freude	1.32	0.74	<.001	0.60	1.20	0.55	<.001	0.58	<.05
	Trauer	2.10	1.18	<.001	<.01	1.87	1.33	<.001	0.12	0.23
	Ueberr.	1.83	1.26	<.001	0.53	1.64	0.88	<.001	0.97	0.12
	Wut	2.07	1.13	<.001	0.15	1.81	0.67	<.001	0.39	<.001
	global	1.84	1.26	<.001	<.001	1.65	0.92	<.001	0.09	<.001
ACC	Angst	0.71	0.36	<.001	<.001	0.74	0.36	<.001	<.001	0.96
	Ekel	0.78	0.33	<.001	<.001	0.79	0.33	<.001	<.001	0.85
	Freude	0.98	0.09	<.001	0.07	0.98	0.11	<.001	<.01	0.58
	Trauer	0.93	0.19	<.001	<.001	0.92	0.21	<.001	<.001	0.81
	Ueberr.	0.88	0.23	<.001	<.001	0.88	0.25	<.001	<.001	0.14
	Wut	0.87	0.26	<.001	<.001	0.87	0.26	<.001	<.001	0.91
	global	0.86	0.27	<.001	<.001	0.86	0.28	<.001	<.001	0.87

Hinweis. KSA = Kolmogorov–Smirnov-Anpassungstest (p-Wert). BF = Brown-Forsythe Test (p-Wert). RT/RSI = Reaktionszeit/ Reaktionssicherheit in Sekunden. RTcor/RSIcor = Reaktionszeit/ Reaktionssicherheit in Sekunden für korrekte Antworten. global = gesamter Test unabhängig der Skalen. ACC = Antwortgenauigkeit, ein Wert von 1 entspricht 100% korrekter Antworten. M = Mittelwert. SD = Standardabweichung. Ueberr. = Überraschung.

Die RT und die RSI waren über alle Affektskalen sowie global rechtsschief, die ACC linksschief verteilt, es lag keine Normalverteilung vor. Abbildung 4 stellt die Verteilung der RT grafisch dar. Weitere Verteilungsplots für RT und RSI nach Affektskala sowie für RSI und ACC für den Gesamttest sind in Anhang E dargestellt. Die Affektskalen wiesen bezüglich der RSI in der Regel Varianzhomogenität auf (Ausnahme: Trauer RSIcor zu Messzeitpunkt T1, Wut RSI zu beiden Messzeitpunkten). Nicht varianzhomogen verhielt sich die RT der Skalen Ekel (RT zu T1, T2; RTcor zu T1), Freude (RTcor zu T2), Trauer (RT und RTcor zu T1, T2), und Wut (RT zu T1, T2; RTcor zu T1). Ebenfalls nicht varianzhomogen war die ACC über alle Skalen (Ausnahme Freude zu T1). Zwischen den Messzeitpunkten T1 und T2 war innerhalb der Testskalen überwiegend Varianzhomogenität gegeben, die Skalen Ekel (RT, RTcor, RSI, RSIcor), Wut (RT, RSI, RSIcor), Angst (RSI, RSIcor) und Freude (RSIcor) stellten die Ausnahmen dar.

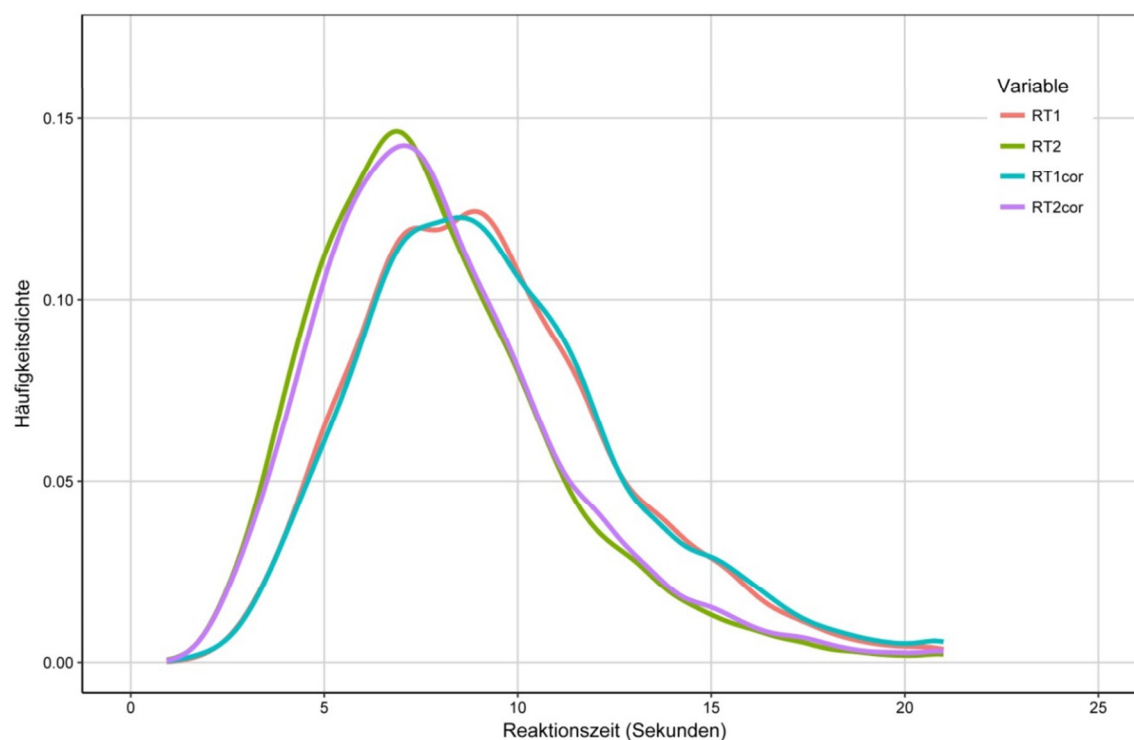


Abbildung 4. Verteilung der Reaktionszeitmittelwerte der Versuchspersonen des PCFAE (Gesamttest) nach Messzeitpunkt. RT1/2 = Reaktionszeit, alle Antworten, Messzeitpunkt T1/T2. RT1/2cor = Reaktionszeit für korrekte Antworten, Messzeitpunkt T1/T2.

Erwachsenenitems wurden mit kürzerer RT korrekt beantwortet (T1: $M = 8.65$, $SD = 1.43$; T2: $M = 7.49$, $SD = 1.17$) als Kinderitems (T1: $M = 10.36$, $SD = 1.91$; T2: $M = 8.74$, $SD = 1.66$). Die varianzanalytische Betrachtung (messwiederholte ANOVA; UV = Alter, within subjects, Faktorstufen Erwachsenenitems, Kinderitems; AV = mittlere RT korrekt beantworteter Items zu Messzeitpunkt T1, T2; unabhängige Betrachtung der Messzeitpunkte; keine Prüfung auf Sphärizität, da nur 2 Faktorstufen) zeigte einen mittleren Alterseffekt (Cohen, 1988) der Items zu beiden Messzeitpunkten (T1: $F(1,153) = 371.44$, $p < 0.001$, $\eta^2 = 0.13$; T2: $F(1,153) = 279.87$, $p < 0.001$, $\eta^2 = 0.06$)¹. Die Verletzung der Normalverteilungsannahme und die Varianzinhomogenität wurden bei der Wahl der statistischen Instrumente für die weitere Analyse und der Interpretation der Ergebnisse berücksichtigt.

Prüfung auf Testwiederholungs- und Übungseffekte

Das mögliche Vorliegen von Testwiederholungs- bzw. Übungseffekten wurde varianzanalytisch überprüft (messwiederholte ANOVA; UV = Messzeitpunkt, within subjects, Faktorstufen T1, T2; AV 1–3 = über alle Affektskalen gemittelte RT, RSI, ACC; keine Prüfung auf Sphärizität, da nur 2 Faktorstufen). Deskriptive Statistiken sind Tabelle 3 zu entnehmen. Die Berechnungsgrundlage für RT und RSI bildeten korrekt beantwortete Items.

Probanden beantworteten Items der Messung T2 im Mittel schneller und sicherer bei gleichbleibender Antwortgenauigkeit. Sowohl die RT ($F(1,153) = 154.53$, $p < .001$, $\eta^2 = 0.11$), als auch die RSI ($F(1,153) = 49.68$, $p < .001$, $\eta^2 = 0.05$) unterschieden sich signifikant zwischen Messzeitpunkt T1 und T2 mit mittlerer (RT) bzw. geringer (RSI) Effektstärke (Cohen, 1988), während kein signifikanter Unterschied zwischen den Messzeitpunkten für die ACC vorlag ($F(1,153) = 0.57$, $p = 0.45$, $\eta^2 = 0.00$).

¹ Die varianzanalytische Betrachtung beschränkt sich auf die Darstellung des Alterseffekts für die für diese Arbeit wichtigste AV der korrekten Reaktionszeit.

Itemanalyse

Trennschärfe der Items. Tabelle 4 berichtet die part - whole korrigierte

Trennschärfe r_{it} der Items des PCFAE je Leistungsmaß¹. Tabelle 5 fasst diese Trennschärfen gemittelt nach Affektskala zusammen (r_{st}). Die Items des PCFAE verhielten sich unter Verwendung des Leistungsmaßes RT für korrekt beantwortete Items zu Messzeitpunkt T1

Tabelle 4

Trennschärfe r_{it} der Items nach Affektskala, Leistungsmaß und Messzeitpunkt.

Skala	Item		RT		RSI		ACC	
	Geschlecht	Alter	T1	T2	T1	T2	T1	T2
Angst	weiblich	Erwachsener	0.59	0.67	0.10	0.50	0.24	0.47
		Kind	0.48	0.71	0.19	0.22	0.26	0.49
	männlich	Erwachsener	0.63	0.67	0.27	0.37	0.43	0.47
		Kind	0.44	0.70	0.28	0.41	0.23	0.29
Ekel	weiblich	Erwachsener	0.69	0.73	0.20	0.46	0.34	0.41
		Kind	0.57	0.78	0.33	0.50	0.23	0.34
	männlich	Erwachsener	0.69	0.71	0.20	0.53	0.33	0.35
		Kind	0.48	0.75	0.38	0.29	0.23	0.39
Freude	weiblich	Erwachsener	0.47	0.60	0.36	0.50	0.02	0.25
		Kind	0.51	0.70	0.34	0.42	0.25	0.22
	männlich	Erwachsener	0.57	0.61	0.25	0.49	0.05	0.14
		Kind	0.56	0.70	0.33	0.55	-0.04	0.17
Trauer	weiblich	Erwachsener	0.67	0.81	0.41	0.36	0.24	0.29
		Kind	0.53	0.64	0.18	0.23	0.09	0.27
	männlich	Erwachsener	0.75	0.79	0.29	0.37	0.30	0.33
		Kind	0.63	0.74	0.25	0.31	0.31	0.32
Überraschung	weiblich	Erwachsener	0.59	0.77	0.27	0.34	0.13	0.09
		Kind	0.56	0.63	0.31	0.31	0.09	0.14
	männlich	Erwachsener	0.51	0.73	0.20	0.42	-0.04	-0.06
		Kind	0.50	0.72	0.30	0.35	0.20	0.16
Wut	weiblich	Erwachsener	0.67	0.76	0.22	0.42	0.08	0.21
		Kind	0.37	0.53	0.14	0.25	0.29	0.42
	männlich	Erwachsener	0.60	0.77	0.28	0.42	0.13	0.13
		Kind	0.61	0.78	0.29	0.16	0.15	0.29

Hinweis. T1/2 = Messzeitpunkt 1/2. RT = Reaktionszeit. RSI = Reaktionssicherheit. ACC = Antwortgenauigkeit. Berechnungsgrundlage für RT und RSI sind korrekt beantwortete Items. R Paket Psych, Skript alpha.

¹ Die Trennschärfenberechnung mit parth - whole Korrektur entspricht der Produkt Moment Korrelation des Leistungsmaßes eines Testitems (part) mit dem Ergebnis des Gesamttest ohne das entsprechende Item (whole). Da dieses mit sich selbst mit einem Wert von 1 korreliert, würde die Itemtrennschärfe ohne Korrektur überschätzt werden (Bühner, 2011).

trennscharf (Trennschärfen überwiegend $r_{it} > 0.50$ oder angenähert). Lediglich das Item *Wut weiblich Kind* ($r_{it} = 0.37$) und *Angst männlich Kind* ($r_{it} = 0.44$) stellten geringfügige Ausreißer dar, erreichten jedoch die Mindestforderung an die Itemtrennschärfe von $r_{it} \geq 0.30$ (Bühner, 2011). Zu Messzeitpunkt T2 betrug die Trennschärfe aller Items $r_{it} \geq 0.60$ mit Ausnahme des Items *Wut weiblich Kind* ($r_{it} = 0.53$).

Die ACC als Leistungsmaß stellte sich zu Messzeitpunkt T1 wenig trennscharf dar. Insbesondere Items der Skalen Freude und Überraschung wiesen teilweise Trennschärfen von annähernd $r_{it} = 0$ auf. Dies ist bei der Skala Freude vermutlich auf einen Deckeneffekt zurückzuführen, da beinahe alle VP die Items richtig beantworteten. Zu Messzeitpunkt T2 erreichten die Skalen Angst, Ekel und Trauer Trennschärfen von $r_{st} \geq 0.30$, während die Skalen Freude, Wut und Überraschung aufgrund schlechter Werte einzelner Items unterhalb der erwünschten Trennschärfe blieben.

Tabelle 5

Gemittelte Trennschärfe r_{st} der Affektskalen nach Leistungsmaß und Messzeitpunkt.

Skala	RT		RSI		ACC	
	T1	T2	T1	T2	T1	T2
Angst	0.54	0.69	0.21	0.38	0.29	0.43
Ekel	0.61	0.74	0.28	0.45	0.28	0.37
Freude	0.53	0.65	0.32	0.49	0.07	0.20
Trauer	0.65	0.75	0.28	0.32	0.24	0.30
Überraschung	0.54	0.71	0.27	0.36	0.10	0.08
Wut	0.56	0.71	0.23	0.31	0.16	0.26

Hinweis. T1/2 = Messzeitpunkt 1/2. RT = Reaktionszeit. RSI = Reaktionssicherheit. ACC = Antwortgenauigkeit. Berechnungsgrundlage für RT und RSI sind korrekt beantwortete Items.

Itemhomogenität. Tabelle 6 berichtet die mittlere Inter-Item-Korrelation (MIC^1) für die Leistungsmaße des PCFAE nach Affektskala sowie für den Gesamttest als Maß der Itemhomogenität. Bezogen auf die korrekte RT zu Messzeitpunkt T1 war die MIC innerhalb

¹Für die Berechnung der MIC wurden zunächst alle Itempaare innerhalb der Affektskalen bzw. des Gesamttests korreliert (Produkt Moment Korrelation) und die resultierenden Korrelationen im Anschluss gemittelt (Cohen & Sverdluk, 2005).

der Affektskalen hoch (min $r = 0.44$, max $r = 0.59$) und bewegte sich für den Gesamttest mit $r = 0.35$ im Idealbereich zwischen $r = 0.20 - 0.40$ (Bühner, 2011). Zu Messzeitpunkt T2 war die MIC für alle Affektskalen mit $r \geq 0.60$ sehr hoch und überschritt für den Gesamttest mit $r = 0.52$ den für die Itemhomogenität gewünschten Bereich.

Bezogen auf die korrekte RSI und die ACC stellte sich die MIC zu Messzeitpunkt T1 für die meisten Skalen des PCFAE gering dar (RSI: MICmin $r = 0.11$, MICmax $r = 0.30$; ACC: MICmin $r = 0.06$, MICmax $r = 0.25$). Zu Messzeitpunkt T2 verhielten sich die Affektskalen bezüglich der RSI größtenteils homogen (MICmin $r = 0.17$, MICmax $r = 0.40$), während sich die Items bezüglich der ACC innerhalb der Affektskalen eher heterogen darstellten (MICmin $r = 0.14$, MICmax $r = 0.34$). Bezogen auf den Gesamttest verfehlte die MIC sowohl für RSI als auch ACC den Idealbereich.

Tabelle 6

Mittlere Inter-Item-Korrelation MIC (r) nach Leistungsmaß und Messzeitpunkt als Maß der Itemhomogenität innerhalb der Affektskalen und des Gesamttests.

Skala	RT		RSI		ACC	
	T1	T2	T1	T2	T1	T2
Angst	0.44	0.61	0.12	0.25	0.25	0.34
Ekel	0.52	0.65	0.11	0.27	0.23	0.29
Freude	0.59	0.66	0.30	0.40	0.06	0.18
Trauer	0.55	0.65	0.23	0.17	0.10	0.18
Überraschung	0.47	0.64	0.17	0.21	0.12	0.14
Wut	0.52	0.61	0.16	0.25	0.11	0.19
Gesamttest	0.35	0.52	0.11	0.18	0.06	0.11
<i>Hinweis.</i>	T1/2 = Messzeitpunkt 1/2. RT = Reaktionszeit. RSI = Reaktionssicherheit. ACC = Antwortgenauigkeit. Berechnungsgrundlage für RT und RSI sind korrekt beantwortete Items.					

Itemschwierigkeit. Tabelle 7 berichtet die zufallskorrigierte Itemschwierigkeit (P_{zk}). Die Korrektur erfolgte für 6 mögliche Antwortoptionen (entsprechend der 6 Basisaffekte) mit einer korrekten Antwort je Affekt. Die Berechnung des Schwierigkeitsindex erfolgte nach der Formel

$$P_{ZK} = \frac{N_R - \frac{N_F}{m-1}}{N}$$

mit N_R = Anzahl Probanden, die das Item richtig gelöst haben,

N_F = Anzahl Probanden, die das Item falsch gelöst haben,

N = Anzahl aller Probanden,

m = Anzahl der Wahlmöglichkeiten. (Bühner, 2011)

Tabelle 7

Zufallskorrigierte Itemschwierigkeit P_{ZK} und mittlere Schwierigkeit der Affektskalen nach Messzeitpunkt.

Item			P_{ZK}		Item			P_{ZK}	
Skala	Geschlecht	Alter	T1	T2	Skala	Geschlecht	Alter	T1	T2
Angst	weiblich	Erwachsener	0.69	0.70	Trauer	weiblich	Erwachsener	0.89	0.88
		Kind	0.76	0.80			Kind	0.79	0.79
	männlich	Erwachsener	0.59	0.63		männlich	Erwachsener	0.89	0.90
		Kind	0.46	0.49			Kind	0.89	0.90
MSk			0.62	0.65	MSk			0.87	0.87
Ekel	weiblich	Erwachsener	0.77	0.73	Überr.	weiblich	Erwachsener	0.87	0.88
		Kind	0.79	0.82			Kind	0.79	0.76
	männlich	Erwachsener	0.76	0.74		männlich	Erwachsener	0.77	0.75
		Kind	0.49	0.58			Kind	0.85	0.86
MSk			0.70	0.71	MSk			0.82	0.81
Freude	weiblich	Erwachsener	0.93	0.94	Wut	weiblich	Erwachsener	0.83	0.84
		Kind	0.91	0.89			Kind	0.69	0.70
	männlich	Erwachsener	0.94	0.93		männlich	Erwachsener	0.89	0.89
		Kind	0.94	0.93			Kind	0.81	0.80
MSk			0.93	0.92	MSk			0.80	0.81

Hinweis. T1/2 = Messzeitpunkt 1/2. MSk = mittlere Schwierigkeit der vorangegangenen Affektskala. Überr. = Überraschung.

Die Itemschwierigkeit des PCFAE bewegt sich zu beiden Messzeitpunkten im leichten bis mittleren Bereich ($P_{ZKmin} = 0.94$, $P_{ZKmax} = 0.46^1$), ohne substanzielle

¹ Eine niedrige Indexzahl kennzeichnet schwierigere Items, die nur von einem geringen Anteil der Personen gelöst werden, während eine hohe Indexzahl leichte Items kennzeichnet, welche von beinahe allen Personen gelöst werden, Wertebereich 0 – 1.

Veränderungen zwischen Messzeitpunkt T1 zu T2. Lediglich 2 Items (*Angst männlich Kind*, *Ekel männlich Kind*) wiesen eine höhere Itemschwierigkeit von $P_{ZK} < 0.50$ auf. Die mittlere Schwierigkeit der Affektskalen variierte zwischen $P_{ZKmin} = 0.93$ und $P_{ZKmax} = 0.62$ ebenfalls im leichten bis mittleren Schwierigkeitsbereich.

Dimensionalität / Explorative Faktorenanalyse. Berichtet werden die Ergebnisse der Hauptachsenanalyse der RT korrekt beantworteter Items zu Messzeitpunkt T1. Es wurden 2 Modelle exploriert, ein *Korrelierte Faktoren Modell* und ein *Bi-Faktor Modell* (Cho, 2016). Die Ladungsmuster der korrekten RSI und ACC zeigten keine inhaltlich gut interpretierbare Struktur.

Prüfung der Voraussetzungen. Die Voraussetzungsprüfung für die Durchführbarkeit der Faktorenanalyse umfasste eine Prüfung auf unverzerrte Zusammenhänge, substanzielle Korrelationen (Eignung der Korrelationsmatrix und Items für die Faktorenanalyse) sowie ausreichende Itemzahl und Itemreliabilität. Die einzelnen Prüfschritte und die Beurteilung der Cutoff-Werte orientierten sich an den Empfehlungen von Bühner (2011) sowie Field, Miles & Field (2012).

Die Bereinigung der Stichprobe zur Sicherstellung unverzerrter Zusammenhänge ist im Kapitel Methode Abschnitt Versuchspersonen beschrieben. Das Vorliegen substanzieller Korrelationen und damit der Eignung der Korrelationsmatrix für eine Faktorenanalyse wurde mithilfe des Kaiser-Meyer-Olkin Koeffizienten (KMO) geprüft. Die Eignung der einzelnen Items für die Durchführung der Faktorenanalyse wurde mithilfe des Measure of Sample Adequacy Koeffizienten (SMA) überprüft. Der Bartlett Test stellt bei Signifikanz sicher, dass alle Korrelationen der Matrix ungleich null sind.

Die auf Eignung geprüfte Korrelationsmatrix der korrekten RT zu Messzeitpunkt T1 wies einen beinahe sehr guten $KMO = 0.89$ auf, der Wertebereich der Items erstreckte sich von guten $SMA_{min} = 0.82$ bis sehr guten $SMA_{max} = 0.92$ Werten. Der Bartlett Test ist

signifikant mit $\chi^2(276) = 1929.79, p < .001$, setzt jedoch eine approximative Normalverteilung der Daten voraus. Diese war mit Blick auf die Werteverteilung und den KSA-Test zu hinterfragen. Daher sollte das Ergebnis des Bartlett Tests kritisch betrachtet werden. Trotz einer Determinante < 0.00001 fanden sich innerhalb der Korrelationsmatrix keine Itemkorrelationen von $r > 0.90$, sodass eine Redundanz von Items zu verneinen war (Field, Miles & Field, 2012).

Die Itemanzahl von 4 Items pro extrahierten Faktor entsprach dem geforderten Mindestmaß (Bühner, 2011). Ebenfalls kritisch war die Reliabilität einzelner Items bzw. deren Kommunalität als Mindestschätzung der Reliabilität von $h^2 < 0.60$ zu werten (siehe Tabellen 9 und 10). Trotz der für diese Bedingungen geringen Stichprobengröße von $N = 154$ waren die von Bühner (2011) aufgezeigten Durchführungskriterien bezüglich der Itemreliabilität noch ausreichend erfüllt, sodass unter Einbezug aller Prüfergebnisse die Faktorenanalyse insgesamt als durchführbar bewertet wurde.

Faktorenextraktion. Die Festlegung der Anzahl zu extrahierender Faktoren erfolgte mittels des Vergleichs verschiedener Extraktionskriterien (Courtney, 2013) in Kombination mit inhaltlichen Überlegungen. Die verwendeten Kriterien und die daraus abzuleitende Anzahl zu extrahierender Faktoren ist in Tabelle 8 dargestellt. Abbildung 5 veranschaulicht zusätzlich den Verlauf der Eigenwerte (Scree-Plot) und stellt die Ergebnisse der

Tabelle 8

Übersicht der empfohlenen Anzahl zu extrahierender Faktoren abhängig vom verwendeten Extraktionskriterium.

Extraktionskriterium	Anzahl zu extrahierender Faktoren
Parallelanalyse (Horn, 1965)	5
Kaiser-Kriterium (Kaiser & Dickmann, 1959)	3
MAP-Test (Velicer, 1976)	3
VSS-Kriterium (Revelle & Rocklin, 1979)	2
Scree-Test (Catell, 1966)	1 oder 3

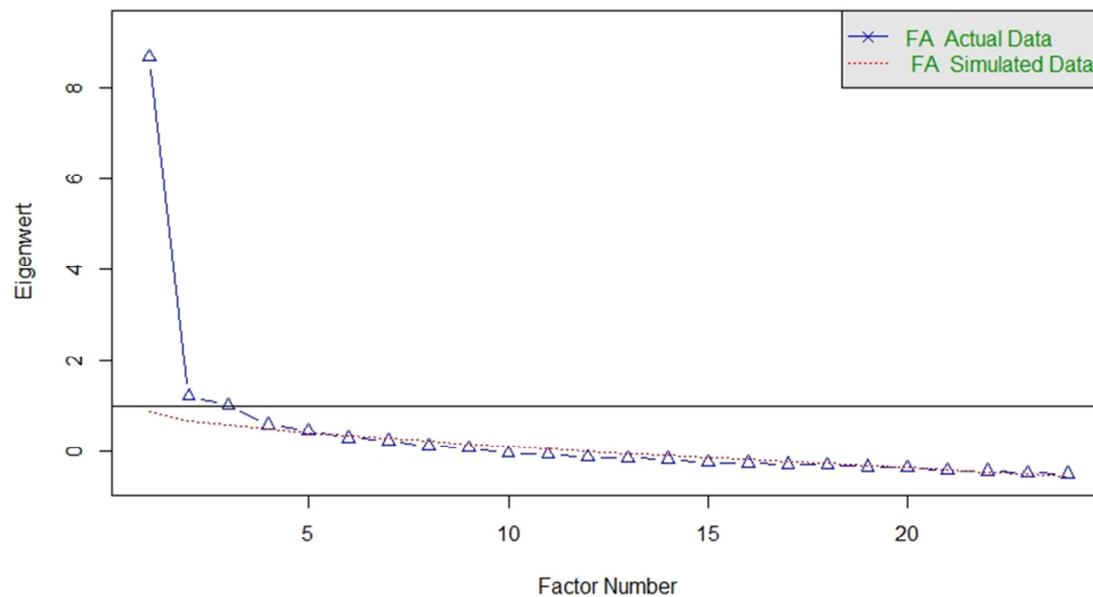


Abbildung 5. Scree-Plot und grafische Darstellung des Ergebnisses der Parallelanalyse für die korrekte Reaktionszeit zu Messzeitpunkt T1.

Parallelanalyse grafisch dar. Extrahiert wurden schlussendlich 6 Faktoren. Implikationen und konstituierende Überlegungen werden diskutiert.

Explorierte Modelle. Die Rotation der unrotierten Komponentenlösung zur Einfachstruktur erfolgte nach dem Oblimin Verfahren ($\gamma = 0^1$). Die Korrelationen der Faktoren sind Anhang G zu entnehmen. Die Faktorladungen und Kommunalitäten der Items der resultierenden Mustermatrix beider explorierter Modelle sind in Tabellen 9 und 10 berichtet. Im *Korrelierte Faktoren Modell* klären die 6 extrahierten Faktoren 57% der Gesamtvarianz auf, im *Bi-Faktor Modell* erklärt der gemeinsame Faktor 49% der Varianz. Aufgrund der Oblimin Rotation wird die Varianzaufklärung der Faktoren überschätzt. Die Eigenwerte der Faktoren im *Korrelierte Faktoren Modell* betragen für Faktor F1 = 0.55, F2 = 0.65, F3 = 1.51, F4 = 0.81, F5 = 1.24, F6 = 8.88, im *Bi-Faktor Modell* F1 = 0.83, F2 = 0.96, F3 = 1.73, F4 = 1.09, F5 = 1.05, F6 = 1.25 sowie für den gemeinsamen Faktor g = 6.67. Beide Modelle werden im Rahmen der Diskussion grafisch dargestellt.

¹ Entspricht der Quartimin Rotation, erlaubt maximale Korrelation der Faktoren untereinander (Jennrich, 2002)

Tabelle 9

Korrelierte Faktoren Modell: Faktorladungen λ mit Nebenladungen und Kommunalitäten der Items, korrekte Reaktionszeit, Messzeitpunkt T1, Oblimin rotierte Lösung.

Skala	Item		Faktor						h ²
	Geschlecht	Alter	F1	F2	F3	F4	F5	F6	
Angst	weiblich	Erwachsener	0.34						0.58
		Kind	0.58						0.49
	männlich	Erwachsener	-	0.32		0.33			0.50
		Kind	-				0.34		0.43
Ekel	weiblich	Erwachsener		0.39					0.57
		Kind		0.47					0.44
	männlich	Erwachsener		0.68					0.73
		Kind		0.52					0.39
Freude	weiblich	Erwachsener			0.77				0.62
		Kind			0.74				0.64
	männlich	Erwachsener			0.73				0.62
		Kind			0.67				0.60
Trauer	weiblich	Erwachsener				0.58			0.63
		Kind				0.74			0.60
	männlich	Erwachsener		0.32		0.34			0.64
		Kind				0.66			0.61
Überraschung	weiblich	Erwachsener	0.50				-		0.50
		Kind					0.70		0.62
	männlich	Erwachsener					0.64		0.53
		Kind					0.55		0.50
Wut	weiblich	Erwachsener						0.59	0.65
		Kind						0.65	0.46
	männlich	Erwachsener						0.53	0.57
		Kind						0.54	0.61

Hinweis. h² = Kommunalität des Faktors

Interpretation der Faktoren:

F1 = Faktor Angst

F2 = Faktor Ekel

F3 = Faktor Freude

F4 = Faktor Trauer

F5 = Faktor Überraschung

F6 = Faktor Wut

Abgetragen wurden Faktorladungen $\lambda > 0.30$, Nebenladungen sind durch rote Schrift gekennzeichnet. Berechnungsgrundlage: R Paket psych, Skript fa.

Tabelle 10

Bi-Faktor Modell: Faktorladungen λ mit Nebenladungen und Kommunalitäten der Items, korrekte Reaktionszeit, Messzeitpunkt T1, Oblimin rotierte Lösung.

Skala	Item		Faktor								
	Geschlecht	Alter	g	F1	F2	F3	F4	F5	F6	h²	
Angst	weiblich	Erwachsener	0.60	0.49							0.58
		Kind	0.42	0.46							0.49
	männlich	Erwachsener	0.57	0.42							0.50
		Kind	0.43	-					0.31		0.43
Ekel	weiblich	Erwachsener	0.64		0.41						0.57
		Kind	0.49		0.45						0.44
	männlich	Erwachsener	0.68		0.46						0.73
		Kind	0.38		-						0.39
Freude	weiblich	Erwachsener	0.64			0.63					0.62
		Kind	0.54			0.52					0.64
	männlich	Erwachsener	0.63			0.58					0.62
		Kind	0.55			0.57					0.60
Trauer	weiblich	Erwachsener	0.59				0.39				0.63
		Kind	0.53				0.48				0.60
	männlich	Erwachsener	0.64	0.31			-				0.64
		Kind	0.65				0.53				0.61
Überraschung	weiblich	Erwachsener	0.54					0.32			0.62
		Kind	0.59					0.55			0.62
	männlich	Erwachsener	0.49					0.39			0.53
		Kind	0.52					0.53			0.50
Wut	weiblich	Erwachsener	0.67						0.57		0.65
		Kind	0.46						0.56		0.46
	männlich	Erwachsener	0.57						0.38		0.61
		Kind	0.64						0.44		0.57

Hinweis. g = gemeinsamer Faktor h² = Kommunalität

Interpretation der Faktoren:

F1 = Faktor Angst

F2 = Faktor Ekel

F3 = Faktor Freude

F4 = Faktor Trauer

F5 = Faktor Überraschung

F6 = Faktor Wut

Abgetragen wurden Faktorladungen $\lambda > 0.30$, Nebenladungen sind durch rote Schrift gekennzeichnet. Berechnungsgrundlage: R Paket psych, Skript fa.

Reliabilitätsanalyse

Tabelle 11 berichtet die berechneten Reliabilitätskoeffizienten für die Leistungsmaße des PCFAE zur Schätzung der internen Konsistenz, der Split-Half Reliabilität und der Test-Retest Reliabilität (zeitliche Stabilität). Um neben der relativen zeitlichen Stabilität des Tests (Prüfung auf gleiche Rangfolge der VP von T1 zu T2 bezüglich der Leistungsmaße) ebenfalls die absolute zeitliche Stabilität unter Berücksichtigung von Testwertverbesserungen der VP bei Testwiederholung zu erfassen, wurde neben der Test-Retest Korrelation ρ_{tt} zusätzlich die Intraklassenkorrelation (ICC2k: gleiche Rater (VP) für jedes Item, zufällige Wahl der Rater, Mittelwerte als Grundlage, absolute Übereinstimmung zw. Messzeitpunkt T1 und T2, justiertes Modell) berechnet (Bühner, 2011; Wirtz & Caspar, 2002).

Tabelle 11

Reliabilitätskoeffizienten der Leistungsmaße des PCFAE nach Messzeitpunkt.

Koeffizient	RT		RSI		ACC	
	T1	T2	T1	T2	T1	T2
Cronbach's α ($\hat{\rho}_t$)	0.93	0.96	0.74	0.84	0.62	0.75
Guttman's λ_4	0.97	0.98	0.85	0.92	0.76	0.86
Revelle's β	0.85	0.92	0.57	0.74	0.40	0.51
McDonald's ω	0.95	0.97	0.90	0.93	0.69	0.81
McDonald's ω_h	0.75	0.82	0.62	0.69	0.30	0.43
Test-Retest $\hat{\rho}_{tt}$	0.77		0.65		0.67	
ICC(2,k)	0.77		0.35		0.60	

Hinweis. T1/2 = Messzeitpunkt 1/2. ICC = Intra-Klassen-Korrelation. RT = Reaktionszeit. RSI = Reaktionssicherheit. ACC = Antwortgenauigkeit. Reliabilitätskoeffizienten für RT und RSI wurden auf Grundlage korrekt beantworteter Items berechnet; R-Paket Psych, Skripte splitHalf, testRetest, omegah, ICC.

Bezogen auf die RT korrekt beantworteter Items zu Messzeitpunkt T1 wird die Reliabilität des PCFAE mittel bis hoch geschätzt, mit einer unteren Grenze der Split-Half

Reliabilität von $\beta = 0.85$ und einer oberen Grenze¹ von $\lambda_4 = 0.97$. Zu Messzeitpunkt T2 wurde die interne Konsistenz und die Split-Half Reliabilität des Tests als sehr hoch geschätzt. Die zeitliche Stabilität ist je nach angewandtem Kriterium als ausreichend ($\rho_{tt} < 0.80$; Jackson, 2011) bis gut ($\text{ICC2k} > 0.70$; Wirtz & Caspar, 2002) zu beurteilen. Die Messgenauigkeit des PCFAE unter den Leistungsmaßen der korrekten RSI und ACC stellt sich uneinheitlich dar. Eine Eignungsbeurteilung der jeweiligen Koeffizienten zur Reliabilitätsschätzung wird diskutiert.

Konstruktvaliditätsanalyse

Die Prüfung der Konstruktvalidität beschränkte sich auf die Prüfung der konvergenten bzw. diskriminanten Validität gegenüber den inhaltlich ähnlichen Selbstauskunftsverfahren 20-Item Toronto Alexithymia Scale (TAS-20) und Skalen zum Erleben von Emotionen (SEE) – Subskalen Erleben von Emotionsüberflutung (SEE EEU) und Erleben von Emotionsmangel (SEE EEM). Aufgrund des Ordinalskalenniveaus der Selbstauskunftsverfahren wurde Spearman's Rangkorrelationskoeffizient (ρ) verwendet. Um ein Vergleichskriterium zu generieren, welches auf Gesamttestebene Antwortgenauigkeit, Fehlerrate und Reaktionszeit bzw. Reaktionssicherheit in ein Leistungsmaß integriert, wurden 2 Leistungsquotienten gebildet:

$$RTquot = \frac{\text{Anzahl korrekter Antworten} + 1}{\text{Anzahl falscher Antworten} + 1} : \text{mittlere Reaktionszeit}$$

$$RSIquot = \frac{\text{Anzahl korrekter Antworten} + 1}{\text{Anzahl falscher Antworten} + 1} : \text{mittlere Reaktionssicherheit}$$

¹ Die obere und untere Grenze der Split-Half Reliabilität ergeben sich aus der günstigsten bzw. ungünstigsten Itemkombination der Testhalbierung. Dies wird im Kapitel Diskussion näher erläutert.

Die Ergebnisse der Korrelation der Leistungsquotienten des PCFAE sowie der Antwortgenauigkeit mit den Selbstauskunftsverfahren sind Tabelle 12 zu entnehmen. Der PCFAE weist diskriminante Validität zum TAS-20 sowie den Skalen des SEE auf. Weder der Leistungsquotient RT_{quot} noch RSI_{quot} korrelieren signifikant mit den beiden genannten Verfahren zur Emotionswahrnehmung. Die größte Korrelation des Leistungsmaßes ACC mit der TAS-20 zu Messzeitpunkt T2 ist annähernd signifikant mit $p = 0.07$. Zur Plausibilitätskontrolle wurde abschließend die Korrelation des TAS-20 mit den SEE Skalen berechnet. Die zeitliche Stabilität der Skalen des SEE und der TAS-20 zwischen Messzeitpunkt T1 und T2 unterlag Schwankungen (T1 zu T2: SEE(EEM) $\rho = 0.73$, $p < 0.001$; SEE(EEU) $\rho = 0.78$, $p < 0.001$; TAS-20 $\rho = 0.75$, $p < 0.001$).

Tabelle 12

Rangkorrelationen (Spearman's ρ) der Leistungsquotienten und ACC des PCFAE mit den Selbstauskunftsverfahren TAS-20 und SEE (Skalen EEU und EEM).

Korrelation (ρ)	TAS	T1			T2		
		RT_{quot}	RSI_{quot}	ACC	RT_{quot}	RSI_{quot}	ACC
TAS		0.06	0.06	0.02	-0.05	-0.04	-0.07
SEE (EEU)	0.53*	0.11	-0.01	-0.06	-0.09	-0.06	-0.10
SEE (EEM)	0.39*	-0.05	0.09	0.08	-0.05	-0.10	-0.07

Hinweis. T1/2 = Messzeitpunkt 1/2. RT_{quot} = Leistungsquotient für die Reaktionszeit. RSI_{quot} = Leistungsquotient für die Reaktionssicherheit. ACC = Antwortgenauigkeit. * = Korrelation signifikant mit $p < .05$.

Diskussion

In der vorliegenden Untersuchung wurde der computergestützte Test zur Erfassung der Affektwahrnehmungskompetenz PCFAE hinsichtlich der Güte seiner Items, der Messgenauigkeit, der zeitlichen Stabilität und der Konstruktvalidität analysiert und seine Dimensionalität exploriert. Hierfür wurden 2 Messungen mit einem zeitlichen Abstand von 4 Wochen und $N = 154$ Versuchspersonen durchgeführt. Betrachtet wurden unter Berücksichtigung des Geschwindigkeits-Genauigkeits-Ausgleichs die Leistungsmaße Reaktionszeit RT korrekt beantworteter Items, Reaktionssicherheit RSI korrekt beantworteter Items sowie die Antwortgenauigkeit ACC.

Die Leistungsmaße des PCFAE sind bei RT und RSI rechtsschief (Tendenz zu kürzeren Reaktionszeiten) und bei ACC linksschief (Tendenz zu mehr korrekten Antworten) verteilt. Eine Normalverteilung ist für die Daten nicht gegeben. Die Verteilung der Leistungsmaße weist auf einen eher einfachen Test hin, welcher schnell und mit einer hohen Anzahl korrekter Antworten zu lösen ist. Die Stichprobe weist eine hohe Fähigkeitsausprägung in der Testaufgabe auf, sodass die zur Verfügung stehende Varianz vergleichsweise gering ausfällt.

Die Leistungsmaße des PCFAE sind sowohl für den Gesamttest als auch innerhalb bestimmter Affektskalen varianzinhomogen. Für den Gesamttest ist dies erwartungsgemäß, da die Skalen unterschiedliche Affekte messen und sich daher in der Varianzausprägung unterscheiden sollten. Innerhalb der Affektskalen wäre jedoch Varianzhomogenität zu erwarten, sofern alle Items den dargestellten Affekt in gleicher Weise messen. Die Mittelwerte und Standardabweichungen je Item zeigen deskriptiv, dass innerhalb der varianzinhomogenen Skalen bei Kinderitems im Vergleich zu Erwachsenenitems längere Reaktionszeiten sowie eine größere Streuung auftreten. Dies betrifft die aversiven Affekte Ekel, Trauer und Wut. Varianzanalytisch wurde zu beiden Messzeitpunkten ein mittlerer

Alterseffekt zwischen Kinder- und Erwachsenenitems nachgewiesen (zu Messzeitpunkt T1 mit Tendenz zu einem großen Effekt), welcher für die gefundene Varianzinhomogenität ursächlich scheint. Dies entspricht ebenfalls dem Befund von Hahn (2016), während Jansen (2018a; unveröffentlicht) für den Affekt Wut innerhalb eines PCFAE-ähnlichen Paradigmas keinen Alterseffekt der Items nachweisen konnte. Neben der möglicherweise mangelnden Affektdarstellungsgüte des Stimulusmaterials kann eine vermutete grundsätzlich verminderte Responsivität auf aversive Affekte in Kindergesichtern den Alterseffekt innerhalb des Stimulusmaterials ebenfalls erklären (Hahn, 2016; Marsh et al., 2005; Seidel et al., 2010). Zwischen Messzeitpunkt T1 und T2 herrschte für das am geeignetsten beurteilte Leistungsmaß der korrekten RT überwiegend Varianzhomogenität innerhalb der Affektskalen. Die gefundene Varianzinhomogenität schwächt im Rahmen dieser Untersuchung die Aussagekraft jener statistischer Verfahren, die auf dem linearen Modell basieren (Wentura & Pospeschill, 2015).

Die VP erkannten die dargebotenen Affekte korrekt im Mittel nach 9,44 Sekunden zu Messzeitpunkt T1 und nach 9,07 Sekunden zu Messzeitpunkt T2 mit einer Reaktionssicherheit von im Mittel 1,92 Sekunden zu T1 abnehmend auf 1,73 Sekunden zu T2 bei gleichbleibender mittlerer Antwortgenauigkeit von 86% richtig gelöster Items. Nach Affektskala wurde Freude am schnellsten und besten erkannt sowie am wenigsten verwechselt, es trat ein Deckeneffekt auf. Die hohe Erkennungsleistung für den Affekt Freude ist ein typischer Befund (Hahn, 2016; Leppänen & Hietanen, 2004; Seidler, 2008, Wegrzyn et al., 2017). Überraschung und Ekel wurden mittelschnell, Angst, Wut und Trauer am langsamsten erkannt. Trauer, Überraschung und Wut wurden vergleichsweise korrekt erkannt, während Angst und Ekel am wenigsten korrekt erkannt wurden.

Die Erkennungsraten und -geschwindigkeiten der Affekte entsprechen den bisherigen Befunden zum PCFAE (Hahn, 2016). Die beobachteten Unterschiede zwischen

den Affektskalen wurden in ihrer Ausprägung ebenfalls für vergleichbare Stimulussets berichtet (KDEF Bilderset: Calvo & Lundqvist, 2008; Goeleven et al., 2008). Die Reaktionszeitunterschiede zwischen den Affekten lassen sich möglicherweise durch eine tiefere und damit ressourcenaufwändigere Verarbeitung aversiver Affekte erklären (Jansen, 2018a; unveröffentlicht). Die Unterschiede der Affekterkennungsraten sind durch die Verwechslung der betreffenden Affekte mit anderen Affekten bedingt. Abbildung 6 stellt diese Verwechslungen grafisch dar. Neben einem vermuteten Anteil bei der Testbearbeitung versehentlich getätigter Falschwahl des betreffenden Affekts sind die prozentual höher vertretenen Verwechslungen der Affekte Angst und Überraschung, Ekel und Wut sowie Ekel und Trauer mit hoher Wahrscheinlichkeit systematischer Natur. Unklar ist, ob diese Verwechslungen grundsätzlich durch mimische Ähnlichkeit der Affekte, durch mangelnde Darstellungsgüte der affektspezifischen Items oder einer Kombination dieser Faktoren

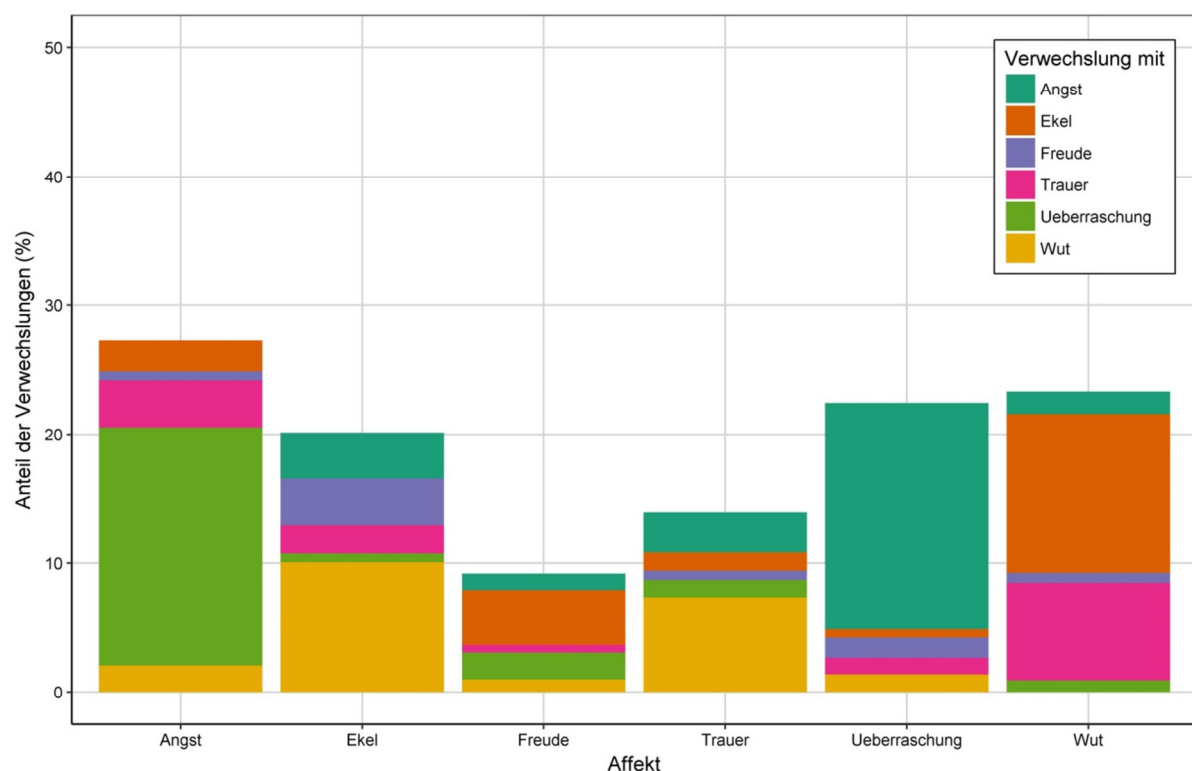


Abbildung 6. Dargestellt ist der prozentuale Anteil gegenüber den Gesamtantworten der Verwechslungen des jeweiligen Affekts mit den restlichen Basisaffekten. Die Berechnung erfolgte auf Affektskalenebene kombiniert für beide Messzeitpunkte.

bedingt sind. Belegt sind systematische Verwechslungen vor allem für die Affekte Angst und Überraschung (Adolphs, 2002; Palermo & Coltheart, 2004; Wegrzyn et al., 2017) sowie Ekel und Wut (Pochedly, Widen & Russell, 2012), während die Verwechslung von Trauer und Wut bisher nicht explizit berichtet wurde. Die gegenüber den restlichen Testitems deutlich höhere Verwechslungsrate der Items *Angst männlich Kind* (45,78%), *Angst männlich Erwachsener* (32,47%) sowie *Ekel männlich Kind* (39,61%) sprechen jedoch für eine Beteiligung der Affektdarstellungsgüte dieser Items an den Verwechslungen.

Testwiederholungseffekte

Die abnehmenden Reaktions- und Reaktionssicherheitszeiten von Messzeitpunkt T1 zu T2 bei gleichbleibender Antwortgenauigkeit weisen auf einen Testwiederholungseffekt hin. Dieser Effekt ist bezüglich der korrekten RSI klein, zeigt jedoch eine Tendenz hin zu einem mittleren Effekt. Bezüglich der korrekten RT ist ein mittlerer Effekt mit Tendenz zu einem großen Effekt gegeben. Die Konzeption des Tests – jedes Item beginnt mit einem neutralen Gesichtsausdruck – reicht folglich nicht aus, Testwiederholungseffekte zu unterbinden. Die Differenzierung der Items in Alter und Geschlecht, und damit 4 innerhalb einer Affektskala wesentlich unterschiedlicher Items, könnte eine unterschiedliche Aufmerksamkeitszuwendung zwischen Messzeitpunkt T1 und T2 bedingt haben. Möglicherweise wurden initial alle Aspekte des Items erfasst. Dies würde eine komplexere kognitive Verarbeitung erfordern und könnte in der Folge zu einer verlängerten RT zu Messzeitpunkt T1 geführt haben. Zu Messzeitpunkt T2 waren die VP mit dem Paradigma vertraut und sie könnten demnach die Aufmerksamkeit selektiver auf die Affekterkennung gerichtet haben, während Alter und Geschlecht des Stimulus nicht weiter berücksichtigt wurden. Die kürzere RSI zu T2 ließe sich ebenfalls durch eine zunehmende Paradigmengewöhnung erklären, welche ein schnelleres Zurechtfinden im Auswahlmenü bedingt haben könnte.

Ein motivationaler Einfluss scheint aufgrund der gleichbleibenden ACC unwahrscheinlich. Hätten die VP zu T1 noch versucht, möglichst genau der experimentellen Aufgabe nachzukommen und zu T2, das Experiment schnell hinter sich zu bringen, hätte dies vermutlich eher zu einer abnehmenden Antwortgenauigkeit im Sinne des Geschwindigkeits-Genauigkeits-Ausgleichs geführt. Ebenso wenig wahrscheinlich ist aufgrund nur eines Übungstrials (T1) ein nachhaltiger Übungseffekt auf die Fähigkeit der Affekterkennung (Drusch et al., 2014; Popova et al., 2014). Die Erinnerung an den Verlauf der affektspezifischen Darstellung stellt, trotz des Beginns der Itempräsentation mit einem neutralen Gesichtsausdruck, neben der erstgenannten Paradigmengewöhnung die wahrscheinlichste Erklärung für den Testwiederholungseffekt dar.

Für den diagnostischen Einsatz ist der gefundene Wiederholungseffekt dann relevant, wenn keine einmalige Eingangsdiagnostik zur Erfassung der Affektwahrnehmungskompetenz durchgeführt werden soll, sondern die Verbesserung der Fähigkeit nach erfolgter Therapie oder nach Verwendung von Trainingstools beurteilt wird. Ein solches Trainingstool soll mit der *Fühl mal App* (Jansen, 2018b; unveröffentlicht) unter Verwendung des Stimulusmaterials des PCFAE etabliert werden. Dessen Wirksamkeit und Konzeption ist auf Grundlage des gefundenen Wiederholungseffekts zu hinterfragen.

Aufgrund der Verwendung gemittelten Stimulusmaterials erzielt der PCFAE im Gegensatz zur Verwendung individueller Einzelbilder gegenüber anderen Inventaren eine geringere Stimulusdiversität von bisher lediglich 4 unterschiedlichen Items pro Affektskala. Für diagnostische Einfachtestungen wird hierdurch eine wenig komplexe und schnelle Durchführbarkeit erreicht. Sobald eine wiederholte Testung vorgesehen ist, sind möglicherweise andere Verfahren aufgrund wesentlich größerer Stimulusvielfalt und der Möglichkeit der Erstellung von Parallelversionen besser geeignet. Mögliche Verbesserungen des PCFAE Stimulusmaterials durch Abstufung der maximalen Affektausprägung werden in

der Folge diskutiert. Die Beurteilung der zeitlichen Stabilität des PCFAE wird durch die Übungeffekte eingeschränkt, es fand eine methodische Kompensation durch Verwendung der Intra-Klassen-Korrelation statt.

Itemanalyse

Im Rahmen der Itemanalyse wurden die 3 Leistungsmaße des PCFAE bezüglich Trennschärfe, Itemhomogenität sowie zufallskorrigierter Itemschwierigkeit überprüft und die Dimensionalität des Tests exploriert.

Trennschärfe und Itemschwierigkeit. Die Trennschärfe als wichtigster Itemkennwert (Fisseni, 2004) gibt an, wie gut ein Item in der Lage ist, zwischen Personen geringer und hoher Merkmalsausprägung zu unterscheiden. Bezogen auf die RT korrekt beantworteter Items zu Messzeitpunkt T1 erfüllen alle Items des PCFAE die Mindestforderung an die Itemtrennschärfe von $r_{it} \geq 0.30$, 22 der 24 Testitems erfüllen die Idealforderung von $r_{it} \geq 0.50$ oder nähern sich dieser an (Bühner, 2011). Zu Messzeitpunkt T2 lagen alle Items wesentlich oberhalb dieser Idealforderung mit hohen Trennschärfen bis zu $r_{it} = 0.78$. Bezogen auf die RSI korrekt beantworteter Items lagen zu Messzeitpunkt T1 nur wenige Items oberhalb der Mindestanforderung an die Trennschärfe. Diese wurde erst zu Messzeitpunkt T2 für die meisten Items erreicht, während 5 Testitems noch unterhalb $r_{it} = 0.30$ lagen. Bezogen auf die ACC lagen zu Messzeitpunkt T1 lediglich 6 Items in der Nähe oder oberhalb der Mindestanforderung an die Trennschärfe, während sich 7 Items als nicht trennscharf erwiesen. Zu Messzeitpunkt T2 erreichten 14 der 24 Testitems eine akzeptable Trennschärfe, allerdings lagen 2 Items noch bei einer nicht akzeptablen Trennschärfe von $r_{it} < 0.10$.

Die gemittelten Trennschärfen der Affektskalen ergaben ein vergleichbares Bild. Während die mittleren Trennschärfen der Affektskalen bezogen auf die korrekte RT zu Messzeitpunkt T1 für alle Skalen bei $r_{it} > 0.50$, zu Messzeitpunkt T2 sogar bei sehr guten $r_{it} >$

0.70 oder angenähert lagen, erreichten bezogen auf die korrekte RSI zu Messzeitpunkt T1 nicht alle Skalen die Mindestanforderung an die Trennschärfe. Für die ACC verblieben zu T1 die Affektskalen Freude und zu T1 und T2 Überraschung im nicht akzeptablen Bereich von $r_{it} = 0.10$ oder geringer, wobei die geringe Trennschärfe der Items der Affektskala Freude auf einen Deckeneffekt zurückzuführen ist. Beinahe alle Teilnehmer konnten diese Items korrekt beantworten.

Der Vergleich der Trennschärfen für die Leistungsmaße des PCFAE zeigt, dass die Reaktionszeit korrekt beantworteter Items als Leistungsmaß sehr gut in der Lage ist, zwischen Personen hoher und niedriger Fähigkeitsausprägung zu unterscheiden. Zwar erfüllt die Reaktionssicherheit korrekt beantworteter Items ebenfalls überwiegend die Mindestanforderungen an die Trennschärfe, jedoch werden akzeptable Werte erst mit Gewöhnung an das Testparadigma zu Messzeitpunkt T2 erreicht. Verbesserungen wären möglicherweise durch dem Test vorgeschaltete Übungsdurchgänge erzielbar, allerdings steht mit der korrekten RT ein Leistungsmaß zur Verfügung, das auch im Erstversuch und ohne Paradigmengewöhnung in der Lage ist, für beinahe alle Items eine sehr gute Trennschärfe zu erreichen. Gleiches gilt auf Affektskalenebene für die gemittelten Trennschärfen. Die ACC eignet sich ohne Einbezug der RT nur schlecht zur Unterscheidung zwischen Personen hoher und niedriger Merkmalsausprägung. Auch im Vergleich zur Urversion des PCFAE – Seidler (2008) berichtet unter Verwendung der ACC für den PCFAE Trennschärfen im Bereich von $r_{it} = 0.27$ bis 0.41 – führt die Verwendung der korrekten RT zu einer erheblich verbesserten Unterscheidung zwischen Personen hoher und niedriger Merkmalsausprägung.

Die zufallskorrigierte Itemschwierigkeit des PCFAE variierte zwischen $P_{ZK} = 0.62/0.65$ (Messzeitpunkt T1/T2) für die schwierigste Affektskala Angst und $P_{ZK} = 0.93/0.92$ (T1/T2) für die leichteste Affektskala Freude. Konsistent mit den Befunden von Hahn (2016), Jansen (2018a; unveröffentlicht), Leppänen & Hietanen, (2004) und Seidler (2008) nimmt

der dargestellte Affekt Einfluss auf die Itemschwierigkeit mit der Reihung der Affektskalen von leicht zu schwer: Freude < Trauer < Überraschung < Wut < Ekel < Angst. Lediglich 2 Items erzielten höhere Itemschwierigkeiten von $P_{ZK} < 0.50$ (*Angst männlich Kind*: $P_{ZK} = 0.46$ zu T1, *Ekel männlich Kind*: $P_{ZK} = 0.49$ zu T1). Der PCFAE ist in seiner aktuellen Version somit weiterhin als leicht bis mittelschwer zu bewerten und erreicht keine substanzielle Veränderung der Testschwierigkeit im Vergleich zu seiner Urversion (Seidler, 2008). Aufgrund der geringen Schwierigkeit ist anzunehmen, dass der PCFAE unter ausschließlicher Verwendung der ACC nur schlecht zwischen Personen mit hoher Fähigkeitsausprägung differenzieren kann. Die Itemschwierigkeit eines Tests sollte im Idealfall zwischen $P_{ZK} = 0.15$ und 0.85 variieren, um eine gute Abdeckung aller Fähigkeitsausprägungen zu erreichen (Bühner, 2011). Ungeachtet dessen ist die Schwierigkeit des PCFAE mit der Schwierigkeit anderer Verfahren vergleichbar. Diese liegt im Mittel bei $P = 0.70 - 0.88$ (Egger et al., 2011).

Abbildung 7 bildet die Items des PCFAE als Funktion von Itemschwierigkeit und Trennschärfe in Bezug auf die korrekte RT als Leistungsmaß ab. Deutlich wird die Affektabhängigkeit der Itemschwierigkeit durch die Bildung lose zusammenhängender Cluster. Kinderitems weisen im Vergleich zu Erwachsenenitems eine Tendenz zu geringeren Trennschärfen und höheren Schwierigkeiten auf (insbesondere die Items WKW, MKA, MKE, WKA, siehe Abbildung; Schema der Itembenennung gemäß Anhang I). Die bei der Erstellung der Kinderitems uneinheitliche Anzahl eingeflossener Originalbilder und die nicht standardisierte Situation bei deren Erstellung (Franz et al., 2018; unveröffentlicht; Hahn, 2016; Rampoldt, 2014) könnte sich ungünstig auf die Affektdarstellungsgüte ausgewirkt haben. Dieser mögliche Einfluss sollte durch die Verwendung der korrekten RT im Gegensatz zur Verwendung der ACC abgeschwächt auftreten, was sich in akzeptablen bis guten Trennschärfen der betreffenden Kinderitems auswirkt. Eine Überprüfung der Items

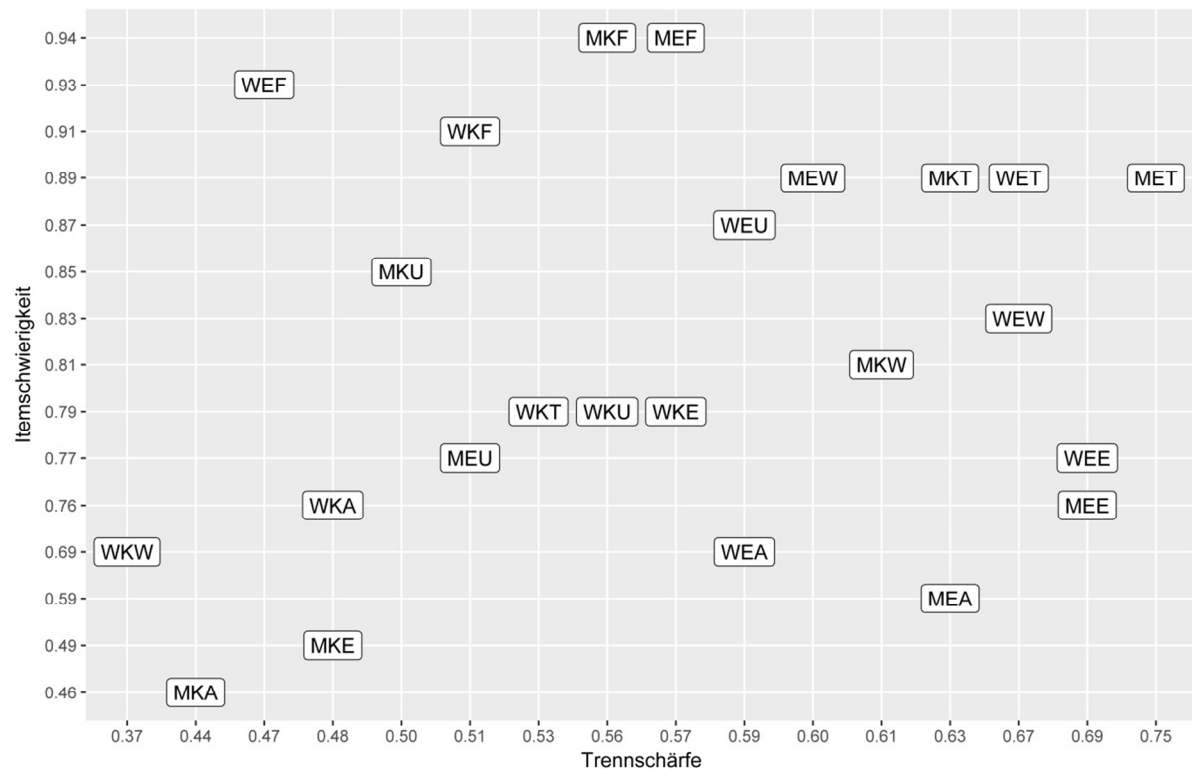


Abbildung 7. Streudiagramm der Items, Funktion Itemschwierigkeit x Trennschärfe für die korrekte RT zu Messzeitpunkt T1. Schema der Itembenennung siehe Anhang I.

könnte die bereits guten Werte jedoch noch weiter verbessern. Um eine breitere Abdeckung hinsichtlich der Itemschwierigkeit zu erreichen, sollten weitere Modifikationen des PCFAE eine Abstufung der maximalen Affektausprägung vorsehen. Zwar kann die Differenzierung der Affekterkennungsfähigkeit unabhängig von der Antwortgenauigkeit über Unterschiede der Reaktionszeit erfolgen, doch würde eine gestufte maximale Affektausprägung die Affektabhängigkeit der Itemschwierigkeit auflösen und die auswertbare Gesamtvarianz des Tests erhöhen. Dies würde eine bessere Fähigkeitsdifferenzierung für Störungsbilder ermöglichen, die mit einer affektspezifischen Einschränkung der Affektwahrnehmung einhergehen oder bei denen lediglich die Genauigkeit der Affekterkennung, nicht aber die Reaktionszeit eingeschränkt ist. (Jusyte, Gulewitsch & Schönenberg, 2017).

Itemhomogenität. Unter Verwendung der korrekten RT stellen sich alle Items des PCFAE innerhalb der Affektskalen vergleichsweise homogen dar. Dies entspricht der Erwartung, da lediglich Geschlecht und Alter, nicht aber der Affekt der Items variieren. Im

Gesamttest ist die Itemhomogenität als ideal zu werten. Unter Verwendung der korrekten RSI zu Messzeitpunkt T1 erfüllen auf Skalenebene lediglich die Skalen Freude und Trauer, unter Verwendung der ACC die Skalen Angst und Ekel die Anforderungen einer optimalen Itemhomogenität zwischen $r = 0.20$ und $r = 0.40$ (Bühner, 2011). Im Gesamttest stellen sich die Items unter den beiden letztgenannten Leistungsmaßen deutlich heterogen dar, sodass die Itemhomogenität unter Verwendung der RSI und ACC als schlecht einzuschätzen ist. Die deutlichen Unterschiede zwischen den Leistungsmaßen RT sowie RSI und ACC lassen vermuten, dass die beiden letztgenannten Maße durch konfundierende Einflüsse (beispielsweise Zeitlassen bei der Affektauswahl, systematische Verwechslungen) verzerrt sind und somit die Affekterkennungsfähigkeit nicht gut abbilden können. Demgegenüber birgt die innerhalb der Affektskalen vergleichsweise hohe Homogenität der Items bezogen auf die korrekte RT die Gefahr der Redundanz von Items, wobei jedoch keine Itemkorrelationen $r > 0.90$ auftraten (Redundanzkennzeichen nach Field, Miles & Field, 2012).

In der zusammenfassenden Gesamtbewertung zeichnet sich die Verwendung der korrekten Reaktionszeit als Leistungsmaß gegenüber der Verwendung der korrekten Reaktionssicherheit oder der Antwortgenauigkeit durch eine gute bis sehr gute Trennschärfe aus. Die korrekte RT ist weiterhin in der Lage, in Fähigkeitsbereichen zu differenzieren, in denen die ACC aufgrund der geringen Testschwierigkeit limitiert ist: Selbst bei gleicher Rate korrekter Antworten bildet sich eine höhere Fähigkeitsausprägung durch schnellere Reaktionszeiten ab. Bei Verwendung der korrekten RT agieren die Items des PCFAE auf Skalenebene erwartungsgemäß sehr homogen, auf Testebene ist die Itemhomogenität im Idealbereich. Die Verwendung der korrekten RT als schnell und einfach verfügbares Leistungsmaß zur Beurteilung der Fähigkeitsausprägung ist der Verwendung der Maße RSI

und ACC deutlich überlegen und sollte als das geeignetere Leistungsmaß des PCFAE den anderen Maßen vorgezogen werden.

Dimensionalität. Die Dimensionalität des PCFAE wurde mittels Hauptachsenanalyse unter Verwendung der Oblimin Rotation exploriert, welche eine maximale Korrelation der Faktoren untereinander zulässt. Eine Replikation der von Seidler (2008) aufgezeigten Methodik zur Dimensionalitätsexploration mittels Hauptkomponentenanalyse war aufgrund der Voraussetzungsverletzung (keine Normalverteilung der Daten) nicht möglich. Weiterhin wurde die Hauptachsenanalyse als das geeignetere Verfahren beurteilt¹. Für das gewählte Rotationsverfahren sprach die Annahme, dass die 6 Basisaffekte im Sinne einer generalisierten Affekterkennungsfähigkeit miteinander in Zusammenhang stehen und der Befund, dass die Basisaffekte entgegen der Annahme von Ekman & Davidson (1994) keine vollständige Unabhängigkeit aufweisen (fazial-motorische Gemeinsamkeiten, systematische Verwechslungen; Palermo, Bonnar & Gosselin, 2002; Pochedly, Widen & Russel, 2012; Wegrzyn et al., 2017).

Die Voraussetzungsprüfung erlaubte die Durchführung der Faktorenanalyse, jedoch ist die Anzahl von lediglich 4 Items pro extrahierten Faktor und die geringe Stichprobengröße in Verbindung mit teilweise geringen Kommunalitäten der Items kritisch zu beurteilen. Grundlage für die Faktorenanalyse stellte die RT korrekt beantworteter Items zu Messzeitpunkt T1 dar. Sowohl RSI als auch ACC wurden ebenfalls zur Dimensionalitätsexploration in Betracht gezogen, erzielten jedoch keine inhaltlich gut interpretierbare Faktorstruktur. Bezugnehmend auf die Ergebnisse der Itemanalyse welche die korrekte RT als das geeignetste Leistungsmaß des PCFAE herausstellt, wurden die auf RSI und ACC bezogenen Lösungen verworfen.

Anhand der Ergebnisse der verschiedenen Extraktionskriterien, welche die auf

¹ Bühner (2011) verweist darauf, dass die Hauptkomponentenanalyse zur Datenreduktion dient. Sollen die Daten mit latenten Faktoren erklärt werden, ist eine explorative Faktorenanalyse durchzuführen.

Datengrundlage optimale Anzahl zu extrahierender Faktoren ermittelten¹ wurden ebenfalls Lösungen mit 3 Faktoren exploriert. Plausibel wäre eine Aufteilung von Alter, Geschlecht und Affekt der Items in ein 3-faktorielles Modell gewesen. Die Begutachtung der Faktorstruktur ergab jedoch eine Zusammenfassung der mimisch ähnlichen Affekte Ekel und Wut sowie Angst und Überraschung unter einem Faktor, was als weiterer Hinweis auf eine Verwechslung dieser Affekte gewertet werden kann.

Inhaltliche Überlegungen berücksichtigten weiterhin die explizite Nennung von 6 Basisaffekten (Ekman, 2003; Ekman & Cordaro, 2011; Ekman, Friesen & Ellsworth, 1982), welche sich bei hinreichender Messgenauigkeit des Tests und der Qualität der Affektdarstellungsgüte der Items trotz mimischer Ähnlichkeiten in eigenen Faktoren darstellen sollten. Weiterhin wurde berücksichtigt, dass der PCFAE die Affekterkennungsfähigkeit für jeden der 6 Basisaffekte unabhängig messen soll, um im klinischen Einsatz Störungsbildern gerecht zu werden, die mit einer Einschränkung der Affektwahrnehmungsfähigkeit für spezifische Affekte einhergehen (Suslow, Junghans & Arolt, 2001; Yalcin-Siedentopf et al., 2014). Folglich wurde entgegen der Hinweise zur Extraktion von nur 3 Faktoren für die Extraktion von 6 Faktoren entschieden. Dies löste die innerhalb der 3-Faktorenlösung aufgetretene Verknüpfung mimisch ähnlicher Faktoren auf, sodass eine nach Basisaffekten aufgeschlüsselte, inhaltlich gut interpretierbare Faktorstruktur entstand.

Exploriert wurden 2 Modelle, ein *Korrelierte Faktoren Modell* und ein *Bi-Faktor Modell* (Cho, 2016) des PCFAE. Beide Modelle zeichnen sich durch substanzielle Faktorladungen ($\lambda > 0.30$) von Items gleicher Affektdarstellung auf denselben Faktor aus.

¹ Bühner (2011) empfiehlt, bei starker erster Komponente und hoher Korrelation der Komponenten untereinander, den MAP Test (Extraktion von 3 Faktoren) anderen Verfahren vorzuziehen. Die Parallelanalyse überschätzt die Anzahl der zu extrahierenden Faktoren häufig. Ebenfalls berücksichtigt werden muss jedoch, dass der MAP Test aufgrund der nicht normalverteilten Datenstruktur und der Korrelation der Faktoren untereinander die Anzahl zu extrahierender Faktoren unterschätzt haben könnte (Bühner, 2011).

Nur wenige Items luden gering auf diesen, als zugehörigen Affekt interpretierten Faktor. Im Bi-Faktor Modell korrelierten alle Items hoch mit dem gemeinsamen Faktor. Lediglich die Items *Ekel männlich Kind*, *Angst männlich Kind* und *Trauer männlich Erwachsener* luden mit $\lambda < 0.30$ auf die affektspezifischen Faktoren. *Angst männlich Kind* lud als systematische Verwechslung substanziell auf den als Überraschung interpretierten Faktor (Palermo, Bonnar & Gosselin, 2002), während für die Nebenladung des Items *Trauer männlich Erwachsener* auf den als Angst interpretierten Faktor bisher keine systematische Verwechslung berichtet wurde. Die Affektdarstellungsgüte des Items muss hinterfragt werden. Gleichzeitig wiesen die beiden erstgenannten Items eine im Vergleich zu den restlichen Testitems höhere Itemschwierigkeit und geringere Trennschärfe auf. Im Korrelierte Faktoren Modell lud das Item *Überraschung weiblich Erwachsener* mit $\lambda = 0.50$ auf den als Angst interpretierten Faktor, das Item *Angst männlich Kind* lud substanziell auf den als Überraschung interpretierten Faktor. Weiterhin lud das Item *Angst männlich Erwachsener* auf den als Trauer interpretierten Faktor. Als Ursache wurden ebenfalls systematische Verwechslungen angenommen. Aufgrund der hohen Verwechslungsraten beider Items ist eine Überprüfung der Affektdarstellungsgüte notwendig.

Abbildungen 8 und 9 stellen die explorierten Modelle des PCFAE dar. Es liegen 6 untereinander korrelierte Faktoren zugrunde, welche die 6 Basisaffekte repräsentierend interpretiert wurden. Diese werden durch die affektspezifischen Items erfasst (Korrelierte Faktoren Modell). Im Bi-Faktor Modell besteht neben den affektspezifischen Faktoren ebenfalls ein generalisierter Faktor, auf welchen alle Items affektunabhängig laden. Dieser wurde als die Affekterkennungsfähigkeit interpretiert. Neben dem Bi-Faktor Modell des PCFAE mit einem nicht hierarchischen gemeinsamen Faktor ist die Annahme einer hierarchischen Beziehung im Sinne eines gemeinsamen Faktors zweiter Ordnung plausibel,

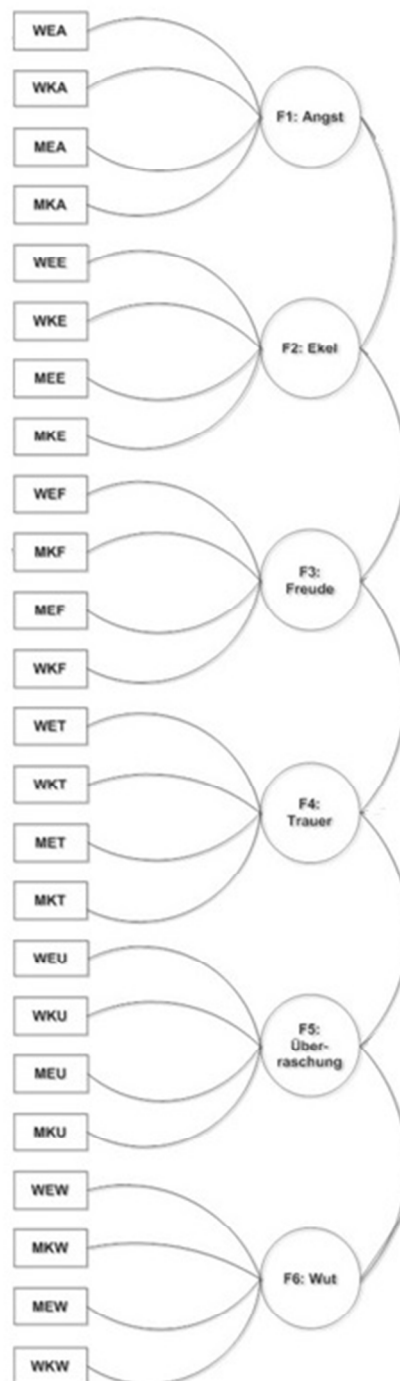


Abbildung 8. Korrelierte Faktoren Modell des PCFAE. F1 – F6 = affektspezifische Faktoren. Vereinfachte Darstellung der Korrelationen der Faktoren F1-F6: Alle Faktoren korrelieren miteinander (siehe Anhang G). Schema der Itembenennung gemäß Anhang I.



Abbildung 9. Bi-Faktor Modell des PCFAE. F1 – F6 = affektspezifische Faktoren. G = generalisierter Faktor. Vereinfachte Darstellung der Korrelationen der Faktoren F1-F6: Alle Faktoren korrelieren miteinander (siehe Anhang G). Schema der Itembenennung gemäß Anhang I.

auf den die affektspezifischen Faktoren laden (Second Order Faktor Modell, Cho, 2016; siehe Anhang H). Ein solches Modell wurde im Rahmen dieser Arbeit nicht exploriert.

Trotz des für eine explorative Faktorenanalyse geringen Stichprobenumfangs und geringster noch zulässiger Itemzahl pro Faktor schlüsseln sich beide explorierten Modelle nach den 6 Basisaffekten auf. Da sich diese Struktur bereits unter ungünstigen Bedingungen deutlich herausstellt, ist es wahrscheinlich, dass sich unter verbesserten Bedingungen eine affektspezifische Struktur weiter erhärtet. Einschränkend ist der Eigenwerteverlauf der Faktoren zu werten. Im Korrelierte Faktoren Modell weisen 3 Faktoren Eigenwerte bedeutend kleiner als 1 auf, im Bi-Faktor Modell ist der Eigenwert eines Faktors wesentlich kleiner als 1. Diese Einschränkung wird durch die Korrelation der Faktoren untereinander verstärkt. Die aufgrund des geringen Eigenwerts bereits weniger wichtigen Faktoren enthalten zusätzlich ebenfalls Informationen der anderen Faktoren. Die affektspezifische Faktorstruktur bildet sich zudem nur unter dem Oblimin Rotationsverfahren aus.

Ungünstig auf die Eigenwerte der betreffenden Faktoren wirkt sich aus, dass aufgrund der systematischen Verwechslungen nicht alle affektspezifischen Items substantiell auf die als zugehörig interpretierten Faktoren laden, sodass im Korrelierte Faktoren Modell ein Faktor nur 2 Items, ein weiterer nur 3 Items mit Faktorladungen von $\lambda < 0.30$ enthält. Im Bi-Faktor Modell enthalten 3 Faktoren nur 3 Items mit Faktorladungen von $\lambda < 0.30$. Eine Optimierung der Affektdarstellungsgüte der Items, die hohe Nebenladungen aufweisen, könnte zu substantiellen Faktorladungen auf den affektspezifischen Faktoren und damit zu einer Steigerung der Eigenwerte führen. Eine mehrfache Präsentation der Testitems mit unterschiedlicher Abstufung der maximalen Affektausprägung würde die Zahl verfügbarer Items des PCFAE erhöhen und sich positiv auf die Reliabilität¹ und somit ebenfalls auf die Eigenwerte der affektspezifischen Faktoren auswirken.

¹ Die Reliabilität eines Tests steigt mit seiner Länge (Bühner, 2011).

Reliabilitätsanalyse

Die Messgenauigkeit des PCFAE ist unter Verwendung der RT korrekt beantworteter Items zu Messzeitpunkt T1 als mittel bis hoch, zu Messzeitpunkt T2 als hoch zu beurteilen. Die Messgenauigkeit unter Verwendung der RSI für korrekt beantwortete Items erreicht nicht die Reliabilitätswerte der RT. Die ACC allein, ohne Einbezug der RT ist als Leistungsmaß ebenfalls wenig reliabel. Bei der Reliabilitätsbeurteilung ist zu beachten, dass die berichteten Koeffizienten die obere (λ_4 , günstigste Split Half Kombination der Items) und untere Grenze (β , ungünstigste Split Half Kombination der Items¹) der möglichen Split-Half Reliabilität des vorliegenden Tests darstellen (Guttman, 1945; Jackson & Agunwamba, 1977; Revelle, 1979), jedoch Parallelität der jeweiligen Testhälften erfordern (Bühner, 2011). Cronbach's α (Cronbach, 1951) schätzt die Reliabilität des Tests nur dann präzise, wenn sowohl eine unidimensionale Skala, als auch τ -Äquivalenz der Items vorliegen. Ansonsten unterschätzt dieser Koeffizient die wahre Reliabilität des Tests systematisch (Cho, 2016; Zinbarg et al., 2005). Cho (2016) stellt dar, dass zur möglichst exakten Reliabilitätsbestimmung eines Tests zunächst seine Struktur abzuklären und anschließend der korrekte Reliabilitätskoeffizient anzuwenden ist. Sollte statt des unidimensionalen Modells (Seidler, 2008) ein mehrdimensionales Modell des PCFAE zutreffen, beschreibt die kongenerische Reliabilität für eine multidimensionale Teststruktur die Reliabilität des PCFAE präziser. Ein geeigneter Koeffizient ist McDonald's ω für das Korrelierte Faktoren Modell bzw. McDonald's ω_h (McDonald, 1978; 1999) für das BiFaktor Modell (Cho, 2016). Bei Zutreffen des Korrelierte Faktoren Modells schätzt ω die Reliabilität des PCFAE demnach als hoch. Innerhalb des Bi-Faktor Modells gilt, dass ω die Reliabilität des gesamten PCFAE schätzt, während ω_h die reliable Varianz, die mit dem generalisierten Faktor

¹ Die ungünstigste Itemkombination bezeichnet dabei die Split Half Kombination bei welcher die Items keine Gemeinsamkeiten aufweisen außer der, die allen Items des Tests gemeinsam ist. β misst somit ähnlich wie ω_h einen gemeinsamen Faktor, ist jedoch nur unter strengen Voraussetzungen mit ω_h identisch (Zinbarg et al., 2005)

assoziiert ist, beurteilt. Wird dieser als die generalisierte Fähigkeit der Affektwahrnehmungskompetenz interpretiert, gibt ω_h demnach an, wie reliabel der PCFAE dieses Kriterium misst (Zinbarg et al., 2006). In diesem Fall wäre die Reliabilität mit $\omega_h = 0.75$ zu T1 als niedrig anzusehen, läge jedoch im Vergleich über der Messgenauigkeit des inhaltlich ähnlichen EK-60F von $\alpha = 0.62$ (Young et al, 2002).

Unter den beschriebenen Voraussetzungen und Einschränkungen kann erst mit Vorliegen einer Modellbestätigung der korrekte Reliabilitätskoeffizient ausgewählt und eine abschließende Aussage über die Messgenauigkeit des PCFAE getroffen werden. Eine direkte Vergleichbarkeit der Reliabilität mit der Urversion des PCFAE ist aufgrund der ungeklärten Dimensionalität des Tests derzeit nicht möglich. Unter den gegebenen Voraussetzungen sind die berechneten Koeffizienten als Hinweis auf die Reliabilität des Tests zu werten. Im Falle eines eindimensionalen Konstrukts wäre eine wesentliche Verbesserung der Reliabilität von $\hat{\rho}_r = 0.75$ auf $\hat{\rho}_r = 0.93$, bzw. $\lambda_1 = 0.82$ auf $\lambda_1 = 0.97$ zu bescheinigen (Seidler, 2008). Eine wesentliche Verbesserung der Reliabilität wäre mit $\omega = 0.95$ unter Annahme des Korrelierte Faktoren Modells zu verzeichnen, während unter Annahme des Bi-Faktor Modells keine Verbesserung der Messgenauigkeit des PCFAE gegenüber seiner Urversion erreicht worden wäre. Die Analyse von Seidler (2008) wäre im Fall eines mehrdimensionalen Konstrukts des PCFAE zu hinterfragen. Mögliche Einschränkungen ergeben sich aus der hohen Homogenität der Items, welche die auf interne Konsistenz bzw. Halbierung des Tests bezogenen Reliabilitätskoeffizienten trivialerweise ebenfalls hoch erscheinen lassen kann (Bühner, 2011). Eine Beurteilung möglicher Redundanzen der Testitems ist angezeigt.

Die zeitliche Stabilität des PCFAE wurde mittels Test-Retest Korrelation berechnet. Diese lag bezogen auf die RT korrekt beantworteter Items knapp unterhalb der empfohlenen Reliabilitätsgrenze von $\hat{\rho}_{tt} = 0.80$ (Jackson, 2011). Die Test-Retest Korrelation lässt jedoch lediglich eine Aussage über die relative zeitliche Stabilität des Tests (also im Idealfall eine

gleiche Rangfolge der VP von Messzeitpunkt T1 zu T2) zu. Das kurze zeitliche Intervall von 4 Wochen zwischen Test und Retest führte zu Testwiederholungseffekten, die die Reliabilitätsschätzung positiv verzerren können (Lienert & Raatz, 1998). Mittels Intra-Klassen-Korrelation (ICC2k) wurde daher beabsichtigt, die Verbesserung der VP im Testergebnis von Messzeitpunkt T1 zu T2 zu berücksichtigen und auf die absolute zeitliche Stabilität rückzuschließen (Bühner, 2011). Eine $ICC > 0.70$ wird als Hinweis auf eine gute Reliabilität gewertet, sodass dem PCFAE bezogen auf die korrekte RT trotz aufgetretener Testwiederholungseffekte eine gute zeitliche Stabilität zuzuschreiben ist. Das verwendete justierte Modell ist zudem unempfindlich gegenüber der Varianzinhomogenität der Daten, da an dieses Modell keine Forderung der Varianzhomogenität gestellt wird. Da ebenfalls keine Generalisierung des Raterurteils (Beurteilung der Stimuli durch die VP) auf die Gesamtpopulation erfolgen soll, ist die Eignung als Reliabilitätsmaß gegeben (Wirtz & Caspar, 2002). Die Leistungsmaße RSI für korrekt beantwortete Items und ACC erfüllen nicht die gewünschte Messgenauigkeit.

In der abschließenden Beurteilung der Reliabilität des PCFAE stellt sich die korrekte RT konsistent mit der bisherigen Testanalyse als das geeignetste Leistungsmaß heraus, während die korrekte RSI und die ACC ohne Einbezug der RT zu stark durch konfundierende Einflüsse verzerrt und damit wenig messgenau sind. Die aktuelle Version des PCFAE erfüllt mit der korrekten RT die Anforderungen an die Messgenauigkeit eines Tests hinsichtlich der zeitlichen Stabilität.

Konstruktvaliditätsanalyse

Die Bestimmung der Konstruktvalidität des PCFAE beschränkte sich auf die Berechnung der Korrelation der Leistungsquotienten $RTquot$ und $RSIquot$ sowie der Antwortgenauigkeit mit den Zielmaßen der inhaltlich ähnlichen Verfahren SEE (Subskalen: Erleben von Emotionsüberflutung EEU sowie Erleben von Emotionsmangel EEM) und TAS-20, in

Ergänzung zu Seidler (2008). Die überprüften Leistungsmaße des PCFAE korrelierten nicht mit den Zielmaßen beider Verfahren. Zur Plausibilitätsprüfung wurden außerdem die angegebenen Skalen des SEE mit der TAS-20 korreliert. Die von Kupfer, Brosig & Braehler (2001) gefundenen Korrelationen zwischen TAS-26 und SEE EEM von $r = 0.41$ sowie TAS-26 und SEE EEU von $r = 0.55$ wurden beinahe exakt repliziert, sodass die gefundenen Zusammenhänge auf den TAS-20 übertragbar scheinen und die Ergebnisse im Zusammenhang mit den Leistungsquotienten des PCFAE plausibel scheinen.

Sowohl der SEE wie auch der TAS-20 erfassen die Fähigkeit, eigene Emotionen zu erkennen, demnach einen introspektiven Prozess. Demgegenüber erfasst der PCFAE die Fähigkeit, die Affekte anderer wahrzunehmen und korrekt zu identifizieren. Auch wenn die Affektwahrnehmung augenscheinlich als Gemeinsamkeit in Erscheinung tritt, ist ein Zusammenhang zwischen der Fähigkeit zur Erkennung eigener und fremder Affekte anhand der gegebenen Daten nicht nachweisbar. Es ist zu berücksichtigen, dass es sich bei den Verfahren SEE und TAS-20 um Selbstauskunftsverfahren handelt, die Ergebnisse gegebenenfalls durch Antwortverzerrungen beeinflusst wurden und Schwankungen zwischen den Messzeitpunkten T1 und T2 in der Selbsteinschätzung der Probanden auftraten. Die Plausibilitätsprüfung schränkt diese Interpretation jedoch ein, der PCFAE weist demnach diskriminante Validität gegenüber den beiden genannten Verfahren auf. Zukünftige Untersuchungen sollten die von Seidler (2008) berichtete konvergente Validität des PCFAE zum EK-60F auf weitere, inhaltlich ähnliche Verfahren ausweiten.

Einschränkungen

Ungünstig ist die uneinheitliche Rechnerausstattung des Instituts zu beurteilen. Obwohl diese zu keiner Einschränkung im Sinne eines statistischen Effekts führte, ist eine einheitliche Darstellung, insbesondere bezüglich der Bildschirmgröße, im Sinne der Standardisierung sinnvoll. Das Intervall zwischen Test und Retest von 4 Wochen stellte sich

im Nachhinein jedoch als zu kurz heraus und führte zu Testwiederholungseffekten.

Zukünftige Untersuchungen sollten ein längeres Intervall zwischen Messzeitpunkt T1 und T2 aufweisen. Die Randomisierung der psychometrischen Instrumente sollte der Konstanthaltung der Versuchsreihenfolge zukünftig vorgezogen werden. Weiterhin schwächte die Varianzinhomogenität verschiedener Skalen die Schätzkraft der auf dem linearen Modell basierenden Methoden. Für die Durchführung der explorativen Faktorenanalyse stellte die gewählte Stichprobengröße das Mindestmaß dar, Implikationen werden in der Folge diskutiert.

Besonders kritisch sind die Unterschiede bei der Erstellung des Stimulusmaterials zu beurteilen. Während die Komposition der Items sich in Alter und Geschlecht gewollt unterscheidet, stellt eine stark heterogene Itemqualität hinsichtlich der Affektdarstellungsgüte eine systematische Fehlerquelle dar, die sich ungünstig auf die Messung der Affektwahrnehmungskompetenz auswirken kann. Deutlich wird dies bei dem Item *Ekel männlich Kind*. Dieses weicht um 21% von der mittleren Schwierigkeit der Affektskala ab, wird wesentlich häufiger verwechselt und weist im Bi-Faktor Modell des PCFAE keine substanzielle Korrelation mit dem Faktor Ekel auf. Obwohl Testitems in ihrer Schwierigkeit variieren sollen, muss sichergestellt sein, dass die Abweichung nicht durch eine mangelnde oder mehrdeutige Affektdarstellung bedingt ist. Für das Item *Angst männlich Kind*, welches am häufigsten verwechselt wurde, wurden bei der Mittelung lediglich 3 Originalbilder verwendet (Hahn, 2016; siehe Anhang A). Beide Items weisen im Vergleich zu den übrigen Stimuli vergleichsweise schlechtere Itemkennwerte auf, welche zwar durchaus im akzeptablen Bereich liegen, jedoch das Verbesserungspotenzial des PCFAE verdeutlichen.

Neben methodischen Unzulänglichkeiten bei der Itemerstellung wird die Affektdarstellungsgüte dadurch eingeschränkt, dass bei der Stimuluserstellung eine künstliche Situation vorlag und die entsprechende Mimik nicht natürlich bzw. unwillkürlich

gezeigt wurde (Rampoldt, 2014). Die natürliche Darstellung des Affekts Freude¹ wird beispielsweise durch die mangelnde Fähigkeit vieler Menschen eingeschränkt, den *Musculus orbicularis oculi lateral* bewusst zu kontrahieren (Pinel, 2007; Seidler, 2008). Ekman & Friesen (1975) erwarteten jedoch durch eine willkürliche Affektdarstellung idealtypischere Stimuli zu produzieren, als dies bei unwillkürlicher Affektdarstellung der Fall wäre. Obwohl die Darstellungsgüte des Stimulusmaterials des PCFAE vollständig mittels FACS Kodierung bestätigt wurde (Franz et al., 2018; unveröffentlicht), kann ein solches Expertenrating weiterhin fehlerbehaftet sein. Es wiegt die methodischen Unzulänglichkeiten bei der Stimuluserstellung nicht auf, sodass möglicherweise keine optimale Affektdarstellungsgüte, vor allem durch die beispielhaft genannten Items erreicht wurde.

Der Beginn der Darstellung jedes Items mit einem neutralen Gesichtsausdruck kann von bestimmten Personengruppen als bedrohlich empfunden werden (Personen mit Borderline Persönlichkeitsstörung: Donegan et al., 2003; Kinder: Thomas et al., 2001). Dies könnte in der klinischen Anwendung eine systematische Fehlerquelle darstellen und das Konzept des neutralen Gesichts insgesamt in Frage stellen. Ein Beginn der Itempräsentation mit einer sehr geringen Affektausprägung ist als Lösung denkbar, zumal ein neutraler Gesichtsausdruck nicht dazu beigetragen hat, Testwiederholungseffekte zu unterbinden. Eine weitere Möglichkeit ist die Verwendung gelassener Gesichtsausdrücke (Tottenham et al., 2009), was jedoch aufgrund der notwendigen vollständigen Neuerstellung des Stimulusmaterials und erneuten Analyse des Tests keine ökonomische Lösung darstellt.

Die quasidynamische Präsentation der Itemsequenz in 700 Millisekunden Schritten stellt einen Kompromiss aus beinahe dynamischem Eindruck und der Möglichkeit dar, schrittweise zu entscheiden, ob der Affekt erkannt wurde. Eine vollständig dynamische, realistische Darstellung des Affektaufbaus mittels Videodarstellung würde innerhalb eines

¹ Bezeichnet als Duchenne Lächeln (Ekman & Davidson, 1993)

sehr viel kürzeren Zeitraumes geschehen. Die Auswertung würde durch die Konfundierung mit der verarbeitungsbedingten Komponente der Reaktionszeit (Nicht-Entscheidungsphase) eine wesentlich komplexere Auswertestrategie erfordern. Diesem Aspekt wohnt ein generelles Problem des PCFAE bzw. der Beurteilung des Geschwindigkeits-Genauigkeits-Ausgleichs inne: Unklar ist, wie man RT und ACC sinnvoll in ein Maß integriert und welche Gewichtung der beiden Komponenten angezeigt ist. Eine mögliche Lösung stellt das Diffusionsmodell dar (Ratcliff, 1978; Voss, Rothermund & Voss, 2004), dessen Komplexität jedoch der Forderung nach einer einfachen Anwend- und Auswertbarkeit des PCFAE entgegensteht.

Die Präsentation der Items bis zu einem Affektapex von 100% Affektausprägung wirkt sich ungünstig auf die Testschwierigkeit aus, reduziert die verfügbare Varianz und erschwert die Diskriminierung der Performanz im Bereich hoher Merkmalsausprägung. Gemeinsam mit den Unterschieden bei der Erstellung des Stimulusmaterials weist dieser Aspekt das größte Verbesserungspotenzial für den PCFAE auf.

Abschließende Wertung, Implikationen und Ausblick

Die Modifizierungsmaßnahmen des PCFAE erzielten hinsichtlich der Itemgüte gegenüber seiner Urversion eine Verbesserung des Paradigmas. Insbesondere die Trennschärfe der Items bzw. der Affektskalen wurde durch die Verwendung der RT korrekt beantworteter Items als Leistungsmaß gegenüber der Verwendung der ACC wesentlich angehoben. Der PCFAE ist innerhalb der Affektskalen stark homogen, während der Gesamttest eine ideale Homogenität aufweist. Insbesondere innerhalb der Skala Freude ist eine Redundanz der Items möglich. Dies lässt eine Reduktion des Testumfangs zugunsten der Einführung von Items mit abgestufter Affektausprägung zu. Weitere Verbesserungen des Tests setzen die Aufklärung der Dimensionalität voraus, sodass eine konfirmatorische Faktorenanalyse der explorierten Modelle des PCFAE unbedingt angezeigt ist. Anschließend

kann eine abschließende Beurteilung der Reliabilität erfolgen. Eine Verbesserung der Messgenauigkeit des PCFAE erscheint durch die bisher erfolgten Modifikationen jedoch wahrscheinlich und es ist zu erwarten, dass die Messgenauigkeit anderer Verfahren übertroffen wird.

Modellkonfirmierende Folgeuntersuchungen zum PCFAE sollten zunächst einen deutlich größeren Stichprobenumfang umfassen, welcher an die Itemanzahl je Faktor (Affektskala) angepasst ist. Bühner (2011) empfiehlt bei 4 Items pro Faktor eine Mindestanzahl von 500 VP. Die mehrfache Präsentation der Items mit verschiedenen Abstufungen der maximalen Affektausprägung würde eine höhere Zahl affektspezifischer Items pro Faktor generieren. Hierdurch würde einerseits die Reliabilität der Faktoren selbst verbessert und gleichzeitig könnte der Stichprobenumfang für die Folgeuntersuchung reduziert werden. Weiterhin sollte die Faktorstruktur des PCFAE mit größerer Itemzahl deutlicher zutage treten und der Eigenwerteverlauf der Faktoren sich erheblich verbessern. Eine entscheidende Voraussetzung hierfür ist jedoch die vorherige Überprüfung und Verbesserung der Affektdarstellungsgüte der Testitems. Eine konfirmatorische Faktorenanalyse sollte das Second Order Faktor Modell (Anhang H) und das von Seidler (2008) explorierte einfaktorielle Modell des PCFAE ebenfalls mit einbeziehen, denn gegen die Aufteilung nach 6 Faktoren sprechen trotz inhaltlicher Plausibilität der Eigenwertverlauf mit starker erster Komponente, die hohen Korrelationen der Faktoren untereinander sowie die hohe Homogenität der Items. Diese Aspekte stellen Hinweise auf ein eindimensionales Konstrukt dar. Eine Modellprüfung muss berücksichtigen, ob die Verwendung von 6 Faktoren die Struktur des PCFAE besser erklären kann und zu einer bedeutend höheren Varianzaufklärung führt, als eine sparsamere eindimensionale Lösung.

Der PCFAE ist als leichter bis mittelschwerer Test zu klassifizieren und hebt sich in seiner Schwierigkeit bisher nicht von vergleichbaren Verfahren ab (Egger et al., 2011).

Aufgrund der hohen Bedeutung der Affekterkennung für das Zusammen- und Überleben scheint der Mensch eine hohe Fähigkeitsausprägung in diesem Bereich aufzuweisen, sodass entsprechende Testverfahren für Gesunde natürlicherweise keine hohen Schwierigkeiten aufweisen sollten, sofern die Items eine maximale Affektausprägung abbilden. Eine Abstufung der maximalen Affektausprägung aller Items auf 66% und 33% bei gleichbleibender maximaler Präsentationsdauer würde sowohl die verfügbare Varianz als auch die Testschwierigkeit erhöhen, sodass im diagnostischen Prozess wesentlich besser im Bereich hoher Fähigkeitsausprägung unterschieden werden könnte (Jansen, 2018a; unveröffentlicht; Phillips et al., 1997).

Die Items des PSYCAFE Sets („Kinderitems“) sollten hinsichtlich der Anzahl und Darstellungsgüte der in den Mittelungsprozess eingeflossenen Originalbilder verbessert werden, um methodische Mängel und eine mangelnde Affektdarstellungsgüte als systematische Fehlerquellen auszuschließen. Dies erfordert eine Neuerstellung der Items unter standardisierten Bedingungen. Die Modifizierung der Items und die Abstufung der maximalen Affektausprägung sind als die wichtigsten notwendigen Verbesserungen zu beurteilen und sollten vor Durchführung der Modellkonfirmierung mit hoher Priorisierung verfolgt werden.

Ebenfalls sollte entschieden werden, welche Art von Testverfahren kreiert werden soll. Verfahren mit hoher Stimulusdiversität (Alter, Geschlecht, Ethnie) sind mit gemittelten Items schwierig umzusetzen, während einfache diagnostische Verfahren schnell und unkompliziert anzuwenden sein sollten. Für den PCFAE bedeutet dies, dass eine Aufteilung nach Alter und Geschlecht nicht nur zu einer Redundanz von Items führen kann, sondern dass die Einfachheit der Durchführung im diagnostischen Prozess einer derzeitigen Durchführungsdauer von ca. 20 Minuten entgegensteht. Die angesprochene Abstufung der maximalen Affektausprägung ermöglicht eine Reduktion der Itemdiversität und des

Testumfangs. Items mit vergleichsweise schlechten Kennwerten oder hohen Verwechslungsraten könnten aus dem Test entfernt werden. Denkbar wäre ebenfalls ein weiterer Mittelungsprozess zur Schaffung geschlechtsneutraler Items oder eine Reduktion der gezeigten Stimuli je Testdurchgang. Dies würde ebenfalls die Schaffung von Paralleltestversionen ermöglichen. Die Darstellung von Affekten speziell in Kindergesichtern kann für bestimmte Anwendungsbereiche¹ sinnvoll sein, daher wäre eine zweckgebundene Aufteilung des PCFAE in eine Version mit Kinderaffekten und eine Version mit Erwachsenenaffekten denkbar. Es sollte eine kritische Evaluierung der Ziele und Einsatzgebiete des PCFAE erfolgen, an der sich zukünftige Modifizierungsmaßnahmen orientieren können.

Bei der Auswahl eines geeigneten Leistungsmaßes für den PCFAE zeigte sich die korrekte Reaktionszeit gegenüber der korrekten Antwortsicherheit und der Antwortgenauigkeit hinsichtlich aller Gütekriterien überlegen. Die RSI scheint verschiedenen konfundierenden Einflüssen zu unterliegen, sodass von der Auswahlgeschwindigkeit des korrekten Affekts nur schlecht auf die Fähigkeitsausprägung der Affektwahrnehmungskompetenz geschlossen werden kann. Der alleinige Einbezug der ACC berücksichtigt nicht den Geschwindigkeits-Genauigkeits-Ausgleich. Bei Störungsbildern, die eine verlängerte RT bei gleichbleibender ACC bedingen (Aebi, 2012), würde unter Verwendung der ACC die Einschränkung der Affektwahrnehmungskompetenz nicht korrekt erfasst werden.

Eine einfache Auswertung des PCFAE kann auf Test- oder Skalenebene mittels des dargestellten Leistungsquotienten $RTquot$ und $RSIquot$ erfolgen, welcher RT, ACC und Anzahl der Fehler in ein gemeinsames Maß integriert. Die Überprüfung der Konstruktvalidität wurde mittels dieser Leistungsquotienten durchgeführt. Diese integrieren zusätzlich zur Reaktionszeit und der Antwortgenauigkeit ebenfalls die Fehlerrate und bilden

¹ Beispielsweise diagnostischer Einsatz im Rahmen des Programms Wir2 welches unter anderem die Affektresonanz von alleinerziehenden Eltern gegenüber Kindern unterstützen soll (Franz, 2014)

damit mehr Informationen ab, als die korrekte RT bzw. die korrekte RSI allein. Die Quotienten erlauben jedoch keine Auswertung des Tests auf Itemebene. Hierfür steht neben der RT korrekt beantworteter Items eine durch Jansen (2018b; unveröffentlicht) vorgeschlagene Auswertungsroutine zur Verfügung. Diese vergibt eine feste Startpunktzahl je Item, unabhängig von der korrekten Beantwortung, welche sich mit zunehmender Reaktionszeit reduziert sowie einen zusätzlichen pauschalen Punktezuschlag für die korrekte Beantwortung des Items. Zukünftige Untersuchungen sollten die Beziehung der Leistungsmaße des PCFAE zueinander näher untersuchen, um zu klären welches Maß eine möglichst valide Erfassung der Affektwahrnehmungskompetenz abbildet. In diesem Zusammenhang ist die Überprüfung der (Konstrukt-) Validität des PCFAE auszuweiten. Sind diese Maßnahmen umgesetzt, so steht mit dem PCFAE ein leistungsfähiger Test zur Erfassung der Affektwahrnehmungskompetenz zur Verfügung, welcher schnell und einfach in der Durchführung und Auswertung ist und bisherige Verfahren konzeptionell und qualitativ übertreffen könnte.

Literaturverzeichnis

Abrosoft Corporation (2002). FantaMorph. www.fantamorph.com, letzter Abruf 18.12.2018.

Addington, J., Barbato, M., 2015. Social cognition in those at high risk of psychosis. In:

Shutt, R.K., Seidman, L.J., Keshavan, M. (Eds.), Social Neuroscience: Mind, Brain, and Society. Harvard University Press.

Adolphs, R. (2002). Recognizing Emotion From Facial Expressions: Psychological and Neurological Mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1(1), 21-62.

Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., Damasio, A.R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*. 433(7021), 68 - 72.

Aebi, M.E. (2012). Facial Affect Recognition Deficits in Students that Exhibit Subclinical Borderline Personality Traits. Cleveland, Bachelor of Arts in Broadcast Journalism Harding University.

Ambadar, Z., Schooler, J. W. & Cohn, J. F. (2005). The Importance of facial dynamics in interpreting subtle facial expressions. *American Psychological Society*, 16(5), 403-410.

Apel, M. & Ludz, P. (1976). Philosophisches Wörterbuch. 6. Auflage, Berlin, Walter de Gruyter.

Aronson, E., Wilson, T. & Akert, R. (2014). Sozialpsychologie. Hallbergmoos, Pearson Deutschland GmbH.

Bagby, R.M. Parker, J.D.A., Taylor, G.J.(1994). The twenty-item Toronto Alexithymia Scale-I. Item Selection and cross-validation of the factor-structure. *Journal of Psychosomatic Research*, Vol. 38(1), 23-32.

-
- Barbato, M., Buchy, L., Bray, S., MacMaster, F. & Deighton, S. (2017). Testing the feasibility of a computerized facial affect recognition training in early psychosis. *Schizophr. Res.* 2017, <http://dx.doi.org/10.1016/j.schres.2017.03.011>.
- Barbeau, E., Joubert, S., & Felician, O. (2009). Traitement et reconnaissance des visages: Du percept à la personne. Marseille: Solal.
- Batchelder, W.H. (1998). Multinomial processing tree models and psychological assessment. *Psychological Assessment*, 10, 331–344.
- Bedi, G., Shiffrin, L. Vadhana, N.P., Nunes, E.V., Foltin, R.W. & Bisage, A. (2016). Effects of levodopa-carbidopa-entacapone and smoked cocaine on facial affect recognition in cocaine smokers. *Journal of Psychopharmacology*, 30(4), 370-377.
- Behr, M., & Becker, M. (2004). Skalen zum Erleben von Emotionen (SEE). Hogrefe.
- Bernaards, C.A. and Jennrich, R.I. (2005) Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. *Educational and Psychological Measurement*, 65, 676–696.
- Beyer von Morgenstern, S., Becker, I. & Sinzig, J. (2013). Improvement of facial affect recognition in children and adolescents with attentiondeficit/ hyperactivity disorder under methylphenidate. *Acta Neuropsychiatrica* 2014, 202-208.
- Bodamer, J. (1947). Die prosop-agnosie. *Archiv für Psychiatrie und Nervenkrankheiten*, 179(1-2), 6-53.
- Boelte, S., Feineis-Matthews, S., Leber, S. Dierks, T., Hubl, D. & Poustka, F. (2002). The development and evaluation of a computerbased program to test and to teach the recognition of facial affect. *Int J Circumpolar Health*, 61(2), 61–68.
- Boll, S & Gamer, M. (2016). Psychopathic traits affect the visual exploration of facial expressions. *Biological Psychology* 117, 194-201.

-
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296–322.
- Bühner, M. (2011). Einführung in die Test- und Fragebogenkonstruktion (3. Auflage). Hallbergemoos, Deutschland, Pearson.
- Buss, D. M. (1999). Evolutionary psychology. The new science of the mind. Boston, MA: Allyn and Bacon.
- Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior research methods*, 40(1), 109-115.
- Canli, D., Özdemir, H. & Kocak, O.M. (2015). Magical ideation associated social cognition in adolescents: signs of a negative facial affect recognition deficit. *Comprehensive Psychiatry* 61, 90-96.
- Chapman, H., Gillespie, S.M. & Mitchell, I.J. (2018). Facial affect processing in incarcerated violent males: A systematic review. *Aggression and Violent Behavior* 38, 123 -138.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Cho, E. (2016). Making Reliability Reliable: A Systematic Approach to Reliability Coefficients. *Organizational Research Methods*, 19 (4), 651 – 682.
- Cigna, M.H., Guay, J.P.& Renaud, P. (2017). Psychopathic traits and their relation to facial affect recognition. *Personality and Individual Differences*, 117, 210 – 215.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2. Auflage). Hillsdale, NJ, Erlbaum.
- Cohen, R. J., & Swerdlik, M. E. (2005). Psychological testing and assessment: An introduction to tests and measurement (6th ed.). New York: McGraw-Hill.

-
- Courtney, M.G.R. (2013). Determining the Number of Factors to Retain in EFA: Using the SPSS R-Menu v2.0 to Make More Judicious Estimations. *Practical Assessment, Research & Evaluation*, 18 (8), 1-14.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16 (3), 297-334.
- Csukly, G., Stefanics, G., Komlosi, S., Czigler, I. & Czobor, P. (2014). Event-Related Theta Synchronization Predicts Deficit in Facial Affect Recognition in Schizophrenia. *Journal of Abnormal Psychology*, 123(1), 178 – 189.
- Dalrymple, K. A., Gomez, J., & Duchaine, B. (2013). The Dartmouth Database of Children's Faces: acquisition and validation of a new face stimulus set. *PloS one*, 8(11), e79131.
- Darwin (1872/1965). The expression of the emotions in man and animals. Chicago: University of Chicago Press.
- Dawel, A., O'Kearney, R., McKone, E. & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience and Biobehavioral Reviews*, 36, 2288 – 2304.
- De Gelder, B., Vroomen, J., Pourtois, G. & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *NeuroReport*, 10(18), 3759-3763.
- Demirel, H., Yesilbas, D., Ozver, I., Yuksek, E., Sahin, F., Aliustaoglu, S. & Emul, M. (2014). Psychopathy and facial emotion recognition ability in patients with bipolar affective disorder with or without delinquent behaviors. *Comprehensive Psychiatry*, 55, 542 – 546.
- Derogatis, L.R.,(1993). Brief Symptom Inventory (BSI), administration, scoring and

-
- procedures manual, third edition. Minneapolis: National Computer Services.Drusch.
- Donegan N.H., Sanislow C.A., Blumberg H.P., Fulbright R.K., Lacadie C., Skudlarski P., et al. (2003). Amygdala hyperreactivity in borderline personality disorder: implications for emotional dysregulation. *Biological Psychiatry*, 54,1284-1293.
- Donell, J. (2015). Alexithymie bei Patienten mit chronischer Multipler Sklerose. Düsseldorf, Heinrich-Heine-Universität.
- Dragulsecu, A.A. (2014). Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files. Retrieved from <https://CRAN.R-project.org/package=xlsx>.
- Duchaine, B.(2008). Famous Faces Test. Prosopagnosia Research Center, Harvard & London University. http://www.faceblind.org/facetests/ff/ff_intro.php, letzter Abruf 14.06.2018.
- Duchenne, B. (1862). Mecanisme de la physionomie humaine ou analyse electro-physiologique de Vexpression des passions, Paris, Bailliere.
- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior research methods*, 42(1), 351-362.
- Egger, H. L., Pine, D. S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K. E., & Angold, A. (2011). The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): a new set of children's facial emotion stimuli. *International journal of methods in psychiatric research*,20(3), 145-156.
- Ekman, P. (1972). Universals and cultural differences in facial expression of emotion. In Cole, J.K. (Ed.), *Nebraska symposium on motivation*, 19, 207-283. Lincoln, NE: University of Nebraska Press.
- Ekman, P. (1973). Cross-cultural studies of facial expression. Darwin and facial expression: *A century of research in review*, 169222, 1.

-
- Ekman, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? *Emotions in the human face*, 39-55.
- Ekman, P. (1988). *Gesichtsausdruck und Gefühl*. Paderborn: Junfermann.
- Ekman, P. (1992). Facial Expressions of Emotion: New Findings, New Questions. *American Psychological Society*, 3(1), 34 – 38.
- Ekman, P. (2003). *Emotions revealed*. New York, NY: Holt Paperbacks.
- Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion review*, 3(4), 364-370.
- Ekman, P. E., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. Oxford University Press.
- Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of communication*, 22(4), 353-374.
- Ekman, P. & Friesen, W.V. (1975). *Unmasking The Face. A Guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.
- Ekman, P. & Friesen, W. (1976) *Pictures of facial affect*. Consulting Psychologists Press, Palo Alto.
- Ekman, P. & Friesen, W. (1978). *Facial action coding system. A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologist Press.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). *Emotion in the Human Face*. Cambridge University Press.
- Ekman, P., Friesen, W.V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., et al. (1987). Universals and cultural differences in the judgments of facial expressions of emotions, *Journal of Personality and Social Psychology*, 53, 712-717.

-
- Ekman, P., Friesen, W.V. & Tomkins, S.S. (1971). Facial Affect Scoring Technique: A First Validity Study. *Semiotica*, 3(1), 37-58.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage publications.
- Fisseni, H. J. (2004). *Lehrbuch der psychologischen Diagnostik: mit Hinweisen zur Intervention*. Hogrefe Verlag.
- Franke, G.H. (1999). Brief Symptom Inventory von L.R. Derogatis (Kurzform des SCL-90-R) – Deutsche Version – Manual. Beltz Test GmbH Göttingen.
- Franz, M. (2014). *wir2. Bindungstraining für Alleinerziehende*. Göttingen, Vandenhoeck & Ruprecht.
- Franz, M., Müller, T., Hahn, S., Lundqvist, D., Rampoldt, D., & Schäfer, R. (2018). Picture-Set of Young Children Affective Facial Expressions (PSYCAFE). (unveröffentlicht)
- Fridlund, A. J. (2014). *Human facial expression: An evolutionary view*. Academic Press.
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., et al. (2009). Universal recognition of three basic emotions in music. *Current biology*, 19(7), 573-576.
- Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., et al. (2009) Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J Psychiatry Neurosci* 34(6),418–432.
- Genova, H.M., Genualdi, A., Goverover, Y., Chiaravalloti, N.D., Marino, C. & Lengenfelder, J. (2017). An investigation of the impact of facial affect recognition impairments in moderate to severe TBI on fatigue, depression, and quality of life. *Social Neuroscience*, 12(3), 303 – 307.
- Greenberg, L.S., Rice, L.N. & Elliott, R.K. (1996). *Facilitating emotional change: The moment-by-moment process*. New York, Guilford.

-
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: a validation study. *Cognition and emotion*, 22(6), 1094-1118.
- Goschke, T. & Dreisbach, G. (2010). Kognitiv-affektive Neurowissenschaft: Emotionale Modulation des Denkens, Erinnerns und Handelns. In U. Wittchen & J. Hoyer (Hrsg.) (2. Aufl.), *Klinische Psychologie und Psychotherapie*. Berlin: Springer.
- Guo K. (2012) Holistic Gaze Strategy to Categorize Facial Expression of Varying Intensities. *PLoS ONE* 7(8): e42585.
- Guttman, L. (1945). A Basis for Analyzing Test-Retest Reliability. *Psychometrika*, 10(4), 255 – 282.
- Hager, J.C., Ekman, P., Friesen, W.V. (2002). Facial action coding system. Salt Lake City, UT: A Human Face.
- Horn, J. L.(1965). A rationale and test for the number of factors in factor analysis. *Psychometrika* 30, 179-185.
- Ingenhag, A., Schäfer, R. & Franz, M. (2007). Test of Perceptual Competence of Facial Affective Expression. Düsseldorf, Institut für Psychosomatische Medizin und Psychotherapie, Universitätsklinikum der Heinrich Heine Universität.
(unveröffentlicht)
- Izard, C. E. (1971). The face of emotion. East Norwalk, CT, US: Appleton-Century-Crofts.
- Jackson, P. H., & Agunwamba, C. C. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: I: Algebraic lower bounds. *Psychometrika*, 42(4), 567-578.
- Jackson, S.L. (2011). Research methods: A Modular Approach (2nd International Edition). Wadsworth, USA, Cengage Learning.
- James, W. (1884). What is an emotion? *Mind* 9, 188-205.
- Jansen, P. (2018a). Psychophysiological markers of facial affect recognition - A combined

-
- EEG and EMG study. Düsseldorf, Institut für Psychosomatische Medizin und Psychotherapie, Universitätsklinikum der Heinrich Heine Universität.
(unveröffentlicht)
- Jansen, P. (2018b). Fühl mal App. Trainingsprogramm zur Affekterkennung. Düsseldorf, Institut für Psychosomatische Medizin und Psychotherapie, Universitätsklinikum der Heinrich Heine Universität. (unveröffentlicht)
- Jansen, P., Schäfer, R. & Franz, M. (2018). Famous Faces Test. (unveröffentlicht)
- Javdani, S., Sadeh, N., Donenberg, G.R., Emerson, E.M., Houck, C. & Brown, L.K. (2017). Affect recognition among adolescents in therapeutic schools: relationships with posttraumatic stress disorder and conduct disorder symptoms. *Child and Adolescent Mental Health*, 22(1), 42 – 48.
- Jennrich, R. I. (2002). A simple general method for oblique rotation. *Psychometrika*, 67(1), 7-19.
- Jusyte, A., Gulewitsch, M.D. & Schönenberg, M. (2017). Recognition of peer emotions in children with ADHD: Evidence from an animated facial expressions task. *Psychiatry Research*, 258, 351 – 357.
- Kaiser, H. F. & Dickman, K. (1959). Analytic determination of common factors. *American Psychologist* 14, 425.
- Kilts, C. D., Egan, G., Gideon, D. A. Ely, T. D. & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage*, 18, 156-168.
- Kosson, D. S., Suchy, Y., Mayer, A. R., & Libby, J. (2002). Facial affect recognition in criminal psychopaths. *Emotion*, 2(4), 398–411.
- Kragel, P. A. & LaBar, K. S. (2013). Multivariate pattern classification reveals autonomic and experiential representations of discrete emotions. *Emotion*,

- 13(4), 681-690.
- Kupfer, J., Brosig, B., & Brähler, E. (2001). Toronto-Alexithymie-Skala-26 (TAS-26). Göttingen, Hogrefe.
- LaBar, K. S., Crupain, M. J. Voyvodic, J. T. & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, 13, 1023-1033.
- Lawrence, M. A. (2016). ez: Easy Analysis and Visualization of Factorial Experiments. Retrieved from <https://CRAN.R-project.org/package=ez>.
- Lazarus, R. S. (1991). Emotion and adaptation. New York: Oxford University Press.
- Lench, H. C., Flores, S. A. & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgement, experience, behavior, and physiology: A metaanalysis of experimental emotion elicitations. *Psychological Bulletin*, 137(5), 834-855.
- Lenhard, W. & Lenhard, A. (2016). Berechnung von Effektstärken. <https://www.psychometrica.de/effektstaerke.html>. Letzter Abruf 20.06.2018, Dettelbach: Psychometrica.
- Leppänen, J.M., Hietanen, J.K. (2004) Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychol Res.*, 69(1-2), 22 - 29.
- Lienert, G. A., & Raatz, U. (1998). Testaufbau und Testanalyse. 6. Auflage, Beltz.
- LoBue, V., & Thrasher, C. (2014). The Child Affective Facial Expression (CAFE) set. Databrary.
- Lozier, L.M., Vanmeter, J.W. & Marsh, A.A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology*, 26, 933 – 945.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces

-
- (KDEF). CD ROM from Department of Clinical Neuroscience, Psychology section, *Karolinska Institutet*, 91, 630.
- Lundqvist, D., & Litton, J. E. (1998). The Averaged Karolinska Directed Emotional Faces—AKDEF, CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institute.
- Mandler, G. (1984). *Mind and body. Psychology of emotion and stress*. New York, Norton.
- Marsh, A. A., Ambady, N. & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Social Cognition*, 31, 745-757.
- McDonald, R. P. (1978). Generalizability in factorable domains: Domain validity and generalizability. 1. *Educational and Psychological Measurement*, 38 (1), 75-79.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- McDougall, W. (1908/1960). *An Introduction to social psychology* (31. Ed.). London: Methuen.
- Meinong, A. (1894). *Psychologisch-ethische Untersuchungen zur Werth-Theorie*. Graz: Leuschner und Lubensky.
- Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, 124, 1241-1252.
- Nachtigall, C & Wirtz, M. (2013). *Wahrscheinlichkeitsrechnung und Inferenzstatistik. Statistische Methoden für Psychologen Teil 2*. 6. Auflage, Weinheim, Beltz Juventa.
- Nemiah, J.C., Freyberger, H. & Sifneos, P.E. (1976). Alexithymia: A view of the psychosomatic process. In: *O. W. Hill (Hrsg.): Modern Trends in Psychosomatic Medicine. Band 3*, Butterworths, London, 430–439.

-
- Newcombe, F. (1979) The processing of visual information in prosopagnosia and acquired dyslexia: functional versus physiological interpretation. *Research in Psychology and Medicine, 1*, 315-322.
- Ortony, A. & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review, 97*(3), 315-331.
- Paiva-Silva, A. I. D., Pontes, M. K., Aguiar, J. S. R., & de Souza, W. C. (2016). How do we evaluate facial emotion recognition?. *Psychology & neuroscience, 9*(2), 153.
- Palermo, R. & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behav Res Methods Instrum Comput., 36*(4), 634-638.
- Parker, J. D., Taylor, G. J., & Bagby, M. (1993). Alexithymia and the recognition of facial expressions of emotion. *Psychotherapy and psychosomatics, 59*(3-4), 197-202.
- Parkinson, B. (1996). Emotions are social. *British Journal of Psychology, 87*, 663–683.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage 16*(2), 331–348.
- Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature, 389*(6650), 495.
- Pierce, J.W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.
- Pinel, J.P.J. (2007). Biopsychologie. München: Pearson Studium.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In: *Plutchik R. & Kellerman H. (Ed.). Theories of emotion*, New York, Academic Press.
- Pochedly, J.T., Widen, S.C. & Russell, J.A. (2012). What emotion does the facial expression of disgust express? *Emotion 12*(6), 1315.

Popova, P., Popov, T.G., Wienbruch, C., Carolus, A.M., Miller, G.A. & Rockstroh, B.S.

(2014). Changing facial affect recognition in schizophrenia: Effects of training on brain dynamics. *NeuroImage: Clinical*, 6, 156 – 165.

Prochnow D., Donell J., Schäfer R., Jörgens S., Hartung HP., Franz M., Seitz RJ. (2011).

Alexithymia and impaired facial affect recognition in multiple sclerosis. *J Neurol.*, 258(9), 1683-1688.

R Core Team (2018). R: A language and environment for statistical computing. Vienna, Austria. <https://CRAN.Rproject.org>.

Raiche, G. (2010). Parallel Analysis and Non Graphical Solutions to the Cattell Scree Test.

<https://CRAN.Rproject.org/package=nFactors>.

Rampoldt, D. (2014). Erstellung eines Picture-Set of Young Children Affective Facial Expression (PSYCAFE). Heinrich-Heine Universität, Düsseldorf.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.

Reischies, F.M. (2007). Emotion und Affekt. In: *Psychopathologie*. Springer, Berlin, Heidelberg.

Revell, W. (1979). Hierarchical cluster-analysis and the internal structure of tests.

Multivariate Behavioral Research, 14(1), 57-74.

Revelle, W. (2018). psych: Procedures for Personality and Psychological Research.

Northwestern University, Evanston, Illinois, USA. Retrieved from

<https://CRAN.Rproject.org/package=psych>.

Revelle, W., Rocklin, T. (1979). Very simple structure – alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, 14 (4), 403-414.

Rigon, A., Voss, M.W., Turkstra, L.S., Mutlu, B. & Duff, M.C. (2017). Relationship between individual differences in functional connectivity and facial-emotion recognition

-
- abilities in adults with traumatic brain injury. *NeuroImage: Clinical*, 13, 370 – 377.
- Rogers, C.R. (1951). *Client-Centered Therapy*. London: Constable.
- Rudolph, U. (2013). *Motivationspsychologie kompakt*. 3. Auflage, Weinheim, Beltz Verlag.
- Sato, W., Fujimura, T., Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. *Int J Psychophysiol.* 70, 70–74.
- Sato, W., Kochiyama, T., Uono, S. (2015). Spatiotemporal neural network dynamics for the processing of dynamic facial expressions. *Nature Publishing Group*, 1–13.
- Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 1, 49-80, New York, Academic Press.
- Scherer, K.R. (1990). *Psychologie der Emotion*. Göttingen, Hogrefe.
- Scherer, K.R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 695 – 729.
- Seidel, E. M., Habel, U., Kirschner, M. Gur, R. C. & Derntl, B. (2010). The impact of facial emotional expressions on behavioral tendencies in women and men. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 500-507.
- Sfärlea, A., Greimel, E., Platt, B., Dieler, A.C. & Schulte-Körne, G. (2018). Recognition of emotional facial expressions in adolescents with anorexia nervosa and adolescents with major depression. *Psychiatry Research*, 262, 586 – 594.
- Smith, M.L., Cottrell, G.W., Gosselin, F., Schyns, P.G. (2005). Transmitting and decoding facial expressions. *PsycholSci*, 16(3), 184 - 189.
- Spearman, Charles, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271–295.

-
- Steiner J.E. (1979) Human facial expression in response to Taste and smell stimulation. In Reese H.W. & Lipsitt L.P. (Eds.) *Advances in child development and behavior*, 13, 257-295. New York: Academic.
- Stephens, C.L., Christie, I. C. & Friedman, B. H. (2010). Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis. *Biological Psychology*, 84, 463-473.
- Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of personality and social psychology*, 54(5), 768.
- Stroth, S., Kampf, D., Frommann, N. & Wölwer, W. (2014). Effects of Training of Affect Recognition on the recognition and visual exploration of emotional faces in schizophrenia. *Schizophrenia Research*, 159, 485 – 490.
- Suslow, T., Junghans, K. & Arolt, V. (2001). Detection of Facial Expressions of Emotions in Depression. *Perceptual and Motor Skills*, 92(3), 857 – 868.
- Taylor, G.J., Bagby, R.M. & Parker, J.D.A. (2003). The 20-Item Toronto Alexithymia Scale IV. Reliability and factorial validity in different languages and cultures. *Journal of Psychosomatic Research*, 55, 277-283.
- Thomas K.M., Drevets W.C., Whalen P.J., Eccard C.H., Dahl R.E., Ryan N.D., Casey B.J. (2001). Amygdala response to facial expressions in children and adults. *Biological Psychiatry*, 49, 309-316.
- Toh, W.L., Castle, D.J. & Rossell, S.L. (2015). Disorders Facial affect recognition in body dysmorphic disorder versus obsessive-compulsive disorder: An eye-tracking study. *Journal of Anxiety Disorders*, 35, 49 – 59.
- Tomkins, S.S. (1962). Affect, imagery, consciousness. Vol. I. The positive affects. New York: Springer.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... &

Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry research*, 168(3), 242-249.

Trautmann, S. A., Fehr, T. & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, 1284, 100-115.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321-327.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, 32, 1206-1220.

Vytal, K. & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Cognitive Neuroscience*, 22(12), 2864-2885.

Watson, J. B. (1919). A schematic outline of the emotions. *Psychological Review* 26, 165-196.

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J. & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS ONE* 12(5), e0177239.

Wehrle, T., Kaiser, S., Schmidt, S. & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, 78(1), 105-119.

Weyers, P., Mühlberger, A., Hefele, C. & Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology*, 43, 450-453.

Weiner, B. (1986). An attributional theory of motivation and emotion. New York, Springer.

Wentura, D., & Pospeschill, M. (2015). Multivariate Datenanalyse: Eine kompakte Einführung. Springer-Verlag.

Wickham, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software*, 21, 1–20.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York: Springer.

Wirth, W. (2013). Grundlagen emotionaler Medienwirkungen. In *Handbuch Medienwirkungsforschung*, 227-246. Springer VS, Wiesbaden.

Wirtz, M. & Caspar, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen, Hogrefe.

Yalcin-Siedentopf, N., Hoernagel, C.M., Biedermann, F., Baumgartner, S., Deisenhammer, E.A., Hausmann, A., et al. (2014). Facial affect recognition in symptomatically remitted patients with schizophrenia and bipolar disorder. *Schizophrenia Research*, 152, 440 – 445.

Young, A.W., Newcombe, F. de Haan, E.H.F., Small, M. & Hay, D.C. (1993). Face perception after brain injury. Selective impairments affecting identity and expression. *Brain*, 116, 941-959.

Young, A., Perrett, D., Calder, A., Sprengelmeyer, R., Ekman, P. (2002). The Ekman-60-Faces-Test. In Young, A., Perrett, D., Calder, A., Sprengelmeyer, R., Ekman, P., *Facial Expression of Emotion – Stimuli and Test. Psychology Manual v1.0, Sektion 1*, pp. 1-19. Bury St. Edmunds, UK: Thames Valley Test Company.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's a, Revelle's b, and McDonald's oH: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70(1), 123-133.

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for oh. *Applied Psychological Measurement*, 30(2), 121-144.

Anhang A

Stimulus- und Ausgangsmaterial des PCFAE

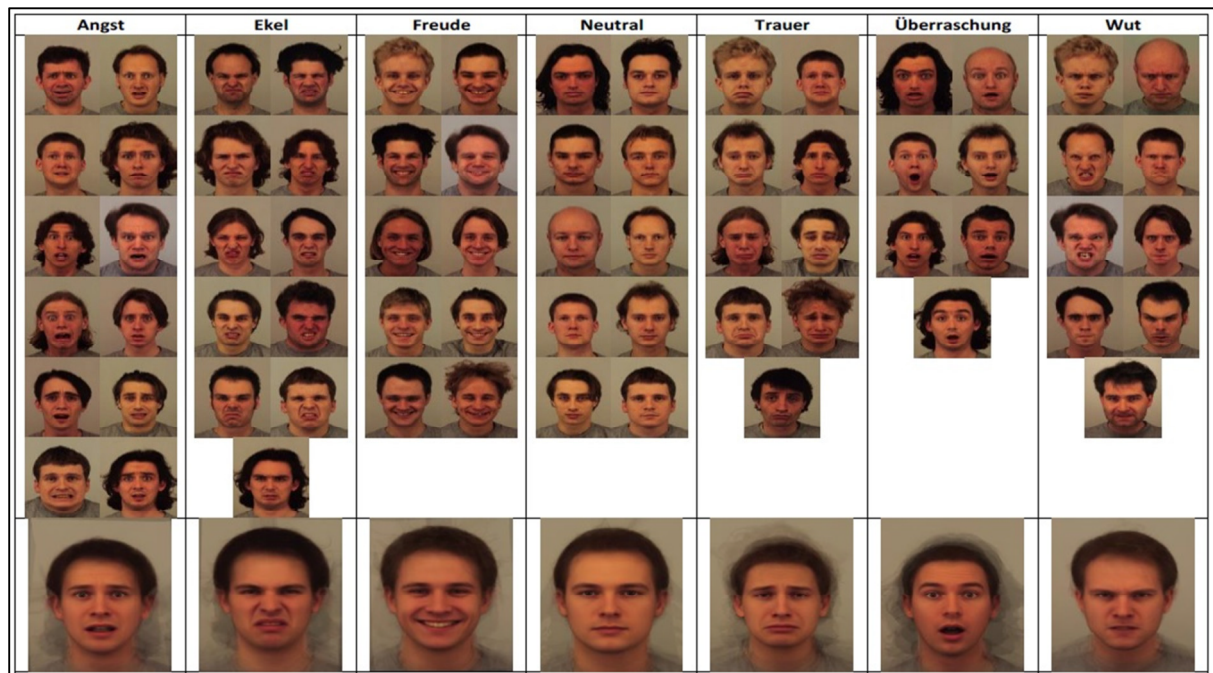


Abbildung A1.

Stimulus- und Ausgangsmaterial, männlich Erwachsene.

Die oberen Bilder zeigen die Originalfotografien des KDEF Stimulussets der prototypischen Affektdarstellung männlicher Erwachsener. Hieraus wurde das affektspezifische Stimulusmaterial für den PCFAE gemittelt (unterste Reihe). Der Morphingprozess des Stimulusmaterials hin zu einem Testitem wurde durch Mittelung eines affektspezifischen Stimulus mit dem neutralen Stimulus erreicht, sodass 30 Einzelbilder zunehmender Affektausprägung entstanden. Abschließend wurde auf die Peripherie eine ovale, schwarze Maske gelegt (siehe Kapitel Einleitung).



Abbildung A2.

Stimulus- und Ausgangsmaterial weiblich Erwachsene. Struktur siehe Abbildung A1.



Abbildung A3.

Stimulus- und Ausgangsmaterial, männlich Kinder. Struktur siehe Abbildung A1.



Abbildung A4. Stimulus- und Ausgangsmaterial, weiblich Kinder. Struktur siehe Abbildung A1.

Anhang B

In der Studie verwendete Probandenunterlagen

Probandeninformation

Sehr geehrter Versuchsteilnehmer,
vielen Dank, dass Sie sich entschieden haben, an unserer Studie zur Affektwahrnehmung teilzunehmen. Wir möchten Sie bitten, sich die folgenden Informationen aufmerksam durchzulesen. Der Versuchsleiter wird Sie zusätzlich mündlich in den Versuchsablauf einweisen und Ihnen eventuelle Fragen gerne beantworten. Bitte bestätigen Sie uns im Anschluss Ihre Teilnahmebereitschaft durch eine Unterschrift auf der beiliegenden Einverständniserklärung.

Die Studie wird vom Klinischen Institut für Psychosomatische Medizin und Psychotherapie unter Leitung von Herr Prof. Dr. Matthias Franz durchgeführt. Die Durchführung des Versuchs obliegt Herrn Peter Richter oder einem Vertreter. Sollten Sie nachträglich nach Abschluss des Versuchs noch Fragen haben, wenden Sie sich bitte an Peter.Richter@med.uni-duesseldorf.de.

Einleitung und Ziel der Studie

Ihnen werden gleich auf einem Bildschirm Gesichter präsentiert. Diese Gesichter verwandeln ihren Ausdruck von einem neutralen hin zu einem emotionalen Ausdruck. Ihre Aufgabe wird es sein, die gezeigten Affekte schnellstmöglich richtig zu bestimmen.

Ziel der Studie ist die Gewinnung von Leistungsdaten für die Emotionserkennung. Wir möchten untersuchen, wie schnell und wie gut Menschen unterschiedliche Affekte erkennen können. Weiterhin möchten wir diese Messung mit einer nachfolgenden zweiten Messung vergleichen und so überprüfen, wie messgenau der vorliegende Test die Fähigkeit zur Gesichtserkennung abbilden kann. Hierfür möchten wir Sie für einen Folgetermin in 4 Wochen einladen.

Die gewonnenen Daten dienen der Weiterentwicklung eines Tests, mit dem die menschliche Fähigkeit Emotionen zu erkennen besser erfasst werden kann. Die Verbesserungen umfassen vor allem eine Leistungskomponente, sprich wie schnell Sie in der Lage sind, die gezeigte Emotion zu erkennen.

Ablauf der Studie

Bevor wir beginnen, werden wir Sie vollständig in die Räumlichkeiten und den Ablauf der Studie einweisen. Anschließend fragen wir einige Ausschlusskriterien von der Untersuchung ab und bitten Sie, uns die freiwillige Teilnahme an der Untersuchung durch Ihre Unterschrift zu bestätigen.

Nachdem alle Vorbereitungen abgeschlossen sind und Sie keine weiteren Fragen zum Ablauf haben sollten, werden Sie gebeten, vor dem Computer Platz zu nehmen. Dort werden wir Sie zunächst bitten, einige Fragen zu beantworten. Diese Fragen dienen der Erfassung ihres psychischen Allgemeinzustandes sowie der generellen Fähigkeit, Gesichter und Emotionen zu erkennen. Wir erheben diese Daten um einschätzen zu können, ob Sie sich in guter gesundheitlicher Verfassung befinden oder ob gegebenenfalls bisher nicht erfasste Gründe für einen Ausschluss von der Untersuchung sprechen.

Die elektronischen Fragebögen umfassen demografische Daten, jeweils einen Fragebogen zum Wohlbefinden, zur Fähigkeit eigene Gefühle zu erkennen sowie zur Fähigkeit Gesichter zu erkennen. Die Einweisung in die jeweiligen Verfahren erfolgt unmittelbar zu Beginn der jeweiligen Testung auf dem Bildschirm.

Im Anschluss beginnen wir mit der eigentlichen Untersuchung. Hierfür werden wir Ihnen eine Abfolge von Bildern präsentieren, in denen sich eine Emotion schrittweise aufbaut. Ihre Aufgabe wird es sein, die präsentierte Emotion so schnell wie möglich zu richtig zu benennen. Eine genaue Erläuterung des Versuchsablaufs und Ihrer Aufgaben erfolgt unmittelbar vor Beginn der

Untersuchung auf dem Bildschirm des PCs. Sie werden vor Beginn die Möglichkeit erhalten, den Ablauf kurz beispielhaft zu üben.

Wiederholung der Studie

Da wir wissen wollen, ob sich die Testergebnisse bei wiederholter Messung ändern, werden innerhalb eines Zeitraums von einem Monat zwei Testungen durchgeführt. Für beide Testungen benötigen wir dieselben Versuchspersonen. Aus diesem Grund bitten wir Sie im Rahmen dieser Probandeninformation um die Hinterlegung Ihrer Kontaktdaten, sodass wir Sie für die Folgeuntersuchung erneut einladen können. Ihre Kontaktdaten werden nach Abschluss des zweiten Versuchsdurchgangs gelöscht, falls Sie dies wünschen.

Risiken und Vorsichtsmaßnahmen

Die Untersuchung beinhaltet keine körperliche Intervention oder Behandlung. Folglich sind nach menschlichem Ermessen keine Risiken durch die Untersuchung zu erwarten.

Ausschlusskriterien

Sollten Sie eines der folgenden Ausschlusskriterien erfüllen, können Sie leider nicht an der Untersuchung teilnehmen. Bitte teilen Sie uns das Vorliegen von Ausschlusskriterien umgehend mit.

- vergangene oder aktuelle psychiatrische, neurologische oder degenerative Erkrankungen
- schwere Gehirnerkrankung, Gehirnverletzung oder vergangene Gehirn-Operation
- akute Einnahme psychoaktiver Medikamente oder Substanzen inklusive Alkoholabusus
- schwere Stoffwechselstörung
- nicht-korrigierte ~~Visusminderung~~
- Einschränkungen der Reaktionsfähigkeit
- Unzureichende deutsche Sprachkenntnisse

Dauer der Studie und Aufwandsentschädigung

Die Gesamtdauer der Studie beträgt an zwei unterschiedlichen Terminen ca. 60 Minuten. Für die abgeschlossene Teilnahme an der Studie (beide Termine) erhalten Sie eine Aufwandsentschädigung in Höhe von 20€ oder als Student der Psychologie an der Heinrich Heine Universität Düsseldorf alternativ eine Bescheinigung über abgeleistete Versuchspersonenstunden von je einer VP Stunde je abgeschlossene 30 Minuten Versuchsteilnahme. Bitten denken Sie in diesem Falle daran, sich die Versuchspersonenstunden zusätzlich bei Professor Musch bestätigen zu lassen.

Wege-Unfall-Versicherung

Während des Experiments und auf dem direkten Weg zum Experiment sowie auf dem Weg vom Experiment nach Hause sind Sie im Rahmen einer Wege-Unfall-Versicherung über die Heinrich Heine Universität Düsseldorf versichert. Der Versicherungsschutz entfällt, wenn die normale Dauer des Weges verlängert bzw. der Weg selbst durch rein private Maßnahmen unterbrochen wird.

Studie:	PCFAE Test
Versicherer:	SV Sparkassen Versicherung über Marsh Medical Consulting GmbH
Anschrift:	Bismarckstraße 2, 32756 Detmold
Telefon:	+49 (0)5231 - 30819 106
Fax:	+49 (0)5231 - 30819 200

Abbildung B1. Probandeninformationen zur Studie.

Einwilligungserklärung zur Studie

„PCFAE-Studie“

Klinisches Institut für Psychosomatische Medizin und Psychotherapie
Universitätsklinikum Düsseldorf (Geb. 15.16)
Moorenstraße 5
40225 Düsseldorf
Tel 0211-8118338
Fax 0211-8116250

Nach ausführlicher schriftlicher und mündlicher Aufklärung über das Ziel der Studie, die Art der hierfür eingesetzten Methoden und Untersuchungsmittel, meine Rechte und Pflichten und den mir zustehenden Versicherungsschutz erkläre ich mich freiwillig zur Teilnahme bereit. Ich hatte Gelegenheit Fragen zu stellen. Diese wurden zufriedenstellend und vollständig beantwortet. Die Teilnahmebereitschaft kann durch mich jederzeit mündlich oder schriftlich widerrufen werden. Sämtliche personenbezogenen und experimentellen Daten werden in diesem Falle gelöscht und vernichtet. Ich erkläre mich damit einverstanden, dass meine Kontaktdaten zur Einladung für die Kontaktaufnahme zur 2. Testung (Retest) genutzt werden und dass die Aufwandsentschädigung erst nach vollständigem Abschluss beider Testungen vergeben wird. Bei Rücktritt von der Studienteilnahme entfällt die Aufwandsvergütung.

Persönliche Angaben:

Name _____
Vorname _____
Straße, Nr. _____
PLZ, Wohnort _____
Telefon _____
E-Mail-Adresse _____

Unter den oben aufgeführten Bedingungen stimme ich der Teilnahme an der wissenschaftlichen Studie zu.

Düsseldorf, den _____ Unterschrift: _____

Abbildung B2.

Erklärung zur freiwilligen Teilnahme an der Studie.

Mir ist bekannt und ich willige ein, dass bei dieser Studie personenbezogene Daten über mich erhoben, gespeichert und ausgewertet werden sollen. Die Verwendung der Angaben erfolgt nach gesetzlichen Bestimmungen und setzt vor der Teilnahme an der Studie folgende freiwillig abgegebene Einwilligungserklärung voraus, d.h. ohne die nachfolgende Einwilligung kann ich nicht an der Studie teilnehmen.

Einwilligungserklärung zum Datenschutz

1) Ich willige ein, dass im Rahmen dieser Studie erhobene Daten in Papierform oder auf elektronischen Datenträgern im Institut für Psychosomatische Medizin und Psychotherapie pseudonymisiert aufgezeichnet und gespeichert werden. Sie werden mit einer Identifikationsnummer versehen und getrennt von meinen persönlichen Angaben in einem EDV-System gespeichert und wissenschaftlich von der Universität Düsseldorf ausgewertet. Die Studienergebnisse werden ohne Bezug zu Ihrer Person veröffentlicht. Ein Personenbezug ist damit für Dritte nicht mehr herstellbar.

2) Ich bin darüber aufgeklärt worden, dass ich meine Einwilligung in die Aufzeichnung, Speicherung und Verwendung meiner Daten jederzeit widerrufen kann. Bei einem Widerruf werden meine Daten unverzüglich gelöscht.

3) Ich willige ein, dass meine Daten nach Beendigung oder Abbruch der Studie 10 Jahre lang aufbewahrt werden. Danach werden meine personenbezogenen Daten gelöscht, soweit dem nicht gesetzliche, satzungsgemäße oder vertragliche Aufbewahrungsfristen entgegenstehen.

Name Studienteilnehmer (Druckbuchstaben)

Datum

Unterschrift

Abbildung B3.

Datenschutzerklärung zur Studie.

Versuchspersonen Code

Der Versuchspersonen Code wird benötigt um Ihre Daten aus den einzelnen Aufgaben zuordnen zu können. Ihre Daten werden so nicht mit Ihrem Namen, sondern stets nur mit Ihrem Versuchspersonen Code in Verbindung gebracht und pseudonymisiert gespeichert.

Erster Buchstabe des Vornamens Ihrer Mutter: _____

Erster Buchstabe des Geburtsmonats Ihrer Mutter: _____

Der Tag Ihres eigenen Geburtstages (2 Ziffern): _____

Erster und letzte Buchstabe Ihres eigenen Geburtsorts: _____

Beispiel: aus Angela, Juli, 07.08.1995, Düsseldorf ergibt sich der Code: AJ07DF

Ihr Versuchspersonen Code lautet: _____

Abbildung B4. Generierung des Versuchspersonencodes.

Anamnesebogen	
Datum: _____	VP-Code: _____
1) Wie alt sind Sie?	
_____ Jahre	
2) Welches Geschlecht haben Sie?	
☐ weiblich ☐ männlich	
3) Was ist Ihr höchster Bildungsabschluss?	
☐ Hauptschule ☐ Realschule ☐ Abitur ☐ FH/Universität	
4) Was ist Ihre Muttersprache?	
☐ Deutsch	
☐ Andere _____	
_____ Jahre Deutschkenntnisse	
5) Sind Sie...	
☐ ledig ☐ geschieden	
☐ verheiratet ☐ verwitwet	

Abbildung B5. Erfassung der demografischen Daten.

Anhang C

In der Studie verwendete psychometrische Instrumente

TAS-20

Name:		Vorname:		Datum:	
Geschlecht: männlich ☐		weiblich ☐		Geburtsdatum:	
Höchster Schulabschluß: keiner ☐		mittlere Reife ☐		Beruf:	
Sonderschule ☐		Abitur/Fachabitur ☐		Mutter-	
Hauptschule ☐		Universität / FH ☐		sprache:	
Ich bin Rechtshänder ☐		Linkshänder ☐			
Waren Sie im letzten viertel Jahr in ärztlicher oder zahnärztlicher Behandlung? Wenn ja, wo?					
Nehmen Sie Medikamente? Wenn ja, welche?					
		Telefon:			

Bitte kreuzen Sie an, inwieweit die folgenden 20 Aussagen auf Sie zutreffen. Sie können den Grad Ihrer Zustimmung oder Ablehnung anhand einer 5-Punkte-Skala einschätzen:

1 = trifft überhaupt nicht zu
 2 = trifft selten zu
 3 = trifft teilweise zu, teilweise auch nicht
 4 = trifft oft zu
 5 = trifft immer zu

1. Mir ist oft unklar, welche Gefühle ich gerade habe.	1---2---3---4---5
2. Es fällt mir schwer, die richtigen Worte für meine Gefühle zu finden.	1---2---3---4---5
3. Ich habe körperliche Empfindungen, die sogar die Ärzte nicht verstehen.	1---2---3---4---5
4. Es fällt mir leicht, meine Gefühle zu beschreiben.	1---2---3---4---5
5. Ich gehe Problemen lieber auf den Grund, als sie nur zu beschreiben	1---2---3---4---5
6. Wenn mich etwas aus der Fassung gebracht hat, weiß ich oft nicht, ob ich traurig, ängstlich oder wütend bin.	1---2---3---4---5
7. Ich bin oft über Vorgänge in meinem Körper verwirrt.	1---2---3---4---5
8. Ich lasse die Dinge lieber einfach geschehen und versuche nicht herauszufinden, warum sie gerade so passiert sind.	1---2---3---4---5
9. Einige meiner Gefühle kann ich gar nicht richtig benennen.	1---2---3---4---5
10. Sich mit Gefühlen zu beschäftigen finde ich sehr wichtig.	1---2---3---4---5
11. Ich finde es schwierig zu beschreiben, was ich für andere Menschen empfinde.	1---2---3---4---5
12. Andere fordern mich auf, meine Gefühle mehr zu beschreiben.	1---2---3---4---5
13. Ich weiß nicht, was in mir vorgeht.	1---2---3---4---5
14. Ich weiß oft nicht, warum ich wütend bin.	1---2---3---4---5
15. Ich unterhalte mich mit anderen nicht so gerne über ihre Gefühle, sondern lieber darüber, womit sie sich täglich beschäftigen.	1---2---3---4---5
16. Ich sehe mir lieber „leichte“ Unterhaltungssendungen als psychologische Problemfilme an.	1---2---3---4---5
17. Es fällt mir schwer, selbst engen Freunden gegenüber meine innersten Gefühle mitzuteilen.	1---2---3---4---5
18. Ich kann mich jemandem sogar in Augenblicken des Schweigens sehr nahe fühlen.	1---2---3---4---5
19. Ich finde, daß das Mir-Klarwerden über meine persönlichen Gefühle wichtig ist, wenn ich persönliche Probleme lösen muß	1---2---3---4---5
20. Durch die Suche nach verborgenen Bedeutungen nimmt man sich das Vergnügen an Filmen oder Theaterstücken.	1---2---3---4---5

TAS-20 (Taylor, Bagby & Parker 1992)

Abbildung C1. 20-Item Toronto Alexithymia Scale (TAS-20).

Skalen zum Erleben von Emotionen (SEE)

Sie finden auf den folgenden Seiten Aussagen über die Wahrnehmung von und den Umgang mit Gefühlen. Bei jeder Aussage können Sie ankreuzen, wie sehr diese für Sie persönlich stimmt. Es gibt keine richtigen oder falschen Antworten – wichtig ist, was für Sie individuell zutrifft.

Bitte versuchen Sie nicht, die Antwort zu suchen, die den „besten Eindruck“ macht, sondern antworten Sie nach Ihrem persönlichen Empfinden. Auch wenn einmal eine Fragestellung nicht exakt auf Sie persönlich zugeschnitten ist, bitte kreuzen Sie jede Frage bestmöglich an.

		stimmt gar nicht	stimmt kaum	stimmt mittel- mäßig	stimmt ziemlich	stimmt völlig
1.	Durch meine Träume werden mir meine Gefühle klarer.	☐	☐	☐	☐	☐
2.	Ich kenne schon Möglichkeiten, wie ich meinen Gefühlszustand steuern kann.	☐	☐	☐	☐	☐
3.	Ich habe einige Gefühle, die ich besser nicht hätte.	☐	☐	☐	☐	☐
4.	Meine Gefühle spiegeln sich oft auch in Körperempfindungen wider.	☐	☐	☐	☐	☐
5.	Meine Selbstbeherrschung lässt zu wünschen übrig.	☐	☐	☐	☐	☐
6.	Bei der Bewältigung von Vergangenem helfen mir meine Phantasien.	☐	☐	☐	☐	☐
7.	Ich stehe zu allen meinen Empfindungen.	☐	☐	☐	☐	☐
8.	Meine Körpersignale bemerke ich schnell.	☐	☐	☐	☐	☐
9.	Gottseidank habe ich meine Gefühle gut im Griff.	☐	☐	☐	☐	☐
10.	Die Wahrnehmung von inneren Verkrampfungen, Entspannungen usw. hilft mir bei Entscheidungen.	☐	☐	☐	☐	☐
11.	Ich wünschte, ich würde nicht immer so von meinen Emotionen hin und her geworfen.	☐	☐	☐	☐	☐
12.	Mein inneres Erleben fühle ich oft nicht.	☐	☐	☐	☐	☐
13.	Oft hilft es mir, mich bei Belastungen Träumereien hinzugeben.	☐	☐	☐	☐	☐
14.	Wenn ich in eine bessere, lebendigere Stimmung kommen will, kann ich mich gut selbst beeinflussen.	☐	☐	☐	☐	☐
15.	Auch wenn ich innerlich brodele, kann ich nach außen hin ruhig wirken.	☐	☐	☐	☐	☐
16.	Ich fühle, was ich fühle, und das ist ok.	☐	☐	☐	☐	☐
17.	Oft wünsche ich, ich könnte meine Gefühle besser wahrnehmen.	☐	☐	☐	☐	☐
18.	Ich finde Tagträume nützlich.	☐	☐	☐	☐	☐
19.	Mein körperliches Befinden entspricht meist meinem seelischen Befinden.	☐	☐	☐	☐	☐

		stimmt gar nicht	stimmt kaum	stimmt mittel- mäßig	stimmt ziemlich	stimmt völlig
20.	Ich bin so voller Gefühle, dass ich mich oft damit nicht ausstehen kann.	☐	☐	☐	☐	☐
21.	Wenn ich aufgeregt bin, weiß ich meist, wie ich mich beruhigen kann.	☐	☐	☐	☐	☐
22.	Bei Entscheidungen verlasse ich mich auf meine leiblichen Empfindungen.	☐	☐	☐	☐	☐
23.	Es ist für mich keine Frage, dass ich ein Recht auf alle meine Gefühle habe.	☐	☐	☐	☐	☐
24.	Ich habe mich selbst immer unter Kontrolle.	☐	☐	☐	☐	☐
25.	Manchmal ärgere ich mich über meine Empfindungen.	☐	☐	☐	☐	☐
26.	Meine Tagträume geben mir Hinweise auf meine Bedürfnisse und Wünsche.	☐	☐	☐	☐	☐
27.	Empfindungen wie Herzklopfen, Magendrücken, Hautkribbeln sind für mich eine gute Orientierung für das, was ich will.	☐	☐	☐	☐	☐
28.	Ich wünschte, ich könnte meine Gedanken manchmal abstellen.	☐	☐	☐	☐	☐
29.	Ich würde gerne innerlich mehr erleben.	☐	☐	☐	☐	☐
30.	Für meine Gefühle schäme ich mich nicht.	☐	☐	☐	☐	☐
31.	Ich habe viele Gefühle in mir, die ich am liebsten beseitigen würde.	☐	☐	☐	☐	☐
32.	Wenn ich mich in einer Situation unwohl fühle, dann merke ich dies z.B. im Magen, an der Haut, an Verspannungen o.ä.	☐	☐	☐	☐	☐
33.	Wenn ich innerlich brodele, merkt meine Umwelt das leider sofort.	☐	☐	☐	☐	☐
34.	Alle meine Gefühle dürfen genauso sein, wie sie sind.	☐	☐	☐	☐	☐
35.	Körperliche Beschwerden empfinde ich oft als Ausdruck seelischen Unbehagens.	☐	☐	☐	☐	☐
36.	Die Signale meines Körpers spüre ich oft nicht.	☐	☐	☐	☐	☐
37.	Zu oft werde ich von meinen Gefühlen überschwemmt.	☐	☐	☐	☐	☐
38.	Meine Beziehungen zu anderen verstehe ich durch meine Träume besser.	☐	☐	☐	☐	☐
39.	Wenn ich will, kann ich meine Emotionen durchaus manipulieren.	☐	☐	☐	☐	☐
40.	Andere können mir in der Regel nicht ansehen, was in mir los ist.	☐	☐	☐	☐	☐
41.	Meine Intuition hat für mich viel mit der Wahrnehmung meines Lebens zu tun.	☐	☐	☐	☐	☐
42.	Was bei mir an Gefühlen hochkommt, kann ich immer zunächst auch annehmen.	☐	☐	☐	☐	☐

Abbildung C2. Skalen zum Erleben von Emotionen (SEE).

BSI

Sie finden nachstehend eine Liste von Problemen und Beschwerden, die man manchmal hat. Bitte lesen Sie jede Frage einzeln sorgfältig durch und entscheiden Sie, wie stark Sie durch diese Beschwerden gestört oder bedrängt worden sind, und zwar während der vergangenen sieben Tage bis heute. Überlegen Sie bitte nicht erst, welche Antwort „den besten Eindruck“ machen könnte, sondern antworten Sie so, wie es für Sie persönlich zutrifft. Machen Sie bitte hinter jede Frage ein Kreuz bei der für Sie am besten zutreffenden Antwort.

Bitte beantworten Sie jede Frage!

Wie sehr litten Sie in den letzten 7 Tagen unter...	überhaupt nicht	ein wenig	ziemlich	stark	sehr stark
1. Nervosität oder innerem Zittern	0	1	2	3	4
2. Ohnmachts- und Schwindelgefühlen	0	1	2	3	4
3. der Idee, dass irgend jemand Macht über Ihre Gedanken hat	0	1	2	3	4
4. dem Gefühl, dass andere an den meisten Ihrer Schwierigkeiten Schuld sind	0	1	2	3	4
5. Gedächtnisschwierigkeiten	0	1	2	3	4
6. dem Gefühl, leicht reizbar oder verärgert zu sein	0	1	2	3	4
7. Herz- oder Brustschmerzen	0	1	2	3	4
8. Furcht auf offenen Plätzen oder auf der Straße	0	1	2	3	4
9. Gedanken, sich das Leben zu nehmen	0	1	2	3	4
10. dem Gefühl, dass man den meisten Menschen nicht trauen kann	0	1	2	3	4
11. schlechtem Appetit	0	1	2	3	4
12. plötzlichem Erschrecken ohne Grund	0	1	2	3	4
13. Gefühlsausbrüchen, denen gegenüber Sie machtlos waren	0	1	2	3	4
14. Einsamkeitsgefühlen, selbst wenn Sie in Gesellschaft sind	0	1	2	3	4
15. dem Gefühl, dass es Ihnen schwerfällt, etwas anzufangen	0	1	2	3	4
16. Einsamkeitsgefühlen	0	1	2	3	4
17. Schermut	0	1	2	3	4
18. dem Gefühl, sich für nichts zu interessieren	0	1	2	3	4
19. Furchtsamkeit	0	1	2	3	4
20. Verletzlichkeit in Gefühlsdingen	0	1	2	3	4
21. dem Gefühl, dass die Leute unfreundlich sind oder Sie nicht leiden können	0	1	2	3	4
22. Minderwertigkeitsgefühlen gegenüber anderen	0	1	2	3	4
23. Übelkeit oder Magenverstimmung	0	1	2	3	4
24. dem Gefühl, dass andere Sie beobachten oder über Sie	0	1	2	3	4

reden					
25. Einschlafschwierigkeiten	0	1	2	3	4
26. dem Zwang, wieder und wieder nachzukontrollieren, was Sie tun	0	1	2	3	4
27. Schwierigkeiten, sich zu entscheiden	0	1	2	3	4
28. Furcht vor Fahrten in Bus, Straßenbahn, U-Bahn oder Zug	0	1	2	3	4
29. Schwierigkeiten beim Atmen	0	1	2	3	4
30. Hitzewallungen oder Kälteschauern	0	1	2	3	4
31. der Notwendigkeit, bestimmte Dinge, Orte oder Tätigkeiten zu meiden, weil Sie durch diese erschreckt werden	0	1	2	3	4
32. Leere im Kopf	0	1	2	3	4
33. Taubheit oder Kribbeln in einzelnen Körperteilen	0	1	2	3	4
34. dem Gefühl, dass Sie für Ihre Sünden bestraft werden sollten	0	1	2	3	4
35. einem Gefühl der Hoffnungslosigkeit angesichts der Zukunft	0	1	2	3	4
36. Konzentrationsschwierigkeiten	0	1	2	3	4
37. Schwächegefühl in einzelnen Körperteilen	0	1	2	3	4
38. dem Gefühl, gespannt oder aufgeregt zu sein	0	1	2	3	4
39. Gedanken an den Tod und ans Sterben	0	1	2	3	4
40. dem Drang, jemanden zu schlagen, zu verletzen oder ihm Schmerz zuzufügen	0	1	2	3	4
41. dem Drang, Dinge zu zerbrechen oder zu zerschmettern	0	1	2	3	4
42. starker Befangenheit im Umgang mit anderen	0	1	2	3	4
43. Abneigung gegen Menschenmengen, z.B. beim Einkaufen oder im Kino	0	1	2	3	4
44. dem Eindruck, sich einer anderen Person nie so richtig nahe fühlen zu können	0	1	2	3	4
45. Schreck- oder Panikanfällen	0	1	2	3	4
46. der Neigung, immer wieder in Erörterungen und Auseinandersetzungen zu geraten	0	1	2	3	4
47. Nervosität, wenn Sie allein gelassen werden	0	1	2	3	4
48. mangelnder Anerkennung Ihrer Leistungen durch andere	0	1	2	3	4
49. so starker Ruhelosigkeit, dass Sie nicht stillsitzen können	0	1	2	3	4
50. dem Gefühl, wertlos zu sein	0	1	2	3	4
51. dem Gefühl, dass die Leute Sie ausnutzen, wenn Sie es zulassen würden	0	1	2	3	4
52. Schuldgefühlen	0	1	2	3	4
53. dem Gedanken, dass irgendetwas mit Ihrem Verstand nicht in Ordnung ist	0	1	2	3	4

Abbildung C3. Brief Symptom Inventory (BSI).

Tabelle C1

Im Famous Faces Test als Stimuli verwendete Persönlichkeiten

Abschnitt 1: Identifizieren	Abschnitt 2: Wiedererkennen
Rowan Atkinson	Franz Beckenbauer *
Franz Beckenbauer	David Beckham
Boris Becker	Justin Bieber *
Justin Bieber	Hilary Clinton
George Clooney	George Clooney *
Leonardo DiCaprio	Leonardo DiCaprio *
Mahatma Ghandi	Albert Einstein
Thomas Gottschalk	Tom Hanks
Günther Jauch	Madonna
John F. Kennedy	Mahatma Ghandi *
Joachim Löw	Günther Jauch *
Angela Merkel	Heidi Klum
Marilyn Monroe	Thomas Müller
Barack Obama	Barack Obama *
Wladimir Putin	Brad Pitt
Stefan Raab	Queen Elizabeth II
Rihanna	Michael Schuhmacher *
Michael Schumacher	Arnold Schwarzenegger *
Arnold Schwarzenegger	Matthias Schweighöfer
Donald Trump	Donald Trump *

Hinweis. * = wiederzuerkennende Persönlichkeiten
vgl. ebenfalls Jansen (2018a)

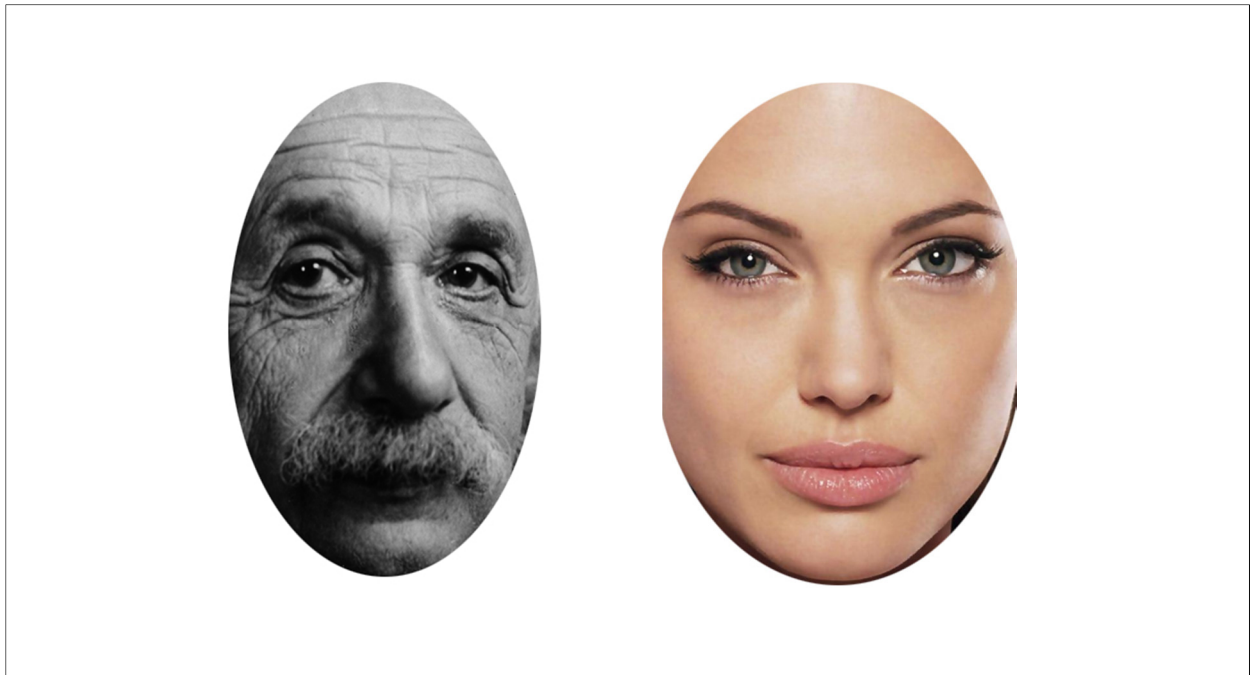
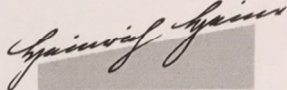


Abbildung C4. Beispielhafte Darstellung zweier Stimuli des Famous Faces Test.

Anhang D

Ethikvotum zur Studie

Medizinische Fakultät
Ethikkommission


HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Heinrich-Heine-Universität Düsseldorf ☎ 40204 Düsseldorf

Herr
Prof. Dr. Matthias Franz
Klinisches Institut für Psychosomatische Medizin
und Psychotherapie

- HIER -

Prof. Dr. med. T. Hohlfeld
Vorsitzender

Prof. Dr. rer. nat. K.-D. Kröncke
Geschäftsführer

Dr. rer. nat. N. Fitzner
Wissenschaftliche Mitarbeiterin

Telefon 0211 / 81-19591
Telefax 0211 / 81-19592
Ethikkommission@
med.uni-duesseldorf.de

Bitte stets angeben:

Studiennummer: 5977R

Registrierungs-ID: 2017044246

**„Reliabilität des Test of Perceptual Competence of Facial Affect
Recognition (PCFAE Test)“**

Ethikkommission an der Medi-
zinischen Fakultät der HHU
Düsseldorf

Kinderklinik Geb. 13.41
Ebene 03 Raum 34
Moorenstr 5
40225 Düsseldorf

Sehr geehrter Herr Prof. Franz,

die Ethikkommission (EK) der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf bestätigt den Erhalt Ihrer modifizierten Unterlagen vom 12.07.2017 (Eingang hier: 26.07.2017)

www.uni-duesseldorf.de

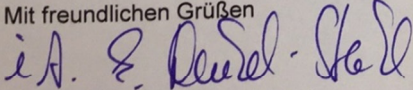
Nachdem Sie die Fragen der EK zufriedenstellend beantwortet bzw. die Auflagen erfüllt haben, bestehen keine ethischen oder rechtlichen Bedenken gegen die Durchführung Ihrer Studie..

Düsseldorf, 10.08.2017

Nach Abschluss des Projektes bitten wir um Übersendung eines knappen Schlussberichtes oder einer abschließenden Publikation.

Für die Durchführung Ihrer Studie wünschen wir viel Erfolg!

Mit freundlichen Grüßen



Prof. Dr. Klaus-Dietrich Kröncke
i. A. der Kommission

Anhang E

Weitere Verteilungsplots des PCFAE

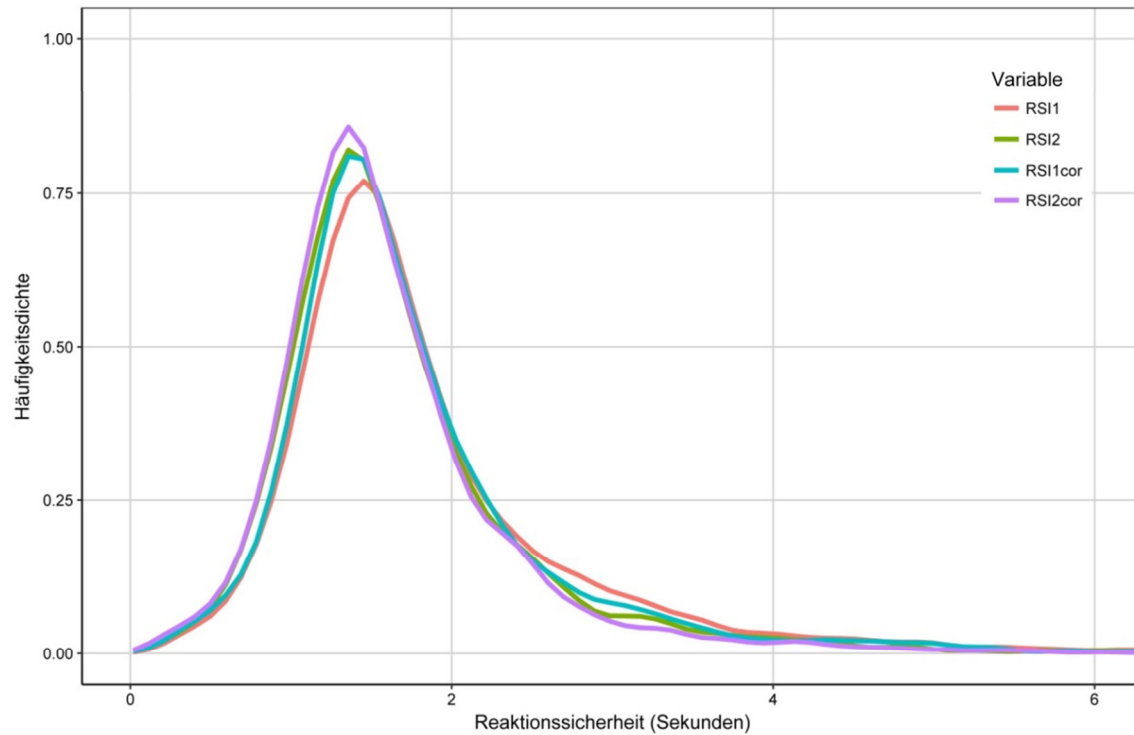


Abbildung E1. Verteilung der Reaktionssicherheitsmittelwerte der Versuchspersonen des PCFAE (Gesamttest) nach Messzeitpunkt. RSI 1/2 = Reaktionssicherheit, alle Antworten, Messzeitpunkt T1/T2. RSI1/2cor = Reaktionssicherheit für korrekte Antworten, Messzeitpunkt T1/T2.

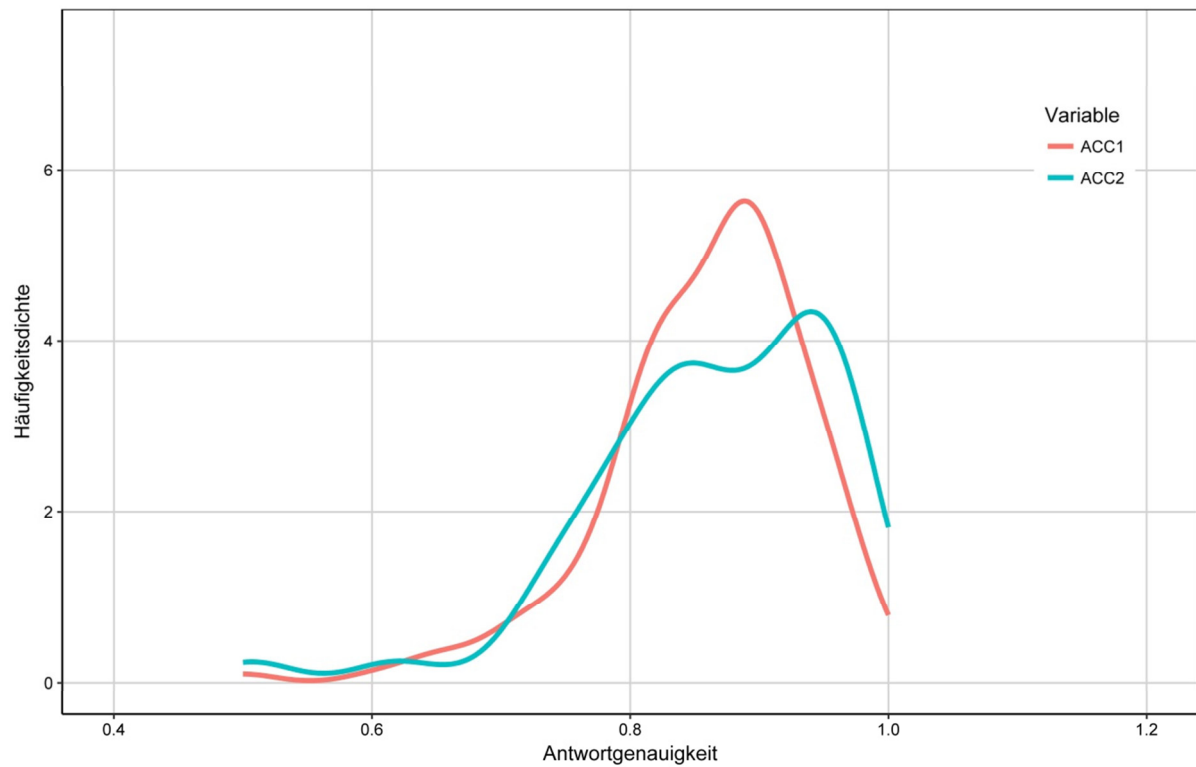


Abbildung E2. Verteilung der Mittelwerte der Antwortgenauigkeit der Versuchspersonen des PCFAE (Gesamttest) nach Messzeitpunkt. ACC1/2 = Antwortgenauigkeit zu Messzeitpunkt T1/T2

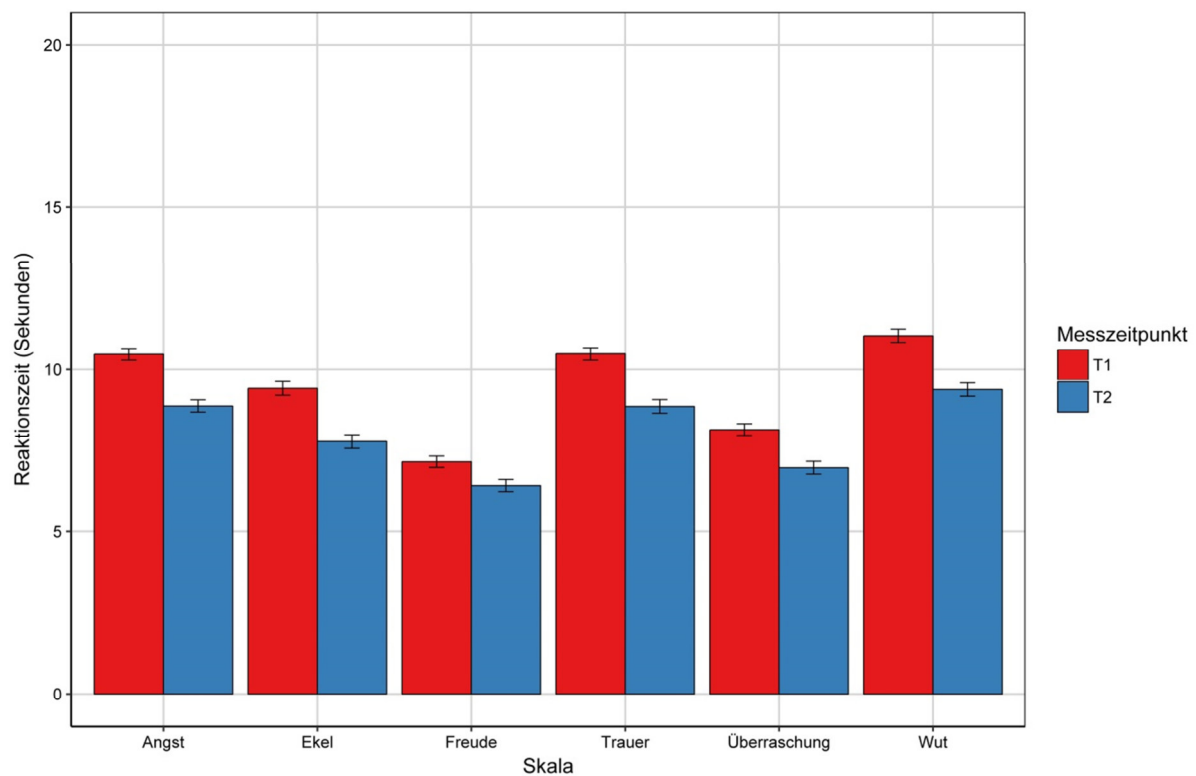


Abbildung E3. Mittelwerte und Standardfehler des Mittelwerts der Reaktionszeit der Versuchspersonen des PCFAE nach Affektskala und Messzeitpunkt. Darstellung für korrekt beantwortete Items.

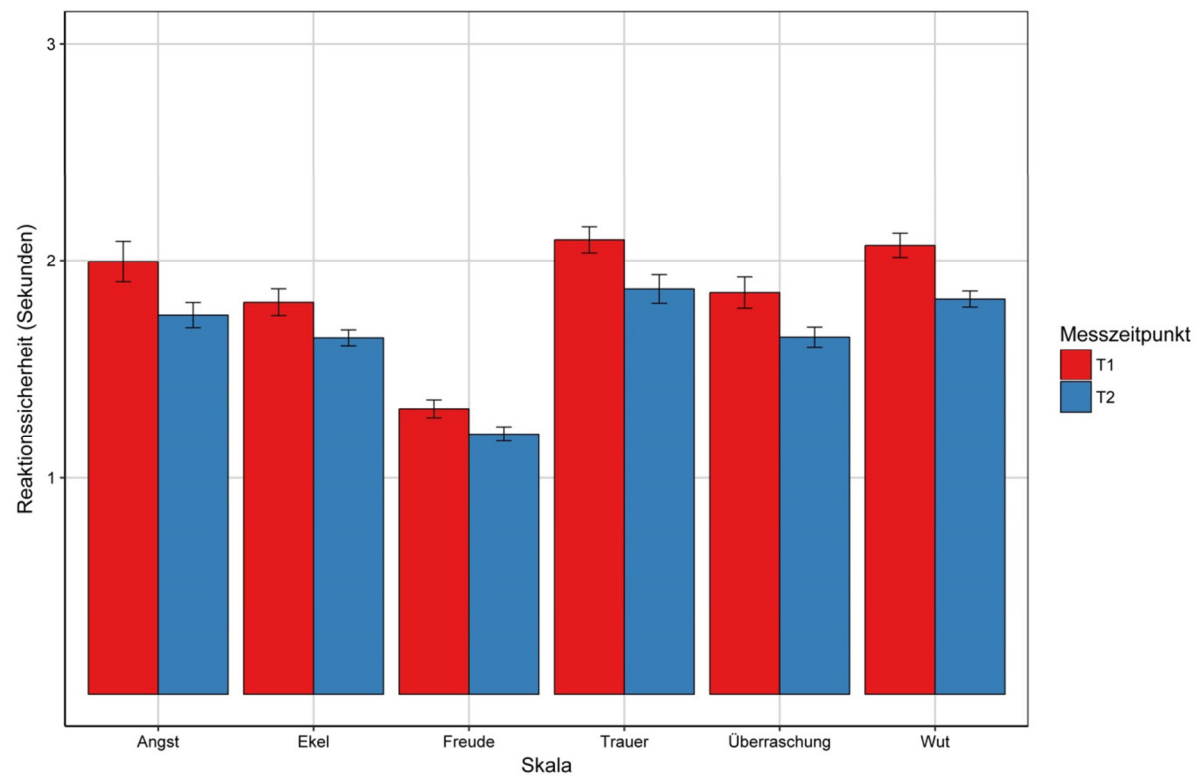


Abbildung E4. Mittelwerte und Standardfehler des Mittelwerts der Reaktionssicherheit der Versuchspersonen des PCFAE nach Affektskala und Messzeitpunkt. Darstellung für korrekt beantwortete Items.

Anhang F

Deskriptive Statistiken der Items des PCFAE

Tabelle F1

Deskriptive Statistiken der Items des PCFAE.

Item			RT (T1)		RT (T2)		RSI (T1)		RSI (T2)		ACC (T1)		ACC (T2)	
Skala	Geschlecht	Alter	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Angst	weiblich	Erwachsen	10.33	2.52	8.95	2.44	1.96	1.44	1.74	1.32	0.77	0.33	0.78	0.33
		Kind	10.54	3.07	8.79	2.83	1.95	2.34	1.75	0.74	0.83	0.29	0.86	0.27
	männlich	Erwachsen	10.54	3.07	9.28	2.90	2.12	2.03	1.74	0.86	0.68	0.37	0.71	0.37
		Kind	10.88	2.96	8.92	2.72	1.82	1.27	1.69	0.97	0.56	0.38	0.58	0.40
Ekel	weiblich	Erwachsen	7.98	2.49	6.59	2.79	1.80	1.07	1.61	0.60	0.84	0.30	0.81	0.31
		Kind	10.10	3.41	7.90	2.99	1.77	0.77	1.57	0.60	0.85	0.27	0.88	0.26
	männlich	Erwachsen	8.09	3.16	7.31	2.94	1.63	0.81	1.64	0.59	0.83	0.29	0.81	0.31
		Kind	11.94	3.42	9.77	3.44	1.97	1.41	1.81	0.93	0.59	0.37	0.67	0.38
Freude	weiblich	Erwachsen	6.86	2.57	6.09	2.70	1.27	0.51	1.23	0.59	0.99	0.08	0.99	0.06
		Kind	8.03	2.75	7.66	3.14	1.38	0.78	1.24	0.56	0.97	0.12	0.95	0.17
	männlich	Erwachsen	6.30	2.21	5.57	2.25	1.34	0.84	1.13	0.50	0.99	0.07	0.98	0.09
		Kind	7.44	2.99	6.41	2.59	1.28	0.80	1.20	0.56	0.99	0.07	0.98	0.10
Trauer	weiblich	Erwachsen	9.47	2.39	8.16	2.58	1.92	0.89	1.75	0.71	0.95	0.15	0.94	0.18
		Kind	12.88	4.16	10.76	4.08	2.43	1.57	2.13	1.61	0.86	0.26	0.85	0.29
	männlich	Erwachsen	8.86	2.48	7.61	2.54	1.95	0.96	1.77	1.56	0.94	0.18	0.95	0.16
		Kind	10.81	3.05	9.18	2.98	2.10	1.13	1.84	1.25	0.95	0.15	0.95	0.16
Überr.	weiblich	Erwachsen	7.70	2.80	6.82	2.61	1.75	0.87	1.68	1.07	0.93	0.18	0.94	0.17
		Kind	8.39	3.40	6.84	2.70	1.97	1.98	1.66	0.87	0.85	0.24	0.83	0.30
	männlich	Erwachsen	7.68	2.72	6.72	3.08	1.85	0.87	1.58	0.68	0.84	0.27	0.82	0.28
		Kind	8.80	2.63	7.37	2.66	1.78	0.97	1.64	0.85	0.91	0.20	0.92	0.20
Wut	weiblich	Erwachsen	9.50	2.84	8.20	2.58	2.11	1.16	1.80	0.73	0.89	0.22	0.90	0.21
		Kind	13.56	3.38	12.19	3.20	2.15	1.09	1.86	0.61	0.77	0.35	0.77	0.33
	männlich	Erwachsen	10.49	2.69	8.63	2.71	1.87	0.68	1.76	0.59	0.94	0.16	0.94	0.18
		Kind	10.93	3.70	9.07	3.45	2.16	1.45	1.84	0.73	0.87	0.23	0.87	0.25

Hinweis. RT = Reaktionszeit korrekt beantworteter Items, RSI = Reaktionssicherheit korrekt beantworteter Items, ACC = Antwortgenauigkeit, T1 = Messzeitpunkt T1, T2 = Messzeitpunkt T2, M = Mittelwert, SD = Standardabweichung.
 Überr. = Überraschung.

Anhang G

Korrelationstabellen zur explorativen Faktorenanalyse

Tabelle G1

Korrelationsmatrix der Faktoren: Korrelierte Faktoren Modell, korrekte Reaktionszeit Messzeitpunkt T1.

		Faktor					
		F1	F2	F3	F4	F5	F6
Faktor	F1	1					
	F2	0.53	1				
	F3	0.33	0.45	1			
	F4	0.34	0.38	0.40	1		
	F5	0.43	0.35	0.26	0.36	1	
	F6	0.46	0.52	0.32	0.24	0.30	1
<i>Hinweis.</i>	Interpretation der Faktoren:						
	F1	= Faktor Angst		F2	= Faktor Ekel		
	F3	= Faktor Freude		F4	= Faktor Trauer		
	F5	= Faktor Überraschung		F6	= Faktor Wut		

Tabelle G2

Korrelationsmatrix der Faktoren: Bi-Faktor Model, korrekte Reaktionszeit Messzeitpunkt T1.

		Faktor					
		F1	F2	F3	F4	F5	F6
Faktor	F1	1					
	F2	0.27	1				
	F3	0.21	0.35	1			
	F4	0.31	0.43	0.34	1		
	F5	0.33	0.35	0.37	0.39	1	
	F6	0.20	0.41	0.30	0.35	0.21	1
<i>Hinweis.</i>	Interpretation der Faktoren:						
	F1	= Faktor Angst		F2	= Faktor Ekel		
	F3	= Faktor Freude		F4	= Faktor Trauer		
	F5	= Faktor Überraschung		F6	= Faktor Wut		

Anhang H

Weitere Modelle des PCFAE

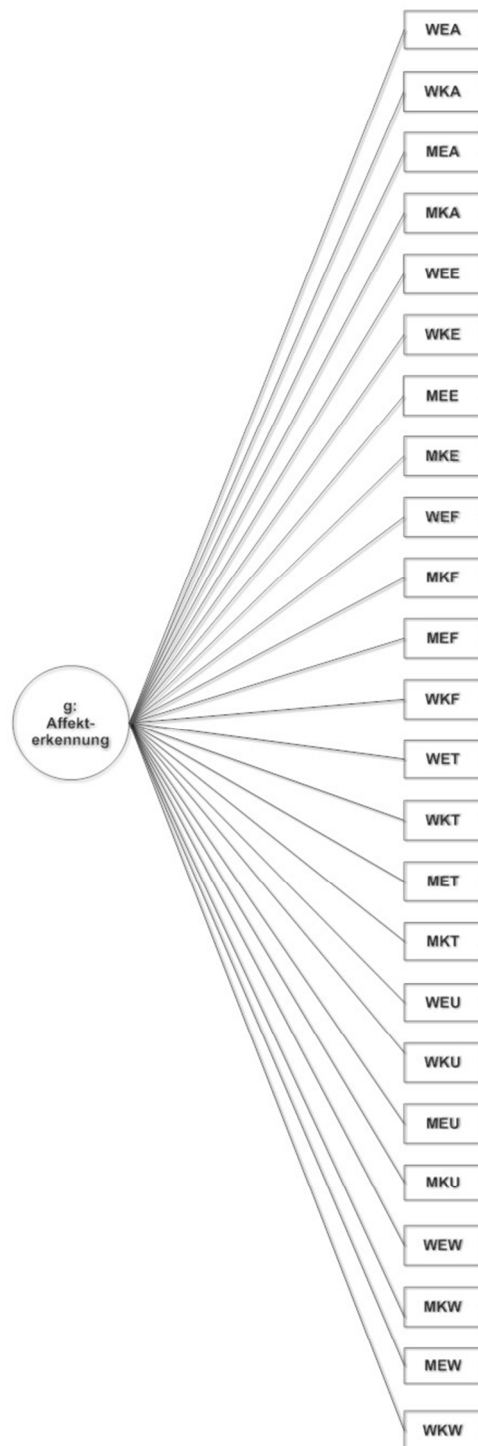


Abbildung G1. Unidimensionales Modell des PCFAE (Seidler, 2008). Schema der Itembenennung gemäß Anhang I.

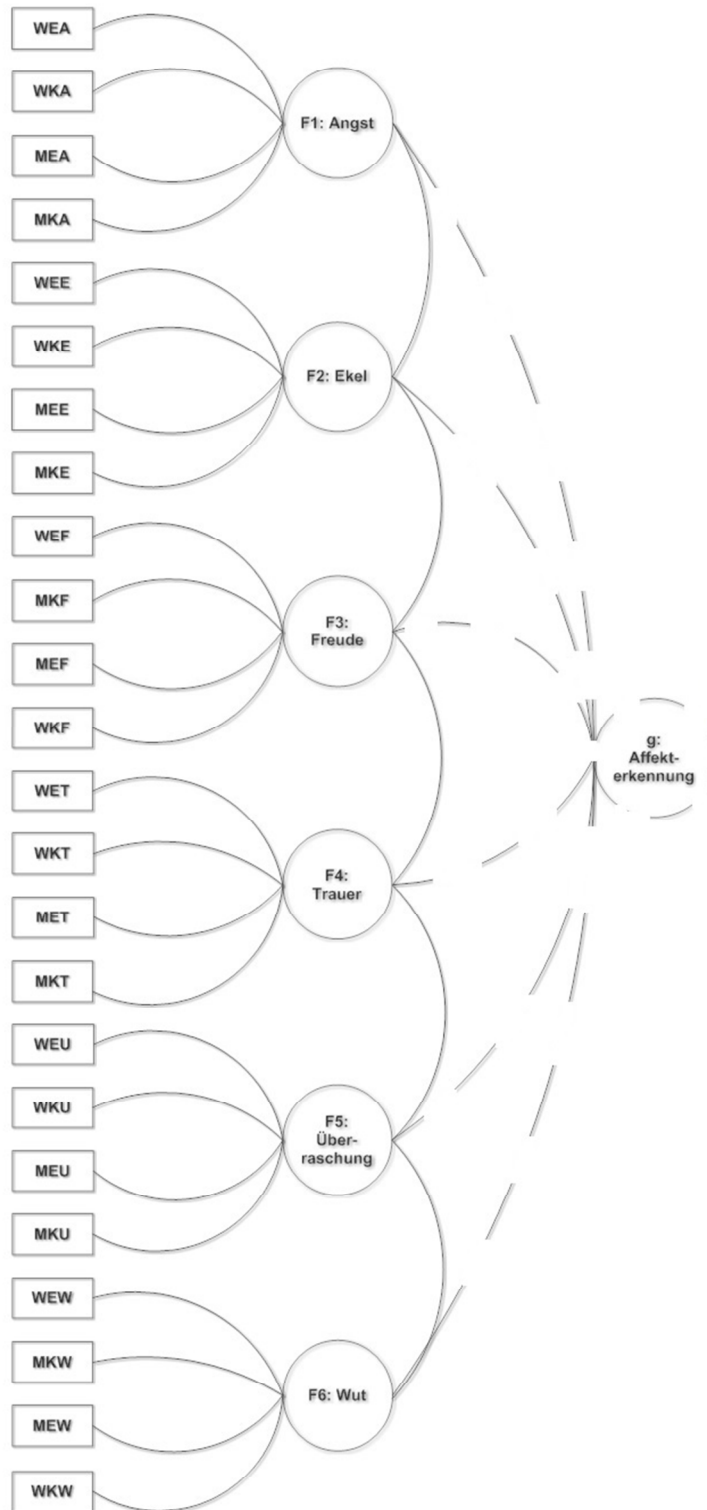


Abbildung G2. Second Order Faktoren Modell des PCFAE. Vereinfachte Darstellung der Korrelationen der Faktoren F1 – F6 (Faktoren erster Ordnung): Alle Faktoren sind untereinander korreliert. Schema der Itembenennung gem. Anhang I. Die Darstellung des übergeordneten Faktors g (Faktor zweiter Ordnung) wurde gestrichelt gewählt, da dieses Modell nicht exploriert wurde und somit keine Hinweise auf die Existenz eines solchen Faktors existieren.

Anhang I

Schema der Itembenennung innerhalb der Abbildungen zum PCFAE

WEA = Angst weiblich Erwachsener

WKA = Angst weiblich Kind

MEA = Angst männlich Erwachsener

MKA = Angst männlich Kind

WEE = Ekel weiblich Erwachsener

WKE = Ekel weiblich Kind

MEE = Ekel männlich Erwachsener

MKE = Ekel männlich Kind

WEF = Freude weiblich Erwachsener

WKF = Freude weiblich Kind

MEF = Freude männlich Erwachsener

MKF = Freude männlich Kind

WET = Trauer weiblich Erwachsener

WKT = Trauer weiblich Kind

MET = Trauer männlich Erwachsener

MKT = Trauer männlich Kind

WEU = Überraschung weiblich Erwachsener

WKU = Überraschung weiblich Kind

MEU = Überraschung männlich Erwachsener

MKU = Überraschung männlich Kind

WEW = Wut weiblich Erwachsener

WKW = Wut weiblich Kind

MEW = Wut männlich Erwachsener

MKW = Wut männlich Kind

Anhang J

Selbständigkeitserklärung / Declaration of Authorship

Hiermit erkläre ich, dass ich die vorliegende Arbeit mit dem Titel

„If he’s happy, do you know it? Analyse des neuen Test of Perceptual Competence of Facial Affective Expression (PCFAE)“

selbstständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Die Stellen der Arbeit sowie eventuell beigefügte Zeichnungen, Skizzen, graphische Darstellungen oder Daten, die anderen Werken entnommen oder an diese angelehnt sind, habe ich unter Angabe der Quelle als Entlehnung kenntlich gemacht.

I hereby declare that I have authored the thesis entitled

„If he’s happy, do you know it? Analyse des neuen Test of Perceptual Competence of Facial Affective Expression (PCFAE)“

No other person’s work has been used without due acknowledgement in this thesis. Where I have quoted from or adapted the work of others, the source is always given. This includes the sources of graphs and data.

Düsseldorf, den 21.12.2018

Mark Peter Richter