

Universität Potsdam
Humanwissenschaftliche Fakultät
Department Psychologie
Lehrstuhl für Emotions- & Biopsychologie



**Vergleich der Verläufe von Schreck- und Hautleitwertreaktionen
in der Akquisitionsphase einer Furchtkonditionierung**

Bachelorarbeit vorgelegt von
Nele Katharina Elbersgerd
zur Erlangung des akademischen Grades
Bachelor of Science
im Studiengang Psychologie

Betreuerin: **Dr. Julia Wendt**
1. Gutachterin: **Dr. Julia Wendt**
2. Gutachterin: **Dipl.-Psych. Miriam Hufenbach**

E-Mail: xxxxxxxxx@uni-potsdam.de
Matrikelnummer: xxxxxx
Einreichungsdatum: 19.05.2021

Inhaltsverzeichnis

| | |
|--|-----|
| Abbildungsverzeichnis | iii |
| Tabellenverzeichnis | iii |
| Abkürzungsverzeichnis | iv |
| Zusammenfassung | v |
| Einleitung | 1 |
| 1 Theorie und Stand der Forschung | 3 |
| 1.1 Das Paradigma der Furchtkonditionierung | 3 |
| 1.2 Indikatoren des Furchtlernens | 5 |
| 1.3 Die Hautleitwertreaktion | 6 |
| 1.4 Die Lidschlusskomponente der Schreckreaktion | 8 |
| 1.5 Verwendung mehrerer Indikatoren | 9 |
| 1.6 Lerntheoretische Perspektiven | 10 |
| 2 Explorative Fragestellung | 15 |
| 3 Methoden | 16 |
| 3.1 Stichprobe | 16 |
| 3.2 Apparatur | 16 |
| 3.3 Durchführung | 18 |
| 3.4 Datenverarbeitung und Reaktionsdefinition | 21 |
| 3.5 Statistische Analyse | 22 |
| 4 Ergebnisse | 29 |
| 4.1 Explorative Datenanalyse | 29 |
| 4.2 Modellbildung | 30 |
| 4.3 Das finale Modell | 33 |
| 4.4 Prüfen der Modellannahmen | 36 |
| 5 Diskussion | 38 |
| 5.1 Interpretation und Einordnung | 38 |
| 5.2 Limitationen und Ausblick | 41 |
| Literaturverzeichnis | 44 |
| Anhang | 50 |
| Selbstständigkeitserklärung | 56 |

Abbildungsverzeichnis

| | | |
|---|--|----|
| 1 | Schematischer Ablauf der Trials | 20 |
| 2 | Schematische Datenstruktur im Rahmen des Mehrebenenmodells | 23 |
| 3 | Mittlere Verläufe von Hautleitwert- und Schreckreaktionen | 31 |
| 4 | Vorhergesagte Verläufe von Hautleitwert- und Schreckreaktionen | 34 |
| 5 | Abbildung der zufälligen Intercepts & Interaktionen | 37 |

Tabellenverzeichnis

| | | |
|---|--|----|
| 1 | Exemplarische Datenstruktur | 24 |
| 2 | Deskriptive Statistiken & Mittelwertunterschiede innerhalb beider abhängiger Variablen | 29 |
| 3 | Unterschiede & Zusammenhänge zwischen den abhängigen Variablen | 32 |
| 4 | Darstellung des Modellbildungsprozesses und der getätigten Modellvergleiche | 33 |
| 5 | Ergebnisse des finalen multivariaten Wachstumsmodells | 35 |

Abkürzungsverzeichnis

ANOVA Analysis of variance (dt.: Varianzanalyse)

ANS Autonomes Nervensystem

AV Abhängige Variable

BMI Body-Mass-Index

CR Conditioned Reaction (dt.: Konditionierte Reaktion)

CS Conditioned Stimulus (dt.: Konditionierter Stimulus)

EDA Elektrodermale Aktivität

EEG Elektroenzephalogramm

EKG Elektrokardiogramm

EMG Elektromyogramm

fMRT funktionelle Magnetresonanztomographie

FPS Furchtpotentierte Schreckreaktion

ITI Intertrial-Intervall

LRT Likelihood-Ratio-Test

KI Konfidenzintervall

MEM Mehrebenenmodell

ML Maximum-Likelihood

REML Restricted Maximum-Likelihood

SAM Self-Assessment Manikin

SCR Skin Conductance Response (dt.: Hautleitwertreaktion)

STR Startle Response (dt.: Schreckreaktion)

UR Unkonditionierte Reaktion

US Unkonditionierter Stimulus

Zusammenfassung

In der Furchtkonditionierungsforschung sind Schreck- und Hautleitwertreaktionen (SCR) zwei der verbreitetsten Indikatoren, um humane Furchtlernprozesse zu untersuchen. Allerdings herrscht Uneinigkeit darüber, inwiefern sie distinkte Lernprozesse widerspiegeln. Die Zwei-Prozess-Theorie nimmt an, dass differentielle Schreckreaktionen eine automatische Ebene des Furchtlernens indizieren, während differentielle SCR Kontingenzwissen auf kognitiver Ebene reflektieren. Diese Arbeit untersucht explorativ, ob sich in einer Furchtakquisition Unterschiede in Verläufen und Diskrimination Konditionierter Stimuli (CS) zwischen SCR und Schreckreaktionen abzeichnen, wenn beide statistisch gleichbehandelt in einem multivariaten Modell ausgewertet werden. Achtunddreißig Versuchspersonen durchlaufen ein differentielles Furchtakquisitionstraining, in dessen Verlauf ihnen zwei geometrische Reize mehrfach präsentiert werden. Einem folgt mit 50 % Wahrscheinlichkeit ein elektrotaktiler Reiz. Nach der Hälfte des Trainings werden Versuchspersonen über die Reizkontingenzen aufgeklärt. Auswertungen erfolgen durch eine multivariate Wachstumskurvenanalyse (als Mehrebenenmodell). Mit dieser können Veränderungen zweier Variablen über die Zeit hinweg und die Variabilität zwischen Versuchspersonen modelliert werden. Wichtigste Ergebnisse zeigen einen Interaktionseffekt zwischen CS und Zeit für die SCR – nicht aber für die Schreckreaktionen –, der substanzial zwischen Versuchspersonen variiert. Mittelwertvergleiche und Visualisierungen verdeutlichen, dass eine Differenzierung zwischen den CS wenn, dann erst nach der Instruktion, auftritt. Die gefundenen Effekte lassen sich schlussfolgernd nicht auf Furchtlernprozesse, sondern auf einen Instruktionseffekt zurückführen. Inferenzen auf die theoretische Indikation mehrerer Lernebenen lassen sich nicht ziehen. Nichtsdestotrotz werden wesentliche Vorteile, aber auch Hürden beim Einsatz von multivariaten Mehrebenenmodellen erläutert. Dieser statistische Ansatz eröffnet neue Möglichkeiten zur Untersuchung und Integration mehrerer Furchtindikatoren und birgt das Potential, einen Beitrag zur Entschlüsselung der Mechanismen potentiell distinkter Furchtlernprozesse zu leisten.

Schlagwörter: Furchtakquisition, Hautleitwertreaktion, Schreckreaktion, Furchtlernen, Mehrebenenmodelle, Multivariate Verfahren

Einleitung

Eine wesentliche Schwierigkeit in der experimentellen Forschung besteht seit jeher darin, Prozesse und Zustände, die man nicht direkt beobachten kann, greifbar zu machen. Es bedarf eines Indikators: eines Hilfsmittels, welches einen Zustand oder eine Entwicklung anzeigen und damit beobachtbar machen kann. Ein Forschungsparadigma, für das sich diese Schwierigkeit in hohem Maße relevant zeigt, ist die Furchtkonditionierung. Mit ihr lassen sich unter anderem Prozesse wie das Furchtlernen untersuchen, bei dem auf Basis von vorherigen Erfahrungen negative Ereignisse antizipiert und vorsorgliche Schutz- oder Abwehrreaktionen in die Wege geleitet werden können. Ein Verständnis für diese Prozesse und des ihnen zugrundeliegenden Furchtsystems ist von hohem Interesse. Zum Beispiel können fehlende Regulationsmechanismen in diesem anpassungsfähigen System mit pathologischen Mustern zusammenhängen, wie sie unter anderem bei Angststörungen festgestellt werden. Ein vollständiges Durchdringen der Mechanismen rings um die Entstehung und Aufrechterhaltung solcher Muster kann Behandlungs- und Forschungsmethoden verbessern und validieren. Anwendungsbereiche wie diese machen die Furchtkonditionierung zu einem wichtigen, experimentellen Paradigma.

Um den Erwerb – oder die Akquise – von Furcht nachvollziehen zu können, werden verschiedene Indikatoren verwendet, wobei zwei physiologische Maße besonders verbreitet sind: die Hautleitwertreaktion und die Schreckreaktion. Bei erfolgreichem Furchtlernen wird typischerweise beobachtet, dass diese Reaktionen auf einen Reiz, der in der Vergangenheit ein negatives Ereignis vorhergesagt hat, höher ausfallen als auf einen, der diese Vorhersage nicht bereitstellt. Genannt wird diese Form des Furchterwerbs auch differentielles Furchtlernen. Die Schreckreaktion besteht übergreifend aus einer körperlichen Kaskade an defensiven Bewegungen in Folge eines Schreckkreizes, wobei das Schließen der Augen (die sogenannte Lidschlusskomponente) beim Menschen am häufigsten untersucht wird. In der Stärke der Schreckreaktion spiegelt sich unter anderem, wie angenehm oder unangenehm ein Reiz empfunden wird. Die Hautleitwertreaktion daneben beschreibt eine erhöhte Leitfähigkeit der Haut durch Schwitzen. Sie gilt vor allem als Maß für Erregung und reagiert besonders bei wichtigen oder neuen Stimuli. In Furchtkonditionierungsstudien wird sie ebenfalls traditionell als Indikator für Lernprozesse eingesetzt.

Eine immer noch aktuelle Fragestellung der Furchtkonditionierungsforschung ist, ob das Furchtlernen auf einem oder mehreren abgrenzbaren Prozessen basiert und welche Rolle das Bewusstsein über gelernte Verbindungen zwischen Reizen dabei spielt. Die Schreck- und die Hautleitwertreaktion werden dabei häufig als Indikatoren zweier verschiedener theoretischer Lernebenen herangezogen. Während die differentielle Schreckreaktion hauptsächlich mit einem automatischen und unbewussten Prozess in Verbindung gebracht wird, gilt eine Differenz in der Hautleitwertreaktion als Indiz eines kognitiven und bewussten Prozesses. Ob diese Unterscheidung gerechtfertigt ist und inwiefern die beiden Maße tatsächlich unterscheidbare Informationen über Furchtlernprozesse bereitstellen, bleibt dabei größtenteils ungeklärt. Im Kontext dieser Kontroverse hat diese Arbeit zum Ziel, die Verläufe der beiden Indikatoren in einer Furchtakquisition mittels eines multivariaten Verfahrens zu untersuchen. Durch diesen Auswertungsansatz können Limitationen herkömmlicher univariater Verfahren zum Teil überwunden und ein umfangreicher Einblick in das Zusammenspiel beider Maße gewonnen werden. Auf Basis dieses explorativen Ansatzes kann möglicherweise zu der Literatur rund um die Unterscheidung zweier konzeptueller Lernprozesse auf

Reaktionsebene beigetragen werden. Außerdem erweitert die Arbeit die kaum bestehende Literatur, die Hautleitwert- und Schreckreaktion als physiologische Indikatoren in direkter Form miteinander vergleicht.

Zu Beginn dieser Arbeit in Kapitel 1 werden wesentliche Konzepte definiert und relevante Forschungsergebnisse eingeordnet. Zunächst wird das grundlegende Paradigma der Furchtkonditionierung und ein Ausschnitt aus traditionellen und zeitgenössischen Konditionierungstheorien und -paradigmen beschrieben. Der anschließende Überblick über Indikatoren des Furchtlernens dient als Grundlage für den angestrebten Vergleich der Schreck- und Hautleitwertreaktionen, deren Methodik und Anwendung im Kontext der Furchtkonditionierung resümiert und verglichen wird. Die Fragestellung, ob Furchtlernen auf einem oder mehreren distinkten Prozessen basiert, wird anhand der sogenannten Ein-Prozess- und Zwei-Prozess-Theorien erläutert und außerdem einige der kontroversen Forschungsergebnisse diesbezüglich angesprochen. Hier spielen die Hautleitwert- und die Schreckreaktion als mögliche Indikatoren verschiedener Ebenen eine entscheidende Rolle. Auf diesen theoretischen Aspekten bauen die Ziele und exploratorischen Fragestellungen dieser Arbeit in Kapitel 2 auf. In Kapitel 3 gehe ich ausführlich auf die Methodik der Datenerhebung und Organisation dieser Arbeit sowie die verwendeten statistischen Analysen ein. Außerdem werden an dieser Stelle die Entscheidungskriterien für den multivariaten Modellbildungsprozess dargelegt. Es folgt die Aufstellung des statistischen Modells und der Ergebnisbericht in Kapitel 4 sowie abschließend die Integration und Diskussion der Ergebnisse in Kapitel 5.

1 Theorie und Stand der Forschung

Wenn wir Furcht empfinden, dann zumeist als Reaktion auf einen potenziell gefährlichen Reiz in unserer direkten Umwelt. Unser Körper begibt sich in einen erregten, defensiven Zustand und erhöht die Verhaltensbereitschaft für Flucht- oder Abwehrreaktionen. Furcht ist damit eine sehr adaptive und evolutionär relevante Emotion, die den Organismus vor Gefahr bewahren und möglichst schnell eine schützende Reaktion initiieren soll. In der psychologischen Forschung wird die Emotion Furcht typischerweise durch eine hohe Erregung und eine negative Valenz charakterisiert (Lang, 1995). Welche Reize in der Umwelt als bedrohlich wahrgenommen werden, ist von verschiedenen Faktoren abhängig. Mittlerweile ist aber klar, dass Lernprozesse eine substanzIELLE Rolle dabei spielen. Unsere vergangenen Erfahrungen lenken unsere Wahrnehmung, Verarbeitung und Reaktion heute. Dieses erfahrungsbasierter Lernen – oder *Konditionierung* – ist eine der Grundlagen für die Plastizität des Furchtsystems: Reize, die zuvor mit furchterregenden Ereignissen in Verbindung gebracht wurden, sind später in der Lage, selbst Furcht hervorzurufen.

In der psychologischen Grundlagen- und Anwendungsforschung werden Experimente zur Furchtkonditionierung seit knapp einem Jahrhundert für viele Fragestellungen eingesetzt. Die Erkenntnisse führen in verschiedenen Disziplinen der Humanwissenschaften zu einem Boom an Abwandlungen des Paradigmas und Weiterentwicklungen in multiplen Forschungszweigen (Merz & Lonsdorf, 2020). Konzeptuell lassen sich Furchtkonditionierungsstudien in experimentelle Phasen unterteilen, die verschiedene Teilprozesse wie Furchterwerb, Extinktion und Wiederauftreten von Furcht zu modellieren versuchen. Innerhalb der Forschung um die Entstehung und Aufrechterhaltung von Angst- und Belastungsstörungen haben Furchtkonditionierungsexperimente entscheidende Einblicke in die Lern- und Gedächtnisprozesse von Menschen liefern können (für eine Übersichtsarbeit siehe Mineka & Oehlberg, 2008). Ebenso relevant sind sie für die (Weiter-)Entwicklung von Therapieelementen bei diesen Störungsbildern. Gerade dieser translationale Wert von Furchtkonditionierungsstudien trägt wesentlich zu ihrer unbestrittenen Relevanz bei.

Die Fülle an Anwendungsgebieten mündete in einer Vielzahl an methodischen und terminologischen Variationen. Um eine bessere Vergleichbarkeit zu gewährleisten und Konsistenz in der Forschung zu fördern, haben Lonsdorf et al. (2017) Empfehlungen zur Methodik und Terminologie in Furchtkonditionierungsstudien zusammengestellt. Es sei zu Beginn darauf hingewiesen, dass methodische und technische Termini dieser Arbeit – soweit nicht anders definiert – im Einklang mit der Handreichung von Lonsdorf et al. (2017) zu verstehen sind.

1.1 Das Paradigma der Furchtkonditionierung

1.1.1 Grundlagen der Klassischen Konditionierung

Die Furchtkonditionierung beruht grundlegend auf dem Paradigma der Klassischen Konditionierung (auch Pavlovsche Konditionierung genannt). In seiner Pionierarbeit berichtete der russische Forscher Pavlov (1927) von einer Konditionierung an einem Hund, an dessen Beispiel das Grundprinzip nachfolgend erläutert wird. Die Klassische Konditionierung nutzt eine bereits bestehende Verbindung zwischen einem sogenannten Unkonditionierten Stimulus (US; hier Futter) und einer Unkonditionierten

Reaktion (UR; hier erhöhter Speichelfluss beim Hund). Ein zunächst neutraler Stimulus – bei Pavlov ein Glockenton – löst wiederum eine neutrale Reaktion aus. In einem sogenannten *Akquisitionstraining* wird der Glockenton wiederholt mit dem Futter (US) gemeinsam präsentiert, worauf der Hund mit erhöhtem Speichelfluss (UR) reagiert. Die Präsentation einiger dieser Paarungen führt dazu, dass der Glockenton zu einem *Konditionierten Stimulus* (engl.: conditioned stimulus; CS) wird. Nach erfolgreichem Lernen ist die alleinige Präsentation des CS ausreichend, um den Speichelfluss auszulösen, der nun als *Konditionierte Reaktion* (engl.: conditioned response; CR) bezeichnet wird. Diese CR ist in ihrer Modalität der UR meist ähnlich – sprich, sie beschreiben das gleiche Verhalten – allerdings ist die CR oft weniger intensiv.

Die Furchtkonditionierung ist eine Variante dieses klassischen Paradigmas. Hier wird ein meist neutraler CS mit einem negativ valenten US (z. B. einem elektrotaktilen Stimulus) gepaart. Diese aversiven US lösen beim Individuum eine meist automatische (Furcht-)Reaktion wie den Schreckreflex aus. Diese Arbeit betrachtet die Akquisitionsphase einer Furchtkonditionierungsstudie. Hier wird mittels eines Akquisitionstrainings (der wiederholten gemeinsamen Präsentation von CS und US) eine konditionierte Reaktion manifestiert. Konzeptuell induziert dieses Training den Prozess des *Furchtlernens*.

1.1.2 Kognitive Erweiterungen

Der Konditionierungsprozess wurde traditionell als simple und automatische Assoziationsbildung angesehen, die reflexiv und unfreiwillig erworben und abgerufen wird. Erst in den 70er Jahren wurde mit der *kognitiven Wende* in der psychologischen Forschung den kognitiven Prozessen ein größerer Stellenwert eingeräumt. Zum Beispiel zeigte Rescorla (1968) in seiner Arbeit, dass die Kontingenz zwischen CS und US eine ausschlaggebende Rolle für den Erwerb einer CR spielt. Zudem lieferte er einen ersten Beleg dafür, dass die CS-US-Assoziation am besten gelernt wird, wenn der CS kurz vor dem US präsentiert wird. Diese Form der zeitlichen Kontinguität wird in der Furchtkonditionierungsforschung als *Verzögerungskonditionierung* bezeichnet. Kognitive Ansätze dieser Art von Rescorla und anderen führten zu der weithin vertretenen Annahme, dass sich nach erfolgreichem Furchtlernen im Individuum die Erwartung an den US bildet, sobald ein CS dargeboten wurde. Diese US-Erwartung rückt die CR in die Position einer provisorischen Schutzreaktion auf Basis der gelernten Erfahrung, dass der US bald folgen wird (Kiesel & Koch, 2012). Anders gesagt, kommt dem CS eine Prädiktorfunktion für den US zu.

Neben dem assoziativen Lernen im Sinne direkt erlebter Erfahrungen werden auch andere Pfade für das Furchtlernen in Betracht gezogen. Rachman (1977) postulierte neben der Konditionierung auch den Furchterwerb über Beobachtung und über Instruktionen. Beim beobachtenden Lernen observiert die Versuchsperson zum Beispiel die Furchtreaktion eines anderen Individuums auf einen Reiz, während sie beim instruktionalen Furchtlernen über die kausale Beziehung zwischen CS und US informiert wird. Auch wenn dieser theoretische Ansatz nicht Gegenstand dieser Arbeit ist, gehe ich auf das instruktionale Furchtlernen kurz ein, da das verwendete Studiendesign Instruktionselemente enthält (Abschnitt 3.3.4).

Instruktionen können die Versuchssubjekte teilweise oder vollständig über die vorhandenen CS-US-Kontingenzen aufklären. Während bei uninstruierten Designs erfahrungsbasierter Furchtlernen über mehrere CS-US-Paarungen hinweg stattfinden kann, führt die explizite Aufklärung über die CS-US-Kontingenz sofort zu einer veränderten Furchtreaktion. Man spricht in diesem Fall von einem *Furchtausdruck* statt von Furchtlernen (Lonsdorf et al., 2017). Diese erzwungene Bewusstwerdung kann einen erheblichen Einfluss auf den Lernprozess haben. Zum Beispiel zeigten Weidemann et al. (2016), dass verbale Instruktionen über das Vorhandensein einer Beziehung zwischen CS und US das Auftreten von

konditionierten Reaktionen (gemessen über die Lidschlusskomponente der Schreckreaktion) im Vergleich u. a. zu einer uninstruierten Gruppe erhöhen. Eine kürzlich erschienene Übersichtsarbeit von Mertens et al. (2018) beleuchtet den moderierenden Effekt verbaler Instruktion auf die verschiedenen Phasen der Furchtkonditionierung tiefergehend (zur instruierten Extinktion sei auch die Übersichtsarbeit von Luck & Lipp, 2016, empfohlen).

1.1.3 Das Differentielle Paradigma

In der Furchtkonditionierungsforschung ist das sogenannte *differentielle* Design eine relevante Erweiterung. Hierbei gibt es einen CS+, der immer oder manchmal mit einem US gepaart wird, und einen CS-, der niemals gemeinsam mit einem US auftritt. Typischerweise löst der CS+ eine stärkere CR aus und wird negativer bewertet als der CS-. Die Quantifizierung der Reaktionen kann zum Beispiel als differentielle CR erfolgen, sprich als Differenz der Reaktionen auf CS+ und CS-. In den überwiegenden Fällen sind CS+ und CS- Reize ähnlicher oder gleicher Modalität (z. B. zwei unterschiedliche visuelle Stimuli). Die paradigmatische Ergänzung bietet mehrere Vorteile. Sie gewährleistet hinsichtlich des Erwerbs der CS-US-Kontingenz, dass der prädiktive Wert eines bestimmten Reizes gelernt und nicht das Vorhandensein eines Reizes an sich als Prädiktor missverstanden wird. Außerdem erlaubt das differentielle Konditionieren eine Kontrolle von möglichen Sensibilisierungseffekten. Das heißt, dass Effekte aufgrund der Modalität der CS vergleichbar zwischen CS+ und CS- sind.

1.2 Indikatoren des Furchtlernens

Emotionen und Lernen haben als psychologische Konzepte eine wesentliche Sache gemeinsam: Auch wenn man eine Vorstellung von ihnen hat und sie relativ gut erkennt, so ist es gar nicht so leicht, sie richtig zu messen. Um das zu tun, müssen sie operationalisiert werden. Beide Konstrukte sind komplexe Phänomene, die mit vielfältigen Reaktionen auf verschiedenen Reaktionsdimensionen auslösen können. Diese können erfasst werden und ermöglichen es, Rückschlüsse auf eine bestimmte Emotion oder einen erfolgten Lernprozess zu ziehen. Furcht ist eine Emotion, die mit einer Aktivierung des Defensivnetzwerks einhergeht. Wenn Menschen Furcht empfinden, gibt es – abhängig von der Proximität des furchtauslösenden Reizes – eine Reihe an körperlichen Systemen, die in Gang gesetzt werden (zum menschlichen Defensivsystem siehe Hamm & Weike, 2005; Hamm et al., 2006; Öhman & Mineka, 2001). Außerdem sind Menschen in der Lage, mitzuteilen, wie sich Furcht anfühlt. Ganz ähnlich ist das beim Lernen. In Kiesel und Koch (2012) wird Lernen definiert als „Prozess, der als Ergebnis von Erfahrungen relativ langfristige Änderungen im Verhaltenspotential erzeugt“ (S. 11). Allerdings reicht eine Änderung im Verhaltenspotential nicht, um Lernen experimentell nachweisen zu können. Es muss eine tatsächliche Änderung im Verhalten stattfinden, damit Rückschlüsse auf einen zugrundeliegenden Lernprozess gezogen werden können.

Um nun Furchtlernen im Labor zu erfassen, kommen grundlegend drei Reaktionssysteme in Frage, die von Bradley und Lang (2000) in der Humanforschung unterschieden werden. Das sind (a) Handlung oder die Ausführung behavioraler Sequenzen (z. B. Flucht), (b) emotionale Sprache und (c) physiologische Reaktionen (z. B. steigende Herzrate). Die Veränderungen in Verhalten, physiologischen Reaktionen oder verbalen Berichten im Verlauf der Zeit sind damit klassische abhängige Variablen in Furchtkonditionierungsexperimenten. Bezogen auf die Furchtkonditionierung werden behaviorale Maße eher selten verwendet, am ehesten in Reaktionszeitexperimenten oder mittels Beobachtung von

Vermeidungsverhalten. Verbale Maße sind häufiger: Sie können erfassen, inwieweit Subjekte explizites Wissen über CS-US-Kontingenzen erwerben („Der Reiz A sagt den Reiz B vorher“) oder wie die Bewertung von Affekt und Erregung sich im Laufe eines Experimentes ändert. Physiologische Maße haben aufgrund ihrer Vergleichbarkeit mit Ergebnissen aus Tierstudien eine besonders wichtige Stellung in der Konditionierungsforschung. Unter den am häufigsten verwendeten Korrelaten sind peripherphysiologische Indikatoren wie die Herzrate, die Hautleitfähigkeit und die Schreckreaktion sowie neurobiologische Maße wie BOLD-Kontraste im fMRT oder Ereigniskorrelierte Potenziale im EEG. Die physiologischen Maße, vor allem diejenigen, die die Aktivität des autonomen Nervensystems (ANS) erfassen, gehen mit Vor- und Nachteilen einher. Ihr größtes Potential liegt darin, dass sie parallel zum Lernprozess erhoben werden können. Damit liefern sie einzigartige Informationen über körperliche Zustände, die zeitlich kontingent zum tatsächlichen Lernen geschehen. Allerdings sind sie – anders als beispielsweise Ratings – nicht oder zumindest nicht gleichermaßen furchtselektiv, was Inferenzen auf einzelne psychologische Prozesse erschwert (Abschnitte 1.3.2 und 1.4.2). Zwei der am häufigsten verwendeten physiologischen CR sind die Lidschlusskomponente der Schreckreaktion und die Hautleitwertreaktion. Die folgenden Abschnitte erklären, wie sie erfasst werden, welche Körperprozesse sie abbilden und warum ihnen eine besondere Bedeutung zukommt.

1.3 Die Hautleitwertreaktion

1.3.1 Aspekte der Methodik

Die Hautleitwertreaktion (engl.: skin conductance response; SCR) ist ein Maß der Elektrodermalen Aktivität (EDA) und quantifiziert eine kurzfristige Änderung in der Leitfähigkeit der Haut in Reaktion auf einen Stimulus oder ein Ereignis. Sie ist ein weit verbreitetes und etabliertes Maß der Psychophysiologie. Die Leitfähigkeit der Haut wird über Schweißdrüsen vermittelt und kann dort leicht und non-invasiv erfasst werden. Ekkrine Schweißdrüsen befinden sich fast überall am Körper, wobei ihre Dichte in der Handinnenfläche und an den Fingern mit am größten ist. Ableitungen der EDA werden präferiert dort vorgenommen (Boucsein et al., 2012; Dawson et al., 2007). In methodischer Hinsicht liegen die Vorteile der EDA darin, dass sie risikofrei, günstig und einfach zu erheben ist im Vergleich zu aufwendigeren physiologischen Methoden wie zum Beispiel dem EEG. Allerdings sind Hautleitwertreaktionen mit ihren 1 bis 4 s Latenzzeit ein relativ langsames Maß (Boucsein et al., 2012; Dawson et al., 2007). Abgesehen von der Reaktionszeit unterscheidet sich die SCR von der Schreckreaktion auch dahingehend, dass kein zusätzlicher Auslösereiz benötigt wird, um sie zu erfassen (siehe Abschnitt 1.4 zur Schreckreaktion). Im differentiellen Furchtkonditionierungsparadigma wird die SCR als Differenzmaß analysiert. Hier zeigte sich in zahlreichen Studien, dass durch erfolgreiches Lernen die Präsentation eines CS+ typischerweise eine höhere phasische Reaktion des Hautleitwerts hervorruft als die Präsentation eines CS- (Dawson & Biferno, 1973; Sevenster et al., 2014). Aufgrund der Latenzzeit der SCR muss bei dem Design von Furchtkonditionierungsstudien darauf geachtet werden, dass die Intertrial-Intervalle (ITI) zwischen zwei präsentierten Stimuli lang genug sind, damit die Hautleitwertreaktion auftreten und wieder abklingen kann (Lonsdorf et al., 2017).

1.3.2 Verwendung als Indikator

Elektrodermale Maße werden in psychologischen Experimenten zur Erfassung vieler latenter Variablen und Prozesse herangezogen. Hautleitwertreaktionen sind als reliable Komponente der Orientierungsreaktion in Gegenwart von neuen, unerwarteten und signifikanten Stimuli (Siddle, 1991) weit verbreitet. Typischerweise wird beobachtet, dass die SCR bei wiederholter Präsentation in ihrer Amplitudenhöhe abnimmt. Dieser Habituationssverlauf geschieht zum Teil langsamer, je signifikanter (wichtiger) der Stimulus für das Individuum ist (Boucsein et al., 2012). Ein Beispiel für diese Signifikanz ist etwa die Relevanz des CS+ als Prädiktor für den US in einer Konditionierungsstudie. Zudem spielen auch Aufmerksamkeitseffekte und autonome Erregung eine nicht zu vernachlässigende Rolle in der Modulation von Hautleitwertreaktionen (für eine Ausführung siehe Dawson et al., 2007).

Die Veränderung der EDA tritt weder in Isolation auf noch ist sie spezifisch für eine bestimmte Art von Reiz oder Kontext. Das bedeutet, dass die SCR oft nur eine Komponente in einer komplexen Reaktionskette des ANS abbildet. Diese kann wiederum durch verschiedene Einflüsse ausgelöst und moduliert werden, von Kontextbedingungen über Stimuluseigenschaften bis hin zum motivationalen und emotionalen Zustand, in dem sich die Versuchsperson befindet. Es ist typisch, dass Reaktionen, die die EDA beeinflussen, auch in andere Maße des ANS wie die Herzrate intervenieren. Tatsächlich müssen aber nicht alle diese Maße miteinander hoch korrelieren. Das liegt daran, dass Individuen immer wieder ein ihnen eigenes, relatives Reaktionsmuster des ANS produzieren und sich diese Muster aber interindividuell unterscheiden. Ein Phänomen, das als *Individual Response Stereotypy* bekannt ist (Engel, 1960) und einen Grund dafür darstellt, dass für die Aktivierung des autonomen Nervensystems in der Forschung mehrere Indikatoren herangezogen werden.

Für sympathische Aktivität gilt die EDA hingegen als sehr spezifisches Maß. Ekkrine Schweißdrüsen und damit die EDA werden – anders als andere Maße des ANS wie die Herzrate – fast ausschließlich durch den Sympathikus kontrolliert (Dawson et al., 2007). Die EDA kann dadurch mit der sympathischen Aktivierung infolge einer Defensivreaktion in Verbindung gebracht werden (Merz & Lonsdorf, 2020). Besonders responsiv zeigt sich die EDA in Situationen mit diskreten furchtauslösenden Stimuli, in denen eine Vermeidung nicht möglich ist (Dawson et al., 2007). Verwendet man die EDA nun als Indikator für Furchtlernen, bleibt jedoch das Problem der geringen Spezifität bestehen. Zwar konnten Studien zeigen, dass die SCR linear mit steigender induzierter Erregung der Stimuli ansteigt (Lang et al., 1990), allerdings spielt sich dieser Effekt unabhängig von der Valenz der Stimuli ab. Im Kontext einer Konditionierung bedeutet das, dass ein CS gepaart mit einem aversiven US (Furchtkonditionierung) dasselbe Reaktionsmuster der EDA auslöst wie ein CS, der mit einem nicht-aversiven US gepaart wird (Hamm & Vaitl, 1996).

Ihre fehlende Spezifität wird der EDA häufig als Nachteil ausgelegt, denn durch ihre multikausale Bedingtheit ist sie kaum spezifisch interpretierbar. Trotz alledem ist ihre Bedeutung nicht von der Hand zu weisen. In den meisten experimentellen Designs hat das Ausmaß der Kontrolle einen entscheidenden Einfluss darauf, wie überzeugend und konkret SCR interpretiert und Inferenzen auf bestimmte psychologische Prozesse gezogen werden können (Dawson et al., 2007).

1.4 Die Lidschlusskomponente der Schreckreaktion

1.4.1 Aspekte der Methodik

Die Schreckreaktion (engl.: startle reflex) beschreibt eine Reihe von protektiven motorischen Reflexen, die in Reaktion auf einen abrupten sensorischen Stimulus (engl.: probe stimulus; im Folgenden: Schreckkreis) ausgelöst werden (Landis & Hunt, 1939). Dieses Reaktionsmuster ist speziesübergreifend feststellbar und seine neuronalen Schaltkreise wurden in Tierstudien bereits eingehend erforscht und zum Teil in der Humanforschung repliziert (für eine Übersichtsarbeit siehe Davis, 2006). In letzterer erweist sich die Lidschlusskomponente (engl.: startle eye blink; im Folgenden allgemein: Schreckreaktion) als erste und stabilste Komponente, um diesen Schutzreflex systematisch zu erfassen (Lang et al., 1990). Die Lidschlusskomponente beschreibt das reflexartige Zusammenkneifen der Augen infolge eines meist akustischen Schreckkreises. Wird dieser im Labor präsentiert, kann die Muskelkontraktion des Musculus Orbicularis Oculi über ein Elektromyogramm (EMG) aufgezeichnet werden. Die Elektroden hierfür werden zumeist auf der Hautoberfläche unterhalb des Auges angebracht. Der Schreckkreis ist meist ein 50 ms kurzes, 90 bis 105 dB lautes, weißes Rauschen und nach seiner Vergabe erfolgt die Schreckreaktion meist innerhalb von 150 ms (Blumenthal et al., 2005). Dadurch, dass diese Reaktion automatisch ausgelöst wird und sich damit der willentlichen Kontrolle entzieht, ist sie resistent gegenüber Effekten kognitiver Verarbeitung und Antworttendenzen, wie sie z. B. in verbalen Berichten auftreten können (Grillon & Baas, 2003; Lipp, 2006). Das ist methodisch ein Vorteil, da andere physiologische Maße wie die Herzrate und die Hautleitfähigkeit durchaus – wenn auch in begrenztem Maß – aktiv durch die Versuchspersonen beeinflusst werden können (z. B. mit Beeinflussung der Atmung; Lonsdorf et al., 2017).

Bei Verwendung in Furchtkonditionierungsstudien ist ein wesentlicher methodischer Vorteil der Schreckreaktion, dass sie circa um das 50-fache schneller nach Reizdarbietung auftritt als die Hautleitwertreaktion. Außerdem weist sie eine kürzere Refraktärzeit auf. Durch diesen zeitlichen Vorteil kann die Schreckreaktion zu jeder Zeit während des Furchtakquisitionstrainings ausgelöst werden und wichtige zeitlich contingente Einblicke liefern (Lipp, 2006). Im Gegensatz zur SCR ist es außerdem möglich, Reaktionen innerhalb der ITI zu erheben und damit eine Basisreaktion unabhängig von der CS-Präsentation zu definieren. Die Erfassung der Schreckreaktion als Furchtmaß benötigt allerdings eine Habituationssphase, in der der Schreckkreis zwei bis zehn Mal präsentiert wird. Damit soll gewährleistet werden, dass die ersten starken Reaktionen auf den Schreckkreis den Verlauf der CR in Reaktion auf einen Stimulus weniger beeinflussen (Blumenthal et al., 2005). Eine Phase wie diese ist bei Erhebung der SCR nicht nötig.

1.4.2 Verwendung als Indikator

In den allermeisten Fällen, in denen die Schreckreaktion experimentell erhoben wird, geht es nicht primär um den motorischen Reflex als solchen, sondern um die psychologischen Prozesse, mit denen er in Verbindung gebracht wird. Die Relevanz der Schreckreaktion für die Furchtkonditionierungsforschung begründet sich durch ihre relative Spezifität zur Erfassung von Furcht. Lang et al. (1990) zeigten, dass die Schreckreaktion durch Affekt moduliert wird. Während angenehme Reize die Schreckreaktion hemmen, wird sie durch unangenehme Stimuli potenziert. Dieses Phänomen ist im Kontext der Furchtkonditionierung als Furchtpotenzierung der Schreckreaktion (FPS) bekannt: Ein bereits konditionierter Stimulus (CS+) löst eine stärkere Schreckreaktion aus als ein nicht-verstärkter CS- (Bradley et al., 2005; Brown

et al., 1951; Hamm & Vaitl, 1996; Hamm & Weike, 2005). Die FPS tritt unabhängig von der Neuheit der Reize auf. Zwar nimmt die Reaktion in absoluten Werten bei mehreren Trials mit der Zeit ab, der Effekt der emotionalen Modulation bleibt jedoch robust bestehen (Bradley & Lang, 2000). Zudem zeigten Cuthbert et al. (1996), dass Erregung die Schreckreaktion auf aversive Reize zusätzlich moduliert. Die Potenzierung der Schreckreaktion zeigt sich bei der Präsentation von hoch erregenden und negativ valenten Stimuli, allerdings nicht bei wenig erregenden und trotzdem negativ valenten Reizen (Cuthbert et al., 1996). Ein daraus resultierender Vorteil ist, dass die Schreckreaktion Furchtlernen verhältnismäßig spezifisch erfassen kann. Im Vergleich mit der Hautleitwertreaktion, die eher valenzunspezifische Erregung reflektiert, wird der Schreckreaktion die Rolle als spezifischerer Furchtindikator zugesucht (Hamm & Vaitl, 1996).

Die neuroanatomische Basis wurde vielfach in tierexperimentellen Studien untersucht. Zahlreiche Studien, u.a. von Davis und Kolleg*innen, lieferten übereinstimmende Evidenz dafür, dass die Potenzierung der Schreckreaktion über die Amygdala vermittelt wird (für eine Übersichtsarbeit siehe Davis, 2006). Die Amygdala ist eine limbische Struktur im medialen Temporallappen und gilt als eine zentrale Schaltstelle des Defensivsystems (Davis, 1992; Hamm et al., 2006). Mit ihren efferenten Projektionen initiiert und mediert die Amygdala Furchtsymptome wie körperliche Erregung oder protektive motorische Reflexe, wozu auch die Schreckreaktion gehört. Bei Läsionen des zentralen Kerns der Amygdala findet die Schreckreaktion auf einen Schreckkreis durchaus statt, allerdings wird sie bei Furchtreizen nicht mehr potenziert (Hitchcock & Davis, 1986, 1987). Das macht die Schreckreaktion zu einem validen Indikator für die Aktivierung des Defensivsystems.

Ein methodischer und theoretischer Nachteil im Vergleich zur SCR ergibt sich durch die Notwendigkeit eines zusätzlichen Schreckkreises zum Auslösen der Schreckreaktion. Es ist bereits gezeigt worden, dass die physikalischen Eigenschaften des Schreckkreises die Schreckreaktion in Intensität und Onset modulieren (Blumenthal et al., 2005). Außerdem wird der Schreckkreis von Versuchspersonen ebenfalls häufig als aversiv wahrgenommen. Ein potentieller Einfluss auf den emotionalen Zustand der Person und demnach auch auf die Intensität des Schreckreflexes kann nicht ausgeschlossen werden (Lipp, 2006). Sjouwerman et al. (2016) konnten außerdem zeigen, dass die Inklusion von Schreckkreisen in einer Furchtkonditionierungsstudie die Furchtakquisition verzögerte (gemessen anhand der SCR). Die Wahl der abhängigen Variable(n) hat demnach durchaus einen nicht zu unterschätzenden Einfluss auf den Lernprozess.

1.5 Verwendung mehrerer Indikatoren

Furchtlernen gilt als Konstrukt mit einem komplexen Aktivierungsnetzwerk. Die eingesetzten Indikatoren können nicht distinkt in die spezifische Erfassung von Kontingenzwissen, konditionierter Reaktion, Aufmerksamkeitsverarbeitung oder emotionalem Zustand eingeordnet werden (Lipp, 2006). Es hat sich während der letzten Forschungsjahrzehnte etabliert, mehrere Indikatoren gleicher oder unterschiedlicher Reaktionsebenen zur Erfassung von Furchtlernen zu verwenden, um ein vollständigeres Bild der Prozesse zu erhalten. Nach Lipp (2006) ging diese wichtige Änderung mit der Realisierung einher, dass einzelne Maße unterschiedliche Spezifitäten aufweisen. Verbale Maße sind oft durch Antwort- und Erinnerungstendenzen verfälscht. Physiologische Maße reflektieren nicht nur Furcht, sondern werden ebenfalls durch Prozesse wie Aufmerksamkeit oder Kontexteffekte beeinflusst. Es wird auch in der

Empirie deutlich, dass Indikatoren verschiedener Ebenen oft nur in geringem Maße kovariieren (Bradley & Lang, 2000; Lipp, 2006). Ein verwandtes Problem beim Verwenden mehrerer Maße ist, dass sich die Indikatoren auch in ihrer Sensitivität unterscheiden. Der direkte Vergleich von unterschiedlichen CR, vor allem über verschiedene Studien hinweg, gestaltet sich als schwieriges Unterfangen. Zusammenfassend bedeutet das: Ein einzelnes Maß kann viele Komponenten abbilden, währenddessen sich die gewünschten Effekte in mehreren Maßen gleichzeitig widerspiegeln können. Die Integration mehrerer Indikatoren genauso wie das Wissen um ihre Interdependenzen sind entscheidende Bausteine der aktuellen und zukünftigen Furchtkonditionierungsforschung.

Die steigende Komplexität, die die Wahl der abhängigen Variablen mit sich bringt, ist auch zunehmend Gegenstand der Forschungsliteratur. Neben den bereits zitierten Empfehlungen und Richtlinien von Lonsdorf et al. (2017) haben auch Ojala und Bach (2020) erst kürzlich eine Übersichtsarbeit über verschiedene Arten konditionierter Reaktionen veröffentlicht. Darin beurteilten sie verschiedene CR unter anderem nach ihrer Qualität, Inferenzen auf gebildete CS-US-Assoziationen zu ermöglichen. Eine Studie von Leuchs et al. (2019) verglich die Indikatoren SCR und Schreckreaktion mit der an Aufmerksamkeit gewinnenden Pupillendilatation als CR einer Furchtkonditionierung. Die Akquisitionsdaten zeigen signifikante, aber eher geringe Korrelationen zwischen allen drei Maßen (trialweise gemittelt; für SCR und FPS $r = .16$, $p < .001$, Leuchs et al., 2019). Sowohl die SCR als auch die FPS unterlagen einer starken Habituation, diskriminierten aber beide signifikant zwischen CS+ und CS- (Leuchs et al., 2019). Relevant für den Einsatz mehrerer Reaktionsmaße ist auch, dass sich verschiedene CR durchaus gegenseitig beeinflussen (Lonsdorf et al., 2017). Das liegt nicht nur daran, dass Trials und ITI in Rücksicht auf die Onset- und Refraktärzeiten der Hautleitwertreaktion zeitlich koordiniert werden müssen, sondern auch an der Wechselwirkung von CR an sich. Dazu gehört zum Beispiel die bereits erwähnte Studie von Sjouwerman et al. (2016), bei der das Verwenden von akustischen Schreckkreisen das Furchtlernen, gezeigt an der SCR, verzögert. Ein weiterer, wichtiger Aspekt ist die umstrittene Annahme, dass unterschiedliche Indikatoren verschiedene inhaltliche Prozesse des Furchtlernens abbilden. Im Folgenden wird dieser Disput in der Forschung näher erläutert.

1.6 Lerntheoretische Perspektiven

1.6.1 Ein oder zwei Systeme?

Ursprünglich wurde der Konditionierungsprozess als simple, beinahe mechanische Bildung von Assoziationen unabhängig von kognitiven Mechanismen verstanden (Lovibond, 2006). Es ist mittlerweile größtenteils unbestritten, dass auch kognitive Mechanismen (wie die Abhängigkeit des Lernens von Kontingenz und Kontinguität, Rescorla, 1968) bei der Erklärung von Furchtlernphänomenen eine nicht zu unterschätzende Rolle spielen. In diesem Zusammenhang werden häufig zwei aktive Systeme angenommen, sobald Menschen Reizkontingenzen beim klassischen Konditionieren ausgesetzt sind: zum einen das automatische, reflexive System und zum anderen das propositionale, kognitive System (Lovibond, 2006). Diese Dualitätsannahme taucht in verschiedenen Ansätzen zu Lern- und Gedächtnisprozessen mit jeweils kleinen Abweichungen immer wieder auf. Begrifflich werden theoretische Differenzierungen zwischen dem expliziten und dem impliziten, dem bewussten und dem unbewussten oder dem kontrollierten und dem automatischen Lernen vorgenommen. Die Idee von zwei *Ebenen des Lernens* beim menschlichen Assoziationslernen reflektiert sich in den verschiedenen dualen Lerntheorien. Doch auch wenn der

lernbasierte Ansatz und die substanziale Bedeutung von Kognitionen in Lernprozessen allgemein akzeptiert bleibt, ist das Vorhandensein dualer oder mehrerer Prozesse an sich ein umstrittenes Thema.

In der Furchtkonditionierungsforschung beschäftigt sich der Diskurs damit, ob überhaupt distinkte Prozesse am Klassischen Konditionieren beteiligt sind. Ein Fokus liegt hier auf dem Konflikt um die Notwendigkeit des Bewusstseins für das Zeigen einer konditionierten Reaktion (für systematische Übersichtsarbeiten siehe Lovibond & Shanks, 2002; Mertens & Engelhard, 2020). Es wird unter anderem darüber disputiert, ob und zu welchem Grad die Konditionierung beim Menschen von einem bewussten Verständnis der gelernten Kontingenzen abhängt und welche Operationalisierung für Bewusstsein hierfür herangezogen werden kann. Maßgeblich wird der Konflikt von zwei Polen bestimmt. Die *Zwei-Prozess-Theorie* vertritt die Auffassung von zwei unabhängigen Prozessen des Lernens (Hamm & Vaitl, 1996; Hamm & Weike, 2005; Öhman & Mineka, 2001; Weike et al., 2007). Zum einen führt ein propositionaler, kognitiver Lernvorgang zu verbalisierbarem Kontingenzbewusstsein. Zum anderen wird auf einer nicht-propositionalen, impliziten Ebene eine autonome CR wie die Schreckreaktion produziert. Die *Ein-Prozess-Theorie* steht dem gegenüber. Vertreter*innen dieser Annahme kritisieren, dass in Studien zum Klassischen Konditionieren beim Menschen kein hinreichender Nachweis für die Notwendigkeit von Bewusstsein – und damit für das Vorhandensein zweier Ebenen – erbracht wurde. Der Ein-Prozess-Ansatz geht von nur einer Ebene aus, die einen propositionalen Lernprozess beschreibt (De Houwer, 2009, 2014; Lovibond & Shanks, 2002; Mertens & Engelhard, 2020; Mitchell et al., 2009). In seiner stärksten Form wird angenommen, dass Kontingenzbewusstsein eine Voraussetzung für das Auftreten einer CR ist. Personen müssen demnach in der Lage sein, explizit die Beziehung zwischen CS und US zu beschreiben, um konditionierte Reaktionen zu zeigen. In einer schwächeren Form der Ein-Prozess-Theorie sind sowohl Kontingenzbewusstsein als auch CR ebenbürtige Endpunkte des propositionalen Lernprozesses, die allerdings stark miteinander korrespondieren.

Die Forschungsergebnisse zu dem Disput sind bis heute ambig. Unter anderem kritisierten Lovibond und Shanks (2002) die Operationalisierung und Erfassung von Kontingenzbewusstsein und bemängelten, dass kein hinreichender Nachweis für einen Dualismus oder ein Konditionieren ohne Bewusstsein erbracht worden sei. Außerdem erweise sich die unterschiedliche Methodik in der Durchführung von Konditionierungsstudien als Schwierigkeit für die Interpretation und Vergleichbarkeit unterschiedlicher Forschungsergebnisse. In ihrer aktuellen Metaanalyse verzeichneten Mertens und Engelhard (2020), dass 39 von 41 inkludierten Studien mindestens einem methodischen Problem unterlagen. Dazu gehören unter anderem die Erfassung von Kontingenzbewusstsein und Entscheidungsfreiheiten der Wissenschaftler*innen in Bezug auf Ausschlusskriterien und Kennwertbestimmung.

Auch wenn grundlegend relevant für die aktuelle Forschung, so wird das Konstrukt *Kontingenzbewusstsein* als solches und der Disput um seine Operationalisierung nicht näher in dieser Arbeit betrachtet. Interessierte Leser*innen seien zum Beispiel an die Arbeiten von Lovibond und Shanks (2002), Wiens und Öhman (2002) oder Mertens und Engelhard (2020) verwiesen. Im nächsten Abschnitt folgt ein Überblick über Forschungsergebnisse, bei denen Hautleitwert- und Schreckreaktionen als Indikatoren verschiedener Lernprozesse herangezogen wurden. Dies dient als Grundlage für die Herleitung der Fragestellung und Analyse in dieser Arbeit.

1.6.2 Schreck- und Hautleitwertreaktion als Indikatoren unterschiedlicher Lernebenen

Insbesondere im Rahmen der Zwei-Prozess-Theorie werden Schreck- und Hautleitwertreaktionen als Indikatoren verschiedener Lernebenen herangezogen (Hamm & Vaitl, 1996; Sevenster et al., 2014). Es wird angenommen, dass die FPS das implizite und emotionale Furchtlernen reflektiert und unabhängiger von kognitivem Wissen über die Stimulusrelationen auftritt (z. B. Hamm & Weike, 2005; Sevenster et al., 2014). Dagegen wird die differentielle SCR grundlegend mit Kontingenzlernen und dem US-Erwartungslernen in Verbindung gebracht (z. B. Hamm & Weike, 2005; Weike et al., 2007). US-Erwartung beschreibt, wie groß die Erwartung der Versuchspersonen an einen US ist, nachdem ein bestimmter CS präsentiert wurde. Allerdings ist der Zusammenhang zwischen SCR und Kontingenzbewusstsein umstritten. Es herrscht konkurrierende Evidenz darüber, ob das Zeigen einer differentiellen SCR vom Kontingenzbewusstsein der Person abhängt (Dawson & Biferno, 1973; Hamm & Vaitl, 1996; Purkis & Lipp, 2001; Weike et al., 2007) oder nicht (Knight et al., 2003; Schultz & Helmstetter, 2010).

Ein Beispiel für die Dissoziation zwischen SCR und Schreckreaktion in einem differentiellen Furchtkonditionierungsparadigma lieferten Sevenster et al. (2014). Sie reproduzierten eine Studie von Schultz und Helmstetter (2010), bei der eine Gruppe auf ein einfaches und eine andere auf ein schwierig zu diskriminierendes Paar von CS+ und CS- konditioniert wurde. Schultz und Helmstetter (2010) fanden eine differentielle US-Erwartung ursprünglich nur in der leichten Bedingung, während beide Gruppen eine differentielle SCR zeigten. Daraus schlussfolgerten sie, dass die differentielle SCR nicht durch Kontingenzbewusstsein beeinflusst wird (Schultz & Helmstetter, 2010). In ihrer Replikationsstudie verwendeten Sevenster et al. (2014) dasselbe Set an Stimuli und erfassten zusätzlich die Schreckreaktion als Indikator. Ihre Ergebnisse zeigten, dass die differentielle SCR nur bei den Versuchspersonen auftrat, die deklaratives Wissen über die CS-US-Kontingenzen aufgebaut hatten (erfasst über Befragungen zur US-Erwartung während des Trainings). Eine differentielle Schreckreaktion zeigte sich unabhängig vom Kontingenzwissen. Diese Daten unterstützen die Zwei-Prozess-Theorie und lassen vermuten, dass die Hautleitwertreaktion im Gegensatz zur Schreckreaktion mit bewusstem Lernen über die Stimuluskontingenzen zusammenhängt. Die Ergebnisse der Studie von Weike et al. (2007) deuten zudem auf eine Abhängigkeit dieser potentiellen Dissoziation von der Methodik der CS-US-Paarungen hin. In einem Vergleich zwischen Spuren- und Verzögerungskonditionierung zeigte sich die differentielle SCR in beiden Gruppen nur bei Personen, die Kontingenzbewusstsein aufwiesen (erfasst über ein Interview nach dem Training). Die konditionierte FPS allerdings trat in der Gruppe mit Verzögerungskonditionierung unabhängig vom Bewusstsein auf, während sie bei der Gruppe mit Spurenkonditionierung nur bei Personen mit Kontingenzbewusstsein zu finden war. Daraus lässt sich schließen, dass auch methodische und zeitliche Aspekte des Furchtakquisitionstrainings einen Einfluss darauf haben, inwiefern Schreck- und Hautleitwertreaktionen als Indikatoren zweier distinkter Lernprozesse herangezogen werden können. Limitiert wird die Interpretierbarkeit dieses Ergebnisses durch die Erfassung von Kontingenzbewusstsein nach dem Akquisitionstraining, wodurch Gedächtniseffekte nicht ausgeschlossen werden können (Mertens & Engelhard, 2020).

Auf der anderen Seite gibt es auch Befunde, die keine Nachweise für Furchtkonditionierung der Schreckreaktion in Abwesenheit von Kontingenzbewusstsein finden und demnach keine Inferenz auf distinkte Prozesse zulassen. In einer Furchtakquisitionsstudie mit Verzögerungskonditionierung von Purkis und Lipp (2001) zeigten sich differentielle Hautleitwertreaktionen nur für die Versuchspersonen,

die Kontingenzbewusstsein aufwiesen. Das gleiche Muster zeichnete sich allerdings auch für die Schreckreaktionen ab, was nahelegt, dass die FPS abhängig vom Kontingenzbewusstsein der Versuchspersonen ist. Ähnliche Ergebnisse erbrachte die Studie von Grillon (2002), bei der ebenfalls nur Versuchspersonen mit Kontingenzbewusstsein differentielle SCR und FPS zeigten. Diese Daten sind aufgrund der fehlenden Diskrepanz zwischen propositionalem Wissen und dem Zeigen einer differentiellen CR konsistent mit dem Ein-Prozess-Ansatz.

Abseits der hauptsächlich peripherphysiologischen Forschung wurde die Dualitätsannahme im Furchtlernen auch auf neurobiologischer Basis in Tierstudien untersucht. In einer der bekanntesten Theorien zur Rolle der Amygdala im Furchtnetzwerk postuliert LeDoux (2003) zwei distinkte Routen der Furchtaktivierung. Auf der sogenannten automatischen *low road* wird der furchtauslösende Input über den Thalamus (Kernstruktur im Zwischenhirn) direkt an die Amygdala geleitet, ohne kortikal verarbeitet zu werden. Die Amygdala wiederum initiiert eine schnelle präventive Reflexreaktion, die sich zum Beispiel im Zusammenkneifen der Augen zeigt (Schreckreaktion). Die *high road* hingegen ist langsamer: Der visuelle Input wird über den Thalamus an den Kortex geleitet und dort analysiert und eingeordnet. Während LeDoux (1996) argumentiert, dass sich der Amygdala-Schaltkreis weitestgehend kognitiver Kontrolle entzieht, werden im Weg über den Kortex kognitive Prozesse mit einbezogen. Diese neurobiologische Unterscheidung legt nahe, dass die Aktivierung des Defensivsystems, vermittelt über die Amygdala, vom Erwerb deklarativen Wissens auf kognitiver Ebene differenziert werden kann (Öhman & Mineka, 2001). Dissoziative Ergebnisse, in denen sich Schreckreaktionen, aber nicht Hautleitwertreaktionen unabhängig vom propositionalen Kontingenzwissen zeigen, sind mit dieser Theorie von LeDoux (2003) vereinbar. In einer Bildgebungsstudie beim Menschen konnte bei Versuchspersonen, die kein Kontingenzbewusstsein zeigten, auch eine erhöhte Aktivität der Amygdala auf den CS+ nachgewiesen werden (Tabbert et al., 2006). Zudem erfassten Tabbert et al. (2006) Hautleitwertreaktionen, die nur in der Gruppe mit Kontingenzbewusstsein signifikante Unterschiede zwischen CS+ und CS- aufwiesen. Diese Ergebnisse unterstützen die Annahme einer Aktivierung des Furchtnetzwerks über die Amygdala in Abwesenheit (und damit unabhängig) von deklarativem Kontingenzwissen. Auch in dieser Studie wurde Kontingenzbewusstsein über einen Fragebogen nach dem eigentlichen Akquisitionstraining erfasst, was als Limitation zu betrachten ist (Mertens & Engelhard, 2020).

Die Forschungsergebnisse und ihre Diskrepanzen zeigen, dass man noch ein gutes Stück davon entfernt ist, die genauen Mechanismen des Furchtlernens zu verstehen. Die Argumentation der Zwei-Prozess-Theorie führt einen zu der grundlegenden Unterscheidung zwischen dem emotionalen und dem kognitiven Lernen (Hamm & Weike, 2005; Öhman & Mineka, 2001). Ersteres beschreibt die Aktivierung des Defensivsystems als Furchtreaktion auf den CS. Diese Aktivierung kann über die direkte Verbindung zwischen Thalamus und Amygdala erklärt werden (LeDoux, 2003). Die Potenzierung der Schreckreaktion als reliabler und relativ spezifischer Furchtindikator wird wiederum über die Amygdala vermittelt (Davis, 2006; Abschnitt 1.4.2). Daneben beschreibt das kognitive Lernen das propositionalen Erlernen des Zusammenhangs zwischen dem CS und dem US. Dieser deklarative Wissenserwerb wird eher mit einer differentiellen SCR assoziiert. Eine höhere Hautleitwertreaktion erfolgt auf den Stimuli, der ein signifikantes Ereignis vorhersagt, unabhängig davon, ob es aversiv ist oder nicht (Hamm & Vaitl, 1996; Hamm & Weike, 2005). Dies deckt sich mit den Forschungsergebnissen, die die SCR mit Orientierungsreaktionen auf neue und signifikante Stimuli in Verbindung bringen (Siddle, 1991; Abschnitt

1.3.2). Eine solch klare Unterscheidung zeigt sich jedoch nicht immer in den Forschungsergebnissen. Es bedarf in jedem Fall weiterer empirischer Forschung, um die Beziehung zwischen verschiedenen Indikatoren des Furchtlernens besser nachvollziehen zu können. Obwohl es viele Studien zu Schreck- und Hautleitwertreaktionen im Furchtlernkontext gibt, existiert dennoch wenig Literatur, die beide Maße direkt miteinander in Beziehung setzt (Ojala & Bach, 2020). Erst kürzlich haben Constantinou et al. (2021) in einer Furchtkonditionierungsstudie u. a. Mehrebenenmodelle verwendet, um SCR-Reaktionen mit US-Erwartungsratings in Verbindung zu setzen. Um zu diesem Forschungsgebiet beizutragen und der Frage nachzugehen, ob sich mögliche distinkte Lernprozesse auf der physiologischen Reaktionsebene zeigen, strebe ich in dieser Arbeit einen Vergleich der beiden Indikatoren Schreck- und Hautleitwertreaktionen an.

2 Explorative Fragestellung

Anhand der Literatur wird deutlich, dass sowohl Schreck- als auch Hautleitwertreaktionen viel verwendete Indikatoren für das Furchtlernen sind. Ob und inwiefern die beiden Maße tatsächlich als Indikatoren verschiedener Lernprozesse herangezogen werden können, bleibt allerdings größtenteils ungeklärt. Um einen anderen Blickwinkel auf diese Diskordanz zu erhalten, umfasst diese Arbeit einen direkten explorativen Vergleich beider Maße als abhängige Variablen in einer Furchtakquisition. Wenn, wie die Zwei-Prozess-Theorie nahelegt, die beiden physiologischen Maße durch verschiedene kognitiv-affektive Prozesse moduliert werden, ließe sich vermuten, dass sich dies in unterschiedlichen zeitlichen Antwortmustern während des Furchtlernens widerspiegelt.

Um die Beziehung zwischen den beiden Indikatoren besser zu verstehen, wird in einem explorativen Ansatz beleuchtet, ob die Schreck- und Hautleitwertreaktionen in der Furchtakquisitionsphase unterschiedliche Lernprozesse auf der Signalebene abbilden. Die Aufmerksamkeit liegt auf der Frage, ob sich unterschiedliche Lernverläufe in Hautleitwert- und Schreckreaktion abzeichnen, wenn beide Maße im Modell gleich behandelt werden. Mit anderen Worten: Wären verschiedene Lernsysteme – sofern sie existieren – auf Reaktionsebene durch den Lernverlauf der beiden Maße unterscheidbar?

Für diesen Ansatz werden die Daten eines differentiellen, partiell verstärkten, verzögerten Furchtakquisitionstraining erhoben und analysiert. Für den Vergleich der beiden abhängigen Variablen wird eine multivariate Wachstumskurvenanalyse im Rahmen einer Mehrebenenmodellierung durchgeführt. Innerhalb dieses multivariaten Modells werden die Veränderungen der beiden Reaktionsmaße über die Trials hinweg als Wachstumskurven modelliert und der Einfluss des Stimulustyps untersucht.

Konkret lassen sich die Ziele dieser Arbeit in zwei leitende exploratorische Fragestellungen einteilen. Zunächst wird untersucht, ob sich unterschiedliche Verläufe in Hautleitwert- und Schreckreaktion abzeichnen, wenn beide Maße im Modell gleich behandelt werden. Bezugnehmend auf das differentielle Furchtlernen wird außerdem betrachtet, ob sich der Einfluss des Stimulustyps bedeutsam zwischen den abhängigen Variablen unterscheidet bzw. ob im Durchschnitt in beiden abhängigen Variablen in gleichem Maße zwischen dem CS+ und dem CS- diskriminiert wird. Diese Leitfragen sind eingebettet in den explorativen Kontext dieser Arbeit, in dem ein multivariates Modell gebaut, analysiert und auf dessen Besonderheiten eingegangen wird.

3 Methoden

3.1 Stichprobe

Die Stichprobe bildet sich aus 40 gesunden Personen zwischen 18 und 30 Jahren (32 weiblich, 7 männlich, 1 divers, $M_{\text{Alter}} = 24.1$ mit $SD_{\text{Alter}} = 3.3$). Ausschlusskriterien im Zuge der Rekrutierung sind zum einen ein BMI, der kleiner als 18.5 kg/m^2 oder größer als 30 kg/m^2 ist. Zum Anderen führt das Vorliegen von kardiovaskulären, neurologischen und Atemwegserkrankungen, eingeschränktes Hör- oder Farbsehvermögen, Klaustrophobie, Schwangerschaft oder die Einnahme von bestimmten Medikamenten (kardioaktiv, antidepressiv, antipsychotisch, antihypertensiv oder anticholinerg) zum Ausschluss. Des Weiteren müssen Teilnehmer*innen ein deutsches Sprachverständnis auf mindestens C1-Niveau nachweisen, um an der Studie teilnehmen zu können. Die $n = 40$ Versuchspersonen haben zum Zeitpunkt der Datenauswertung mindestens die Akquisitionsphase vollständig durchlaufen. Nicht enthalten in dieser Stichprobe sind Personen, die rekrutiert worden sind, aber noch vor ($n = 8$) oder während ($n = 1$) des Akquisitionstermins die Teilnahme beendeten. Ein Abbruch nach dem vollständigen Abschluss der Akquisitionsphase ist für diese Arbeit irrelevant und es wurden keine Versuchspersonen aufgrund dessen exkludiert. Auf Basis der Datenaufbereitung der Hautleitwert- und Schreckreaktionen (Abschnitt 3.4) wurden zwei der 40 Versuchspersonen ausgeschlossen, sodass die finale Analyse mit $n = 38$ durchgeführt wurde.

Für den Abschluss der insgesamt vierwöchigen Untersuchung erhalten Teilnehmer*innen eine Aufwandsentschädigung in Form von pauschal 50 Euro oder 6.5 bzw. 10.5 Versuchspersonenstunden bei Zuordnung zur Kontrollgruppe respektive zur Interventionsgruppe (für einen Abbruch nach der Akquisitionsphase wird eine anteilige Entschädigung vergeben). Die Rekrutierung der teilnehmenden Personen erfolgt über Aushänge an der Universität, Online-Plattformen und über Beiträge in Sozialen Netzwerken. Jede potentiell geeignete Person durchläuft zu Beginn ein Telefoninterview zur Feststellung der Eignung und wird dabei über den Ablauf der Untersuchung aufgeklärt. Zum ersten Labortermin geben alle Versuchspersonen ihr schriftliches Einverständnis für die Datenverarbeitung und Durchführung der Studie. Genehmigt wurde das Projekt durch die Ethikkommission der Universitätsmedizin Greifswald. Aufgrund der Ausschlusskriterien und der Einordnung in ein Großprojekt (Abschnitt 3.3.1), welches den zeitlichen Rahmen dieser Arbeit weit überschreitet, ist die Stichprobe als eine Gelegenheitsstichprobe zu verstehen.

3.2 Apparatur

3.2.1 Stimuli

Die konditionierten Stimuli bilden mit einem blauen Quadrat (Hexadezimalcode: #00AFF4) und einem orangenen Kreis (#E46B0C) zwei visuelle geometrische Reize in gleicher Helligkeit. Präsentiert werden diese auf einem schwarzen Hintergrund in der Mitte des Computerbildschirms, jeweils mit einer Dauer von 6 s (für eine Darstellung der Trials siehe Abbildung 1). Die Intertrial-Intervalle zwischen den Stimuli zeigen ein weißes Fixationskreuz mittig des Bildschirms und variieren in ihrer Länge gleichmäßig zwischen 14, 15 und 16 s. Die Reizpräsentation wird über die Software Presentation (Version 20.3, Fa.

Neurobehavioral Systems, Inc.) realisiert. Welcher der beiden Stimuli den CS+ und welcher den CS- darstellte, wird über Teilnehmer*innen hinweg ausbalanciert (blaues Quadrat als CS+ $n = 21$).

Der CS+ wird partiell durch einen elektrotaktile Reiz (US) verstärkt, dessen Intensität vor dem Training individuell mit der Versuchsperson eingestellt wird, bis der Reiz als *unangenehm, aber nicht schmerhaft* bewertet wird ($M = 6 \text{ mA}$, $SD = 3.7$, Range [1.3; 21.0]). Die Stimuluselektrode (Stabelektrode, Fa. Technomed GmbH) wird auf der Innenseite des rechten Beins posterior circa 3 cm über dem Knöchel platziert. Kontrolliert wird die elektrische Stimulation über einen DS7A Constant Current Stimulator (Fa. Digitimer, Hertfordshire, UK). Der Ausgangstrom liegt zwischen 1 und 100 mA bei einer Stimulationsdauer von 1 ms. Um Habituationssprozesse zu vermeiden, wird die Anzahl der präsentierten US so gering wie möglich gehalten ($M = 7.1$ Durchgänge, $SD = 4.2$).

Als akustischer Schreckkreis dient ein 50 ms langes, 95 dB weißes Rauschen, welches binaural über Kopfhörer (Modell ATH-PRO700MK2, Fa. Audio-Technica Ltd., Inc.) präsentiert wird. Die Schreckkreise werden 4.5 oder 5.5 s nach jedem CS-Onset und in der Hälfte aller ITI (d.h. 16 mal) 7.5 s nach CS-Offset verabreicht.

3.2.2 EDA- und EMG-Ableitungen

Alle physiologischen Ableitungen werden über ein BIOPAC MP160 Verstärker-System und die AcqKnowledge 5.0.2 Software realisiert (Fa. BIOPAC Systems, Inc., Goleta, CA, USA). Eine Erdungselektrode aus Silicon (Fa. TerniMed) wird am linken Oberarm der Versuchsperson befestigt. Alle Elektroden sind wiederverwendbar und werden nach jeder Messung gesäubert und hängend bei Raumtemperatur gelagert.

Für die EDA-Ableitung wird die nichtdominante Hand der Versuchspersonen zunächst mit Wasser gereinigt. Es werden zwei Ag/AgCl gesinterte Biopotentialelektroden (8 mm Kontaktflächen-durchmesser; Fa. Easycap GmbH) mit isotonischem Elektrodenkontaktgel (0.5 % NaCl, GEL101, Fa. BIOPAC Systems, Inc.) gefüllt und anschließend mit doppelseitigen Kleberingen palmarseitig über der Hypothenarmuskulatur angebracht. Die Ableitung der EDA erfolgt mit einer EDA100C Komponente (Fa. BIOPAC Systems, Inc.) exosomatisch über ein Konstant-Spannungsverfahren (0.5 V). Die EDA-Daten werden kontinuierlich mit einer Abtastrate von 2 000 Hz und einem Verstärkungsfaktor von 5 $\mu\text{S/V}$ abgeleitet und durchlaufen einen 10 Hz Tiefpassfilter.

Die Schreckreaktion wird über ein EMG des Orbicularis Oculi Muskels gemessen. Dafür wird die Haut unter dem linken Auge der Versuchsperson mit einem wasserfeuchten Tuch gereinigt. Anschließend werden zwei Ag/AgCl Elektroden (5 mm Kontaktflächendurchmesser; Fa. Schuler Medizintechnik GmbH) mit Elektrodenkontaktecreme (Electrode Cream, Fa. CareFusion) gefüllt. Diese werden mit doppelseitigen, zugeschnittenen Kleberingen, um sie möglichst nah aneinander zu positionieren, unter dem linken Auge angebracht. Die erste Elektrode liegt dabei circa 0.5 bis 1 cm unterhalb des Auges lotrecht zur Pupille, die zweite in laterale Richtung daneben (parallel zum Lidverlauf), circa 1 cm vom äußeren Lidwinkel entfernt. Das EMG-Signal wird mit der EMG100C Komponente (Fa. BIOPAC Systems, Inc.) bei einer Abtastrate von 2 000 Hz erhoben. Die Rohdaten des EMG durchlaufen den verstärker-eigenen Bandpassfilter, der Frequenzen zwischen 10 und 500 Hz passieren lässt, und werden um den Faktor 5 000 verstärkt.

3.2.3 Zusätzliche Erhebungen

Aufgrund des Großprojekts, in dessen Rahmen diese Erhebung läuft, werden zusätzlich zu den obig beschriebenen Ableitungen auch andere abhängige Variablen erhoben, die in dieser Arbeit nicht weiter betrachtet werden. Vor dem Akquisitionstermin füllen die Versuchspersonen einen Fragebogen aus, der verschiedene Elemente u. a. aus dem State-Trait-Angstinventar (Laux et al., 1981) oder dem Brief Symptom Inventory (Derogatis & Melisaratos, 1983) enthält. Während der Akquisitionsphase wird außerdem ein Elektrokardiogramm (EKG) aufgezeichnet. Hierfür werden zwei mit Elektrodenkontaktcreme (Electrode Cream, Fa. CareFusion) gefüllte Ag/AgCl-Elektroden (10 mm Kontaktflächendurchmesser; Fa. Schuler Medizintechnik GmbH) mit doppelseitigen Kleberingen am rechten Unterarm (ca. 2 cm unterhalb der Armbeuge) und am linken Bein (ca. 2 cm proximal des Knöchels) angebracht. Die Ableitung erfolgt endosomatisch ebenfalls über das BIOPAC-System (Fa. BIOPAC Systems, Inc.). Zusätzlich bewerten die Proband*innen vor und nach dem Akquisitionstraining die Valenz und Erregung der präsentierten Stimuli und beantworten nach der Messung kurze Fragen zu ihrem Kontingenzbewusstsein und ihrer Lebensführung.

3.3 Durchführung

3.3.1 Einordnung in das Forschungsprojekt

Die Daten werden im Rahmen eines größeren Forschungsprojekts, finanziert von der Deutschen Forschungsgesellschaft, erhoben. Dieses untersucht die Auswirkungen einer Intervention zum Herzratenvariabilität-Biofeedback auf das Extinktionslernen bei gesunden Menschen und bei Personen mit diagnostizierter Panikstörung, Agoraphobie oder beidem. Die Erhebungen laufen an der Universität Potsdam und an der Universität Greifswald unter der Projektleitung von Dr. Wendt und Dipl.-Psych. Hufenbach. Im Rahmen der Studie werden die Versuchspersonen über einen Zeitraum von vier Wochen begleitet. Nach einem Labortermin zur Furchtakquisition werden Herzratenparameter der Teilnehmer*innen einmal pro Woche für vier Wochen erfasst. Die gesunden Versuchspersonen der Interventionsgruppe führen während dieser Zeit täglich Übungen zum Stressmanagement mit dem mobilen HRV-Biofeedbackgerät Qiu (Fa. BioSign GmbH) durch. Am Ende der vier Wochen erfolgt ein Labortermin zur Furchtextinktion. Für diese Arbeit werden lediglich die Daten der Akquisitionsphase (am ersten Labortermin) betrachtet und ausschließlich die in Potsdam erhobene Stichprobe der gesunden Versuchspersonen. Weder die Versuchsperson noch die Versuchsleitung weiß zum Zeitpunkt der Akquisitionsphase, welcher der beiden Experimentalgruppen die Person zugeordnet wird. Diese Zuordnung erfolgt erst nach abgeschlossenem Akquisitionstraining und ist somit irrelevant für den Inhalt dieser Arbeit.

3.3.2 Anmerkungen zum Erhebungskontext

Die Datenerhebung begann im Juni 2020 in den peripherphysiologischen Laborräumen der Universität Potsdam und hält zum Zeitpunkt des Verfassens dieser Arbeit noch an. Die Fenster des Experimentalraums sind für den Zeitraum der Erhebung vollständig abgedunkelt, der Raum ist geräuscharm. Die Beleuchtung wird während der Experimentalabschnitte gedimmt, wobei aufgrund technischer Schwierigkeiten die Lichtintensität über die Untersuchungsmonate hinweg geringen Schwankungen unterliegt. Um mögliche Interferenzen durch laute Geräusche zu unterbinden, wird die Klimaanlage während der Messung ausgeschaltet. Die Raumtemperatur beträgt im Mittel 22 °C ($SD = 1$) bei einer mittleren

Luftfeuchtigkeit von 43.2% ($SD = 10$) .

Vor dem Labortermin erhalten alle Versuchspersonen eine E-Mail mit Instruktionen, die vor der Messung einzuhalten sind. Die Proband*innen werden instruiert, am Tag vor dem Termin der Furchtakquisition ihre gewöhnliche Schlafroutine zu verfolgen, auf intensives körperliches Training und Alkohol mindestens 24 h vor der Messung zu verzichten und im Zeitraum von zwei Stunden vor der Messung keine Mahlzeiten, Koffein oder Nikotin zu sich zu nehmen. Außerdem werden sie gebeten, auf Make-Up im Augenbereich und Strumpfhosen im Sinne der Elektrodenanbringung zu verzichten.

Die Datenerhebung findet während der globalen Covid-19 Pandemie statt und unterliegt somit verschärften Hygienebestimmungen, die im Hygienekonzept des Bereichs Emotions- und Biopsychologie der Universität Potsdam vom 09.05.2020 festgehalten wurden. Die Versuchspersonen und die Versuchsleitung sind verpflichtet, sich vor Beginn der Messung intensiv die Hände zu waschen. Bis auf das Anbringen und Abnehmen der Elektroden wird ein Mindestabstand von 1.5 m zur Versuchsperson eingehalten. Die Versuchsleiter*innen tragen während des kompletten Erhebungszeitraums einen Mund-Nasen-Schutz sowie zusätzlich ein Gesichtsvisier, wenn der Mindestabstand unterschritten wird. Die Versuchspersonen selbst tragen zu Beginn und zum Ende der Termine ebenfalls einen Mund-Nasen-Schutz, während der Messungen wird aufgrund der EMG-Ableitung am Auge darauf verzichtet. Die maximale Anzahl an Personen in einem Raum wird auf zwei beschränkt, parallele Messungen im angrenzenden EEG-Labor finden nur zu seltenen Ausnahmen statt. Zudem wird eine ausreichende Desinfektion und Reinigung aller mehrfach verwendeten Materialien sowie ausgiebiges Lüften zwischen den Messungen akribisch eingehalten.

3.3.3 Vor dem Akquisitionstraining

Zu Beginn werden die Versuchspersonen in den abgedunkelten Laborraum geführt und nehmen auf einem Laborstuhl Platz, dessen Rückenlehne sich in 2 m Entfernung zu einem Computerbildschirm (Modell VG275Q, Fa. ASUSTeK Computer, Inc.) befindet. Die Proband*innen werden über den Ablauf der Untersuchung informiert und geben ihr schriftliches Einverständnis. Nach Anbringen der EKG-Elektroden findet eine sechsminütige Ruhemessung der Herzratenvariabilität statt. Die Versuchspersonen werden verbal instruiert, entspannt und möglichst bewegungsarm mit geschlossenen Augen im Laborstuhl zu sitzen. Danach werden die EDA- und EMG-Elektroden (Abschnitt 3.2) angebracht. EKG-, EDA- und EMG-Daten werden bis zum Ende des Akquisitionstrainings ununterbrochen aufgezeichnet, was ca. 40 min entspricht. Es folgt eine Habituation der Schreckreaktion, bei der die Versuchspersonen ein weißes Fixationskreuz in der Mitte des schwarzen Computerbildschirms fixieren sollen, während ihnen der akustische Schreckkreis vier Mal präsentiert wird. Sie werden instruiert, dabei keine besondere Aufgabe zu haben.

Danach erfolgt die individuelle Einstellung des elektrotaktilem Reizes. Zunächst werden die Versuchspersonen über das Vorgehen aufgeklärt, im Anschluss wird die Stimulationselektrode angebracht. Für den Kalibrierungsprozess erhalten die Versuchspersonen eine Computermaus und eine 5-stufige Skala in laminierter Form (von *Gar nicht unangenehm* bis *Schmerhaft*; Anhang A1). Auf dem Computerbildschirm erscheint eine Instruktion in Textform mit dem Wortlaut: „Drücke die linke Maustaste für den Reiz. Dieser folgt dann 3 Sekunden später“. Die Versuchspersonen lösen den elektrotaktilen Reiz im Einstellungsprozess immer eigenständig aus und schätzen nach jedem Reiz auf der ausgeteilten Skala ein, wie unangenehm er für sie war. Die Tür zum Kontrollraum, in dem die Versuchsleitung die

Reizeinstellung verändert und dokumentiert, bleibt dabei offen, sodass eine direkte Kommunikation mit den Versuchspersonen möglich ist. Die Intensität des elektrotaktile Reizes startet bei 1 mA und wird sequentiell in Abhängigkeit des Ratings erhöht oder verringert, bis die Personen den Reiz als *unangenehm, aber nicht schmerhaft* einschätzen, was auf der Skala dem Rating 4 entspricht (Anhang A2 enthält die Tabelle zur Kalibrierung). Diese Intensität wird als final eingestellt und dokumentiert.

Anschließend werden den Versuchspersonen auf dem Computerbildschirm die beiden visuellen Stimuli (Abschnitt 3.2.1) nebeneinander präsentiert. Die Proband*innen werden aufgefordert, die Stimuli verbal hinsichtlich ihrer Valenz und Erregung einzuschätzen. Dazu dienen Self-Assessment-Manikin Skalen für Valenz und Erregung (SAM; Bradley & Lang, 1994), denen aufsteigend absolute Zahlen von 1 bis 9 zugeordnet werden (Anhang B).

3.3.4 Akquisitionstraining

Für das Akquisitionstraining erhalten die Versuchspersonen verbal die Instruktion, ruhig mit offenen Augen sitzen zu bleiben und den Computerbildschirm aufmerksam zu beobachten. Sie werden darauf hingewiesen, dass ihnen neben den visuellen Reizen auch akustische Stimuli und elektrotaktile Reize dargeboten werden können. Außerdem wird ihnen mitgeteilt, dass der nächste Abschnitt mit akustischen Reizen beginne, auf die sie nicht zu achten brauchten und dass sie keine besondere Aufgabe hätten. Das Furchtakquisitionstraining beginnt mit vier Schreckreizen und besteht danach aus insgesamt 32 Trials. Beide CS werden jeweils 16 Mal in ausbalancierter Reihenfolge in Viererblöcken präsentiert, sodass nicht mehr als zwei gleiche Reize direkt aufeinander folgen. Das Verhältnis der Stimuli innerhalb der Blöcke bleibt dabei gleich. Auf den CS+ folgt in 50 % aller Fälle ein US (d.h. acht verstärkte Trials), während der CS- niemals verstärkt wird. Das Design entspricht einer Verzögerungskonditionierung mit einem Interstimulus-Intervall von 5.95 s, sodass der US kurz vor dem CS-Offset verabreicht wird.

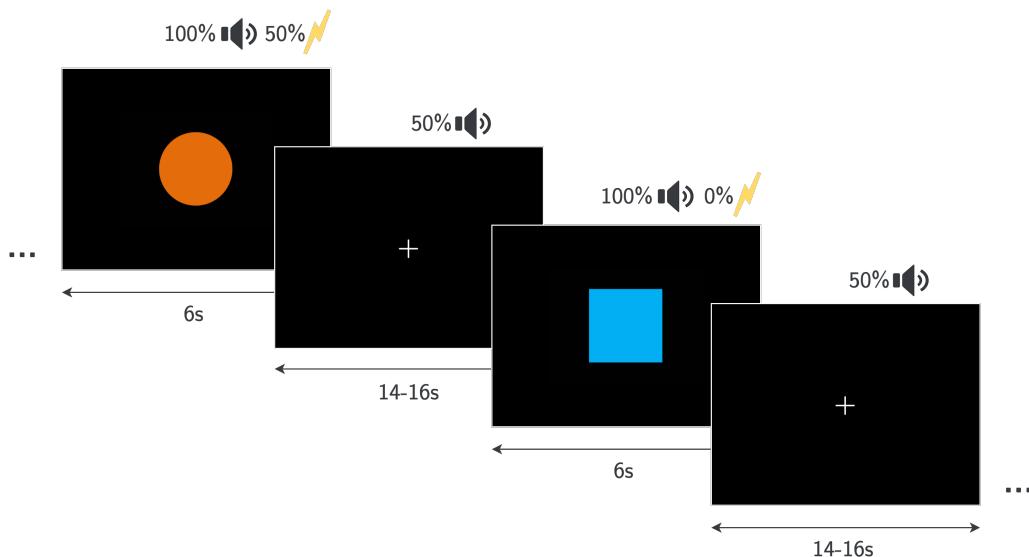


Abbildung 1. Schematischer Ablauf der Trials (exemplarisch: orangener Kreis als CS+). Die Prozente entsprechen den Wahrscheinlichkeiten für die Vergabe von Schreckreizen und elektrotaktile Reizen in den jeweiligen Trials.

Die Versuchspersonen werden auf jeweils einer von vier verschiedenen Trialreihenfolgen getestet,

deren wesentliche Unterschiede darin liegen, mit welchen Stimuli das Akquisitionstraining beginnt (CS+ oder CS-) und ob der erste präsentierte CS+ verstärkt wird oder nicht. Die Reihenfolgen werden über Versuchspersonen hinweg ausbalanciert. In der Hälfte des Trainings nach dem 16. Trial wird eine vollständige Instruktion auf dem Bildschirm präsentiert. Mit dem Wortlaut „Wenn ein unangenehmer Reiz am Bein erfolgt, dann immer nur bei diesem Symbol.“ und einer Abbildung des CS+ werden die Versuchspersonen über die CS-US-Kontingenzen aufgeklärt. Grund für die Darbietung der Instruktion ist einerseits die Fokussierung des Großprojekts auf den Extinktionstermin, für den eine gelernte CS-US-Kontingenz unabdingbar ist, und andererseits die anteilige Messung der Patient*innen-Stichprobe.

Das Akquisitionstraining insgesamt dauert ca. 12 min. Im Anschluss werden die Proband*innen erneut um eine Einschätzung aller präsentierten Reize (CS+, CS-, Auslösereiz, US) auf den SAM-Skalen (Anhang B) gebeten. Es folgt ein kurzes Bewusstseins-Interview, bei dem erfasst wird, ob die Versuchspersonen eine Regel in Bezug auf die CS-US-Kontingenzen verbalisieren können. An dieser Stelle wird von der Versuchsleitung sichergestellt, dass die Kontingenzen spätestens nach Abschluss des Trainings bekannt ist. Dieses halb-standardisierte Interview soll eine erfolgreiche Furchtakquisition bei allen Versuchspersonen gewährleisten, um den für das Großprojekt relevanten Extinktionstermin durchführen zu können. Nach einigen kurzen Fragen zu Mahlzeiten, Koffein- und Alkoholkonsum vor der Messung werden die Versuchspersonen an die Interventionsleitung übergeben. Dort werden sie über ihre Zugehörigkeit zu Kontroll- bzw. Interventionsgruppe aufgeklärt und im Zuge des Großprojekts weiter begleitet.

3.4 Datenverarbeitung und Reaktionsdefinition

3.4.1 Hautleitwertreaktion

Die Vorverarbeitung sowohl der EDA- als auch der EMG-Daten erfolgt in MATLAB (Version R2020b, MathWorks, Inc., Natick, USA). Die EDA-Daten werden offline auf 100 Hz heruntergetaktet. Die Vorverarbeitung und Kennwertbestimmung der Daten erfolgt mit der Ledalab Software V3.4.9 (Benedek & Kaernbach, 2010) nach veröffentlichten Richtlinien (Boucsein et al., 2012). Dabei wird automatisch im Zeitfenster von 0.9 bis 5 s nach CS-Onset eine klassische Tiefpunkt-bis-Hochpunkt Differenz gebildet, wobei eine Mindestamplitude von 0.001 µS als Wertungskriterium dient. Die Hautleitwertreaktion entspricht damit der jeweiligen Amplitudenhöhe in µS. Dadurch, dass der Schreckkreis frühestens 4.5 s nach CS-Onset präsentiert wird und eine SCR auf diesen erst eine Sekunde ihren Onset hätte, interferiert die Vergabe des Schreckkreises nicht mit der Erfassung der SCR in diesem Zeitfenster. Nach der Verarbeitung in Ledalab werden die Gesamtreaktionen der einzelnen Versuchspersonen manuell inspiziert. In $n = 2$ Fällen sind keine SCR-Reaktionen im Akquisitionstraining erfasst worden, vermutlich aufgrund des Ablösens der Elektroden während der Messung. Diese beiden Versuchspersonen werden aus allen Folgeanalysen ausgeschlossen. Nullreaktionen werden vergeben, wenn die Amplitude der SCR das Mindestkriterium nicht erfüllt. Die übrigen 38 Datensätze enthalten 18% (223 von 1 216) Nullreaktionen. Fehlende Reaktionen gibt es aufgrund der automatischen Auswertung nicht. Auf Basis der Empfehlung eines durchgeföhrten Box-Cox-Tests für die SCR-Daten wird eine log-Transformation dieser Daten vorgenommen (Box & Cox, 1964).

3.4.2 Schreckreaktion

Die EMG-Daten werden offline in MATLAB durch einen Hochpassfilter mit 60 Hz und einen Tiefpassfilter mit 400 Hz geschickt (Butterworth Filter zweiten Grades, Blumenthal et al., 2005) und durchlaufen zusätzlich einen Kerbfilter bei 50 Hz, um Interferenzen durch Netzbrummen zu reduzieren. Das Signal wird gleichgerichtet und mit einem Tiefpassfilter von 15.9 Hz und einer Zeitkonstante von 10 ms geglättet. Die Kennwertbestimmung erfolgt manuell, indem innerhalb von 20 bis 120 ms nach dem Schreckkreis der Fußpunkt der Reaktion und innerhalb von 150 ms das lokale Maximum als Spitze festgelegt werden (Blumenthal et al., 2005). Die Schreckreaktion ergibt sich aus der Differenz (d. h. der Amplitudenhöhe) in μ V. Trials, die eine exzessive Grundlinienaktivität oder ein zufälliges Blinzeln kurz vor oder zu Beginn des interessierenden Zeitfensters zeigen, werden als *fehlend* definiert. Von allen Trials sind 13 (1.1 %) fehlend (wobei keine Versuchsperson mehr als zwei fehlende Werte aufweist). Für Trials, in denen eine Reaktion nicht distinkt erkennbar ist bzw. vollständig ausbleibt, wird eine Nullreaktion vergeben. Die Anzahl der Nullreaktionen beläuft sich auf 19 (1.6 %), wobei anzumerken ist, dass kein Kriterium für eine Mindestreaktionshöhe einer vorhandenen Reaktion angelegt wird.

3.5 Statistische Analyse

3.5.1 Multivariates Wachstumsmodell mit Mehrebenenstruktur

Der statistische Vergleich findet durch ein multivariates Wachstumsmodell (engl.: Growth Curve Analysis) im Rahmen eines Mehrebenenmodells statt. Mehrebenenmodelle (MEM) haben sich in den letzten Jahren psychophysiologischer Forschung zunehmend durchgesetzt, da sie für messwiederholte Daten als flexibler und präziser gelten. Als Erweiterung der regulären Regression können mit ihnen Daten analysiert werden, denen eine Mehrebenenstruktur zugrunde liegt (man spricht häufig von einer Hierarchie oder Verschachtelung der Daten). Im Fall dieser Arbeit besteht die Mehrebenenstruktur darin, dass für jede Versuchsperson eine Vielzahl an Messungen (mehrere Trials) erhoben wird und die einzelnen Beobachtungen einer Person nicht als unabhängig voneinander anzusehen sind.

Üblicherweise verwendete statistische Verfahren wie die messwiederholte ANOVA oder Regressions mit Ordinary-Least-Squares Schätzungen (dt.: Methode der kleinsten Quadrate) können für diese Datenstruktur nicht aufkommen und führen zu verfälschten Schätzparametern und Signifikanztests (Nezlek, 2012). MEM währenddessen erlauben den Regressionsparametern wie dem Intercept (dt.: Achsenabschnitt) und den Slopes (dt.: Steigungen) zwischen den Ebenen zu variieren. Damit schätzen sie Regressionskoeffizienten nicht nur als feste Effekte (engl.: fixed effects), sondern modellieren sie als zufällige Effekte (engl.: random effects). Die festen Effekte sind ähnlich wie in herkömmlichen Regressionsmethoden zu verstehen. Mit der zusätzlichen Schätzung der Zufallsfehler dieser festen Koeffizienten ist es den Regressionsparametern der einzelnen Personen im Modell nun aber *erlaubt*, von diesem mittleren Effekt abzuweichen. Die Modellierung als zufällige Effekte repräsentiert die Annahme, dass die geschätzten Regressionsparameter Stichproben aus einer zugrundeliegenden Verteilung von Parametern in der Grundgesamtheit sind. Die simultane Schätzung von Zufallsfehlern nicht nur auf Ebene der Reaktion (Ebene 1), sondern auch auf Ebene der einzelnen Versuchsperson (Ebene 2) ist ein wesentlicher Vorteil von MEM. Vor allem in der Psychophysiolgie, wo es zwischen verschiedenen Personen zu enormen Schwankungen in Reaktionsmustern kommen kann, zeigen MEM die notwendige Flexibilität, bessere Erklärungen für ihre Datenmuster bereitzustellen zu können. Zusätzlich bieten MEM den Vorteil, für eine

fehlende Äquidistanz von Messzeitpunkten aufzukommen sowie mit fehlenden Datenpunkten umzugehen (ungleich üblicher ANOVA-Methoden) und damit reliablere Schlussfolgerungen zu liefern.

Im Kontext von messwiederholten Daten, denen eine zeitliche Struktur zugrunde liegt, erlauben diese Modelle die Schätzung von sogenannten *Wachstumskurven*. Dieser eher breite Term umfasst im MEM-Kontext die Untersuchung von intraindividuellen Veränderungen und den Prädiktoren, die diese beeinflussen können (Curran et al., 2010). Eine tiefgehende Erläuterung der Konzepte von Wachstumskurven- und Mehrebenenanalysen geht über den Umfang dieser Arbeit weit hinaus. Eine englischsprachige Einführung in die Thematik geben zum Beispiel Bagiella et al. (2000) und Nezlek (2012) und Page-Gould (2016), als deutschsprachige Einführung eignet sich Nezlek et al. (2006). Für einen tieferen Einblick seien die Bücher von Bickel (2007) und Grimm et al. (2017) und Raudenbush und Bryk (2002) empfohlen. Um einen Vergleich zweier abhängiger Variablen zu implementieren, wird das univariate Wachstumsmodell auf ein multivariates erweitert (nach dem Vorbild von Curran et al., 2012; MacCallum et al., 1997). Zusätzlich zu der Schätzung von Wachstumsmodellen für jede der beiden abhängigen Variablen erlaubt der multivariate Ansatz die Untersuchung ihrer Kovarianzen. Im Folgenden erläutere ich die in dieser Arbeit umgesetzte Modellstruktur sowie das allgemeine Vorgehen, die Annahmen und die Entscheidungskriterien zum Aufstellen und Schätzen des Modells.

3.5.2 Modellstruktur & Variablen

Die Daten werden als Zwei-Ebenen-Modell analysiert mit der Reaktion auf Ebene 1 und der Versuchsperson auf Ebene 2. Das heißt, dass jede Reaktion der Versuchsperson zugeordnet wird, von der sie stammt. Einen Überblick über die Datenstruktur liefert Abbildung 2.

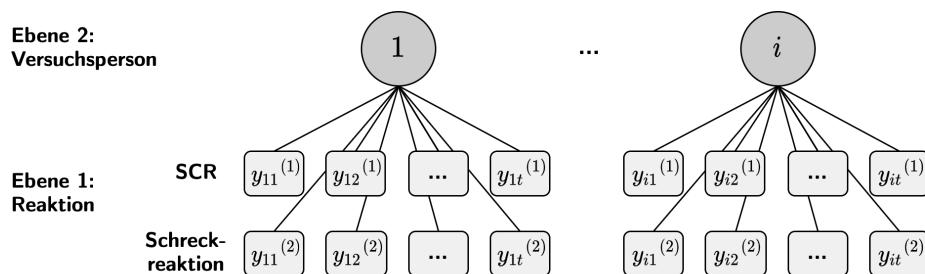


Abbildung 2. Schematische Datenstruktur im Rahmen des Mehrebenenmodells mit zwei Ebenen und zwei abhängigen Variablen (i : Versuchsperson, t : Trial).

Um ein Wachstumsmodell aufzustellen, benötigt es eine Zeitvariable. In dieser Arbeit wird die Trialnummer als kontinuierlicher Prädiktor herangezogen, um Veränderungen über die Zeit zu erfassen. Bei der Modellbildung wird getestet, ob das Datenmuster der beiden Reaktionsvariablen durch eine lineare oder eine quadratische Zeitvariable besser erklärt werden kann. Dafür wird im Modell ein Polynom-Term zweiten Grades linear-additiv in das Modell inkludiert und sein Effekt auf die Anpassungsgüte überprüft. Um der Multikollinearität der Zeitprädiktoren vorzubeugen, wird die Trialnummer zentriert und anschließend für den quadratischen Modellterm quadriert. Damit liegt der Mittelpunkt der Trialvariable zwischen dem 8. und 9. Trial (analog zu Kristjansson et al., 2007). Der Stimulustyp wird als zweistufiger kategorialer Prädiktor dummykodiert mit dem CS- als Referenzkategorie. Da jede Versuchsperson sowohl auf den CS+ als auch auf den CS- in beiden Reaktionsmaßen reagiert,

gilt der Stimulustyp als ein sogenannter zeitvarianter Prädiktor, der auf Reaktionsebene des Modells hinzugefügt wird. Das multivariate Wachstumsmodell enthält schließlich ebenfalls eine Variable, die kodiert, zu welchem Reaktionsmaß eine Reaktion gehört. Es untersucht den individuellen Verlauf der beiden abhängigen Variablen und kann zusätzlich die Kovariationen zwischen den Koeffizienten dieser beiden Wachstumsmodelle schätzen (Grimm et al., 2017). Die Besonderheit ist, dass für jede Person ein eigenes Wachstumsmodell für die Hautleitwert- und die Schreckreaktion geschätzt wird, indem die jeweilige Reaktionsvariable in Abhängigkeit von der Zeit und gegebenenfalls vom Stimulustyp bestimmt wird. Die Modellbildung erfolgt mit den unstandardisierten Werten beider Variablen. Die Tabelle 1 zeigt exemplarisch die Struktur des Datensatzes in langem Format einschließlich der Kodierung und Zentrierung der Prädiktorvariablen.

Tabelle 1

Exemplarische Datenstruktur

| VP | Reihenfolge | Trial | lin | quad | CS | AV | Reaktion |
|-----|-------------|-------|------|-------|-----|-----|----------|
| 1 | 3 | 1 | -7.5 | 56.25 | 0 | SCR | 0.0059 |
| 1 | 3 | 1 | -7.5 | 56.25 | 0 | STR | 61.421 |
| 1 | 3 | 1 | -7.5 | 56.25 | 1 | SCR | 1.5482 |
| 1 | 3 | 1 | -7.5 | 56.25 | 1 | STR | 29.184 |
| 1 | 3 | 2 | -6.5 | 42.25 | 0 | SCR | 0.0024 |
| 1 | 3 | 2 | -6.5 | 42.25 | 0 | STR | 49.147 |
| 1 | 3 | 2 | -6.5 | 42.25 | 1 | SCR | 0.2949 |
| ... | ... | ... | ... | ... | ... | ... | ... |

Anmerkungen. VP: Versuchsperson, Reihenfolge: Zuordnung der Trialreihenfolge, lin: linearer zentraler Zeitterm, quad: quadrierter zentraler Zeitterm, CS: Stimulustyp (1: CS+, 0: CS-), AV: abhängige Variable (SCR: Hautleitwertreaktion, STR: Schreckreaktion).

3.5.3 Allgemeines Vorgehen bei der Modellbildung

Der Modellbildungsprozess läuft in mehreren Schritten ab. Nach einer deskriptiven und visuellen Analyse der Reaktionsvariablen wird in Schritt 0 zunächst ein unkonditioniertes Modell aufgestellt, welches keinerlei Prädiktoren enthält. Dieses sogenannte *Nullmodell* dient dazu, die Variabilität zu erfassen, die auf Unterschiede der Versuchspersonen zurückgeführt werden kann.

$$\begin{aligned} \text{Ebene 1} \quad y_{it}^{(k)} &= \pi_{0i}^{(k)} + e_{it}^{(k)} \\ \text{Ebene 2} \quad \pi_{0i}^{(k)} &= \beta_{00}^{(k)} + r_{0i}^{(k)} \end{aligned} \quad (1)$$

In diesem Nullmodell ist y die Reaktion der Versuchsperson i zum Zeitpunkt t auf der Reaktionsvariable k . Der Index k beschreibt für $k = 1$ die Hautleitwertreaktion und für $k = 2$ die Schreckreaktion. Der Modellterm $\pi_{0i}^{(k)}$ ist der individuelle Intercept der Person i auf der Variable k . Das Residuum $e_{it}^{(k)}$ erfasst die Abweichung der Reaktion y vom individuellen Mittelwert, wiederum jeweils für beide k . Ebene 2 enthält die Gleichung, mit der der Intercept als zufälliger Effekt modelliert wird. Das Residuum $r_{0i}^{(k)}$ gibt an, wie der personenspezifische Intercept vom globalen (festen) Intercept $\beta_{00}^{(k)}$ abweicht. Diese Mehrebenengleichungen lassen sich auch in einem gemischten Format (d. h. innerhalb einer Gleichung für

beide abhängigen Variablen) kombinieren. Um das zu realisieren, werden zwei zusätzliche Dummyvariablen benötigt – eine für jedes Reaktionsmaß. Diese kodieren, zu welcher abhängigen Variable eine Reaktion zugeordnet werden kann (MacCallum et al., 1997).

$$\delta_1 = \begin{cases} 1 & \text{wenn } k = 1 \text{ (Hautleitwertreaktion)} \\ 0 & \text{wenn } k = 2 \end{cases} \quad \text{und} \quad \delta_2 = \begin{cases} 1 & \text{wenn } k = 2 \text{ (Schreckreaktion)} \\ 0 & \text{wenn } k = 1 \end{cases} \quad (2)$$

Mit der Wichtung über die Dummyvariablen ist es möglich, ein einziges Modell zu schätzen und dabei separate Wachstumsmodelle für die jeweiligen Reaktionsmaße sowie ihre geteilte Varianz zu bestimmen. Dabei kann das Modell trotz einer multivariaten Zwei-Ebenen-Struktur mit derselben Linearkombination geschätzt werden. Für ein gegebenes $y_{it}^{(1)}$ zum Beispiel gewährleistet die Variable δ , dass der zweite Summand der Summe aus der Gleichung (3) entfällt. In der Summenschreibweise werden damit für $m = 1$ die Regressionsparameter der Hautleitwertreaktion ($k = 1$) und für $m = 2$ die der Schreckreaktion ($k = 2$) geschätzt.

$$y_{it}^{(k)} = \sum_{m=1}^2 \delta_m \left(\beta_{00}^{(m)} + r_{0i}^{(m)} + e_{it}^{(m)} \right) \quad (3)$$

In Schritt 1 wird die Struktur der zufälligen Effekte bestehend aus zufälligen Intercepts und Slopes für jedes Reaktionsmaß festgelegt. Dafür wird zunächst die maximale Struktur festgestellt (d.h. jeder Regressionskoeffizient wird als zufällig modelliert; Barr, 2013). In der maximalen Struktur sind neben zufälligen Intercepts außerdem zufällige Slopes für beide Zeitterme, den Stimulustyp und seine Interaktionen mit den Zeittermen enthalten (Brauer & Curtin, 2018). Außerdem schätzt das Modell die Kovarianzen zwischen den einzelnen Effekten. Es kann vorkommen, dass das Modell für die Datengrundlage, auf der es beruht, zu komplex ist. Dann können Konvergenzprobleme oder sehr hohe Korrelationen zwischen den zufälligen Effekten auftreten (Brauer & Curtin, 2018; Matuschek et al., 2017). Sind zufällige Effekte nah bei null, kann das bedeuten, dass zu wenig Varianz in den Daten vorliegt, um diese zu schätzen. Sehr hohe Korrelationen können eine Überanpassung anzeigen. In diesen Fällen – und ebenso bei Konvergenzproblemen – wird empfohlen, die Komplexität zu verringern (Brauer & Curtin, 2018; Matuschek et al., 2017). Brauer und Curtin (2018) stellen in ihrem Artikel einen Leitfaden für diesen Zweck zur Verfügung. Sollte einer der genannten Fälle bei der maximalen Struktur zufälliger Effekte auftreten, werden die Empfehlungen von Brauer und Curtin (2018) zu Rate gezogen. Im Ergebnisteil werden die auf deren Basis getroffene Entscheidungen ausführlich dargelegt. Potentielle Schritte zur Behebung umfassen das Heraufsetzen der maximalen Iterationen, das Herausnehmen bestimmter – potentiell weniger relevanter – zufälliger Effekte oder das Unterdrücken von Kovarianzen zwischen ihnen (Barr, 2013; Brauer & Curtin, 2018). Modelle mit verschiedenen zufälligen Effekten werden über das Restricted Maximum-Likelihood (REML) Verfahren berechnet, da es präziser für die Schätzung der Varianzkomponenten ist (für eine Gegenüberstellung von ML und REML siehe Raudenbush & Bryk, 2002).

Ist die Struktur der zufälligen Effekte gewählt, kann in Schritt 2 die passende Struktur der Zeitterme für die Daten gewählt werden. Diese festen Effekte werden sequentiell über Vorwärtsselektion hinzugefügt und ihr Einfluss auf die Anpassungsgüte mittels Modellvergleichen – in diesem Fall Likelihood-

Ratio Tests (LRT) – evaluiert. Der LRT vergleicht ein einfacheres Modell mit einem komplexeren. Die Nullhypothese, dass die Anpassungsgüte der beiden Modelle keinen Unterschied aufweist, wird abgelehnt, wenn die χ^2 -verteilte Teststatistik auf einem 5 %-Niveau signifikant wird. Eine bedeutend bessere Anpassung des komplexeren Modells resultiert in seiner Weiterverwendung. Bei einer Insignifikanz wird das sparsamere Modell beibehalten. Zusätzlich liefert das R-Paket `nlme` approximierte p -Werte und Konfidenzintervalle für feste Effekte auf Basis von bedingten t -Tests (Pinheiro & Bates, 2000). Beim Auswählen der Struktur der festen Effekte präferieren Pinheiro und Bates (2000) die Verwendung der p -Werte über LRT-Vergleiche. In dieser Arbeit ziehe ich beide Verfahren zur Beurteilung der Signifikanz heran. Sollte das Signifikanzergebnis unterschiedlich sein, wird das Ergebnis des t -Tests priorisiert. Ausgehend von dem Nullmodell mit der zufälligen Effekt-Struktur aus Schritt 1 wird zunächst der lineare Zeitterm (lin) als Prädiktor hinzugefügt. Mit Aufnahme des quadrierten Terms (quad) wird getestet, welche Zeitstruktur eine bessere Anpassungsgüte an die Daten aufweist. Beim Vergleich von Modellen mit unterschiedlicher Struktur der festen Effekte werden die Parameter iterativ mit dem Maximum-Likelihood (ML) Verfahren geschätzt.

Sobald auch die Struktur der Zeitterme ausgewählt ist, kann der zeitvariante Prädiktor Stimulustyp in das Modell aufgenommen werden. In Schritt 3 wird diese Variable auf Ebene 1 eingefügt. Mittels LRT wird separat für jede der abhängigen Variablen überprüft, ob die Interaktion von Stimulustyp mit den Zeittermen zu einer besseren Anpassungsgüte führt. Ist diese Verbesserung signifikant, erhält das komplexere Modell den Vorzug. Für die Vergleiche von Modellen mit unterschiedlicher fester Effekt-Struktur wird erneut das ML-Schätzverfahren verwendet.

Das maximal mögliche Modell enthält für jede der abhängigen Variablen einen linearen (lin) und einen quadratischen Zeitterm (quad) sowie den Prädiktor Stimulustyp (CS) und seine Interaktionen mit den Zeittermen. Die zufällige Effekt-Struktur ist maximal, d. h. dass jeder Regressionskoeffizient als zufällig modelliert wird.

$$y_{it}^{(k)} = \sum_{m=1}^2 \delta_m \left[\left(\beta_{00}^{(m)} + \beta_{10}^{(m)} \text{lin}_{it} + \beta_{20}^{(m)} \text{quad}_{it} + \beta_{30}^{(m)} \text{CS}_{it} \right. \right. \\ \left. \left. + \beta_{40}^{(m)} \text{lin}_{it} \text{CS}_{it} + \beta_{50}^{(m)} \text{quad}_{it} \text{CS}_{it} \right) \right. \\ \left. + \left(r_{0i}^{(m)} + r_{1i}^{(m)} \text{lin}_{it} + r_{2i}^{(m)} \text{quad}_{it} + r_{3i}^{(m)} \text{CS}_{it} \right. \right. \\ \left. \left. + r_{4i}^{(m)} \text{lin}_{it} \text{CS}_{it} + r_{5i}^{(m)} \text{quad}_{it} \text{CS}_{it} \right) + e_{it}^{(m)} \right] \quad (4)$$

Dieses Modell erweitert das Nullmodell in Gleichung (3) um zusätzliche Parameter. Die Variable y ist wiederum die Reaktion der Versuchsperson i zum Zeitpunkt t auf Reaktionsvariable k . Über die Dummy-Variablen aus Gleichung (2) lässt sich die multivariate Form auch hier über eine Summe darstellen, wobei das Modell für $m = 1$ Parameter der SCR schätzt und für $m = 2$ die der Schreckreaktion. Die Effekte β werden für jedes Reaktionsmaß k separat berechnet, ebenso die zufälligen Effekte r und der Residuen-Term e auf Ebene 1. Die Prädiktorvariablen variieren über Person i und Messzeitpunkt t , allerdings unterscheiden sie sich nicht für die verschiedenen abhängigen Variablen, weswegen diese Modellterme keinen oberen Index haben. Durch die neuen Modellterme beschreibt β_{00} nun den mittleren Intercept zwischen dem 8. und 9. CS– Trial (durch die Zentrierung von Trial und dem CS– als Referenzkategorie). Der Parameter β_{10} definiert für jedes k die mittlere (momentane) lineare Änderungsrate für den CS–,

β_{20} die mittlere quadratische Slope für den CS-. Der mittlere Effekt des Stimulustyps wird durch β_{30} beschrieben, β_{40} schätzt den mittleren Unterschied in der linearen Änderungsrate und β_{50} den mittleren Unterschied in der quadratischen Änderungsrate zwischen CS+ und CS-. Die Modellterme r sind die zufälligen Effekte, die die β -Koeffizienten auf Ebene 2 als zufällig modellieren.

Eine vereinfachte Schreibweise bietet die Darstellung über Matrizen (Gleichung 5). Angenommen, \mathbf{X}_1 sei eine Matrix der festen Effekte und $\mathbf{B}^{(k)}$ eine Matrix der Koeffizienten dieser Effekte für jedes k sowie \mathbf{X}_2 eine Matrix aller als zufällig modellierten Effekte und $\mathbf{R}^{(k)}$ eine Matrix der zugehörigen Koeffizienten, dann bietet Gleichung (6) eine minimalistische Schreibweise des maximal möglichen Modells dieser Arbeit aus Gleichung (4).

$$\mathbf{B}^{(k)} = \begin{bmatrix} \beta_{00}^{(k)} \\ \beta_{10}^{(k)} \\ \beta_{20}^{(k)} \\ \beta_{30}^{(k)} \\ \beta_{40}^{(k)} \\ \beta_{50}^{(k)} \end{bmatrix}, \quad \mathbf{X}_1 = \begin{bmatrix} 1 \\ \text{lin}_{it} \\ \text{quad}_{it} \\ \text{CS}_{it} \\ \text{CS}_{it} \times \text{lin}_{it} \\ \text{CS}_{it} \times \text{quad}_{it} \end{bmatrix}, \quad \mathbf{R}^{(k)} = \begin{bmatrix} r_{0i}^{(k)} \\ r_{1i}^{(k)} \\ r_{2i}^{(k)} \\ r_{3i}^{(k)} \\ r_{4i}^{(k)} \\ r_{5i}^{(k)} \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} 1 \\ \text{lin}_{it} \\ \text{quad}_{it} \\ \text{CS}_{it} \\ \text{CS}_{it} \times \text{lin}_{it} \\ \text{CS}_{it} \times \text{quad}_{it} \end{bmatrix} \quad (5)$$

$$y_{it}^{(k)} = \sum_{m=1}^2 \delta_m \left[(\mathbf{B}^{(m)})^T \mathbf{X}_1 + (\mathbf{R}^{(m)})^T \mathbf{X}_2 + e_{it}^{(m)} \right] \quad (6)$$

Im Einklang mit üblichen Konventionen wird das finale multivariate Wachstumsmodell mit dem REML-Verfahren geschätzt und anschließend die Ergebnisse berichtet. Alle statistischen Analysen erfolgen in der freien Programmiersprache R für statistische Berechnungen (Version 4.0.3, R Core Team, 2020). Die Modellschätzung erfolgt mit dem Paket `nlme` (Pinheiro et al., 2020). Weitere verwendete Pakete im Zuge der Datenaufbereitung und Visualisierung sind `tidyverse` (Wickham et al., 2019), `MASS` (Venables & Ripley, 2002) und `extrafont` (Chang, 2014).

Die Beantwortung der explorativen Fragestellungen wird lose in zwei Abschnitte geteilt. Zunächst wird über die Modellbildung deutlich, ob bei beiden abhängigen Variablen die gleiche Zeittermstruktur signifikant wird. Hierüber wird ersichtlich, durch welche Wachstumskurven sich die Reaktionsmaße am besten modellieren lassen. Zur Betrachtung des differentiellen Furchtlernens wird untersucht, welchen Einfluss die Interaktionen von Stimulustyp und Zeittermen auf die Anpassungsgüte haben. Durch die Möglichkeit, Effekte als zufällig zu modellieren, können explorativ Aussagen über die Unterschiedlichkeit von Versuchspersonen auf einzelnen Regressionsparametern getroffen werden (in Abhängigkeit von der Modellbildung). Außerdem lassen sich vereinzelt Korrelationen zwischen zufälligen Parametern der abhängigen Variablen prüfen.

3.5.4 Zu prüfende Annahmen

Zunächst wird angenommen, dass die Residuen auf Ebene 1 ($e_{it}^{(k)}$) multivariat normalverteilt sind mit einem Erwartungswert von 0.

$$\begin{bmatrix} e_{it}^{(1)} \\ e_{it}^{(2)} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e^{(1)}}^2 & 0 \\ 0 & \sigma_{e^{(1)}, e^{(2)}} \end{bmatrix} \right) \quad (7)$$

Eine weitere Annahme betrifft die multivariate Normalverteilung der zufälligen Komponenten auf Ebene 2 mit Σ_T als unstrukturierter Varianz-Kovarianzmatrix.

$$r_{0i}^{(k)}, r_{1i}^{(k)}, r_{2i}^{(k)}, r_{3i}^{(k)}, r_{4i}^{(k)}, r_{5i}^{(k)} \sim \mathcal{N}(0, \Sigma_T) \quad (8)$$

An den Best-Practice-Empfehlungen in Meteyard und Davies (2020) orientiert, werden Annahmen über die Residuen am finalen Modell geprüft. Die Residuen jeder abhängigen Variable werden auf Linearität und multivariate Normalverteilung untersucht. Außerdem werden die Residuen jeder Reaktionsvariable separat auf Homoskedastizität geprüft. Für diese Annahmen werden Q-Q-Diagramme, Histogramme und Streudiagramme zwischen Residuen und angepassten Werten grafisch evaluiert. Eine weitere Annahme von MEM bezieht sich auf die Abwesenheit von Multikollinearität der Prädiktoren. Durch die Zentrierung der Zeitterme und dem Stimulustyp als binäre Variable wird diese Annahme als unverletzt betrachtet.

4 Ergebnisse

4.1 Explorative Datenanalyse

Der gesamte Mittelwert der log-transformierten Hautleitwertreaktionen in diesem Datensatz beträgt $0.15 \log \mu\text{S}$ ($SD = 0.26$). Für die Schreckreaktion liegt der Mittelwert bei $67.57 \mu\text{V}$ ($SD = 61.94$). Tabelle 2 beinhaltet eine Übersicht über deskriptive Statistiken über den Verlauf des Akquisitionstrainings separat für beide abhängige Variablen. Eine signifikante CR (d. h. auf den CS+ wird bedeutend stärker reagiert als auf den CS-) ist über alle Trials hinweg für die Hautleitwertreaktionen zu beobachten und zeigt sich in den einzelnen Trials 10 bis 14 der zweiten Trainingshälfte. Für die Schreckreaktionen sind in Trial 12 und 13 die Reaktionen auf den CS+ signifikant von denen auf den CS- verschieden sowie ebenfalls gemittelt über den Trainingszeitraum. Abbildung 3 zeigt die durchschnittlichen Verläufe beider abhängigen Variablen über die Dauer des Akquisitionstrainings. Die Tabelle 3 zeigt bivariate Unterschiede und Zusammenhänge zwischen den abhängigen Variablen zu allen Messzeitpunkten auf. Die erste Spalte beinhaltet bivariate Korrelationen von Schreck- und Hautleitwertreaktionen (insgesamt und aufgeschlüsselt nach Messzeitpunkten). Gemittelt über das ganze Training gehen höhere SCR

Tabelle 2

Deskriptive Statistiken & Mittelwertunterschiede innerhalb beider abhängiger Variablen

| Trial | SCR | | | STR | | |
|-------|--------------------|--------------------|------------------------|--------------------|--------------------|------------------------|
| | CS+ ^a | CS- ^a | Differenz ^b | CS+ ^a | CS- ^a | Differenz ^b |
| insg. | 0.18 (0.28) | 0.12 (0.23) | 0.06*** | 70.1 (59.7) | 65.1 (64.0) | 5.0** |
| 1 | 0.31 (0.38) | 0.27 (0.31) | 0.04 | 93.1 (61.1) | 100.4 (91.8) | -7.3 |
| 2 | 0.14 (0.24) | 0.15 (0.23) | -0.01 | 92.0 (60.6) | 85.4 (64.2) | 6.6 |
| 3 | 0.17 (0.21) | 0.15 (0.21) | 0.02 | 94.0 (75.2) | 80.5 (60.5) | 13.5 |
| 4 | 0.14 (0.25) | 0.11 (0.21) | 0.03 | 73.2 (55.3) | 76.7 (61.4) | -3.5 |
| 5 | 0.16 (0.29) | 0.18 (0.33) | -0.02 | 80.7 (72.5) | 83.1 (66.1) | -2.4 |
| 6 | 0.20 (0.31) | 0.18 (0.29) | 0.02 | 61.6 (47.4) | 70.4 (75.0) | -8.8 |
| 7 | 0.08 (0.12) | 0.09 (0.17) | -0.01 | 66.7 (55.6) | 68.2 (63.1) | -1.5 |
| 8 | 0.15 (0.31) | 0.09 (0.23) | 0.06 | 68.1 (56.7) | 64.2 (69.3) | 3.9 |
| 9 | 0.34 (0.43) | 0.26 (0.32) | 0.08 | 86.0 (77.3) | 63.2 (67.3) | 22.8 |
| 10 | 0.17 (0.23) | 0.10 (0.20) | 0.07* | 59.1 (49.1) | 57.1 (63.5) | 2.0 |
| 11 | 0.19 (0.26) | 0.06 (0.12) | 0.13** | 63.8 (60.0) | 54.2 (56.2) | 9.6 |
| 12 | 0.15 (0.23) | 0.05 (0.10) | 0.10** | 69.1 (60.8) | 51.4 (52.1) | 17.7* |
| 13 | 0.18 (0.28) | 0.05 (0.09) | 0.13** | 64.3 (53.3) | 44.5 (38.0) | 19.8** |
| 14 | 0.24 (0.33) | 0.10 (0.24) | 0.14* | 47.5 (44.9) | 47.1 (48.6) | 0.4 |
| 15 | 0.12 (0.21) | 0.06 (0.10) | 0.06 | 51.7 (45.6) | 48.0 (62.7) | 3.7 |
| 16 | 0.16 (0.27) | 0.06 (0.15) | 0.10 | 50.4 (49.3) | 45.4 (43.9) | 5.0 |

Anmerkungen. SCR: Hautleitwertreaktion, STR: Schreckreaktion, insg.: über alle Trials hinweg.

^a Mittelwerte (für SCR in $\log \mu\text{S}$. für STR in μV) mit Standardabweichungen in Klammern.

^b CS+/CS- Differenzen; Mittelwerte auf Unterschiede getestet mittels gepaarten *t*-Tests; * $p < .05$, ** $p < .01$, *** $p < .001$.

tendenziell mit höheren Schreckreaktionen einher (kleiner Effekt von $r = .13, p < .001$). Auch aufgeteilt in einzelne Messzeitpunkte zeigt sich in der Tendenz ein positiver Zusammenhang zwischen den beiden Reaktionsmaßen, auch wenn er nur im ersten ($r = .25, p = .03$) und im neunten Trial ($r = .30, p = .01$) statistisch bedeutsam ist. In Anhang C ist zusätzlich ein Streudiagramm zwischen den Reaktionen beider abhängiger Variablen zur Visualisierung abgebildet. Die zweite Spalte der Tabelle 3 zeigt die Korrelationen zwischen den Differenzwerten (CS+ minus CS-) für beide Reaktionsmaße. Der insgesamt kleine, positive Zusammenhang von ($r = .15, p < .001$) weist darauf hin, dass eine stärkere konditionierte Reaktion auf dem einen Maß mit einer stärkeren CR auf dem anderen Maß einhergeht. Betrachtet man die einzelnen Trials ist die Korrelation zum Trial 9 – direkt nach der Instruktion – positiv und signifikant ($r = .46, p = .004$). In Trial 3 und 8 sind die Korrelationen der Differenzwerte signifikant negativ. Hier gehen stärkere CR auf dem einen Maß mit geringeren CR auf dem anderen Maß einher. Die letzte Spalte der bivariate Tabelle vergleicht die Mittelwerte der CS+/CS- Differenzen insgesamt und aufgeteilt nach Trials. Dafür wurden die Reaktionen einmalig vorab standardisiert, sodass die Differenzen als z -Werte vergleichbar werden. Über das ganze Akquisitionstraining hinweg unterscheiden sich die CS+/CS- Differenzen für die beiden Maße signifikant voneinander, wobei die z -Differenzen der SCR im Durchschnitt größer sind als die der STR. In einzelnen Trials sind diese Unterschiede nur in Trial 14 evident. Auch hier weist die SCR eine größere Differenz auf.

Eine Überprüfung der Normalverteilung der abhängigen Variablen mittels Shapiro-Wilk-Tests (Shapiro & Wilk, 1965) ergab, dass weder global noch für die einzelnen Messzeitpunkte eine Normalverteilung angenommen werden kann (Anhang D liefert eine tabellarische Übersicht der Teststatistiken für beide abhängigen Variablen). Es sei darauf hingewiesen, dass dies nicht einschränkend für die Modellbildung ist. Auf eine Analyse der Ausreißer sowie der fehlenden Werte wird an dieser Stelle verzichtet, da eine Vorverarbeitung im Rahmen der Mehrebenenmodellierung nicht notwendig ist.

4.2 Modellbildung

Aufgrund der verhältnismäßig kleinen Stichprobengröße und der Komplexität des Modells wird über die Funktion `lmeControl` (nach Grimm et al., 2017) in R eine Kontrollfunktion definiert, die unter anderem die maximale Anzahl an Iterationen auf 500 erhöht. Die Konfidenzintervalle und p -Werte der Modellparameter werden innerhalb des Pakets `nlme` über eine Normalapproximation der Verteilung der geschätzten Effekte berechnet (für die mathematische Herleitung siehe Pinheiro & Bates, 2000). Eine Übersicht über den im Folgenden beschriebenen sequentiellen Modellbildungsprozess sowie über alle berechneten Modellvergleiche liefert Tabelle 4.

Das in Schritt 0 erstellte Nullmodell (Modellname `null`) schätzt einen Intercept für die Hautleitwertreaktion von $0.15 \log \mu\text{S}$ ($\beta_{00}^{(1)}, SE = 0.022$, Konfidenzintervall $KI [0.108, 0.194]$). Für die Schreckreaktion beträgt die Schätzung $67.19 \mu\text{V}$ ($\beta_{00}^{(2)}, SE = 8.14, KI [51.24, 83.13]$). Die zufälligen Effekte zeigen mit einer Standardabweichung von 0.13 für die SCR ($KI [0.101; 0.166]$) und 49.7 für die Schreckreaktion ($KI [39.5; 62.5]$) substanziale Varianz der Versuchspersonen in ihren Intercepts auf den jeweiligen Reaktionsvariablen.

In Schritt 1 wird das Nullmodell in seiner zufälligen Effekt-Struktur auf den maximal möglichen Aufbau wie in Gleichung (6) erweitert (Modellname `ran1`). Dieses Modell konvergiert mit den hier verwendeten Daten nicht, sodass pragmatische Entscheidungen zur Reduzierung der Komplexität getroffen

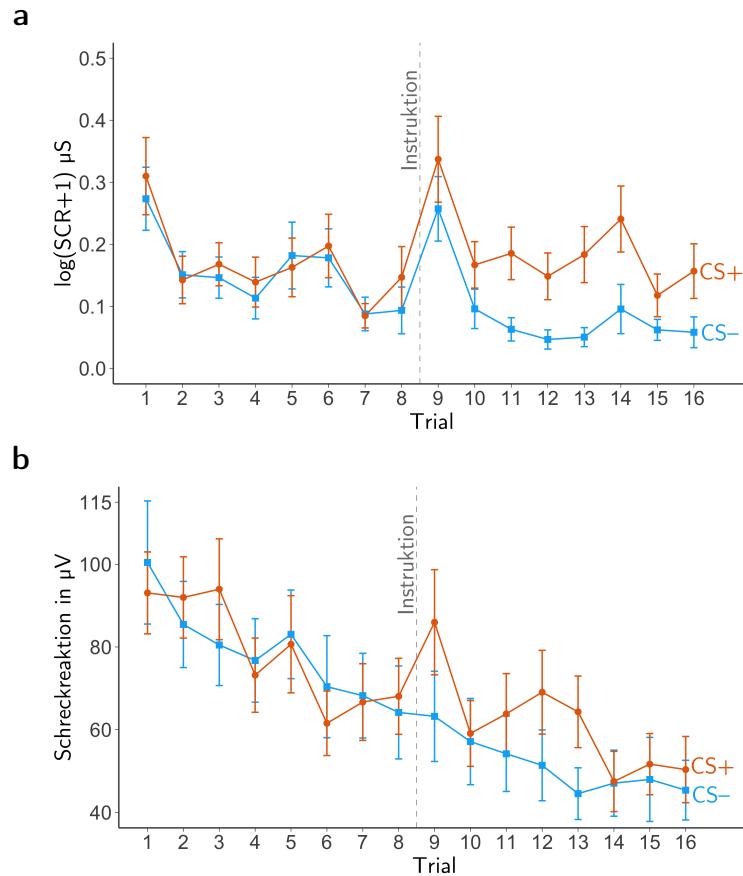


Abbildung 3. Mittlere Hautleitwert- (a) und Schreckreaktionen (b) auf den CS+ und CS- über den Verlauf des Experiments. Die Fehlerbalken repräsentieren SE.

wurden. Lose orientiert an den Empfehlungen von Brauer und Curtin (2018) wird die Struktur der zufälligen Effekte sequentiell reduziert – zunächst, bis Konvergenz erzielt wird. Wenn die Interaktion von sogenannten *within-subject* Prädiktoren (innerhalb Versuchspersonen) von Interesse ist, soll diese laut Barr (2013) nach Möglichkeit zufällig modelliert und einfache zufällige Effekte fallen gelassen werden. Die Entscheidungen werden daher im Sinne der Untersuchung des Furchtlernens getroffen, für das insbesondere die Interaktion *CS* × *lin* (sprich die Differenzierung der Stimuli über die Zeit) wichtig ist. Unter dieser Begründung sind die Entscheidungen zur zufälligen Effektstruktur sowohl daten- als auch theoriegetrieben. Zunächst wird auf den zufälligen Effekt von *quad* ($r_{2i}^{(k)}$) für beide abhängigen Variablen verzichtet. Nachdem auch dieses Modell keine Konvergenz erreicht, wird der zufällige Slope für den linearen Zeitterm ($r_{1i}^{(k)}$) entfernt, doch auch dann konvergiert das Modell nicht. Erst ein zusätzliches Herausnehmen des zufälligen Anstiegs für den Stimulustyp ($r_{3i}^{(k)}$) und seine Interaktion mit dem quadratischen Zeitterm ($r_{5i}^{(k)}$) führen zur Konvergenz des Modells. Übrig bleiben für die Struktur der zufälligen Effekte also der Intercept ($r_{0i}^{(k)}$) und die Interaktion *CS* × *lin* ($r_{4i}^{(k)}$; Modellname *ran5*).

Im nächsten Schritt wird die Struktur der Zeitterme gewählt. Für beide abhängigen Variablen zeigen sich die linearen Terme als signifikante Prädiktoren (SCR: $\beta_{10}^{(1)} = -0.007$, $SE = 0.002$, $p < .001$; STR: $\beta_{10}^{(2)} = -3.22$, $SE = 0.35$, $p < .001$). Ein LRT zeigt, dass die Hinzunahme die Anpassungsgüte deutlich verbessert ($\chi^2(2) = 84.25$, $p < .001$). Bei den quadratischen Zeittermen zeichnet sich ein heterogeneres Bild ab. Für die SCR ist der neue Term kein signifikanter Prädiktor ($\beta_{20}^{(1)} = 0.0003$,

Tabelle 3

Unterschiede & Zusammenhänge zwischen den abhängigen Variablen

| Trial | Korrelation zwischen SCR & STR ^a | Korrelation der CS+/CS- Differenzwerte ^a | t-Teststatistiken der standardisierten Differenzwerte ^b |
|-------|---|---|--|
| insg. | .13*** | .15*** | -2.88** |
| 1 | .25* | .24 | -1.01 |
| 2 | .08 | -.08 | 0.77 |
| 3 | .04 | -.34* | 0.66 |
| 4 | .01 | -.02 | -0.60 |
| 5 | .09 | .28 | 0.18 |
| 6 | .08 | .28 | -0.94 |
| 7 | .01 | -.07 | -0.22 |
| 8 | .04 | -.36* | -0.84 |
| 9 | .30** | .46** | 0.11 |
| 10 | .11 | .32 | -1.64 |
| 11 | .06 | .00 | -1.60 |
| 12 | .16 | .18 | -1.07 |
| 13 | .14 | -.07 | -0.86 |
| 14 | .02 | .15 | -2.11* |
| 15 | .08 | .04 | -0.91 |
| 16 | -.03 | .10 | -1.37 |

Anmerkungen. * $p < .05$, ** $p < .01$, *** $p < .001$. ^a Pearsons Korrelationskoeffizient r .^b Die Reaktionen wurden global z -standardisiert ($M = 0$, $SD = 1$) und die Differenzwerte CS+/CS- für SCR und STR gebildet. Diese wurden mit gepaarten t -Tests auf Mittelwertunterschiede geprüft.

$p = .39$), wohl aber für die Schreckreaktion ($\beta_{20}^{(2)} = 0.109$, $SE = 0.05$, $p = .028$). Der quadratische Zeitterm wird daher für die SCR aus dem Modell entfernt (Modellname *time2.1*). Ein anschließender LRT zeigt, dass das Hinzufügen von *quad* für die Schreckreaktion die Anpassungsgüte gegenüber dem Modell mit nur linearen festen Effekten verbessert ($\chi^2(1) = 4.86$, $p = .028$).

Nach der Festlegung der Zeitstruktur wird in Schritt 3 der Prädiktor Stimulustyp in das Modell aufgenommen. Für beide abhängige Variablen ist der einfache *CS*-Term signifikant (SCR: $\beta_{30}^{(1)} = 0.058$, $SE = 0.0124$, $p < .001$; STR: $\beta_{30}^{(2)} = 5.19$, $SE = 1.866$, $p = .005$). Die Anpassungsgüte dieses Modells (Modellname *cs1*) gegenüber dem einfacheren (*time2.1*) ist bedeutend besser ($\chi^2(2) = 29.68$, $p < .001$). Als nächstes wird nun auch der feste Effekt der Interaktion vom Stimulustyp mit dem linearen Term eingefügt. Prinzipiell verbessert das Hinzufügen dieses festen Effekts die Anpassungsgüte des Modells ($\chi^2(2) = 11.36$, $p = .003$). Betrachtet man die beiden abhängigen Variablen separat voneinander, zeigt sich, dass die Interaktion nur für die SCR signifikant ist ($\beta_{40}^{(1)} = 0.0084$, $SE = 0.0029$, $p = .004$), nicht aber für die Schreckreaktion ($\beta_{40}^{(2)} = 0.713$, $SE = 0.457$, $p = .119$). Für letztere wird der Term daher im Sinne der Sparsamkeit entfernt (Modellname *cs2.1*). Das Hinzufügen der Interaktion *CS* \times *quad* als feste Effekt ist für keine der abhängigen Variablen von Bedeutung (SCR: $\beta_{50}^{(1)} = 0.0002$, $p = .41$; STR: $\beta_{50}^{(2)} = -0.074$, $p = .46$) und verbessert auch die Anpassungsgüte nicht ($\chi^2(2) = 1.25$, $p = .536$). Demnach wird diese nicht in das finale Modell mit aufgenommen.

Tabelle 4

Darstellung des Modellbildungsprozesses und der getätigten Modellvergleiche

| Schritt | Name | einfacheres | Feste | Zufällige | Schätz- | Anpassungsgüte | | LRT | |
|---------|---------|-------------------|----------------------|----------------------------------|---------|----------------|----|-----|---------------------------|
| | | Modell | Effekte ^a | Effekte ^a | methode | logL | df | df | χ^2 |
| 0 | null | | $\beta_{00}^{(k)}$ | $r_{0i}^{(k)}$ | REML | -6079.4 | 7 | - | - |
| 1 | ran1 | | " | alle | REML | | | | keine Konvergenz erreicht |
| 1 | ran2 | | " | $-r_{2i}^{(k)}$ | REML | | | | keine Konvergenz erreicht |
| 1 | ran3 | | " | $-r_{1i}^{(k)}$ | REML | | | | keine Konvergenz erreicht |
| 1 | ran4 | | " | $-r_{3i}^{(k)}$ | REML | | | | keine Konvergenz erreicht |
| 1 | ran5 | null | " | $-r_{5i}^{(k)}$ | REML | -5976.4 | 14 | 7 | 206*** |
| 2 | time1 | ran5 ^b | $+\beta_{10}^{(k)}$ | $r_{0i}^{(k)}$ $r_{4i}^{(k)}$ | ML | -5934.3 | 16 | 2 | 84.25*** |
| 2 | time2 | time1 | $+\beta_{20}^{(k)}$ | " | ML | -5931.5 | 18 | 2 | 5.59 |
| 2 | time2.1 | time1 | $-\beta_{20}^{(1)}$ | " | ML | -5931.9 | 17 | 1 | 4.86* |
| 3 | cs1 | time2.1 | $+\beta_{30}^{(k)}$ | " | ML | -5917.0 | 19 | 2 | 29.68*** |
| 3 | cs2 | cs1 | $+\beta_{40}^{(k)}$ | " | ML | -5911.3 | 21 | 2 | 11.36** |
| 3 | cs2.1 | cs1 | $-\beta_{40}^{(2)}$ | " | ML | -5912.6 | 20 | 1 | 8.93** |
| 3 | cs3 | cs2.1 | $+\beta_{50}^{(k)}$ | " | ML | -5911.9 | 22 | 2 | 1.25 |
| | fn | - | $-\beta_{50}^{(k)}$ | " | REML | -5927.0 | 20 | - | - |

Anmerkungen. $N = 2419$ Reaktionen (Ebene 1), $n = 38$ Versuchspersonen (Ebene 2); LRT: Likelihood-Ratio Test ggü. einfacherem Modell mit * $p < .05$, ** $p < .01$, *** $p < .001$; df: Freiheitsgrade.

^a bei positivem Vorzeichen wird dieser Term in diesem Schritt zum Modell hinzugefügt; bei negativen Vorzeichen wird er entfernt; Terme ohne Vorzeichen beschreiben die aktuelle, vollständige Struktur.

^b für den Vergleich erneut mit ML-Verfahren geschätzt.

4.3 Das finale Modell

Das finale multivariate Wachstumsmodell wird für die Schreckreaktionen mit Polynomen ersten und zweiten und für die Hautleitwertreaktion nur mit Polynomen ersten Grades modelliert. Es enthält feste Effekte für den Stimulustyp mit dem CS- als Referenzkategorie sowie eine Interaktion von $CS \times lin$ für die SCR. Die Struktur der zufälligen Effekte für die Versuchspersonen besteht aus zufälligen Intercepts und der als zufällig modellierten Interaktion $CS \times lin$ für beide Reaktionsmaße. Das final aufgestellte Modell kann als Teilmenge des maximalen Modells aus Gleichung (6) beschrieben werden mit analoger Notation:

$$\mathbf{B}^{(k)} = \begin{bmatrix} \beta_{00}^{(k)} \\ \beta_{10}^{(k)} \\ \beta_{20}^{(2)} \\ \beta_{30}^{(k)} \\ \beta_{40}^{(1)} \end{bmatrix}, \quad \mathbf{X}_1 = \begin{bmatrix} 1 \\ lin_{it} \\ quad_{it} \\ CS_{it} \\ CS_{it} \times lin_{it} \end{bmatrix}, \quad \mathbf{R}^{(k)} = \begin{bmatrix} r_{0i}^{(k)} \\ r_{4i}^{(k)} \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} 1 \\ CS_{it} \times lin_{it} \end{bmatrix} \quad (9)$$

$$y_{it}^{(k)} = \sum_{m=1}^2 \delta_m \left[\left(\mathbf{B}^{(m)} \right)^T \mathbf{X}_1 + \left(\mathbf{R}^{(m)} \right)^T \mathbf{X}_2 + e_{it}^{(m)} \right] \quad (10)$$

Die auf Basis des Modells vorhergesagten Verläufe für beide Reaktionsmaße sind in Abb. 4 verdeutlicht. Tabelle 5 zeigt die Ergebnisse des finalen Modells, darunter die Schätzungen, Standardfehler und Konfidenzintervalle der festen Parameter, die im finalen Modell enthalten sind. Für die Hautleitwertreaktion beträgt der geschätzte Intercept $0.065 \log \mu\text{S}$ ($SE = 0.029$). Er beschreibt die durchschnittliche Höhe der Reaktion auf den CS- genau in der Mitte des Akquisitionstrainings. Diese mittlere Reaktion variiert substanzial zwischen Versuchspersonen mit einer Standardabweichung von 0.13 rings um den Gesamt-Intercept (zufälliger Intercept). Abbildung 5 zeigt die Intercepts der Versuchspersonen als absolute Werte (5a) und als Abweichungen in Standardabweichungen skaliert vom Gesamt-Intercept (5c). Hier wird deutlich, dass für ca. ein Drittel der Versuchspersonen negative Intercepts vom Modell vorhergesagt werden. Die SCR werden durch lineare Wachstumskurven modelliert, wobei sie im Mittel über den Verlauf des Akquisitionstrainings um $0.018 \log \mu\text{S}$ abnehmen ($SE = 0.004$). Der Stimulustyp hat einen signifikanten Einfluss auf die Höhe der SCR. Die Reaktionen auf den CS+ sind durchschnittlich $0.058 \log \mu\text{S}$ höher als auf den CS- ($SE = 0.012$). Für die Interaktion zwischen Stimulustyp und linearem Term schätzt das Modell eine Slope von 0.009 ($\beta_{40}(1)$, $SE = 0.003$). Im Mittel unterscheidet sich die

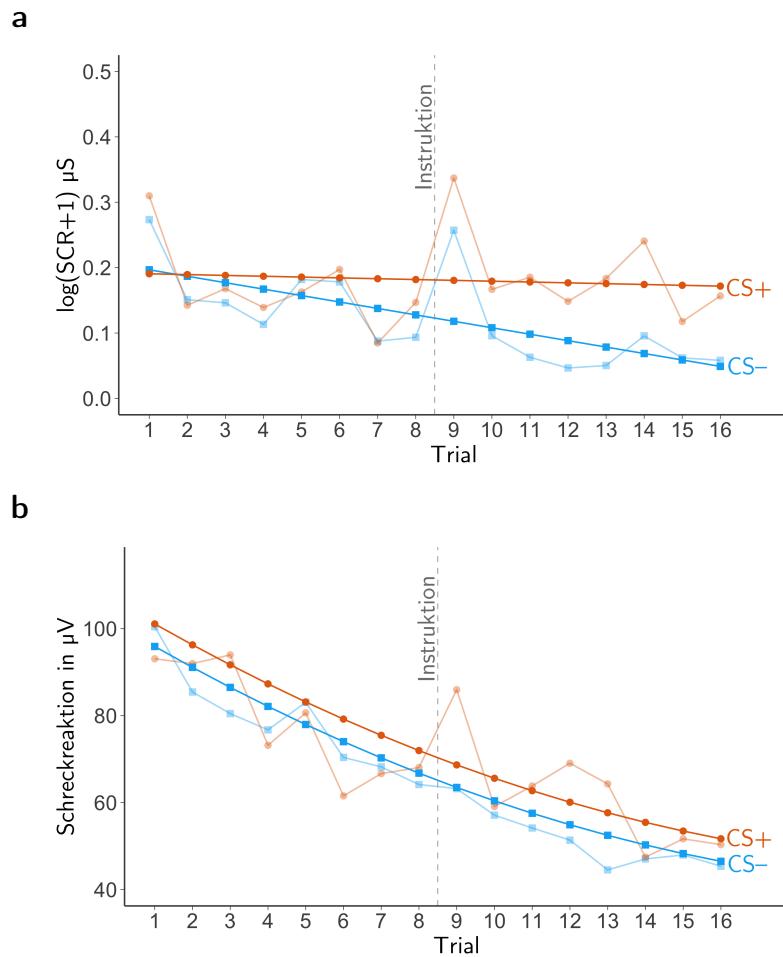


Abbildung 4. Vorhergesagte Verläufe von Hautleitwert- (a) und Schreckreaktionen (b) auf Basis des multivariaten Wachstumsmodells, verblasst sind die Rohverläufe.

lineare Änderungsrate für den CS+ also um $0.009 \mu\text{S}$ gegenüber der des CS-. Abbildung 4 zeigt die geschätzten Verläufe nach Stimulustyp getrennt. Hier wird ersichtlich, dass der Unterschied durch einen allmählichen Abfall der Reaktionen auf den CS- bei relativ gleichbleibender CS+ Reaktion zustande kommt. Der beschriebene Interaktionseffekt variiert ebenfalls zwischen Versuchspersonen ($SD = 0.006$). Die Abweichung der Interaktionsschätzungen für die Versuchspersonen vom festen Effekt $\beta_{40}^{(1)}$ in Abbildung 5e (zentriert bei null) zeigen keine eindeutige Tendenz in Relation zu den zufälligen Intercepts. Wenn auch nicht signifikant, deutet allerdings die geschätzte Korrelation zwischen Intercept und Interaktion ($r = -.35$) darauf hin, dass ein höherer Intercept tendenziell mit einem kleineren Unterschied in dem linearen Anstieg zwischen CS+ und CS- einhergeht.

Tabelle 5

Ergebnisse des finalen multivariaten Wachstumsmodells

| Feste Effekte | | | | | |
|-------------------|--------|----------------|----------------------------|-----------------|--------------|
| | Beta | SE | 95 % KI | t | p |
| SCR | | | | | |
| Intercept | 0.065 | 0.029 | [−0.0008; 0.121] | 2.24 | .025 |
| lin | −0.018 | 0.004 | [−0.027; −0.010] | −4.35 | < .001 |
| CS | 0.058 | 0.012 | [0.034; 0.083] | 4.72 | < .001 |
| CS×lin | 0.009 | 0.003 | [0.003; 0.014] | 3.00 | .003 |
| STR | | | | | |
| Intercept | 59.92 | 8.49 | [43.28; 76.57] | 7.06 | < .001 |
| lin | −3.30 | 0.35 | [−3.98; −2.60] | −9.37 | < .001 |
| quad | 0.11 | 0.05 | [0.01; 0.21] | 2.20 | .028 |
| CS | 5.19 | 1.87 | [1.53; 8.85] | 2.78 | .006 |
| Zufällige Effekte | | | | | |
| | SD | 95 % KI | Korrelationen ^a | | |
| | | | Intercept (SCR) | Intercept (STR) | CS×lin (SCR) |
| Intercept (SCR) | 0.130 | [0.10; 0.17] | — | | |
| Intercept (STR) | 49.98 | [39.8; 62.7] | .14 | — | |
| CS×lin (SCR) | 0.006 | [0.004; 0.009] | −.35 | −.04 | — |
| CS×lin (STR) | 1.341 | [0.99; 1.81] | −.05 | −.48* | −.29 |

Anmerkungen. Konfidenzintervalle, Signifikanz und p-Werte berechnet über Normalapproximationen der Verteilung der geschätzten Effekte. Die daraus angenähere *t*-Teststatistik hat $df = 2374$ Nenner-Freiheitsgrade.

^a bei * signifikant auf 5 % Niveau.

Für die Schreckreaktion beschreibt der geschätzte Intercept mit $59.92 \mu\text{V}$ ebenfalls die mittlere Reaktion zwischen dem 8. und 9. CS- Trial ($SE = 8.49$). Als zufällig modelliert variiert dieser Intercept mit einer geschätzten Standardabweichung von 49.98 zwischen Versuchspersonen. Abbildungen 5b und 5d zeigen die Versuchspersonen-Intercepts als Abweichungen vom mittleren Intercept. Hier liegen die geschätzten Intercept nur für zwei knapp unter Null. Der Koeffizient $\beta_{10}(2)$ mit der Schätzung $−3.3$ ($SE = 0.35$) reflektiert nun die momentane Änderungsrate der Schreckreaktion für den CS- in der Mitte des Akquisitionstrainings. Die quadratische Änderungsrate für den CS- ist positiv ($\beta_{20}(2) = 0.11$,

$SE = 0.05$). Daraus und in Verbindung mit Abbildung 4 wird deutlich, dass die Wachstumskurven für die Schreckreaktion als nach oben geöffnete, gestauchte Parabeln modelliert werden. Auch hier hat der Stimulustyp einen substanzialen Einfluss. Im Mittel sind die Schreckreaktionen der Versuchspersonen auf den CS+ $5.19 \mu\text{V}$ höher als auf den CS- ($SE = 1.87$). Allerdings ist die Interaktion – sprich der Unterschied zwischen CS+ und CS- in der linearen Änderungsrate – im Gegensatz zur SCR statistisch nicht bedeutsam. Über die Zeit hinweg wird der Unterschied zwischen CS+ und CS- daher als konstant modelliert (Abb. 4). Die Varianz, die Versuchspersonen in dieser Interaktion aufzeigen, ist wiederum Teil des Modells. Für den zufälligen Effekt zeigt sich, dass die Interaktionseffekte zwischen Versuchspersonen mit einer Standardabweichung von 1.34 variieren. In der Abbildung 5f lässt sich durch Hinzunahme der nach Größe geordneten Interceptschätzungen aus 5d ein Abwärtstrend vermuten, den die geschätzte Korrelation von $r = -.48$ zwischen Intercept und Interaktion bestätigt. Für Versuchspersonen mit höherer Reaktion auf den CS- in der Mitte des Trainings werden eher kleinere Unterschiede im linearen Anstieg zwischen CS+ und CS- vorhergesagt. Der Anhang E enthält unterstützend paarweise Streudiagramme zwischen den geschätzten zufälligen Effekten innerhalb und zwischen den beiden Reaktionsmaßen. In der multivariaten Betrachtung zeigt sich ein positiver, kleiner und nicht signifikanter Zusammenhang ($r = .14$). Bei den beiden Interaktionseffekten $CS \times lin$ schätzt das Modell eine negative Korrelation von $r = -.29$ zwischen den abhängigen Variablen, die allerdings wiederum nicht signifikant ist. Tendenziell geht damit ein größerer Interaktionsterm auf dem einen Reaktionsmaß mit einer kleineren Interaktion auf dem anderen Maß einher.

4.4 Prüfen der Modellannahmen

Alle Abbildungen für die Prüfung der Modellannahmen sind in Anhang F enthalten. Das Streudiagramm zwischen Residuen und angepassten Werten für die höchste Gruppierungsebene zeigt für die SCR einen klaren Bodeneffekt bei kleinen angepassten Werten. Dies scheint eine wesentliche Einschränkung des Modells zu sein, die durch die verhältnismäßig hohe Anzahl an Nullreaktionen und Reaktionen nahe null zustande kommt. Residuen sind bei diesen Reaktionen zwangsläufig positiv, sodass hier eine Varianzeinschränkung vorliegt. Für die Schreckreaktionen ist diese Verletzung der Varianzhomogenität gering ausgeprägt. Tendenziell ist die Residuenvarianz für kleine angepasste Werte (≤ 50) etwas geringer, für die Werte ab 50 lässt sich jedoch eine zufällige gleichmäßige Streuung beobachten. Für die Normalverteilungsannahme der Residuen auf Ebene der Versuchsperson wird eine visuelle Analyse der Häufigkeitsverteilung durch Histogramme und Q-Q-Diagramme für beide abhängigen Variablen durchgeführt. Sowohl die Histogramme als auch die Q-Q-Diagramme zeigen eine substanziale Abweichung von der Normalverteilung für die Residuen. Für die Hautleitwertreaktionen ist diese stärker als für die Schreckreaktion. Die Häufigkeitsverteilungen deuten daraufhin, dass die Residuen für beide Reaktionsmaße zum Teil starke Ausreißer aufweisen. Außerdem scheint die ausgeprägte Spitze der Verteilung um Null ein Hauptmerkmal für die Abweichung von der Normalverteilung zu sein.

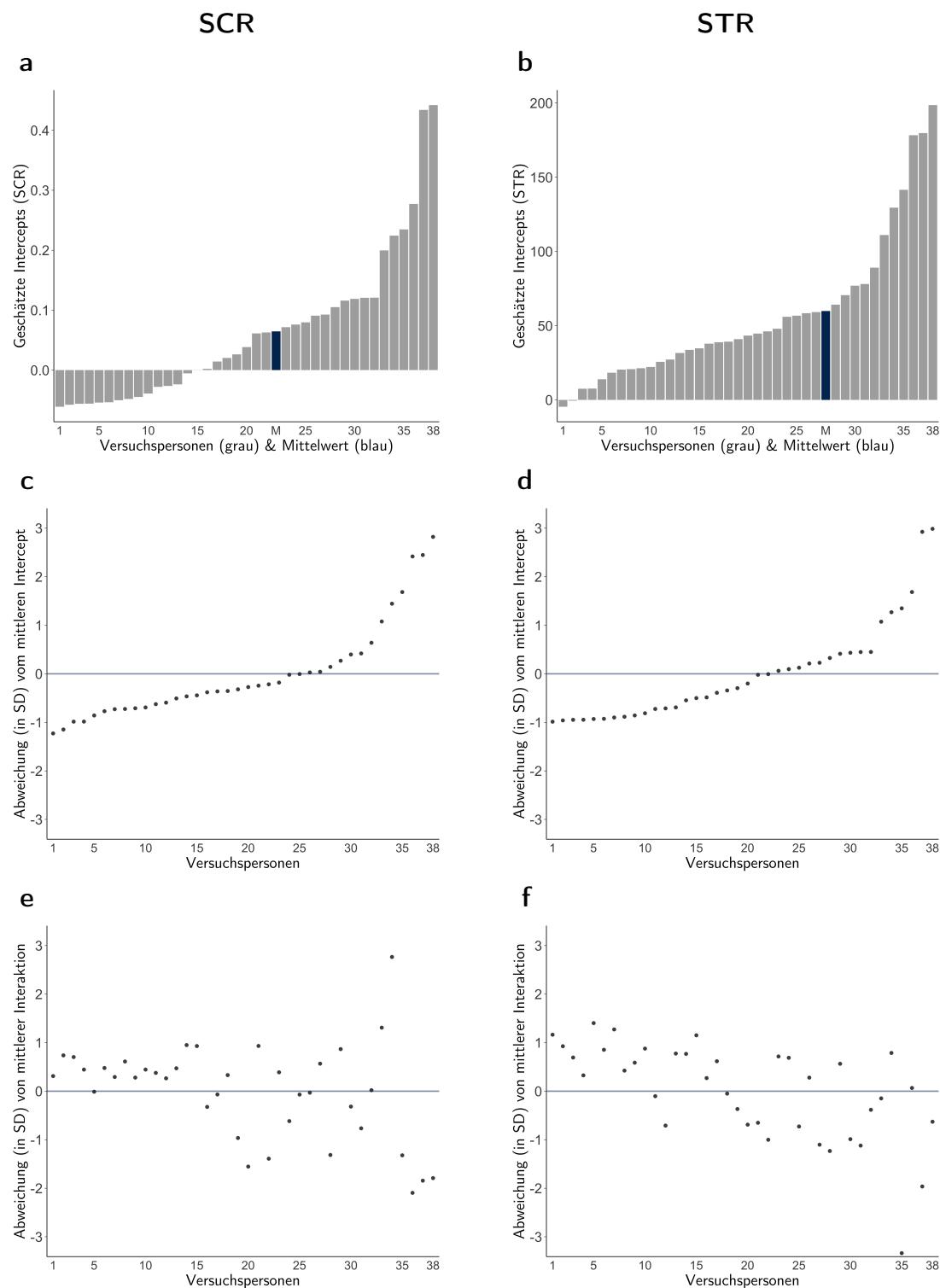


Abbildung 5. (a) und (b): Versuchspersonen-Intercepts des Modells mit dem mittleren (globalen) Intercept in blau. (c) und (d): Abweichungen dieser Intercepts vom mittleren Intercept (zentriert bei null) skaliert in Standardabweichungen. (e) und (f): Abweichungen der Interaktionsschätzungen für die einzelnen Versuchspersonen (skaliert in SD). Abbildungen motiviert durch Meteyard und Davies (2020).

5 Diskussion

Das Ziel dieser Arbeit war der Vergleich der Verläufe von Schreck- und Hautleitwertreaktionen in einer Furchtakquisition durch ein multivariates Mehrebenenmodell. Für diesen Zweck wurde ein Modell aufgestellt, um die Wachstumskurven der Indikatoren zu beschreiben und das differentielle Furchtlernen im Akquisitionstraining zu untersuchen. Die wichtigsten Erkenntnisse lassen sich wie folgt zusammenfassen: Zunächst kann aus den Ergebnissen kein Furchtlerneffekt, sondern lediglich ein Instruktionseffekt geschlussfolgert werden. Die Modellergebnisse sind damit zwar theoretisch und empirisch erwartbar, stellen aber für die Furchtlernttheorie keine wesentlich neuen Erkenntnisse bereit. Trotz alledem bieten Modelle wie diese substanzielle Vorteile und Chancen, in zukünftiger Forschung umfassendere Einblicke in die Prozesse des Furchtlernens zu gewinnen.

5.1 Interpretation und Einordnung

Im Fokus der Ergebnisse steht das aufgestellte Modell, aus dem zwei grundlegende Informationen über die Unterschiedlichkeit der Reaktionsvariablen hervorgehen. Zunächst wurden die mittleren Verläufe beider Indikatoren über das Training hinweg modelliert. Während die SCR durch einen linearen Verlauf am besten beschrieben wurden, erwies sich für die Schreckreaktionen eine quadratische Modellierung als adäquater. Bei der Betrachtung der Rohverläufe wird deutlich, dass beide Reaktionsvariablen Habituationssprozessen unterliegen. Das heißt, dass die absoluten (bzw. log-transformierten) mittleren Reaktionen über die Zeit hinweg abnehmen. Zusätzlich zur linearen Abnahme wurde bei der Schreckreaktion eine Linkskrümmung modelliert. Diese mag zwar signifikant zum Modell beitragen, hat aber im praktischen Sinne wenig Bedeutung für den Verlauf der Habituation, da der Grad der Krümmung gering ist gemessen an der absoluten Größe der Reaktionen. Habituationen der Hautleitwert- und Schreckreaktionen insgesamt sind eine übliche Beobachtung (z. B. Bradley et al., 1993) und lassen sich durch die verwendeten Modellparameter angemessen beschreiben.

Der zweite – und bedeutsamere – Unterschied zeigt sich im Modell durch die Interaktion der linearen Zeitkomponente Trial und dem Stimulustyp. Auffällig ist hier, dass der CS+/CS- Unterschied über die Zeit hinweg bei der SCR als signifikanter Einfluss ins Modell einging, bei der Schreckreaktion jedoch nicht. Die Interaktion kommt bei der SCR dadurch zustande, dass die Reaktionen auf den CS- im Mittel abnehmen, während die Reaktionen auf den CS+ als nahezu konstant über die Trials modelliert werden. Diese Diskrimination spiegelt sich auch in den zeitpunktweisen Mittelwertvergleichen wider. Hier zeichnet sich bei den SCR in der zweiten Hälfte des Akquisitionstrainings zu Teilen ein substanzialer Mittelwertunterschied ab (für Trial 10 bis 14). Für die Schreckreaktion hat die Interaktion währenddessen nicht signifikant zur Anpassungsgüte beigetragen. Zwar ist im Mittel die Reaktion auf den CS+ größer, allerdings ist ein Diskriminationseffekt über die Zeit nicht deutlich nachweisbar. Auch wenn in der zweiten Trainingshälfte eine Unterschiedlichkeit der Reaktionen auf die Stimuli zu zwei Zeitpunkten gegeben ist, so verdeutlichen die Modellergebnisse, dass ein Differenzierungseffekt über die Zeit fern von Eindeutigkeit bzw. von Unsicherheit geprägt zu sein scheint. Der Vergleich der standardisierten Differenzen zwischen den Reaktionsmaßen bestätigt, dass auf dem Maß der SCR im Durchschnitt stärker zwischen CS+ und CS- diskriminiert wird als auf der Schreckreaktion.

Im Großen und Ganzen liefern diese Ergebnisse keine Nachweise für ein differentielles Furchtlernen. Eine Differenzierung, die sich in unterschiedlicher Stärke bei den beiden Reaktionsvariablen zeigt, tritt ausschließlich erst *nach* dem Zeigen der Instruktion (zwischen dem achten und neunten Trial des jeweiligen CS) auf. Das lässt vermuten, dass die beobachteten Differenzierungen nicht auf eine über die Zeit erworbene Assoziation zurückgeführt werden können, sondern stattdessen durch die Kontingenzinstruktion bedingt sind. Nach einer aufklärenden Instruktion spricht man bei einer Differenzierung zwischen CS+ und CS- von einem Furchtausdruck (Lonsdorf et al., 2017) und nicht mehr von Furchtlernen. Vergangene Studien haben ähnliche Instruktionseffekte bereits in Furchtkonditionierungsstudien gezeigt. In der bereits referenzierten Studie von Weidemann et al. (2016) beispielsweise wurde eine Gruppe vollständig über die CS-US-Kontingenz aufgeklärt und zeigte sofort differentielle Schreckreaktionen, während eine uninstruierte Gruppe keinerlei CR erwarb. Bei einer dritten Gruppe, die nur über das Vorhandensein, nicht aber die Art der Kontingenz instruiert wurde, konnte über den Verlauf des Trainings ein zunehmender Furchtlerneffekt beobachtet werden. Die Ergebnisse dieser Arbeit sind in Einklang zu bringen mit der Idee, dass eindeutige Kontingenzinstruktionen die CR verstärken oder zum Teil sogar auslösen können.

Vergleicht man qualitativ den Instruktionseffekt auf den beiden Reaktionsvariablen, fällt auf, dass der Furchtausdruck bei der Hautleitwertreaktion sehr eindeutig gezeigt wird. Die Evidenz für eine Differenzierung bei der Schreckreaktion ist aber weniger eindeutig und schwächer ausgeprägt. Hier ist auffällig, dass trotz einer vollständigen Instruktion über die CS-US-Kontingenzen in der Schreckreaktion keine beständige Differenzierung zwischen CS+ und CS- beobachtet wird. Ein Blick auf die zufälligen Effekte der Schreckreaktion macht deutlich, dass ein höherer Intercept (Reaktion auf den CS- in der Mitte des Trainings) mit einem kleineren Interaktionseffekt dieser Versuchsperson einhergeht. Die Tendenz, dass eine höhere Reaktivität in der Mitte des Trainings mit einem geringeren Differenzierungsgrad zusammenhängt, ist möglicherweise interessant für zukünftige Forschung. Der Effekt der Instruktion schlägt sich auch in den Korrelationen zwischen den abhängigen Variablen nieder. Eine substanzial positive Korrelation zwischen den CS+/CS- Differenzwerten der beiden Maße wird nur direkt nach der Instruktion beobachtet. Eine größere Differenzierung in der SCR geht in diesem Messzeitpunkt also mit einer größeren Differenzierung in der Schreckreaktion einher. Die Instruktion erzielt also – einem Beschleuniger gleich – unmittelbar eine (mehr oder weniger starke) Diskrimination zwischen CS+ und CS- über die nachfolgenden Trials. Überraschend ist, dass bei der SCR im Trial nach der Instruktion zusätzlich zu dem erwarteten Anstieg beim CS+ auch die Reaktion auf den CS- deutlich ansteigt.

Eine mögliche Erklärung für die beobachteten Ergebnisse ist der Einfluss des Untersuchungsdesigns. Insgesamt scheinen die Stimuli vor der Instruktion wenig prädiktive Aussagekraft zu besitzen. Die niedrige Verstärkungsrate von 50 % und die dadurch geringe Anzahl an US könnten die Ursache dieser Unsicherheit sein. In der Literatur werden experimentelle Situationen als *schwach* bezeichnet, wenn sie durch große Komplexität, Unsicherheit und Ambiguität gekennzeichnet sind (Lissek et al., 2006). Es wird vermutet, dass diese schwachen Situationen sensitiver gegenüber interindividuellen Unterschieden in den Reaktionen sind (Lonsdorf et al., 2017). Die hohe postinstruktionale Reaktion auf den CS- bei der SCR und die Uneindeutigkeit in der Differenzierung bei der Schreckreaktion können Hinweise auf eine von Unsicherheit geprägte Situation für die Versuchspersonen sein. Außerdem zeigen die signifikanten zufälligen Effekte ebenfalls, dass Versuchspersonen sich in ihrem Differenzierungsgrad über die Zeit wesentlich unterscheiden.

Ein weiterer potentieller Einflussfaktor knüpft an die geringe Verstärkerrate an: Im Untersuchungsdesign wurden die Versuchspersonen auf einer von vier verschiedenen Trialreihenfolgen getestet. In zwei der vier Reihenfolgen wird der erste CS+ nicht verstärkt. Die erste Präsentation einer CS-US-Paarung kann somit bei einigen erst zum dritten oder vierten Trial erfolgen. Diese späte Präsentation birgt das Potential, das kognitive Lernen zu erschweren, und kann in unterschiedlichen Lernkurven für die verschiedenen Gruppen resultieren (Lonsdorf et al., 2017). Das Phänomen einer verspäteten oder abgeschwächten Assoziationsbildung in Folge einer unverstärkten Exposition mit dem CS+ ist unter dem Namen *latente Inhibition* bekannt. Es ist empirisch gezeigt worden, dass ein konsequenzloses Zeigen des späteren CS+ das Auftreten von differentiellen SCR in einer Furchtakquistion verzögert (z. B. Siddle et al., 1985). In ihrem Review zur latenten Inhibition beim Menschen fassen Vaitl und Lipp (1997) zusammen, dass der Effekt der latenten Inhibition umso stärker ist, je mehr Präexpositionstests gezeigt werden. Beim elektrodermalen Reaktionsmaß sei ein expliziter Inhibitionseffekt im differentiellen Design erst nach mehr als 15 Trials zu beobachten. Dieser Wert überschreitet zwar weit die Anzahl an präexponierten CS in der hier untersuchten Studie, allerdings kann ein Einfluss der Trialreihenfolge nicht explizit ausgeschlossen werden. Ebenso führt die Analyse der Gelegenheitsstichprobe dazu, dass eine Ausbalancierung möglicher Reihenfolgeeffekte als kritisch zu betrachten ist.

Im Endeffekt spiegeln die Ergebnisse dieser Analyse die Besonderheiten des experimentellen Paradigmas wider, liefern aber keine Erklärungen für zugrundeliegende Lernprozesse oder das Zusammenspiel der beiden Faktoren in furchtlerntheoretischer Hinsicht. Im Sinne der Zwei-Prozess-Theorie, die von einer kognitiven und einer automatischen Ebene des Furchtlernens ausgeht (z. B. Hamm & Vaitl, 1996; Öhman & Mineka, 2001), wäre die Instruktion als Einfluss auf die kognitive Ebene zu verstehen. Unter der Annahme, dass die SCR ein Indikator dieser Ebene ist, sollte die Aufklärung über die CS-US-Kontingenz vor allem mit einem verstärkenden Effekt auf die differentielle SCR einhergehen. Mit dieser Annahme lassen sich die Ergebnisse dieser Arbeit grundlegend vereinen. Währenddessen besteht auf einer automatischen Ebene – mit der Schreckreaktion als Indikator – im Mittel auch nach der Instruktion Unsicherheit. Die grundlegende Einschränkung jeglicher Schlussfolgerungen aus diesen Daten ist das vollkommene Fehlen eines differentiellen Furchtlernens. Auch wenn die Modellunterschiede zwischen SCR und Schreckreaktion und der Einfluss der Instruktion potentiell vereinbar mit einer Zwei-Prozess-Theorie sind, macht es der fehlende Lerneffekt unmöglich, Aussagen über Furchtlernprozesse zu tätigen. Die Daten können in diesem Hinblick keine Indizien für die eine oder die andere Theorie zum Furchtlernen bereitstellen. Dieser Punkt wirkt sich auch auf die Einordnung und Beantwortung der leitenden Fragestellungen dieser Arbeit aus. Für die erste Frage untersuchte ich, ob sich unterschiedliche Verläufe in Hautleitwert- und Schreckreaktion abzeichnen, wenn beide Maße im Modell gleich behandelt werden. Zusammenfassend sind beide Verläufe durch Habituationssprozesse gekennzeichnet, wobei der Unterschied der polynomialem Modellierung als vernachlässigbar interpretiert wird. Qualitativ ist ersichtlich, dass die Reaktionen auf den CS– über das Training hinweg bei beiden Maßen abnehmen, während bei dem CS+ nur die Schreckreaktionen als kontinuierlich fallend modelliert werden. Unterschiedliche mittlere Verläufe zeichnen sich dadurch im Modell durchaus ab, liefern aber wie gesagt keinen theoretischen Beitrag. Hier knüpft die zweite Fragestellung an, die beinhaltet, ob sich der Einfluss des Stimulus Typs bedeutsam zwischen den abhängigen Variablen unterscheidet. In der Tat konnten diskrete Unterschiede in der CS+/CS– Diskrimination beobachtet werden, die oben bereits detailliert dargelegt wurden. Allerdings wurde auch gezeigt, dass die beobachtete

Diskrimination der Stimuli auf Reaktionsebene nicht als Indikator einer gelernten Assoziation, sondern als Furchtausdruck in Folge einer Kontingenzinstruktion zu verstehen ist (Lonsdorf et al., 2017). Damit gibt es zwar Unterschiede im Einfluss des Stimulustyps auf SCR und Schreckreaktionen, jedoch keine Möglichkeit, Rückschlüsse auf verschiedene Furchtlernprozesse zu schließen.

Trotz der geringen theoretischen Aussagekraft liefert das Modell durch seinen explorativen Charakter spannende Ein- und Ausblicke. Der zufällige Effekt der Interaktion $lin \times CS$ verdeutlicht, dass bei den Versuchspersonen signifikante Unterschiede in der CS+/CS- Differenzierung über die Zeit hinweg auftreten – und das auf beiden Reaktionsvariablen. Versuchspersonen unterscheiden sich also darin, ob und wie stark sie Furchtlernen bzw. einen Furchtausdruck zeigen. Es ist bekannt, dass Individuen in ihrer physiologischen Reaktivität voneinander abweichen und ihnen eigene physiologische Antwortmuster aufweisen (z. B. Engel, 1960; Lacey & Lacey, 1958). Herkömmliche Auswertungsmethoden wie die rmANOVA kommen mit der Verwendung von Gruppenmittelwerten für diese Idiosynkrasie nicht auf (z. B. Kristjansson et al., 2007), Mehrebenenmodelle jedoch können diese Variationen systematisch analysieren. Der explorative Charakter dieser Arbeit zeigt, dass diese möglicherweise durch die schwache Lernsituation hervorgehobenen interindividuellen Unterschiede sich durch das Modell erfassen und beschreiben lassen. Im multivariaten Teil des Modells – nämlich die Korrelationen zwischen den zufälligen Effekten verschiedener Reaktionsmaße – ist zwar kein Zusammenhang statistisch signifikant. Im Allgemeinen lassen sich hingegen aber Tendenzen ableiten, dass z. B. eine höhere Responsivität der SCR in der Mitte des Trainings *potentiell* mit einer höheren Responsivität der Schreckreaktion einhergeht (Korrelation der zufälligen Intercepts). Im Folgenden werden Limitationen des Designs und des Modells dargelegt sowie ein Ausblick für die Anwendung von MEM gegeben.

5.2 Limitationen und Ausblick

Bis hierhin wurde deutlich, dass das Ausbleiben eines Furchtlerneffekts zu großen Teilen durch das Studiendesign bedingt ist. Dass die Daten für diese Arbeit im Rahmen eines Großprojekts mit anderer Zielsetzung als die der hiesigen Fragestellung erhoben worden sind, ist damit eine wichtige Limitation. Um konkreter Aussagen über Unterschiede der beiden untersuchten Reaktionsmaße im Furchtlernen treffen zu können, schlage ich eine Untersuchung mit folgenden Designoptimierungen vor: In zukünftigen Studien könnten Hautleitwert- und Schreckreaktionen in einem uninstruierten Akquisitionstraining mit höherer Verstärkerrate (z. B. 75 %) erhoben werden. Akquisitionsdaten dieser Art wären aufschlussreich, um der eigentlichen Fragestellung dieser Arbeit näher auf den Grund zu gehen. Bei unterschiedlichen Trialreihenfolgen könnte in künftiger Forschung auch der Faktor der Präexposition mit dem späteren CS+ als Modellvariable aufgenommen werden, um mögliche Einflüsse auf die Furchtlernverläufe explizit zu untersuchen. Im Sinne der Untersuchung, ob distinkte Furchtlernprozesse auf Reaktionsebene unterscheidbar sind, könnte das Hinzufügen von expliziten Bewusstseinsmaßen eine Option darstellen. Subjektive Maße zum Kontingenzbewusstsein, die verbales Wissen über die CS-US-Kontingenz operationalisieren und erfassen, können im Mehrebenenmodellrahmen die Forschung erweitern. Allerdings ist nicht außer Acht zu lassen, dass die Hinzunahme von abhängigen Variablen, die die Versuchsperson zusätzlichen Reizen aussetzen (wie der Schreckkreis oder Befragungselemente), potentiell auch das Furchtlernen beeinflussen (am Beispiel des Schreckkreises: Sjouwerman et al., 2016). Hier könnten zum Beispiel zwei verschiedene Gruppen mit den gleichen Akquisitionsparametern erhoben

werden, von denen jedoch bei einer nur die Schreck- und die Hautleitwertreaktion erfasst werden, bei der anderen aber zusätzlich Befragungselemente zum Kontingenzbewusstsein im Training eingebettet sind. So ließe sich der Einfluss der *Befragung* immerhin mit modellieren. Generell gilt, dass in der Zukunft multivariate Wachstumsmodelle genutzt werden können, um verschiedene Furchtindikatoren in ein einziges Modell zu integrieren. Der Vergleich nicht nur von Hautleitwert- und Schreckreaktionen, sondern auch von Maßen wie der Herzrate und der Pupillendilatation kann potentiell Einblick in die Interdependenzen des peripherphysiologischen Furchtsystems verschaffen. Korrelationen zwischen den zufälligen Effekten bieten die Möglichkeit, geteilte Variation zwischen den abhängigen Variablen zu erfassen (Curran et al., 2012; MacCallum et al., 1997). Eine Einschränkung auf Modellebene ist hierbei aber grundlegend, dass eine steigende Komplexität des Modells auch mit der Notwendigkeit einhergeht, eine größere Stichprobe zu erheben.

Beim Einsatz von MEM gibt es – wie immer in der empirischen Wissenschaft – eine Reihe von Entscheidungen, die bis zum finalen Ergebnisbericht getroffen werden müssen. Zudem gibt es erst seit kurzem erste Schritte in Richtung einer Best-Practice für die Anwendung von MEM in der psychologischen Forschung (Meteyard & Davies, 2020). Ich bin mir darüber bewusst, dass die Entscheidungen, die im Zuge der Modellbildung dieser Arbeit getroffen wurden, nur einen Weg unter vielen möglichen kennzeichnen. Angesichts dessen können meine Entscheidungen einen Einfluss auf die gefundenen Ergebnisse haben. Zunächst ist hier die Aufbereitung der Daten anzubringen. Während sich bei der SCR für eine logarithmische Transformation entschieden wurde, sind die Schreckreaktionen als Rohdaten in die Analyse eingegangen. Die Richtlinien von Blumenthal et al. (2005) verdeutlichen, dass bis jetzt keine Best-Practice-Lösung für die Transformation und Standardisierung von Schreckreaktionen festgelegt wurde. Bradford et al. (2015) verglichen in Hinblick darauf rohe, standardisierte und prozentuale Quantifizierungen von Schreckreaktionen mit der Schlussfolgerung, dass die ersten beiden vergleichbare Ergebnisse hervorbringen. Für die SCR wird demgegenüber eine Transformation (log oder Quadratwurzel) empfohlen, um die Reaktionsdaten zu normalisieren (Boucsein et al., 2012). Diese Empfehlungen bewegen sich auf einer univariaten Analyseebene und es bleibt zunächst offen, ob für multivariate Auswertungen im Sinne der Vergleichbarkeit anders gehandelt werden müsste. Die Frage der Vergleichbarkeit spielt auch in der Interpretation der Modellergebnisse eine entscheidende Rolle: Der tatsächliche Vergleich der beiden abhängigen Variablen in dieser Arbeit beruht hauptsächlich auf qualitativen Modellierungsunterschieden. Um Unterschiede in *Regressionskoeffizienten* zwischen zwei abhängigen Variablen auf unterschiedlichen Skalen zu quantifizieren, könnten standardisierte Regressionskoeffizienten in Frage kommen. Vor- und Nachteile verschiedener Standardisierungsverfahren zu diskutieren überschreitet den Rahmen dieser Arbeit, allerdings birgt dieser Ausblick gegebenenfalls wichtige Implikationen für multivariate Verfahren im Furchtkonditionierungskontext und sollte in zukünftige Forschung eingebunden werden. Im Kontext der Reaktionsaufbereitung muss auch hervorgehoben werden, dass einer der wesentlichen Vorteile der Schreckreaktion – nämlich ihre Erfassung ohne die Präsentation experimentell relevanter Stimuli – in dieser Auswertung ungenutzt blieb. Im Sinne der Gleichbehandlung wurden die Schreckreaktionen aus den ITI als zusätzliche Informationsquelle außer Acht gelassen, was potentiell als Einschränkung der Datenqualität wertbar ist.

Im Kontext des Modellbaus ist die Auswahl der Kovarianzstruktur anzusprechen, die als unstrukturiert (die Standardeinstellung im *n1me*-Paket) gewählt wurde. Bei Wachstumskurven wird manchmal

eine autoregressive Korrelationsstruktur angenommen (Bagiella et al., 2000). Um womöglich auch bei kleineren Datensätzen Konvergenz zu erzielen, wären Annahmen über eine passendere Kovarianzstruktur eine relevante Erweiterung. Eine wichtige Limitation des gebauten Modells ist die Verletzung der Normalverteilungsannahme der Residuen. Maas und Hox (2004) konnten zeigen, dass diese Verletzung auf der höchsten Gruppierungsebene wenig bis gar keinen Effekt auf die Schätzung der festen Effekte, jedoch aber einen Einfluss auf die Schätzung der Standardfehler der zufälligen Effekte hat. An den Residuen ist hier auch ersichtlich, dass das Modell nicht gut mit Nullreaktionen umgehen kann. Es gibt einen klaren Bodeneffekt (v. a. bei der SCR, wo der Anteil an Nullreaktionen größer ist) und zum Teil sagt das Modell negative Reaktionen vorher. Vor allem bei den vorhergesagten Intercepts für die Hautleitwertreaktion wird dieser Bias deutlich. Alternative Ansätze zur klassischen Tiefpunkt-bis-Hochpunkt Definition der SCR können hier einen wertvollen Beitrag leisten. In der traditionellen Reaktionsdefinition führen überlagerte SCR zu einer Unterschätzung der Reaktionsamplitude (Benedek & Kaernbach, 2010). Mittels eines Entfaltungsansatzes schlagen Benedek und Kaernbach (2010) die Aufteilung von elektrodermaler Aktivität in phasische und tonische Komponenten vor (CDA) und zeigen, dass die dadurch verbesserte Tiefpunkt-bis-Hochpunkt Definition zu weniger Nullreaktionen und insgesamt höheren Amplituden der SCR führt. In zukünftigen Studien könnten Bodeneffekte mit diesem Wertungsverfahren minimiert werden. Analysetechnisch seien bei einer solchen Einschränkung ebenso Bootstrapping-Verfahren vorgeschlagen, um reliablere Schätzungen der Parameter zu erhalten (Hox et al., 2018). Allerdings benötigen diese eine substanzial größere Stichprobe und gehen inhaltlich über den Rahmen dieser Arbeit hinaus.

Leider ist das Modell nur mit einer kleinen Teilmenge der möglichen zufälligen Effekte konvergiert. Eine Ursache dafür kann die bereits angesprochene – gemessen an der Modellkomplexität kleinen – Stichprobe sein (Meteyard & Davies, 2020). Vollständigere Einblicke in interindividuelle Unterschiede der Versuchspersonen auf den Zeitkomponenten ließen sich vielleicht bei einer größeren Stichprobe (mind. ≥ 50) gewinnen. Potentiell bietet sich für zukünftige Forschung an, in einem explorativen Ansatz bereits erhobene Daten mit MEM zu reanalysieren, um spätere konfirmatorische Studien anzustoßen.

Abschließend wird trotz bestehender Limitationen deutlich, dass Mehrebenenmodelle nicht zu unterschätzende Vorteile für die Auswertung solch komplexer Daten haben, wie sie in Furchtkonditionierungsstudien zu finden sind. Zum Beispiel diskutieren Ney et al. (2018) in ihrem Artikel verschiedene Probleme und Limitationen in der herkömmlichen statistischen Auswertung von Furchtextinktionsdaten. Ihre Schlussfolgerung legt dar, dass statistische Modellierung (z. B. durch MEM realisiert) eine wichtige Weiterentwicklung ist, um unter dem Schatten der Replikationskrise die Reliabilität und Validität von Ergebnissen und Aussagen zu erhöhen (Ney et al., 2018). In diesem Sinne bin ich der Meinung, dass auch wenn die in dieser Arbeit präsentierten Ergebnisse lerntheoretisch kaum neue Informationen bereitstellen, so doch die vorgestellte Methodik vielleicht Forscherinnen und Forscher zu motivieren vermag, im Kontext der Furchtkonditionierung konkreten multivariaten Fragestellungen nachzugehen. Durch die simultane Auswertung von Hautleitwert- und Schreckreaktionen können zukünftige konfirmatorische Studien einen Beitrag zum Diskurs um die Ein- und Zwei-Prozess-Theorien leisten. Letztendlich birgt der Einsatz integrativer Auswertungsverfahren das Potential und die Hoffnung, bestehende Limitationen langfristig zu überwinden, unsere Theorien rund um das Furchtlernen weiterzuentwickeln und seine Mechanismen zu durchdringen.

Literaturverzeichnis

- Bagiella, E., Sloan, R. P. & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, 37, 13–20. <https://doi.org/10.1111/1469-8986.3710013>
- Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00328>
- Benedek, M. & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1), 80–91. <https://doi.org/10.1016/j.jneumeth.2010.04.028>
- Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!* The Guilford Press.
- Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V. & Boxtel, A. V. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, 42(1), 1–15. <https://doi.org/10.1111/j.1469-8986.2005.00271.x>
- Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., ... Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, 49(8), 1017–1034. <https://doi.org/10.1111/j.1469-8986.2012.01384.x>
- Box, G. E. P. & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211–243. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
- Bradford, D. E., Starr, M. J., Shackman, A. J. & Curtin, J. J. (2015). Empirically based comparisons of the reliability and validity of common quantification approaches for eyeblink startle potentiation in humans. *Psychophysiology*, 52(12), 1669–1681. <https://doi.org/10.1111/psyp.12545>
- Bradley, M. M. & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Hrsg.), *Cognitive neuroscience of emotion* (S. 242–276). Oxford University Press.
- Bradley, M. M., Moulder, B. & Lang, P. J. (2005). When good things go bad: The reflex physiology of defense. *Psychological Science*, 16(6), 468–473. <https://doi.org/10.1111/j.0956-7976.2005.01558.x>
- Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Bradley, M. M., Lang, P. J. & Cuthbert, B. N. (1993). Emotion, novelty, and the startle reflex: Habituation in humans. *Behavioral Neuroscience*, 107(6), 970–980. <https://doi.org/10.1037/0735-7044.107.6.970>
- Brauer, M. & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3), 389–411. <https://doi.org/10.1037/met0000159>
- Brown, J. S., Kalish, H. I. & Farber, I. E. (1951). Conditioned fear as revealed by magnitude of startle response to an auditory stimulus. *Journal of Experimental Psychology*, 41(5), 317–328. <https://doi.org/10.1037/h0060166>
- Chang, W. (2014). *extrafont: Tools for using fonts*. <https://CRAN.R-project.org/package=extrafont>

- Constantinou, E., Purves, K. L., McGregor, T., Lester, K. J., Barry, T. J., ... Eley, T. C. (2021). Measuring fear: Association among different measures of fear learning. *Journal of Behavior Therapy and Experimental Psychiatry*, 70, 101618. <https://doi.org/10.1016/j.jbtep.2020.101618>
- Curran, P. J., McGinley, J. S., Serrano, D. & Burfeind, C. (2012). A multivariate growth curve model for three-level data. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf & K. J. Sher (Hrsg.), *APA handbook of research methods in psychology, Vol. 3. Data analysis and research publication* (S. 335–358). American Psychological Association. <https://doi.org/10.1037/13621-017>
- Curran, P. J., Obeidat, K. & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, 11(2), 121–136. <https://doi.org/10.1080/15248371003699969>
- Cuthbert, B. N., Bradley, M. M. & Lang, P. J. (1996). Probing picture perception: Activation and emotion. *Psychophysiology*, 33(2), 103–111. <https://doi.org/10.1111/j.1469-8986.1996.tb02114.x>
- Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience*, 15(1), 353–375. <https://doi.org/10.1146/annurev.ne.15.030192.002033>
- Davis, M. (2006). Neural systems involved in fear and anxiety measured with fear-potentiated startle. *American Psychologist*, 61(8), 741–756. <https://doi.org/10.1037/0003-066x.61.8.741>
- Dawson, M. E., Schell, A. M. & Filion, D. L. (2007). The electrodermal system. In J. Cacioppo, L. Tassinary & G. Berntson (Hrsg.), *Handbook of psychophysiology* (3. Aufl., S. 159–181). Cambridge University Press.
- Dawson, M. E. & Biferno, M. A. (1973). Concurrent measurement of awareness and electrodermal classical conditioning. *Journal of Experimental Psychology*, 101(1), 55–62. <https://doi.org/10.1037/h0035524>
- De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, 37(1), 1–20. <https://doi.org/10.3758/lb.37.1.1>
- De Houwer, J. (2014). Why a propositional single-process model of associative learning deserves to be defended. In J. Sherman, B. Gawronski & Y. Trope (Hrsg.), *Dual-process theories of the social mind* (S. 530–541). The Guilford Press.
- Derogatis, L. R. & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13(3), 595–605. <https://doi.org/10.1017/s0033291700048017>
- Engel, B. T. (1960). Stimulus-response and individual-response specificity. *Archives of General Psychiatry*, 2(3), 305–313. <https://doi.org/10.1001/archpsyc.1960.03590090061010>
- Grillon, C. (2002). Associative learning deficits increase symptoms of anxiety in humans. *Biological Psychiatry*, 51(11), 851–858. [https://doi.org/https://doi.org/10.1016/S0006-3223\(01\)01370-1](https://doi.org/https://doi.org/10.1016/S0006-3223(01)01370-1)
- Grillon, C. & Baas, J. (2003). A review of the modulation of the startle reflex by affective states and its application in psychiatry. *Clinical Neurophysiology*, 114(9), 1557–1579. [https://doi.org/10.1016/s1388-2457\(03\)00202-5](https://doi.org/10.1016/s1388-2457(03)00202-5)
- Grimm, K. J., Ram, N. & Estabrook, R. (2017). *Growth modeling: Structural equation and multilevel modeling approaches*. The Guilford Press.
- Hamm, A. O. & Vaitl, D. (1996). Affective learning: Awareness and aversion. *Psychophysiology*, 33(6), 698–710. <https://doi.org/10.1111/j.1469-8986.1996.tb02366.x>

- Hamm, A. O. & Weike, A. I. (2005). The neuropsychology of fear learning and fear regulation. *International Journal of Psychophysiology*, 57(1), 5–14. <https://doi.org/10.1016/j.ijpsycho.2005.01.006>
- Hamm, A. O., Weike, A. I. & Melzig, C. A. (2006). Wenn Furcht und Angst entgleisen: Zur Pathologie des menschlichen Defensivsystems. *Psychologische Rundschau*, 57(3), 154–164. <https://doi.org/10.1026/0033-3042.57.3.154>
- Hitchcock, J. & Davis, M. (1986). Lesions of the amygdala, but not of the cerebellum or red nucleus, block conditioned fear as measured with the potentiated startle paradigm. *Behavioral Neuroscience*, 100(1), 11–22. <https://doi.org/10.1037/0735-7044.100.1.11>
- Hitchcock, J. & Davis, M. (1987). Fear-potentiated startle using an auditory conditioned stimulus: Effect of lesions of the amygdala. *Physiology & Behavior*, 39(3), 403–408. [https://doi.org/10.1016/0031-9384\(87\)90242-3](https://doi.org/10.1016/0031-9384(87)90242-3)
- Hox, J. J., Moerbeek, M. & van de Schoot, R. (2018). *Multilevel analysis: Techniques and applications* (3. Aufl.). Routledge.
- Kiesel, A. & Koch, I. (2012). *Lernen: Grundlagen der Lernpsychologie* (1. Aufl.). Verlag für Sozialwissenschaften.
- Knight, D. C., Nguyen, H. T. & Bandettini, P. A. (2003). Expression of conditional fear with and without awareness. *Proceedings of the National Academy of Sciences*, 100(25), 15280–15283. <https://doi.org/10.1073/pnas.2535780100>
- Kristjansson, S. D., Kircher, J. C. & Webb, A. K. (2007). Multilevel models for repeated measures research designs in psychophysiology: An introduction to growth curve modeling. *Psychophysiology*, 44, 728–736. <https://doi.org/10.1111/j.1469-8986.2007.00544.x>
- Lacey, J. I. & Lacey, B. C. (1958). Verification and extension of the principle of autonomic response-stereotypy. *The American Journal of Psychology*, 71(1), 50. <https://doi.org/10.2307/1419197>
- Landis, C. & Hunt, W. (1939). *The startle pattern*. Farrar & Rinehart.
- Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5), 372–385. <https://doi.org/10.1037/0003-066x.50.5.372>
- Lang, P. J., Bradley, M. M. & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, 97(3), 377–395. <https://doi.org/10.1037/0033-295x.97.3.377>
- Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C. (1981). *Das State-Trait-Angstinventar (STA)*: *Theoretische Grundlagen und Handanweisung*. Beltz Testgesellschaft.
- LeDoux, J. E. (1996). *The emotional brain*. Simon & Schuster.
- LeDoux, J. E. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology*, 23, 727–738. <https://doi.org/10.1023/a:1025048802629>
- Leuchs, L., Schneider, M. & Spoormaker, V. I. (2019). Measuring the conditioned response: A comparison of pupillometry, skin conductance, and startle electromyography. *Psychophysiology*, 56, e13283. <https://doi.org/10.1111/psyp.13283>
- Lipp, O. V. (2006). Human fear learning: Contemporary procedures and measurement. In M. G. Craske, D. Hermans & D. Vansteenwegen (Hrsg.), *Fear and learning: From basic processes to clinical implications* (S. 37–51). American Psychological Association. <https://doi.org/10.1037/11474-002>

- Lissek, S., Pine, D. S. & Grillon, C. (2006). The strong situation: A potential impediment to studying the psychobiology and pharmacology of anxiety disorders. *Biological Psychology*, 72(3), 265–270. <https://doi.org/10.1016/j.biopspsycho.2005.11.004>
- Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., ... Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience & Biobehavioral Reviews*, 77, 247–285. <https://doi.org/10.1016/j.neubiorev.2017.02.026>
- Lovibond, P. F. (2006). Fear and avoidance: An integrated expectancy model. In M. G. Craske, D. Hermans & D. Vansteenwegen (Hrsg.), *Fear and learning: From basic processes to clinical implications* (S. 117–132). American Psychological Association. <https://doi.org/10.1037/11474-006>
- Lovibond, P. F. & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes*, 28(1), 3–26. <https://doi.org/10.1037/0097-7403.28.1.3>
- Luck, C. C. & Lipp, O. V. (2016). Instructed extinction in human fear conditioning: History, recent developments, and future directions. *Australian Journal of Psychology*, 68(3), 209–227. <https://doi.org/10.1111/ajpy.12135>
- Maas, C. J. & Hox, J. J. (2004). The influence of violations of assumptions on multilevel parameter estimates and their standard errors. *Computational Statistics & Data Analysis*, 46(3), 427–440. <https://doi.org/10.1016/j.csda.2003.08.006>
- MacCallum, R. C., Kim, C., Malarkey, W. B. & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, 32(3), 215–253. https://doi.org/10.1207/s15327906mbr3203_1
- MATLAB. (2020). *Version 9.9 (R2020b)*. The MathWorks Inc.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H. & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>
- Mertens, G., Boddez, Y., Sevenster, D., Engelhard, I. M. & Houwer, J. D. (2018). A review on the effects of verbal instructions in human fear conditioning: Empirical findings, theoretical considerations, and future directions. *Biological Psychology*, 137, 49–64. <https://doi.org/10.1016/j.biopspsycho.2018.07.002>
- Mertens, G. & Engelhard, I. M. (2020). A systematic review and meta-analysis of the evidence for unaware fear conditioning. *Neuroscience & Biobehavioral Reviews*, 108, 254–268. <https://doi.org/10.1016/j.neubiorev.2019.11.012>
- Merz, C. J. & Lonsdorf, T. B. (2020). Methodische Anmerkungen und Anwendungsbereiche der Furcht-konditionierung in verschiedenen psychologischen Disziplinen. *Psychologische Rundschau*, 71(3), 273–287. <https://doi.org/10.1026/0033-3042/a000427>
- Meteyard, L. & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, 112, 104092. <https://doi.org/10.1016/j.jml.2020.104092>

- Mineka, S. & Oehlberg, K. (2008). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychologica*, 127(3), 567–580. <https://doi.org/10.1016/j.actpsy.2007.11.007>
- Mitchell, C. J., Houwer, J. D. & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, 32(2), 183–198. <https://doi.org/10.1017/s0140525x09000855>
- Ney, L. J., Wade, M., Reynolds, A., Zuj, D. V., Dymond, S., ... Felmingham, K. (2018). Critical evaluation of current data analysis strategies for psychophysiological measures of fear conditioning and extinction in humans. *International Journal of Psychophysiology*, 134, 95–107. <https://doi.org/10.1016/j.ijpsycho.2018.10.010>
- Nezlek, J. B. (2012). Multilevel modeling for psychologists. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf & K. J. Sher (Hrsg.), *APA handbook of research methods in psychology, Vol. 3. Data analysis and research publication* (S. 219–241). American Psychological Association. <https://doi.org/10.1037/13621-011>
- Nezlek, J. B., Schröder-Abé, M. & Schütz, A. (2006). Mehrebenenanalysen in der psychologischen Forschung. *Psychologische Rundschau*, 57(4), 213–223. <https://doi.org/10.1026/0033-3042.57.4.213>
- Öhman, A. & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108(3), 483–522. <https://doi.org/10.1037/0033-295x.108.3.483>
- Ojala, K. E. & Bach, D. R. (2020). Measuring learning in human classical threat conditioning: A review of translational, cognitive and methodological considerations. *Neuroscience and Biobehavioral Reviews*, 114, 96–112. <https://doi.org/10.1016/j.neubiorev.2020.04.019>
- Page-Gould, E. (2016). Multilevel Modeling. In J. Cacioppo, L. Tassinary & G. Berntson (Hrsg.), *Handbook of Psychophysiology* (4. Aufl., S. 662–678). Cambridge University Press. <https://doi.org/10.1017/9781107415782.030>
- Pavlov, I. P. (1927). *Conditioned reflexes*. Oxford University Press.
- Pinheiro, J. & Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team. (2020). *nlme: Linear and Nonlinear Mixed Effects Models* [R package version 3.1-149]. <https://CRAN.R-project.org/package=nlme>
- Purkis, H. M. & Lipp, O. V. (2001). Does affective learning exist in the absence of contingency awareness? *Learning and Motivation*, 32(1), 84–99. <https://doi.org/10.1006/lmot.2000.1066>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Rachman, S. (1977). The conditioning theory of fearacquisition: A critical examination. *Behaviour Research and Therapy*, 15(5), 375–387. [https://doi.org/10.1016/0005-7967\(77\)90041-9](https://doi.org/10.1016/0005-7967(77)90041-9)
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2. Aufl., Bd. 1). Sage Publications.
- Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, 66(1), 1–5. <https://doi.org/10.1037/h0025984>

- Schultz, D. H. & Helmstetter, F. J. (2010). Classical conditioning of autonomic fear responses is independent of contingency awareness. *Journal of Experimental Psychology: Animal Behavior Processes*, 36(4), 495–500. <https://doi.org/10.1037/a0020263>
- Sevenster, D., Beckers, T. & Kindt, M. (2014). Fear conditioning of SCR but not the startle reflex requires conscious discrimination of threat and safety. *Frontiers in Behavioral Neuroscience*, 8(32). <https://doi.org/10.3389/fnbeh.2014.00032>
- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Siddle, D. A. (1991). Orienting, habituation, and resource allocation: An associative analysis. *Psychophysiology*, 28(3), 245–259. <https://doi.org/10.1111/j.1469-8986.1991.tb02190.x>
- Siddle, D. A., Remington, B. & Churchill, M. (1985). Effects of conditioned stimulus preexposure on human electrodermal conditioning. *Biological Psychology*, 20(2), 113–127. [https://doi.org/10.1016/0301-0511\(85\)90020-1](https://doi.org/10.1016/0301-0511(85)90020-1)
- Sjouwerman, R., Niehaus, J., Kuhn, M. & Lonsdorf, T. B. (2016). Don't startle me – Interference of startle probe presentations and intermittent ratings with fear acquisition. *Psychophysiology*, 53(12), 1889–1899. <https://doi.org/10.1111/psyp.12761>
- Tabbert, K., Stark, R., Kirsch, P. & Vaitl, D. (2006). Dissociation of neural responses and skin conductance reactions during fear conditioning with and without awareness of stimulus contingencies. *NeuroImage*, 32(2), 761–770. <https://doi.org/10.1016/j.neuroimage.2006.03.038>
- Vaitl, D. & Lipp, O. V. (1997). Latent inhibition and autonomic responses: A psychophysiological approach. *Behavioural Brain Research*, 88(1), 85–93. [https://doi.org/10.1016/S0166-4328\(97\)02310-3](https://doi.org/10.1016/S0166-4328(97)02310-3)
- Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4. Aufl.). Springer.
- Weidemann, G., Satkunarajah, M. & Lovibond, P. F. (2016). I think, therefore eyeblink: The importance of contingency awareness in conditioning. *Psychological Science*, 27(4), 467–475. <https://doi.org/10.1177/0956797615625973>
- Weike, A. I., Schupp, H. T. & Hamm, A. O. (2007). Fear acquisition requires awareness in trace but not delay conditioning. *Psychophysiology*, 44(1). <https://doi.org/10.1111/j.1469-8986.2006.00469.x>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wiens, S. & Öhman, A. (2002). Unawareness is more than a chance event: Comment on Lovibond and Shanks (2002). *Journal of Experimental Psychology: Animal Behavior Processes*, 28(1), 27–31. <https://doi.org/10.1037/0097-7403.28.1.27>

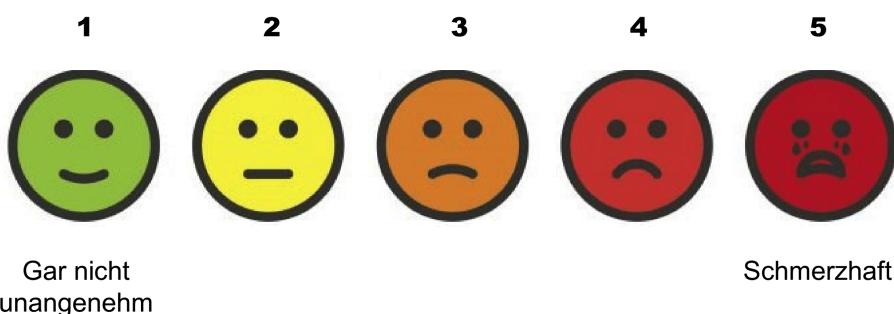
Anhang

Anhang A

Materialien zur Bestimmung der US-Intensität: (1) Skala zur Einschätzung der US-Intensität, (2) Tabelle zur Kalibrierung der individuellen Reizintensität

(1)

Skala zur Beurteilung der Unangenehmheit des elektrotaktilen Reizes



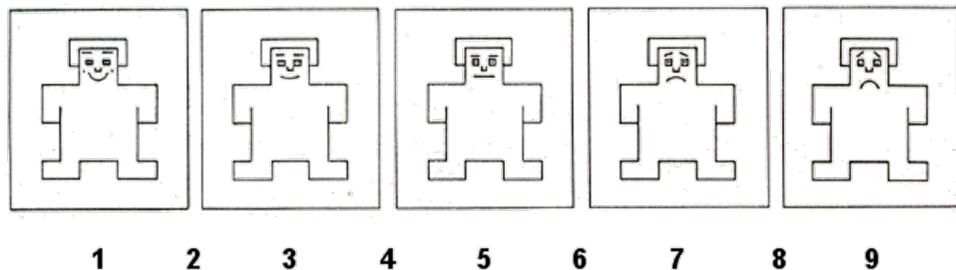
(2) Ausgangsintensität: 0,10 x 10

| Veränderung der Reizintensität | | | | | | | | | |
|--------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 1 | 1,5 | 2 | 2,5 | 3 | 3,5 | 4 | 4,5 | 5 |
| 1. DG | + 0,08 | + 0,05 | + 0,05 | + 0,04 | + 0,03 | + 0,02 | + 0,00 | - 0,01 | - 0,02 |
| 2. DG | + 0,20 | + 0,10 | + 0,10 | + 0,08 | + 0,05 | + 0,03 | + 0,00 | - 0,01 | - 0,02 |
| 3. DG | + 0,30 | + 0,20 | + 0,10 | + 0,08 | + 0,05 | + 0,03 | + 0,00 | - 0,01 | - 0,02 |
| 4. DG | + 0,30 | + 0,20 | + 0,10 | + 0,08 | + 0,05 | + 0,03 | + 0,00 | - 0,01 | - 0,03 |
| 5. DG | + 0,30 | + 0,20 | + 0,10 | + 0,08 | + 0,05 | + 0,03 | + 0,00 | - 0,01 | - 0,03 |
| ab 6. DG | + 0,30 | + 0,20 | + 0,10 | + 0,08 | + 0,05 | + 0,03 | + 0,00 | - 0,02 | - 0,03 |

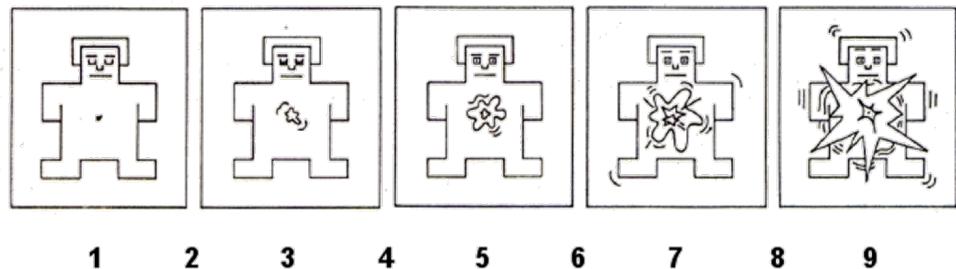
Anhang B

SAM-Skalen zur Einschätzung von Valenz und Erregung (Bradley & Lang, 1994), deutsch adaptierte Fassung

ANGENEHM ←→ **UNANGENEHM**

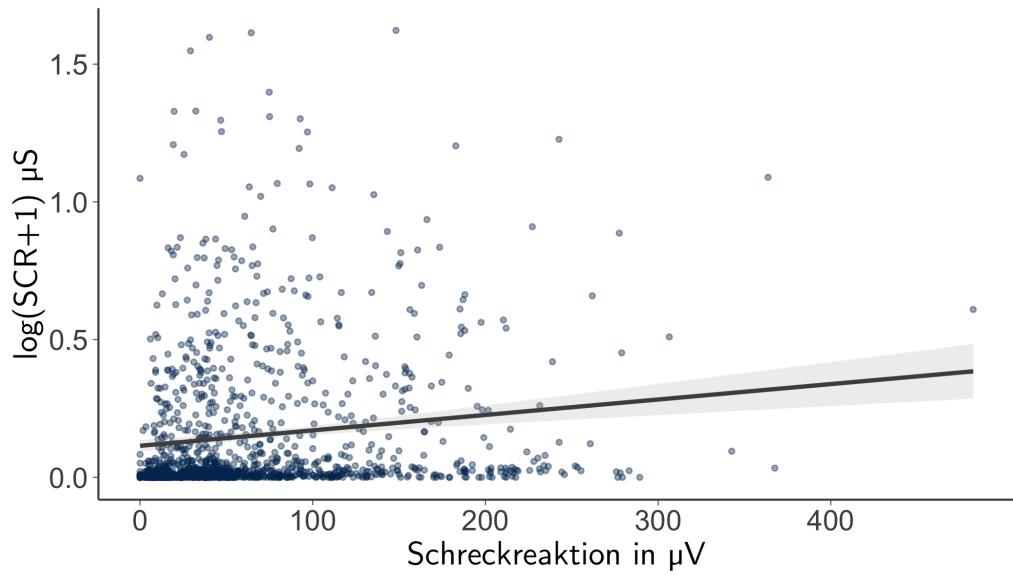


RUHIG → **AUFREGEND**



Anhang C

Streudiagramm zwischen Hautleitwert- und Schreckreaktionen. Die Gerade entspricht der Vorhersage eines einfachen linearen Modells mit 95 %-Konfidenzgrenze



Anhang D

Teststatistiken der Shapiro-Wilk-Tests für beide abhängigen Variablen

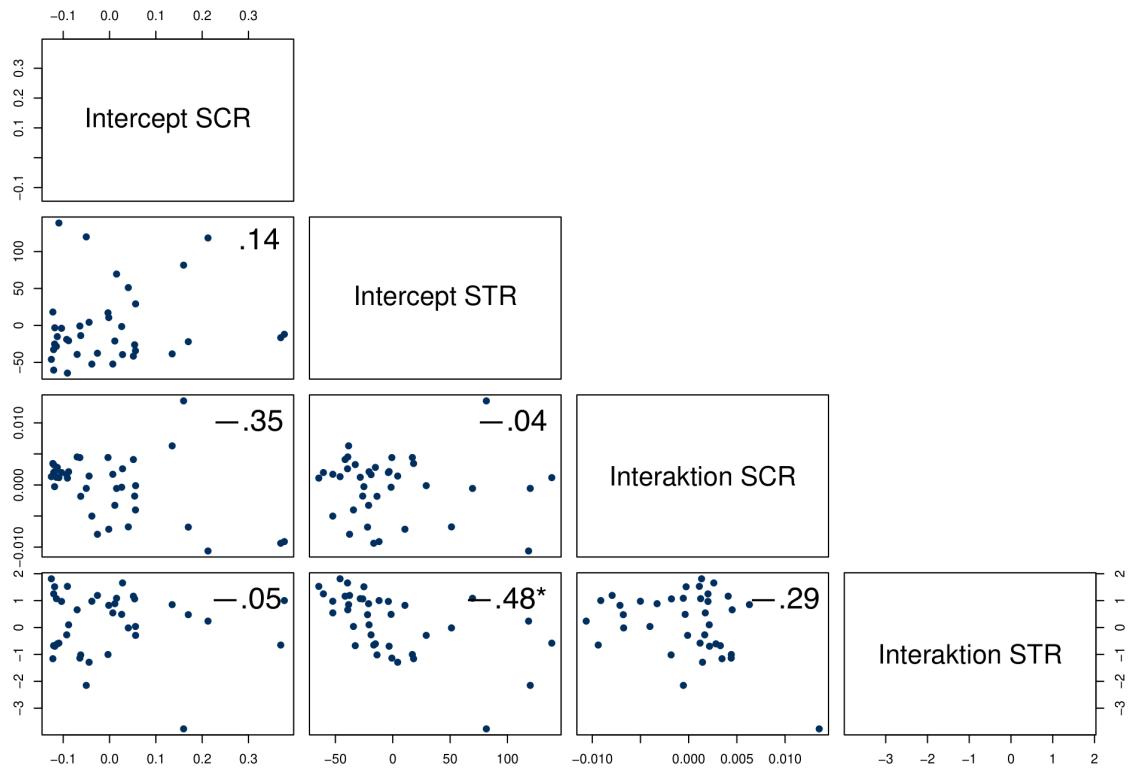
| Trial | SCR | STR |
|-----------|---------|---------|
| insgesamt | 0.64*** | 0.84*** |
| 1 | 0.81*** | 0.81*** |
| 2 | 0.68*** | 0.91*** |
| 3 | 0.76*** | 0.89*** |
| 4 | 0.60*** | 0.89*** |
| 5 | 0.61*** | 0.87*** |
| 6 | 0.67*** | 0.77*** |
| 7 | 0.64*** | 0.85*** |
| 8 | 0.49*** | 0.82*** |
| 9 | 0.78*** | 0.82*** |
| 10 | 0.66*** | 0.79*** |
| 11 | 0.64*** | 0.79*** |
| 12 | 0.59*** | 0.83*** |
| 13 | 0.59*** | 0.86*** |
| 14 | 0.62*** | 0.80*** |
| 15 | 0.59*** | 0.75*** |
| 16 | 0.53*** | 0.82*** |

Anmerkungen. SCR: Hautleitwertreaktion, STR: Schreckreaktion;

* $p < .05$, ** $p < .01$, *** $p < .001$.

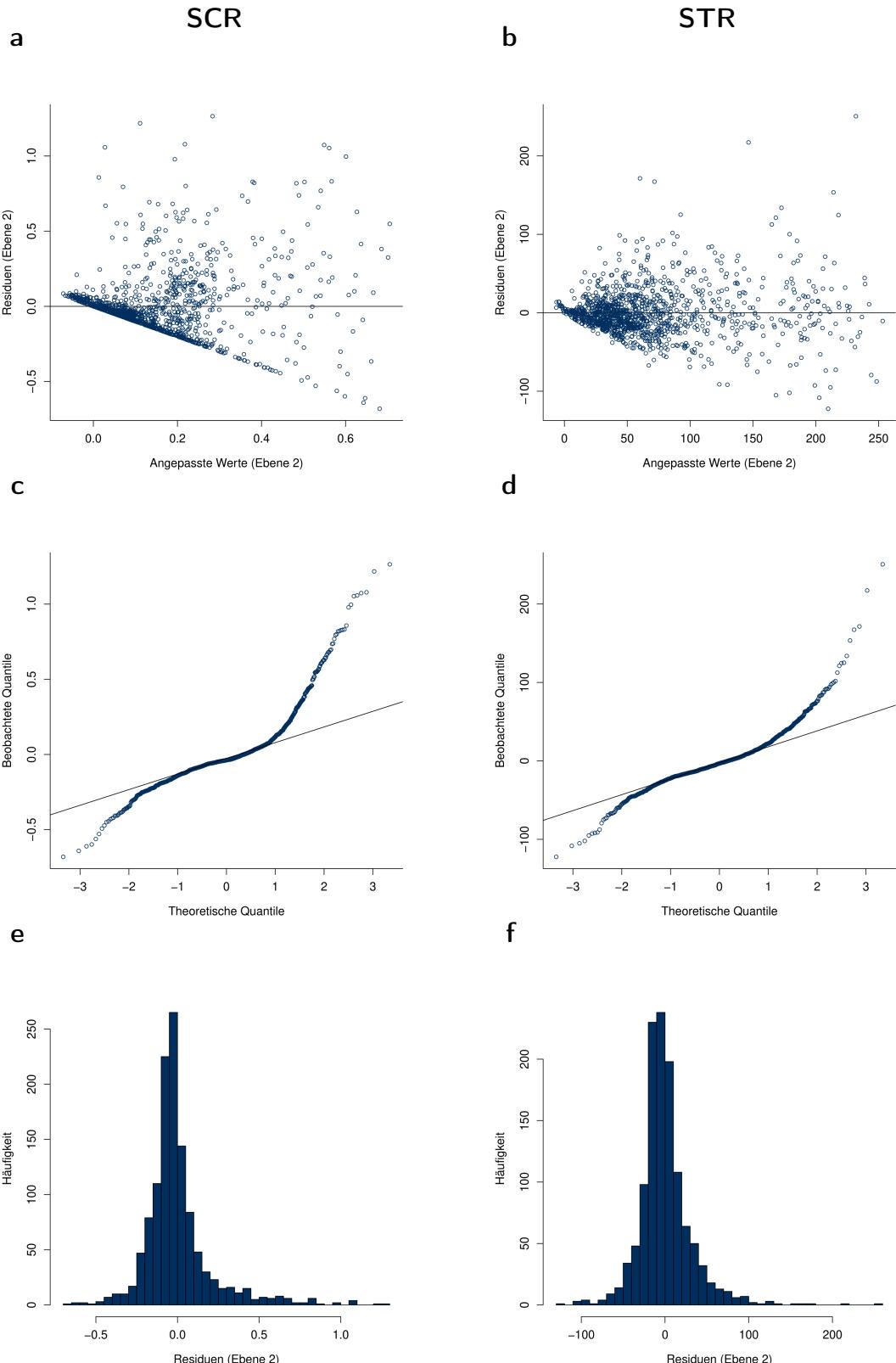
Anhang E

Streudiagramme (und Korrelationen) zwischen den geschätzten zufälligen Effekten beider abhängiger Variablen



Anhang F

Streudiagramme zwischen Residuen der höchsten Gruppierungsebene und angepassten Werten für beide AV (a, b) sowie Q-Q-Diagramme (c, d) und Histogramme (e, f) dieser Residuen für beide AV



Selbstständigkeitserklärung

Hiermit erkläre ich,

Neele Katharina Elbersgerd, geboren am xx.xx.xxxx

an Eides statt, dass ich die vorliegende Arbeit

*Vergleich der Verläufe von Schreck- und Hautleitwertreaktionen
in der Akquisitionsphase einer Furchtkonditionierung*

selbstständig und ohne Benutzung von anderen als den angegebenen Hilfsmitteln und Quellen angefertigt habe. Alle Inhalte, die ich aus anderen veröffentlichten oder unveröffentlichten Quellen dem Wortlaut oder dem Sinne nach entnommen habe, sind kenntlich gemacht und im Literaturverzeichnis aufgeführt. Diese Arbeit wurde nicht im Rahmen eines anderen Prüfungsverfahrens eingereicht.

Ort, Datum

Unterschrift
