

**Der Einfluss von Misserfolgsstress auf die
Intelligenztestleistung in Abhängigkeit von dem Leistungsmotiv**

Nicole Eiseler

Diplomarbeit als Teil der Diplomprüfung für Psychologen
und Psychologinnen

Universität Regensburg

17.05.2010

Inhaltsverzeichnis

0. Einleitung	6
1. Leistungsmotivation, Intelligenz und Misserfolgsstress	8
1.1 Leistungsmotivation	8
1.1.1 Begriffserklärungen: Motiv und Motivation.....	8
1.1.2 Das Leistungsmotiv, Gütemaßstab und Anspruchsniveau.....	9
1.1.3 Das dualistische Modell der Leistungsmotivation.....	10
1.1.4 Das Risiko-Wahl-Modell nach John Atkinson.....	11
1.1.5 Kritik und weiterführende Theorien zum Risiko-Wahl-Modell.....	15
1.1.6 Die Messung des Leistungsmotivs	18
1.1.7 Die vier Leistungsmotivkomponenten.....	19
1.2 Intelligenz	20
1.2.1 Was ist Intelligenz?	20
1.2.2 Intelligenzstrukturmodelle	22
1.2.3 Die Generalfaktorentheorie nach Spearman.....	24
1.2.4 Die Primary Mental Abilities nach Thurstone.....	24
1.2.5 Das Zweifaktorenmodell nach Cattell	25
1.2.6 Das Berliner Intelligenzstrukturmodell nach Jäger	26
1.2.7 Das Drei-Ebenen-Modell nach Carroll	26
1.2.8 Intelligenz – ein komplexes Phänomen	27
1.3 Der Einfluss von Misserfolgsstress in Leistungssituationen	28
1.3.1 Der Begriff Stress.....	29
1.3.2 Das transaktionale Stresskonzept nach Lazarus	30
1.3.3 Techniken der Stresserzeugung.....	31
1.3.4 Die Auswirkungen einer Misserfolgsinstruktion auf die Leistung	33
1.4 Empirische Befunde zum Einfluss der Leistungsmotive und Stress auf die Leistung	37
1.4.1 Empirische Befunde zu Stressinstruktionen und Leistungsmotivation.....	37
1.4.2 Integration der Erklärungsversuche der empirischen Befunde	40
2. Hypothesen und Fragestellungen	45
2.1 Einfluss der Stressinstruktion	45
2.2 Leistungsmotivation – Wohlbefinden, gefühltes Stresslevel und Bewertung der Intelligenztests	45
2.3 Intelligenztestleistungen und Leistungsmotivation	47
2.3.1 Intelligenztestleistung und Leistungsmotivation unter neutraler Bedingung.....	47
2.3.2 Intelligenztestleistung und Leistungsmotivation unter Stress.....	48
2.3.3 Intelligenztestleistung und die Höhe des Leistungsmotivs	49

3. Methode	51
3.1 Beschreibung der Stichprobe	51
3.1.1 Geschlecht	51
3.1.2 Alter	52
3.1.3 Händigkeit	52
3.1.4 Studienfach und Semesteranzahl	53
3.1.5 Schulabschluss und Abiturnote	54
3.1.6 Berufliche Tätigkeit und berufliche Ausbildung	54
3.2 Untersuchungsablauf	55
3.3 Untersuchungsmethode	58
3.3.1 Das Regensburger Leistungs-Motiv-Inventar (RLMI-E)	58
3.3.1.1 Beschreibung des RLMI-E	58
3.3.1.2 Testanalyse des RLMI-E	60
3.3.2 Der Grundintelligenztest Skala 2 (CFT 20-R)	61
3.3.2.1 Beschreibung des CFT 20-R	61
3.3.2.2 Testanalyse des CFT 20-R	62
3.3.3 Das neue Leistungsprüfsystem (LPS-neu)	63
3.3.3.1 Beschreibung des LPS-neu	63
3.3.3.2 Testanalyse des LPS-neu	67
3.3.4 Fragebogen zum aktuellen Wohlbefinden	70
3.3.4.1 Beschreibung des Fragebogens zum aktuellen Wohlbefinden	70
3.3.4.2 Itemanalyse der Fragebögen zum aktuellen Wohlbefinden	71
3.3.4.3 Skalenanalyse der Fragebögen zum aktuellen Wohlbefinden	73
3.3.4.4 Faktorenanalyse der Fragebögen zum aktuellen Wohlbefinden	74
3.3.5 Evaluationsfragebogen des CFT 20-R und des LPS-neu	74
3.3.5.1 Beschreibung des Evaluationsfragebogens	74
3.3.5.2 Itemanalyse der Evaluationsfragebögen	75
3.3.5.3 Skalenanalyse der Evaluationsfragebögen	76
3.3.5.4 Faktorenanalyse der Evaluationsfragebögen	77
3.3.6 Fragebogen zur subjektiven Beurteilung der Stressreaktion	77
3.3.6.1 Beschreibung des Stressreaktions-Fragebogens	77
3.3.6.2 Testanalyse des Stressreaktions-Fragebogens	78
3.3.6.3 Faktorenanalyse des Stressreaktions-Fragebogens	79
3.3.7 Demografischer Fragebogen	80
3.4 Statistische Datenauswertung	80
3.4.1 Verwendete statistische Verfahren	80
3.4.2 Prüfung der Homogenität der zu vergleichenden Gruppen	82
4. Ergebnisse	85
4.1 Einfluss der Stressinstruktion	85
4.1.1 Stressinstruktion und Wohlbefinden	85

4.1.2	Stressinstruktion und gefühltes Stresslevel.....	88
4.1.3	Zusammenhänge zwischen Wohlbefinden, gefühltem Stresslevel und Bewertung.....	89
4.2	Leistungsmotivation.....	90
4.2.1	Verteilung der Leistungsmotivation in der Stichprobe	90
4.2.2	Leistungsmotivation und Wohlbefinden	93
4.2.3	Leistungsmotivation und gefühltes Stresslevel.....	96
4.2.4	Leistungsmotivation und Bewertung.....	97
4.3	Die Intelligenztestleistungen	100
4.3.1	Die Ergebnisse der Intelligenztests in der Stichprobe	100
4.3.2	Konstruktvalidierung	103
4.3.3	Zusammenhänge der Intelligenztests und der Stressinstruktion.....	104
4.4	Intelligenztestleistung und Leistungsmotivation	105
4.4.1	Intelligenztestleistung und Leistungsmotivation unter neutraler Bedingung.....	105
4.4.2	Intelligenztestleistungen der Leistungsmotiv- und experimentellen Gruppen	107
4.4.3	Intelligenztestleistung und Leistungsmotivation unter Stress.....	112
4.4.4	Intelligenztestleistung und die Höhe des Leistungsmotivs	118
4.4.5	Zusammenhänge zwischen Leistungsmotiven und Intelligenztestleistungen	123
5.	Diskussion	124
5.1	Diskussion der Ergebnisse.....	124
5.1.1	Einfluss der Stressreaktion	124
5.1.2	Leistungsmotivation – Wohlbefinden, gefühltes Stresslevel und Bewertung der Intelligenztests	126
5.1.3	Intelligenztestleistungen und Leistungsmotivation	129
5.1.3.1	Intelligenztestleistungen und Leistungsmotivation unter neutraler Bedingung	129
5.1.3.2	Intelligenztestleistungen der Leistungsmotiv- und experimentellen Gruppen.	130
5.1.3.3	Intelligenztestleistungen und Leistungsmotivation unter Stress.....	131
5.1.3.4	Intelligenztestleistungen und die Höhe des Leistungsmotivs.....	133
5.1.3.5	Resümee bezüglich der Intelligenztestleistungen und Leistungsmotivation ...	134
5.2	Diskussion der Methode	139
5.2.1	Untersuchungsmethode.....	140
5.2.1.1	Das Regensburger Leistungs-Motiv-Inventar	140
5.2.1.2	Die Intelligenztests CFT 20-R und LPS-neu	140
5.2.1.3	Die Einschätzungsfragebögen	141
5.2.2	Stichprobe	143
5.2.3	Untersuchungsablauf	146
5.2.4	Ausblick.....	150
6.	Zusammenfassung	153
	Literaturverzeichnis	155

Erklärung	164
Anhang	165
Anhang A Instruktion und Debriefing	165
A-1 Instruktion zum LPS-neu für die Experimentalgruppe	165
A-2 Debriefing der Kontrollgruppe	166
A-3 Debriefing der Experimentalgruppe und Stressreaktions-Fragebogen	167
Anhang B Fragebögen	168
B-1 Fragebogen zur Person	168
B-2 Fragebogen zum aktuellen Wohlbefinden	169
B-3 Evaluationsfragebogen	170
Anhang C Skalenwerte des RLMI-E und CFT 20-R	171
Anhang D Faktorenanalyse des Fragebogens zum aktuellen Wohlbefinden und des Evaluationsfragebogens	173
Anhang E Histogramme	176
Anhang F Ergebnisse der Varianzanalysen	181
Anhang G Abbildungen der Mittelwertsvergleiche	191

0. Einleitung

An diesen Morgen war es soweit – 118 Kandidaten gemeinsam in einem schwülen, heißen Raum, um das Landesexamen zur erhofften Zulassung zum Gymnasium zu bestehen. Den Stoff hat er sich akribisch eingeprägt und doch kämpft der Junge auf einmal mit Herzklopfen, Nervosität, Angst und Schwindel.

Dann versuchte er, sich auf die Fragen und auf die Antworten zu besinnen, aber alles ging ihm durcheinander. Er sah nur immer wieder die große, grüne Tischfläche, die drei alten, ernsten Herren in Gehröcken, das aufgeschlagene Buch und seine zitternd daraufgelegte Hand. Herrgott, was mochte er für Antworten gegeben haben! (Hermann Hesse, 1970, S. 24)

Hermann Hesse (1906/1970) schreibt in seiner Erzählung „Unterm Rad“ von dem Schicksal eines begabten Jungen, der unter dem hohen Leistungsstress erst das Auge für seine kindlichen Wünsche und Naturverbundenheit verliert und im Verlauf der Geschichte an seine physischen Grenzen stößt, bis er am Ende psychisch zusammenbricht. Der Autor klagt in seiner Erzählung das „zermürende Rad“ der leistungsorientierten Gesellschaft, das überfordernde Erziehungsverhalten und das Schulwesen um 1900 an. Die beschriebene Schlüsselszene beschreibt, wie Angst und Furcht in einer wichtigen Leistungssituation den klugen und lerneifrigen Protagonisten negativ beeinträchtigen. Natürlich ist dies kein Phänomen der damaligen Zeit, sondern findet sich auch heute in jeglichen Test- und Leistungssituationen wieder. Wie kommt es dazu, dass einige Personen unter stresserzeugenden Leistungssituationen durch starke emotionale, physiologische oder kognitive Reaktionen negativ beeinflusst werden und andere nicht?

Die individuellen Unterschiede der Auswirkungen von Stress auf die Leistung sind von großer theoretischer und praktischer Bedeutung für das Berufs- und Arbeitsleben sowie für Schule und Studium und wurden daher in verschiedenen Bereichen der Psychologie häufig diskutiert. So kommt es in Hesses Erzählung dazu, dass der Junge zwar in der beschriebenen Situation mit der Angst zu kämpfen hat, insgesamt jedoch Zweitbester des Landesexamens wird. Aufschlussreiche Antworten auf das Phänomen, dass Leistungsstress bei unterschiedlichen Personen und unterschiedlichen Gegebenheiten einen hemmenden oder fördernden Einfluss auf Leistung haben kann, wurden unter anderem in der Leistungsmotivationsforschung von Atkinson (1957), McClelland, Atkinson, Clark und Lowell (1953) und Heckhausen (1963) gefunden. Sie führten ein solches Verhalten, wie das des Protagonisten in Hesses Erzählung auf die unterschiedliche Wirkung und Ausprägung der Leistungsmotivkomponenten „Hoffnung auf Erfolg“ und „Furcht vor Misserfolg“ zurück. In einer Vielzahl von Studien wurde der Einfluss beider Komponenten in Leistungs- bzw. Stresssituationen untersucht, wobei sich

unterschiedliche Ergebnisse bezüglich der Leistungsbeeinflussungen ergaben. So scheint es offensichtlich, dass es sich hier nicht um einen einfachen, geradlinigen Zusammenhang, sondern vielmehr um einen komplexen Wechselwirkungsprozess zwischen Person und Situation handelt.

Mit der vorliegenden Untersuchung soll der komplexe Einfluss der Leistungsmotivausprägung auf dem Hintergrund einer gleichzeitigen Validierung eines neu überarbeiteten Intelligenztest, dem Leistungsprüfsystem (LPS-neu) nach Kreuzpointner (2010), untersucht werden. Dabei wird überprüft, in welchem Ausmaß das Testsetting bzw. eine abweichende Instruktion in Form einer Stresssituation abhängig von individuellen Merkmalsunterschieden die Leistung in einem Intelligenztest beeinflusst.

Das Wissen um die Hintergründe solcher Zusammenhänge kann in der Intelligenztestforschung sowie in schulischen und beruflichen Bereichen Anwendung finden. In der Intelligenztestforschung sind solche Kenntnisse beispielsweise für das Gestalten und das exakte Einhalten von Instruktionen oder für die Interpretation von Intelligenztestergebnissen in Bereichen wie Personalauswahlverfahren, wo diese Tests ihre Verwendung finden, von großer Bedeutung. Weiterhin könnte dieses Wissen dazu dienen, Testbedingungen in Schule, Studium sowie im Arbeitsleben zu optimieren und auf individuelle Persönlichkeitseigenschaften abzustimmen. Natürlich sind solche Forderungen in der Realität nur sehr schwer umsetzbar und mit hohen Kosten verbunden. Trotzdem sind fundierte Kenntnisse darüber wichtig und notwendig, um das Ziel der Gestaltung optimaler Test- und Leistungssituationen zu erreichen.

In folgendem Kapitel 1 soll zunächst ein Überblick über die unterschiedlichen angesprochenen Gebiete mithilfe von Begriffsbestimmungen, postulierten Modellen sowie empirischen Befunden gegeben werden. Anfangs wird der theoretische Hintergrund der Leistungsmotivation erörtert. Es folgt eine Darstellung bedeutender Modelle und Erkenntnisse der Intelligenzforschung und weiterhin werden die für diese Arbeit wichtigen Aspekte bezüglich Stress in Leistungssituationen herausgearbeitet. Abschließend werden bedeutende empirische Befunde, Ergebnisse und Erklärungsmodelle zu den Zusammenhängen der einzelnen Thematiken vorgestellt, welche in Kapitel 2 in der Formulierung der Hypothesen und Fragestellungen für diese Arbeit münden. Danach erfolgt eine Erläuterung des methodischen Vorgehens. Dabei werden die untersuchte Stichprobe und die Operationalisierung der verschiedenen Variablen beschrieben, wobei die verwendeten Testverfahren genauer vorgestellt werden (Kapitel 3). Die aus dieser Untersuchung gewonnenen Ergebnisse werden in Kapitel 4 in Form von Tabellen und Diagrammen dargestellt, welche im fünften Kapitel eingehend diskutiert und mit deren Hilfe die Hypothesen überprüft werden. Es schließt sich eine Zusammenfassung der wichtigsten Erkenntnisse dieser Arbeit an.

1. Leistungsmotivation, Intelligenz und Misserfolgsstress

1.1 Leistungsmotivation

Die empirische Auseinandersetzung mit dem Einfluss des Leistungsmotivs auf Intelligenztests und Stresssituationen setzt zunächst voraus, dass Klarheit darüber besteht, was in der psychologischen Forschung unter dem Begriff Leistungsmotivation verstanden wird.

1.1.1 Begriffserklärungen: Motiv und Motivation

Um das Konstrukt des Leistungsmotivs zu begreifen, ist es empfehlenswert, sich vorab mit den Begriffen *Motivation* bzw. *Motiv* auseinanderzusetzen. Schon die lateinische Abstammung des Wortes „Motivation“ von „movere“ (dt.: „bewegen“) veranschaulicht, dass Motivation und Bewegung eng miteinander verknüpft sind (Rudolph, 2003). Wird diese Wortherkunft mit den zahlreichen wissenschaftlichen Diskussionen um den Begriff verglichen, taucht auch hier „bewegen“, etwas in Bewegung setzen, Antrieb oder beispielsweise in Atkinsons (1964/1975) Worten ausgedrückt, die Lokomotion des Verhaltens, ständig auf. Die Motivationspsychologie beschäftigt sich also allgemein mit zielgerichtetem Verhalten und während zu Beginn der Forschungsentwicklung Motivation als Folge reiner Trieb- und Spannungsreduktion beschrieben wird, berücksichtigen spätere Erklärungsmodelle auch kognitive und emotionale Aspekte (Heckhausen, 1963). Eine präzise Definition zum Begriff Motivation stammt von dem renommierten Motivationspsychologen Heinz Heckhausen (1980):

Sie sind es, so wird unterstellt, die eine Handlungsfolge in Gang setzen, auf ein Ziel richten, auf dem Wege dahin steuern und mit der Zielerreichung wieder schwinden. (Heckhausen, 1980, S. 30)

Motivation gilt als kurzfristige Aktivierung, da sie von Situationsgegebenheiten gelenkt wird (Heckhausen, 1963). Das Motiv hingegen ist, wie es Rheinberg ausdrückt, „als eine personenspezifische Konstante gedacht, hinsichtlich derer sich Menschen unterscheiden“ (Rheinberg, 2000, S. 62). Durch ein Motiv wird ein angestrebter Zielzustand oder Anreiz bewertet und erhält so Bedeutung (Schneider & Schmalt, 2000), was den Prozess der Motivation auslösen und zu einer Handlung führen kann (Heckhausen, 1980). Der Anreiz kann dabei intrinsischer Natur sein, d. h. in der Tätigkeit oder dem Ergebnis selbst liegen, oder aus einem extrinsischen Reiz wie einer materiellen Belohnung oder Selbst- und Fremdbewertung bestehen (Heckhausen & Heckhausen, 2006).

Ein nennenswerter Durchbruch in der Motivationspsychologie ist Murray mit seinem Werk „Explorations in Personality“ zu verdanken, indem er 1938 eine Klassifikation

grundlegender psychogener Bedürfnisse bzw. Motive veröffentlichte. Zwei wichtige Schlüsselbegriffe dieser Arbeit sind „need“, welche als Kraft, unbefriedigende Situationen zielgerichtet zu verändern, beschrieben wird und „press“, womit er den Druck meint, den eine objektiv registrierte Umwelt bzw. subjektiv interpretierte Umwelt auf ein Individuum ausübt (Rheinberg, 2000). Neben den Bedürfnissen oder Motiven nach sozialem Anschluss, Machtausübung, Unabhängigkeit oder Ordnung postulierte er auch das für diese Studie interessante Bedürfnis nach Leistung (Murray, 1938). Das so genannte *need of Achievement* beschreibt Murray (1938) wie folgt:

To accomplish something difficult. To master, manipulate or organize physical objects, human beings, or ideas. To do this as rapidly, and as independently as possible. To overcome obstacles and attain a high standard. To excel one's self. To rival and surpass others. To increase self-regard by the successful exercise of talent. (Murray, 1938, S. 164)

1.1.2 Das Leistungsmotiv, Gütemaßstab und Anspruchsniveau

Obwohl Murray einen wichtigen Beitrag zur der Leistungsmotivationsforschung leistete, gelten McClelland et al. (1953), Atkinson (1957) und Heckhausen (1963) als die Begründer des theoretischen Konstrukts des Leistungsmotivs. McClelland et al. definieren leistungsmotiviertes Verhalten kurz als Auseinandersetzung „with some standard of excellence“ (McClelland et al., 1953, S. 110). Verhalten ist somit nur dann leistungsmotiviert, wenn es darauf abzielt, die eigene Tüchtigkeit zu bewerten, indem versucht wird, einen *Gütemaßstab* zu erreichen oder zu übertreffen (Rheinberg, 2000). Die Ursprünge des Leistungsmotivs eines Individuums, wie eine Studie von Winterbottom (1953) zum Erziehungsverhalten von Müttern in Zusammenhang zum Leistungsmotiv und Alter des Kindes zeigen konnte, sind Lernerfahrungen in einer frühen Kindheitsphase. Laut McClelland und Mitarbeitern (1953) ist das Leistungsmotiv im weiteren Verlauf des Lebens modellierbar und gilt, wie alle Motive, als personenspezifische Konstante, durch die der Prozess der Leistungsmotivation angeregt wird.

Der Begriff des Gütemaßstabs ist ein durch Kultur und Epochen beeinflusstes Bezugssystem (Schneider & Schmalt, 2000), durch welches ein Sachverhalt erst einen Stellenwert erhält (Heckhausen, 1974). Der Vergleich mit eigener vorhergehender Leistung (individuelle Bezugsnormen), mit einer Gruppe (soziale Bezugsnormen), mit der Sache selbst (sachliche Bezugsnormen) oder extern festgelegten Bezugsnormen kann laut Heckhausen (1974, S. 49) als Gütemaßstab dienen. Unterschiedliche Personen können eine Leistung verschieden bewerten und je nach Lebensbereichen können sich auch die Gütemaßstäbe eines Individuums unterscheiden (Lukesch, 2006).

Das Erleben des Erfolgs oder Misserfolgs selbst ist von der individuellen Zielsetzung einer Person an zukünftige eigene Leistung – dem so genannten *Anspruchsniveau* abhängig (Hoppe, 1930). Abbildung 1.1 zeigt den Ablauf einer typischen An-

spruchsniveau-Situation: Ausgehend von der Erfahrung einer gleichen bzw. vergleichbaren Leistungssituation wird eine Zielvorstellung gebildet (Atkinson, 1964/1975). Die Zieldiskrepanz ist nach Atkinson (1964/1975, S. 168) durch den Abstand zwischen der vorhergehenden Leistung und der neu gesetzten Zielvorstellung determiniert. Er nimmt an, dass nach Erfolg die Zieldiskrepanz positiv ist, d. h. das Anspruchsniveau wird nach Erfolg erhöht, wie bereits Hoppe (1930) mit verschiedenen Aufgabentypen zeigen konnte. Bei Misserfolg wird das Anspruchsniveau in der Regel gesenkt (Hoppe, 1930). Nachdem die Leistungssituation absolviert ist, wird das Ergebnis mit der gesetzten Zielvorstellung verglichen, woraufhin nach Atkinson (1964/1975) eine psychologische Reaktion folgt – das Erleben von Affekten in Form von Erfolg oder Misserfolg. Ein Erfolgserlebnis verbunden mit Stolz und Freude stellt sich ein, wenn die Leistung über dem Anspruchsniveau liegt bzw. ihm entspricht (Heckhausen, 1963). Beim Nichterreichen des Ziels tritt Misserfolg in Form von Scham oder Ärger ein (a. a. O.). Befindet sich die Person erneut in einer Leistungssituation, so wirken die aus früheren Erfahrungen antizipierten Gefühlszustände motivierend (a. a. O.).

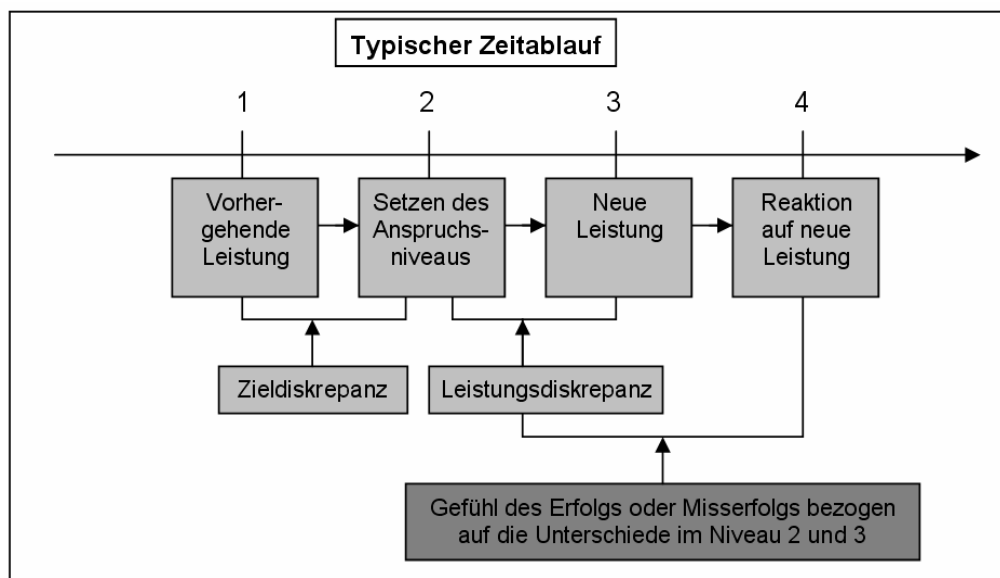


Abbildung 1.1: Vier Hauptphasen einer Anspruchsniveau-Situation (Lewin, Dembo, Festinger & Sears, 1944; zitiert nach Atkinson, 1964/1975, S. 168)

1.1.3 Das dualistische Modell der Leistungsmotivation

Gegenläufig zu der formulierten Anspruchsniveautheorie nach Atkinson (1964/1975) entsprachen einige empirische Befunde nicht den Voraussagen. Es ergaben sich paradoxe Reaktionen auf Leistungsergebnisse, d. h. das Anspruchsniveau wurde nach Erfolg gesenkt und nach Misserfolg angehoben. In einer Studie von Hoppe (1930) änder-

ten Versuchspersonen (Vpn) nach Misserfolg ihr Anspruchsniveau typischerweise leicht mit mäßiger negativer Zieldiskrepanz, andere reagierten aber auch atypisch, d. h. sie setzten ihr Anspruchsniveau sehr niedrig oder hoben es nach Misserfolg an. Die atypische Reaktion bezeichnete Hoppe noch als geringen „Mut zur Wirklichkeit“ (Hoppe, 1930, S. 41), während Atkinson erklärt, dass eine Leistungssituation auch gefürchtet und „the motiv to avoid failure“ (Atkinson, 1957, S. 360) angeregt werden kann. Die Unterteilung des Leistungsmotivs in die zwei Aspekte „hope of success“ und „fear of failure“ wurde auch bereits durch McClelland et al. (1953, S. 214) vorgenommen.

Die Aspekte *Hoffnung auf Erfolg* (HE) und *Furcht vor Misserfolg* (FM) sind „gegen-sätzliche Gefühlserwartungen“, die das eigentlich Motivierende einer Handlung darstellen und zu einem Konflikt zwischen aufsuchender und meidender Tendenz führen können (Heckhausen, 1963, S. 13). Ob eine Person eine Leistungssituation aufsucht oder vermeidet, wird nicht nur durch die Motive eines Individuums, sondern durch Wechselwirkungsprozesse zwischen Person und Situation beeinflusst (Rheinberg, 1976). Das allgemein bekannteste Verhaltensmodell, welches dies berücksichtigt, ist Atkinsons Risiko-Wahl-Modell und dieses wird deshalb anschließend ausführlich behandelt.

1.1.4 Das Risiko-Wahl-Modell nach John Atkinson

Das mathematische Risiko-Wahl-Modell von John Atkinson (1957) bestimmt seit den 60er Jahren die Leistungsmotivforschung (Heckhausen & Heckhausen, 2006). Das Modell versucht Voraussagen über Aufgabenwahl, Leistung und Ausdauer in Leistungssituationen zu machen (Schneider, 1973). Geleitet wurde die Entwicklung des Modells von den Annahmen Lewins, dass Verhalten immer als Funktion der Wechselwirkung von Person- und Umweltvariablen angesehen werden muss und als ein Annäherungs-Vermeidungs-Konflikt aufgefasst werden kann (Lewin, 1938, 1946; zitiert nach Heckhausen, 1980, S. 175) und zwar als ein Konflikt zwischen den Motiven Erfolg aufzusuchen und Misserfolg zu meiden. Atkinson (1964/1975) fasst Leistungsmotivation als Funktion der resultierenden Tendenz dieses Konflikts in Abhängigkeit von Person- und Situationsvariablen auf. Die individuelle Motivausprägung und zwei situative Parameter, der Anreiz und die subjektive Wahrscheinlichkeit, werden durch Atkinson (1957) – im Sinne der „Erwartungs-mal-Wert-Theorien“ (Heckhausen & Heckhausen, 2006) – multiplikativ miteinander verbunden. Die Verknüpfungen der je drei Variablen ergeben im atkinsonschen Modell (1964/1975) die Tendenzen, Erfolg aufzusuchen (T_e) und Misserfolg zu meiden (T_m) und werden wie folgt dargestellt:

$$T_e = M_e \times W_e \times A_e$$

$$T_m = M_m \times W_m \times A_m$$

Das Erfolgs- und Misserfolgsmotiv stellen die stabilen Personenvariablen dar (McClelland, 1953), welche jede Person in unterschiedlicher Ausprägung besitzt (Atkinson, 1964/1975). Atkinson (1957) definiert das Erfolgsmotiv (Me bzw. HE) als die Fähigkeit, nach einer erfolgreichen Handlung mit Stolz zu reagieren und Leistungsziele anzustreben, da diese aus früheren Erfahrungen mit Stolz antizipiert wurden. Das Misserfolgsmotiv (Mm bzw. FM) wird als die Fähigkeit beschrieben, Scham oder Betroffenheit in Verbindung mit Misserfolg zu empfinden; zurückgehend auf vorangegangene Misserfolge wird Scham mit Leistungssituationen in Verbindung gebracht, woraus das Vermeiden von Leistungssituationen resultiert (Weiner, 1972/1976b).

Die grundlegenden situativen Variablen des Modells bezeichnet Atkinson (1957) als *subjektive Wahrscheinlichkeiten* und definiert sie als Stärke der kognitiven Erwartung, dass sich nach einer Handlung tatsächlich Erfolg bzw. Misserfolg einstellt. Die subjektive Wahrscheinlichkeit von Erfolg (We) und die subjektive Wahrscheinlichkeit von Misserfolg (Wm) betrachtet er als komplementäre Größen (Heckhausen & Heckhausen, 2006). Sie ergeben in ihrer Summe 1 ($We + Wm = 1$), d. h. mit wachsender Erfolgserwartung sinkt die Misserfolgserwartung (Atkinson, 1964/1975). Die *Anreize*, als zweite Situationsvariablen, repräsentieren die wahrgenommene Attraktivität von Erfolg (Ae) und Misserfolg (Am) in einer Leistungssituation (Atkinson, 1957). Der Anreiz des Misserfolgs entspricht dem unangenehm antizipierten Gefühl Scham und ist laut Atkinson negativ, wodurch stets auch eine negative Misserfolgstendenz resultiert.

Wichtige Implikationen des atkinsonschen Modells (1964/1975) sind, dass Anreiz und Wahrscheinlichkeit stets zusammenwirken, d. h. der Anreiz kann als Funktion der Schwierigkeit einer Aufgabe aufgestellt werden kann. Mit sinkender subjektiver Erfolgswahrscheinlichkeit bzw. je schwieriger die Aufgabe ist, desto höher ist der Erfolgsanreiz ($Ae = 1 - We$). Umgekehrt ist bei schwierigen Aufgaben oder geringer subjektiver Erfolgswahrscheinlichkeit der Anreizwert des Misserfolgs ($Am = -Wm$) bzw. das Gefühl von Scham bei einem Misserfolg niedrig (Atkinson, 1957). Die subjektive Wahrscheinlichkeit kann laut Feather (1965) neben dem Anreiz auch durch die Situation determiniert sein, d. h. sie steigt nach Erfolg an und sinkt nach Misserfolg.

Befindet sich eine Person in einer Leistungssituation, werden stets beide Tendenzen angeregt (Heckhausen & Heckhausen, 2006). Ob sich die Person gemäß dem erwähnten Annäherungs-Vermeidungs-Konflikt von der Situation zu- oder abwendet, lässt sich aus der Summe der Erfolgs- und der Misserfolgstendenz erschließen (Atkinson, 1957). Es ergibt sich die *resultierende Tendenz* (Tr), wobei die Misserfolgstendenz nur als hemmende Größe konzipiert wurde (Schneider & Schmalt, 2000):

$$Tr = Te + Tm, \text{ bzw. ausführlicher: } Tr = (Me \times We \times Ae) + (Mm \times Wm \times Am)$$

Mit Hilfe des Modells lassen sich spezifische Voraussagen über das Verhalten von Personen je nach individueller Motivausprägung machen. Atkinson nimmt unter anderem an, wenn das Erfolgsmotiv stärker als das Misserfolgsmotiv ($M_e > M_m$) ausgeprägt ist, sollte die resultierende Motivation stets positiv sein und in einer Annäherung an Leistungssituationen resultieren. Je stärker das Misserfolgsmotiv im Vergleich zum Erfolgsmotiv ($M_m > M_e$) ist, desto negativer die resultierende Tendenz und somit sollte diese Person die Wahl einer Aufgabe meiden. Gemäß Heckhausen (1980) drückt die *Netto-Hoffnung* (NH), berechnet aus der Differenz der zwei Motive, die überwiegende Richtung eines der Motive aus. Weiterhin kann aus der Summe beider Komponenten die *Gesamtmotivation* (GM) ermittelt werden (Heckhausen, 1980). Im weiteren Text wird die Personengruppe mit stark ausgeprägtem Erfolgsmotiv als Personen mit HE und diese mit stärkerem Misserfolgsmotiv als Personen mit FM bezeichnet, wobei betont wird, dass damit nur die überwiegende Motivkomponente beschrieben wird und eine Person, wie oben erläutert, stets durch beide Motive geprägt ist.

Inwiefern ist es jedoch in unserer leistungsorientierten Gesellschaft überhaupt möglich, dass Personen mit dominierendem Misserfolgsmotiv jegliche Leistungssituationen meiden? Atkinson bemerkt 1957 und legt später deutlicher dar, dass die Tendenz, eine Leistungssituation aufzusuchen, natürlich auch von anderen „Quellen positiver Handlungsmotivation“ (Atkinson, 1964/1975, S. 398), wie den Motiven nach Anschluss oder Macht, beeinflusst werden kann. Er fasst die Wirksamkeit dieser extrinsischen Faktoren (Tex) zusammen und addiert sie zur resultierenden Tendenz hinzu ($T_r = T_e + T_m + T_{ex}$), betont aber, dass sie mit dem Leistungsbedürfnis an sich nichts zu tun haben.

Eine rechnerische Annahme Atkinsons (1957) ist, dass bei Aufgaben mittlerer Schwierigkeit ($W_m = W_e = .50$) die Tendenz Erfolg zu erzielen und Misserfolg zu meiden und somit die Affekte Stolz bzw. Scham maximal sind. Demnach sollten laut Atkinson Erfolgsmotivierte zu Aufgaben mittlerer Schwierigkeit neigen und Misserfolgsmotivierte eher sehr leichte oder schwere Aufgaben wählen. McClelland (1956) konnte mit Aufgaben zur Risikowahl bei Ringwurf- oder Worterinnerungsaufgaben bei Kindern zeigen, dass Personen mit hoch ausgeprägtem Erfolgsmotiv eher zu einer mittleren Aufgabenschwierigkeit tendieren. Misserfolgsmotivierte Studenten zeigen in McClellands Experimenten oder auch bei einem Torwurfexperiment von Schneider (1973) nur eine größere Varianz im Vergleich zu Erfolgsmotivierten. Das Ausweichen auf extreme Schwierigkeiten Misserfolgsmotivierter konnten Schneider und Meise (1973) mit dem gleichem Torwurfspiel oder Atkinson und Litwin (1960) bei einem Ringwurfspiel mit Schülern verdeutlichen. Unterschiedliche Reaktionen der Studenten und Schüler, vermutete Schneider (1973), liegen daran, dass die Aufgabe für Studenten womöglich nicht persönlich relevant war und das Leistungsmotiv unzureichend anregte.

Der vermutete Grund, dass Personen mit hohem Misserfolgsmotiv die besonders leichten Aufgaben wählen, ist laut Atkinson (1957) ein Schutz vor Misserfolg durch die geringe Misserfolgswahrscheinlichkeit. Die besonders schweren Aufgaben wählen diese laut Atkinson aufgrund des geringen Gefühls von Selbstverschulden bei Mislingen. McClelland (1956) vermutet, dass die Erfolgsmotivierten durch größeres Selbstvertrauen mittlere vor leichten Aufgaben und mittlere vor sehr schwierigen Aufgaben aufgrund geringer Zuweisung auf eigenes Können bei hohem Schwierigkeitsgrad bevorzugen. Er stellt eine gewagte Übertragung seiner Ergebnisse auf Unternehmer und die wirtschaftliche Entwicklung auf und folgert, dass eine Beziehung zwischen Unternehmer mit höherem Erfolgsmotiv und mittlerer Risikopräferenz und einem gutem Wachstum einer Firma besteht (McClelland, 1956). Die Modellvorhersagen können also auf Alltagssituationen übertragen werden, wie auch eine Studie von Mahone (1960) zeigte, nach der erfolgsoversichtliche Studenten realistischere Berufe mittlerer Schwierigkeit wählten und ihre eigenen Fähigkeiten besser einschätzten. Misserfolgsängstliche hingegen entscheiden sich für sehr leichte oder schwere Berufe.

Die Vorhersagen des atkinsonschen Modells betreffen bis dahin nur die Aufgabenpräferenz, der Autor stellt allerdings auch Ideen zur Ausdauer in einer Leistungssituation auf. Atkinson (1964/1975) nimmt an, dass der negative Anreiz eines drohenden Misserfolgs stets eine negative Tendenz ergibt und somit auch während der Ausführung einer Handlung hemmend wirkt. Für Personen mit FM konnte daher im Vergleich zu Personen mit überwiegend HE schlechtere Leistungen und eine geringere Ausdauer in einem Ringwurfspiel gemessen werden (Atkinson & Litwin, 1960). French und Thomas (1958) beschäftigten sich mit den Unterschieden der Intensität und Dauer der Aufgabenbearbeitung bezüglich der Stärke der Leistungsmotivausprägung. Hochleistungsmotivierte fanden im Vergleich zu Niedrigmotivierten öfter die Lösung für ein komplexes Problem und arbeiteten länger daran, obwohl beide Gruppen im Mittel gleich schnell zur richtigen Lösung kamen. Während Personen mit niedriger GM häufiger die Aufgabe abbrachen, nutzen hochleistungsmotivierte Personen die volle zur Verfügung stehende Zeit. Die Korrelation zwischen Intelligenz und Leistung wurde allerdings nur für die Personen mit hoher GM signifikant. French und Thomas (1958) gehen davon aus, dass der energetisierende Effekt der Motivation und die Zielgerichtetheit der Hochleistungsmotivierten dazu führen, die wichtigen Informationen zu erkennen und härter sowie effektiver zu arbeiten. Auch eine Studie von Atkinsons (1958) bestätigte, dass hochleistungsmotivierte Studenten eine bessere Performance bei Aufgaben mittlerer Schwierigkeit zeigten und ausdauernder bei der Sache sind.

1.1.5 Kritik und weiterführende Theorien zum Risiko-Wahl-Modell

Nach ausführlichen Studien wurde auch Kritik an Atkinsons Modell geübt: Unter anderem erbrachte Schneider (1973) den Nachweis, dass erfolgsmotivierte Vpn Aufgaben mit einer Erfolgswahrscheinlichkeit kleiner als .50 am stärksten bevorzugen. Das Präferenzmaximum liegt laut seinen und weiteren Ergebnissen bei Aufgaben mit der Erfolgswahrscheinlichkeit um .40 und somit bei Aufgaben mit leicht erhöhten Anforderungen. Auch Köppl (2006) und Schleich (2007) konnten widerlegen, dass Personen mit überwiegend HE mittlere Schwierigkeitsgrade bevorzugen. In ihren Arbeiten wählten erfolgsorientierte Vpn eher hohe Schwierigkeitsgrade und die Misserfolgsorientierung hängt mit einer eher niedrigen Anspruchsniveausetzung zusammen.

Weiterhin konnten auch andere Vorhersagen des Modells, wie das Ausweichen auf extrem schwierige oder extrem leichte Aufgaben bei Misserfolgsmotivierten, nicht bestätigt werden. Atkinson postulierte mit seinem Modell eine deutliche Präferenz der extremen Pole, wobei er keine differenzierte Erklärung für die Präferenz der Personen mit überwiegendem Misserfolgsmotiv vornimmt. Studien zur Risikowahl konnten nur tendenziell bestätigen, dass Misserfolgsmotivierte zu leichteren oder schwereren Aufgaben im Vergleich zu Erfolgsmotivierten neigen, wobei diese Unterschiede statistisch nicht signifikant wurden (McClelland, 1956; Schneider, 1973).

Aufgrund dieser und weiterer Kritikpunkte kam es zu vielen Erweiterungsversuchen des Modells, beispielsweise von Atkinson selbst (1964/1975) durch den Aspekt der *Trägheit*. Die Trägheitstendenz bezeichnet einen Motivationsrest, eine konkurrierende Handlungstendenz, welche beispielsweise durch nicht erreichte oder nicht vollendete Handlungen entstanden ist. Bei der Wiederaufnahme der alten Tätigkeit oder von verwandten Tätigkeiten des gleichen Motivs treten Trägheitstendenzen auf und werden deshalb der Erfolgstendenz hinzuaddiert (Schneider, 1973). Weiner (1976b) fügte später zusätzlich der Misserfolgstendenz eine negative Trägheitstendenz hinzu, da nach einer unterbrochenen Handlung auch Misserfolgstendenzen bestehen bleiben können. Durch diesen Aspekt wurde dem Modell von Atkinson die Möglichkeit, Leistungsepisoden zu betrachten, hinzugefügt (Heckhausen & Heckhausen, 2006). Eine Wiederaufnahme oder ein Misserfolg sollte Personen mit HE zu besseren Leistungen motivieren und für Misserfolgsmotivierte eher einen Motivationsverlust und somit eine Verschlechterung der Leistungen bedeuten. Eine Untersuchung von Schneider (1973, S. 35) konnte diese Hypothesen nicht bestätigen: Am Anfang schnitten Erfolgs- und Misserfolgsmotivierte gleich ab und sie verbesserten ihre Leistung stärker unter Misserfolg als unter Erfolg. Ein zweites Experiment unterschied zusätzlich zwischen Leistungsgüte, Bearbeitungszeit und der Ausdauer (Schneider, 1973, S. 43). Dabei konnte gezeigt werden, dass nach einer Misserfolgsinduktion Erfolgsmotivierte schneller arbeiteten

und sich häufiger verbesserten, während Misserfolgsfürchtige sich verschlechterten, d. h. langsamer arbeiteten und häufiger Fehler machten. Nach Erfolg verschlechterten sich Misserfolgs- und Erfolgsmotivierte in ähnlichem Maße. Weiterhin beendeten Personen generell nach Misserfolg die Aufgabe häufiger, wobei mehr Misserfolgsfürchtige im Vergleich zu Erfolgshoffende zu einer nicht leistungsorientierten Aufgabe wechselten (Schneider, 1973, S. 63). Gegen das Konzept der Trägheit spricht auch hier, dass Erfolgsmotivierte in der Misserfolgsbedingung ebenfalls öfter abbrachen, wobei diese in der Bedingung laut Konzept eine stärkere Aufsuchenstendenz zeigen sollten. Diese Effekte können allerdings auch nach Theorien zum Anspruchsniveau oder andere Erklärungsmodellen interpretiert werden (Heckhausen & Heckhausen, 2006).

Neben den kurzfristigen Folgen können sich die Erfolgs- und Misserfolgserfahrungen je nach Leistungsmotiv auch auf langfristige Sicht beispielsweise auf das Attributionsmuster oder das Fähigkeitsselbstkonzept unterschiedlich auswirken. Eine Zusammenfassung der Studien von Weiner (1976a, 1972/1976b) zeigte, wie sich das Verhalten in Leistungssituationen von erfolgs- und misserfolgsmotivierten Personen unterscheidet und wie die Gefühle von Erfolg und Misserfolg auch auf die unterschiedlichen kausalen Schemata der Ursachenzuschreibung zurückgeführt werden können. Daraus entwickelte Weiner (1972/1976b) eine attributionstheoretische Fassung der Leistungsmotivtheorie, welche die atkinsonschen Modellparameter mit vier Ursachenzuschreibungen von Erfolg und Misserfolg (vgl. Tabelle 1.1), gegliedert in personenabhängige Faktoren (internal, external) und Stabilitätsfaktoren (stabil, variabel), verbindet. Weiner (1972/1976b) fand heraus, dass Vpn mit starker HE selbstwertdienlicher attribuieren und Erfolg den internalen Faktoren Begabung und Anstrengung und Misserfolg den variablen Faktoren mangelnder Anstrengung und Pech zuschreiben. Während Erfolgsmotivierte dadurch eine Leistungssituation mit positiven Selbstbewertungsaffekten verbinden, ist sie für Misserfolgsmotivierte weniger attraktiv, da sie Erfolg eher external mit Aufgabenleichtigkeit oder Glück und Misserfolg internal, stabil (Mangel an Begabung) attribuieren (Weiner, 1972/1976b). Eine Studie von Sauer und Gattringer (1986) belegte, dass Leistungsmotivation und die Fähigkeitsattribution neben der Intelligenz der Vpn in geringem Maß zur Varianzaufklärung in einem Schulleistungstest sowie der Schulnoten beitragen.

Tabelle 1.1: Vier Ursachenzuschreibungen von Erfolg und Misserfolg

Stabilität	Personenabhängigkeit	
	Internal	External
	Stabil	Begabung
Variabel	Anstrengung	Zufall (Glück/Pech)

Heckhausen (1972, 1975) integriert in seinem Selbstbewertungsmodell die drei Prozesskomponenten der Leistungsmotivation: Zielsetzung bzw. Anspruchsniveau, Ursachenzuschreibung und Selbstbewertung; er versucht so ausführlicher zu erklären, wie sich die motivationalen Unterschiede der Leistungsmotive über Dauer aufrechterhalten. Er geht davon aus, dass Erfolgsmotivierte die Handlungsdirektive verfolgen, ihre eigene Tüchtigkeit zu steigern und neue Kompetenzen zu erwerben. Sie wählen angesichts ihrer Erfolgsoptimismus anspruchsvolle, aber realistische Ziele. Weiterhin stellt sich durch die nach Weiner beschriebenen Attributionsmuster auch im Falle eines nur ausgewogenen Verhältnisses von Erfolg und Misserfolg bei erfolgsmotivierten Personen der dritte Aspekt – die positiven Selbstbewertungsempfindungen – ein (Rheinberg, 2000). Misserfolgsmotivierte verfolgen gegenteilig die Direktive, Selbstwertbelastungen zu reduzieren und da sie nach dem postulierten Attributionsmuster Leistungssituationen eher mit negativen Selbstwertempfindungen verbinden, werden sie diese eher meiden (Heckhausen & Heckhausen, 2006). Somit können Motive nicht mehr nur als unveränderbares Konstrukt gesehen werden, sondern als ein durch Selbstbegründungen entstehendes, kontinuierlich stabilisierendes System. Für die Anwendung der Motivationspsychologie war dieses Modell besonders fruchtbar. So wurden Motivationstrainings entwickelt, die an den drei Punkten, realistische Zielsetzung, günstigere Attributionsmuster und Selbstwertempfindungen, ansetzten und so nachweislich die Erfolgsoptimismus steigerten (Krug & Hanel, 1976).

An dieser Stelle kann der Bogen zu einem nahe stehendem Forschungsgebiet, dem Selbstkonzept der eigenen Begabung, gespannt werden. Für HE stellte sich ein positiver und für FM ein negativer Zusammenhang mit Selbstkonzepten der Problembewältigung, der Standfestigkeit, der Selbstwertschätzung (Kornprobst, 2005) oder dem Selbstkonzept der eigenen Begabung (Sowa, 2009) heraus. In einer anderen Studie wurde das Selbstkonzept, durch vermeintliche Inkompetenz in einer Aufgabe der sozialen Wahrnehmung, verunsichert. Hochleistungsmotivierte neigten stärker dazu, Selbstergänzungsmöglichkeiten symbolischer Art (Entwurf eines persönlichen Idealprofils) sowie realer Art (soziales Trainingsprogramm) zu wählen (Rheinberg, Schwarz & Singer, 1987). Rheinberg et al. meinten, wenn in einem Bereich eine hohe Leistungsmotivationsausprägung besteht, könne das Anstreben von Erfolg auch als Aufrechterhaltung des eigenen Selbstkonzepts verstanden werden. Fähigkeitsselfkonzepte können demnach leistungsmotiviertes Verhalten beeinflussen. Ein Modell von Meyer (1976) verband das Konzept der eigenen Begabung mit der Informationsgewinnung. Die Theorie der Informationsmaximierung machte die Begabungseinschätzung zum Motivationsziel (Meyer, 1976). Demnach sind Erfolgsmotivierte stärker an Informationen über ihre Fähigkeiten interessiert und wählen mittlere Schwierigkeiten, da die Unsicherheit

über ihr Leistungsergebnis dort am größten ist. Ein Experiment von Schneider (1973) ergab, dass die Entscheidungszeit bei Vpn in einem Torwurfspiel auf die Frage, ob sie das Tor treffen oder nicht bei mittlerer Schwierigkeit am größten ist, wo auch 50 % der Ja-Antworten liegen, allerdings unabhängig vom Leistungsmotiv.

Parallel zur Entwicklung der unterschiedlichen Modelle zur theoretischen Konzeptualisierung des Leistungsmotivs, brachte die Forschung zur Messung des Leistungsmotivs ebenfalls eine Vielzahl von Verfahren hervor.

1.1.6 Die Messung des Leistungsmotivs

In dem Werk „Explorations in Personality“ klassifizierte Murray (1938) nicht nur die Motive, sondern er veröffentlichte auch einen Test, der aus heutiger Sicht die Basis zur Messung von Motiven bildet (Schneider & Schmalt, 2000). Basierend auf der Annahme, dass Menschen dazu neigen, unbewusste bzw. verdrängte Absichten und Gefühle auf andere Personen oder Situationen zu projizieren, entwickelte Murray (1938) den Thematischen Auffassungstest (TAT). Anhand von Bildern und zusätzlichen Leitfragen sollen die Testpersonen spontan Geschichten erzählen und diese Handlungen, Gedanken und Gefühle zu den gezeigten Personen sollen nach Murray Aufschluss über eigene Bedürfnisse bzw. Motive geben. So ein projektives Verfahren lässt den Testpersonen Raum für offene, kreative Antworten und umgeht dabei die von Murray kritisierte Form der Fragebögen, in denen die Versuchsperson (Vp) ihre Antworten beispielsweise nach sozialer Erwünschtheit verstellen können.

Mit der Konkretisierung des TAT durch McClelland und et al. (1953) gelang der Schritt zur Messung des Leistungsmotivs als überdauernde Personenkomponente. Sie verwendeten in einer Studie leistungsbezogene Bilder und variierten vor dem Schreiben der Geschichten die Anregebungsbedingungen durch sechs unterschiedliche Instruktionen vorgelegter Aufgaben (McClelland et al., 1953). Durch die Gegenüberstellung der Ergebnisse der entspannten und der leistungsorientierten Anregebungsbedingung entstand ein standardisierter Auswertungsschlüssel, mit dem die TAT-Geschichten nun auf das Vorliegen von Affekten bei der Auseinandersetzung mit einem Gütemaßstab objektiver analysiert werden konnten (Schmalt & Schneider, 2000).

Während in dem TAT der McClelland-Gruppe das Erfolgs- und Misserfolgsmotiv zu einem Punktwert zusammengefasst wurde, betonte Heckhausen (1963) die Notwendigkeit, sie getrennt voneinander zu messen. Anfänglich wurde versucht, das Misserfolgsmotiv durch Ängstlichkeitsfragebögen zu erfassen (Atkinson, 1964/1975). Erst Heckhausen (1963) verwendete einen TAT mit Erfolg und Misserfolg anregenden Bildern und wertet diese nach einem Auswertungsschlüssel getrennt nach HE und FM aus.

Trotz seiner großen Bedeutung ist der TAT aus Sicht der klassischen Testtheorie stark kritikanfällig. Er gilt als zeitaufwendig, umständlich, weist geringe Re-Testreliabilität und Split-half-Reliabilität auf und benötigt ein hohes Maß an Expertentum in der Auswertung aufgrund der geringen Auswertungsobjektivität (Schneider & Schmalt, 2000). Die Debatte zur Objektivität lässt viele Forscher auf Fragebogenmethoden ausweichen, wobei auffällt, dass beide Verfahren etwas anderes zu messen scheinen, da die TAT-Werte nicht mit den Fragebögen korrelieren (Heckhausen, 1963). Dabei wird die Wichtigkeit beider Methoden und eine Koexistenz von Verfahren zur Erfassung *expliziter* und *impliziter* Motive deutlich (Heckhausen & Heckhausen, 2006). Explizite Motive, welche von der Person selbst bewusst formuliert und durch das Fragebogenverfahren erfasst werden können, zeigen das Selbstbild der Person. Die Handlung selbst wird aber auch durch implizite Motive bestimmt, welche der Person weniger bewusst sind.

Auf der Basis der fruchtbaren Kritik wurden verschiedenste neue Verfahren entwickelt beispielsweise das Leistungsmotiv-Gitter (LM-Gitter) durch Schmalt (1976) – ein semi-projektives Verfahren, d. h. eine Kombination der motivanregenden TAT-Bilder und einem Fragebogen mit leistungsthematischen Aussagen.

Das für diese Studie verwendete Regensburger-Leistungs-Motiv-Inventar (RLMI) ist aus mehrjähriger Entwicklung entstanden und beinhaltet, ähnlich wie Schmalts Leistungsmotiv-Gitter, ein vorgegebenes Antwortschema (siehe Kapitel 3.3.1). In dem RLMI werden allerdings zur Anregung keine Bilder gezeigt, sondern es werden Situationen mit leistungsthematischen Inhalten beschrieben. Das Verfahren misst die bereits durch Atkinson postulierten Aspekte des Leistungsmotivs HE und FM sowie zwei zusätzliche Komponenten *Furcht vor Erfolg* (FE) und *Hoffnung auf Misserfolg* (HM).

1.1.7 Die vier Leistungsmotivkomponenten

Die Idee der Erfolgsfurcht, d. h. Personen die Erfolg mit negativen Folgen assoziieren und vor dem Erreichen eines Ziels schlechte Leistungen erbringen (Canavan-Gumpert, Garner & Gumpert, 1968, S. 8), stammt ursprünglich aus dem klinischen Bereich. Horner (1968) führte basierend auf ihren Forschungsarbeiten, bei denen Frauen im Vergleich zu Männern Erfolg häufiger mit negativen Folgen und Affekten verbanden, das Motiv FE in die Debatte um das Konzept des Leistungsmotivs ein. Nachuntersuchungen konnten die Geschlechtsunterschiede nicht konsistent nachweisen, was nach Heckhausen (1980) vermutlich auf den unterschiedlichen Inhalt der Erfolgsfurcht von Frauen und Männern zurückzuführen ist. Während Frauen die erwarteten extrinsischen Nebenfolgen, wie Verlust von sozialem Anschluss und Weiblichkeit, fürchten, stellen Männer eher Studien- und Berufserfolge in Frage (Heckhausen, 1980). Canavan-Gumpert et al. (1978) konnten für beide Geschlechter nachweisen, dass ein steigender

Erfolg und somit die Weiterentwicklung eigener Fähigkeiten bei Erfolgsfürchtigen Angst und ein charakteristisches selbstsabotierendes Verhalten auslösen kann.

Trotz des Widerspruchs zu den heutigen Idealen einer Leistungsgesellschaft macht es Sinn, den Aspekt FE systematisch nach Vorbild anderer dualistischer Motivationsmodelle mit der Gegenüberstellung HM zu ergänzen. Hierzu argumentieren Lukesch, Kornprobst, Köppl und Peters-Häderle (2008), dass durchaus auch Misserfolg für Personen, die beispielsweise keine Verantwortung übernehmen oder ihr negatives Selbstkonzept aufrechterhalten wollen, erstrebenswert sein kann.

Lukesch und Peters-Häderle (2007) trugen die Annahmen der vier Leistungsmotiv Aspekte zusammen und entwickelten mit dem RLMI ein Verfahren, das neben den klassischen Motiv Aspekten HE und FM auch die FE und HM messbar macht.

Bezüglich der von Horner (1968) eingeführten Diskussion zu den geschlechtsspezifischen Unterschieden der Leistungsmotivkomponenten fielen die Ergebnisse von Untersuchungen nicht eindeutig aus. Peters (2001), Kornprobst (2005) und Eder (2006) konnten zeigen, dass sich für Männer in der HE Skala und Frauen in der FM Skala (Kornprobst, 2005; Peters, 2001) und FE Skala (Peters, 2001) signifikant höhere Werte ergaben, während andere Studien keine geschlechtsspezifischen Unterschiede fanden (Griffone, 1977; Orttner, 2006). Unterschiede sind einerseits durch Besonderheiten der jeweiligen Stichprobe, andererseits durch bereichsspezifische oder bereichsübergreifende Analysen zu erklären (Orttner, 2006).

1.2 Intelligenz

Das Ziel dieses Abschnittes ist es, einen Überblick darüber zu geben, wie in der differentiellen Psychologie Intelligenz definiert wird, welche Entwicklung die Intelligenztestforschung genommen hat und welche bedeutenden Intelligenzstrukturmodelle die Forschung prägten.

1.2.1 Was ist Intelligenz?

Wird versucht, eine Definition für den Begriff *Intelligenz* zu finden, wird deutlich, dass kaum ein anderer Begriff der Psychologie so unterschiedlich aufgefasst wird. Roth (1998a) äußerte, dass es selbst nach 100 Jahren Forschung unmöglich ist, Intelligenz allgemein akzeptiert zu definieren. Auch als Steinberg und Dettermann (1986) in einem Symposium renommierten Experten des Forschungsbereichs die Frage stellten : „What is intelligence?“, konnte keine Einigung erreicht, aber ein Rahmen zur Einordnung der kontroversen Definitionen geschaffen werden. Merkmale, welche die Intelligenzdefinitionen 1986 und auch in einem Symposium 65 Jahre zuvor (Thorndike et al., 1921) am häufigsten beinhalteten, waren die höheren kognitiven Prozesse wie logisches

Schlussfolgern, Problemlösen, Entscheidungsprozesse und Inferenzen (Sternberg & Berg, 1986). Während aber 1986 Metakognitionen, Wissen, kultureller Kontext und die Interaktion zwischen Wissen und Prozessen an zweiter Stelle eine Rolle spielten, nannten die Forscher im Jahre 1921 diese Faktoren eher selten. So wie die Definitionen, was Intelligenz sei, variieren, so verschieden sind auch die Ansichten über ihre Funktionen, wobei sich die Einteilung in drei Funktionsbereiche – biologisch, individuell-psychologisch und sozial-psychologisch, durch mehrere Arbeiten zieht (Meer, 1998; Roth, 1998a; Sternberg & Dettermann, 1986). Biologische Funktionen meinen die Steuerung von Verhalten und die Anpassung an eine ständig wandelnde Umwelt. Die individuell-psychologischen Funktionen wie Wahrnehmung, Gedächtnis und Denken steuern die optimale Zielerreichung. Sozio-psychologische Funktionen umfassen die in einer Gesellschaft oder Kultur als relevant und erfolgreich angesehenen Funktionen. Zusammenfassend meinen Häcker und Stapf (2009), gemein sei den verschiedenen Begriffs- und Funktionsbestimmungen, dass sie Intelligenz als:

[...] die Fähigkeit bezeichnen, sich in neuen Situationen auf Grund von Einsichten zurechtzufinden oder Aufgaben mit Hilfe des Denkens zu lösen, ohne dass hierfür Erfahrung, sondern vielmehr die Erfassung von Beziehung das Wesentliche ist. (Häcker & Stapf, 2009, S. 447)

Neben dem Streben nach einer allgemeingültigen Definition wurden laut Sternberg (1984) noch zwei weitere Wege zur Erklärung von Intelligenz genutzt, welche zu unterschiedlichen theoretischen Tests und Modellen führten.

Die dominierende Herangehensweise, um das Phänomen Intelligenz in der Wissenschaft geeignet zu erfassen (Neisser et al., 1996), lag in der objektiven Messung und statistischen Analyse der Struktur der menschlichen Intelligenz – der psychometrischen Methode (Roth, 1998a). Mittels Faktorenanalysen entstanden pluralistische Konstrukte über zugrunde liegende Dimensionen der Intelligenz, wobei die Verfahren sowie die aufgestellten Strukturtheorien stark variieren. Einerseits dominiert in einigen Modellen ein *Generalfaktor g* (Spearman, 1904), während andere ihren Fokus auf die Charakterisierung unabhängiger Fähigkeiten legen (Thurstone, 1938). Die heute wohl meist akzeptierte Sichtweise ordnet die verschiedenen Fähigkeiten in eine Hierarchie, welche meist einem Generalfaktor untergeordnet sind (Carroll, 1993; Jäger, 1973). Das Bestreben dieser Intelligenzforscher bestand vor allem in „finding and establishing laws of cognition than in measuring individual differences“ (Carroll, 1993, S. 68).

Ein weiterer Weg zu dem Verständnis von Intelligenz war die Suche nach kognitiven Prozessen, die zu Intelligenzleistungen führen (Roth, 1998a). Schwerpunkt der Richtung der Informationsverarbeitung lag in der Erfassung kognitiver Prozeduren wie Inferenzen, heuristische Methoden oder Problemlösestrategien (Meer, 1998), die zu einer Veränderung des Wissens führen. Aus dieser Sicht ist es einseitig, Intelligenz nur

aus der erbrachten Leistung zu messen, da diese „nicht statisch, sondern höchst dynamisch“ (Meer, 1998, S. 174) ist. Dies kritisieren ebenfalls Forscher aus dem pädagogischen Bereich wie Guthke (1974), die das Konzept der Lernfähigkeit als Maß des Entwicklungspotentials einer Person bevorzugen. Diese Richtung ist insofern bedeutend, da sie Trainingsvorschläge macht, um die Stärken einer Person zu fördern und Schwächen zu kompensieren und somit auf Fördermöglichkeiten für das schulische und berufliche Leben aufmerksam macht (Sternberg, 2003). Bereits 1933 betonte Nemzek in einem Überblick über die Konstanz von Intelligenz, dass Re-Test Resultate von der Umwelt, der Ungewohntheit der Testsituation, von Coaching und Training abhängig ist. Petty und Harrell (1977) konnten zeigen, dass kognitive Tests durch die Erhöhung der Motivation, des Wissens über Tests sowie Coaching und die Verringerung von Angst in einer signifikanten Steigerung der Intelligenztestleistungen und Verringerung der Fehler resultieren. Andere „dynamische Wege“ versuchen Intelligenz durch ablaufende, neuronale Prozesse zu erschließen oder beschäftigen sich mit der Entwicklung der Intelligenz (Piaget, 1967).

Diese Vielschichtigkeit in den Debatten um Intelligenz veranlasste das Board of Scientific Affairs der American Psychology Association, 1996 eine Arbeitsgruppe zu gründen, die einen einvernehmlichen Basisbericht des momentanen Stands zur Intelligenzforschung erarbeiten sollte (Neisser et al., 1996). Neben einem Überblick zu bisherigen Intelligenzmodellen und Tests erörtern die Forscher die Frage zwischen der genetischen Determiniertheit von Intelligenz und des Einflusses der Umwelt. Aus heutiger Sichtweise determinieren beide Faktoren gleichermaßen die Entwicklung der Intelligenz (Neisser et al., 1996). Abschließend gehen sie auf Unterschiede zwischen ethnischen bzw. kulturellen Gruppen sowie zwischen Männern und Frauen ein.

Es ist offensichtlich, dass die Diskussionen um Intelligenz zu weit gefächert sind, als dass im Rahmen dieser Untersuchung näher auf alle einzelnen Auffassungen eingegangen werden kann. Es soll jedoch ein Überblick über einflussreiche und speziell für die verwendeten Messinstrumente dieser Arbeit bedeutende Intelligenzmodelle gegeben werden.

1.2.2 Intelligenzstrukturmodelle

Die Entwicklung der Intelligenzmodelle geht einher mit der Entwicklung geeigneter Testverfahren, dabei gilt Binet als „Vater aller modernen Intelligenztests“. Aufgrund eines Auftrags des Erziehungsministeriums zur Entwicklung einer diagnostischen Methode für die Erkennung schwach begabter Kinder erarbeiteten Binet und Simon (1911) eine Testbatterie, welche die erste Normierung zur Einstufung der Intelligenz eines Individuums durch das so genannte Intelligenzalter enthielt. Das Intelligenzalter gibt an,

ob das Kind eine alterstypische Leistung erbracht hat, d. h. es sagt aus, welchem Lebensalter die aktuell erbrachte Leistung entspricht. Stern (1912) setzt Lebens- und Intelligenzalter zur Bestimmung des Intelligenzquotienten (IQ) so miteinander in Beziehung, dass somit auch die Bestimmung der Intelligenz Erwachsener ermöglicht wurde. Wechsler (1965), der den IQ als Abweichungsquotient der individuellen Intelligenz von der Intelligenz der eigenen Alterstufe der Grundgesamtheit definierte, schlug die heute übliche Normierung auf einen Mittelwert von 100 und einer Standardabweichung von 15 vor. Binet war es auch, der im Gegensatz zu einer umfassenden Intelligenz mehrere relativ unabhängige Fähigkeiten unter dem Begriff Intelligenz summierte (Eysenck, 2004). Einer, der diese Sichtweise weiterführte, war Thurstone (1938), während die allumfassende Intelligenzidee von Spearman (1904) vertreten wurde. Nachdem Spearman (1904, S. 205) in seiner „Correlational Psychology“ die Faktorenanalyse in die Intelligenzforschung mit dem Ziel, einzelne Attribute der Intelligenz zu unabhängigen Dimensionen zusammenzufassen und zu charakterisieren, einführte, folgte eine Vielzahl von Intelligenzstrukturmodellen, welche in Art und Zahl der Faktoren variieren.

Unter einer *Intelligenzstruktur* versteht Jäger (1973, S. 22) „das System der theoretischen Begriffe, mit dem der im Menschen liegende Anteil der Bedingungen intelligenten Verhaltens erfasst wird.“, wobei das System die Ordnung und Gesetzmäßigkeiten der einzelnen Fähigkeiten beschreibt. Das mathematische Verfahren der Faktorenanalyse extrahiert basierend auf Interkorrelationen der Testergebnisse zugrunde liegende, übergeordnete Faktoren bzw. Dimensionen, welche jeweils einen Teil der Varianz der Daten erklären (Roth, 1998b). Die *Dimensionen* bezeichnen die einzelnen Komponenten intelligenten Verhaltens, welche alle Menschen in verschiedenen Ausprägungen besitzen (Jäger, 1973, S. 24) und teilweise nur gering miteinander korrelieren. Diese Komponenten sind unterschiedlich umfassend, d. h. sie können durch die Aufgabenart (Thurstone, 1938), durch die Unterschiede psychologischer Prozesse (Guilford, 1963) und über das intelligente Verhalten hinaus Persönlichkeitseigenschaften oder auch Motivation und Erfahrungen charakterisiert sein. Wichtig dabei ist laut Jäger (1973) nicht die Identifizierung aller möglichen Faktoren, sondern das Hervorheben bedeutender Faktoren und deren Einordnung in die drei Geltungsbereiche eines Modells – der Universalität, der Generalität und der zeitlichen Konstanz. Die Universalität bezeichnet die Allgemeingültigkeit der Faktoren eines Modells für ein Individuum oder eine gesamte Population, die Generalität meint die Gültigkeit der Faktoren in verschiedenen Situationen und Variablenstichproben (Jäger, 1973). Die zeitliche Konstanz betrifft laut Jäger (1973) hingegen den dynamischen Aspekt – den Zeitraum der Gültigkeit eines Faktors für eine Person oder eine Gruppe.

1.2.3 Die Generalfaktorentheorie nach Spearman

Die Ergebnisse Spearmans (1904) faktoranalytischer Berechnungen ergaben das Bild, dass seine verwendeten Testreihen positiv miteinander korrelierten und auf einem Faktor luden. Daraus entnahm Spearman, dass jeder intellektuellen Leistung der so genannte Generalfaktor oder *g*-Faktor zugrunde liegt, der ein Maß der allgemeinen „mental power“ (Spearman, 1904, S. 21) einer Person darstellt. Da seine ermittelten Korrelationen nicht maximal und die Testergebnisse einer Person bzw. von zwei Personen mit gleichem *g*-Faktor schwankten, formulierte er neben dem dominierenden Generalfaktor spezifische *s*-Faktoren (Roth, 1998b). Die *s*-Faktoren umfassen die charakteristischen Bedingungen eines Tests wie den Testaufbau und können die Leistungsergebnisse beeinflussen (Guthke, 1974). Spearmans (1904) Formulierung der Zweifaktorentheorie blieb nicht lange unkritisiert, da sich durchaus neben *g* weitere gleichrangige Faktoren ermitteln ließen, welche nicht als *s*-Faktoren abgetan werden konnten (Roth, 1998b). Dennoch gründen die meisten einflussreichen Theorien auf Spearmans Generalfaktor. Der Grundintelligenztest CFT 20-R, welcher als Kontrollvariable in dieser Arbeit diente, ist als Maß für den *g*-Faktor nach Spearman bzw. der General-Fluid-Ability nach Cattell (1963), auf welche später noch eingegangen wird, aufgebaut (Weiß, 2006).

1.2.4 Die Primary Mental Abilities nach Thurstone

Ein Kritiker von Spearmans Ansichten war Thurstone (1938), der den Standpunkt von mehreren unabhängigen Faktoren der Intelligenz vertrat. Er schlägt eine verbesserte Analysetechnik – die multiple Faktorenanalyse – vor, welche auf Gruppenfaktoren basiert (Neisser et al., 1996). Demnach sollten für jede Fähigkeit eine Gruppe mehrerer unterschiedlicher Subtests in die Messung von Intelligenz aufgenommen werden (Eysenck, 2004). Die Tests ordnet Thurstone (1938) nach der Ähnlichkeit der Anforderungen in die Gruppen numerische, verbale und räumliche bzw. geometrische Fähigkeiten. Aus den Interkorrelationsmatrizen der Testergebnisse extrahiert Thurstone die Faktoren so, dass sie eine psychologisch bedeutsame Fähigkeit beschreiben (Roth, 1998b). Dabei geht er von Anfang an von einer begrenzten Anzahl von Faktoren aus (Thurstone, 1938) und entwickelt mit der Faktorrotation ein Verfahren, welches es möglich macht, eine vorher festgelegte Struktur zu erfüllen (Roth, 1998b). Die Idee hinter dieser so genannten *Einfachstruktur* war es, die Faktoren so lange zu rotieren, bis jeder Test auf einem Faktor möglichst hohe und auf den restlichsten möglichst geringe Faktorladungen aufwies (Thurstone, 1938). Auf der Grundlage dieses Verfahrens erhielt Thurstone (1938, S. 79-91) sieben psychologisch interpretierbare Faktoren: *spatial ability* (räumliches Sehen und Vorstellungsvermögen), *perceptual speed* (Wahr-

nehmungsgeschwindigkeit), *numerical ability* (numerische Fähigkeit), *verbal realtions* (verbales Verständnis), *verbal fluency* (Wortflüssigkeit), *memory* (Gedächtnis) und *reasoning* (logisches Erkennen von Regeln oder Prinzipien) und die ungenau definierten, provisorischen Fähigkeiten induktives und deduktives Denken, welche ebenfalls dem Reasoning zuzuordnen sind. Ein allgemeiner Faktor der Intelligenz wie bei Spearman ging aus Thurstones Berechnungen nicht hervor. Das Konzept der sieben Grundfaktoren, auch als Primary Mental Abilities bezeichnet, konnte Thurstone (1938) in einigen Nachuntersuchungen mit dem gleichem Testmaterial bestätigen. Er räumte aber ein, dass bereits durch die Änderung des Schwierigkeitsgrads der Aufgaben oder der Altersgruppe sich das Strukturkonzept verschieben kann. Ein Test des deutschen Sprachraums, welcher aufbauend auf den Primärfaktoren entwickelt wurde, ist das Leistungsprüfsystem LPS nach Horn (1962/1983), dessen überarbeitete neue Fassung in dieser Arbeit verwendet wurde. Kritisiert wurde an Thurstones Modell, dass Thurstones festgesetzte Anzahl der extrahierten Faktoren durch homogenere Stichproben sowie die Anzahl der benutzten Subtests zustande kommt und durch die Wahl mehrerer Tests mit ähnlichen Anforderungen Faktoren auch künstlich erzeugt werden können (Lukesch, 1998).

Da sowohl Thurstone als auch Spearman ihre Modelle teilweise empirisch belegen konnten, liegt es nahe, eine Kombination beider Konzepte – ein hierarchisches Modell – zu integrieren (Eysenck, 2004). Hierarchisch beschreibt eine Beziehung der Faktoren unterschiedlicher Ebenen. Die einzelnen Testaufgaben können auf einer zweiten Ebene zu den von Thurstone postulierten Faktoren zusammengefasst werden und diese unterliegen wiederum dem Generalfaktor. Dabei sind die speziellen Faktoren von den Faktoren höherer Ordnung abhängig, andererseits bedingen die Faktoren höherer Ordnung die spezifischen (Roth, 1998b). Dieses Konzept der Intelligenz ist in vielen neueren Modellen enthalten und aktuell wohl am ehesten akzeptiert (Sternberg, 2003).

1.2.5 Das Zweifaktorenmodell nach Cattell

Der Lösungsvorschlag von Cattell (1963) ist eine zweidimensionale Hierarchie der Intelligenz. Das Zweifaktorenmodell beinhaltet zwei gleichgewichtete Intelligenzfaktoren zweiter Ordnung – die *fluide general ability* g_f und die *crystallized general ability* g_c (Cattell, 1963, S. 2). Fluide Intelligenz beschreibt die allgemeine Fähigkeit höherer mentaler Prozesse und das Lösen komplexer Probleme. Nach Cattell (1963) ist diese Fähigkeit größtenteils biologisch determiniert und nimmt mit dem Alter ab. Die Ergebnisse des bereits im Zusammenhang mit Spearman's Generalfaktor erwähnten Intelligenztests CFT 20-R, können laut Weiß (2006) aufgrund eines „culture-fair intelligence tests“ (Cattell, 1963, S. 5) auch als fluide Intelligenz interpretiert werden. Das Ausmaß der Erlern-

ten Fähigkeiten, des Allgemeinwissens und des Wissens im verbalen und numerischen Bereich, welches durch die Anwendung fluider Intelligenz angeeignet wurde, nennt Cattell (1963) kristalline Intelligenz. Der zweite Faktor wird durch Lernerfahrung abhängig von kulturellen sowie familiären Einflüssen geprägt und bleibt über die Lebensspanne weitestgehend konstant (Cattell, 1963).

1.2.6 Das Berliner Intelligenzstrukturmodell nach Jäger

Jäger verfolgte, neben anderen Intelligenzforschern, das Ziel, die bis dahin entstandene, unüberschaubare Vielfalt der „Faktoreninflation“ (Guthke, 1974, S. 61) in eine Klassifikationsstruktur einzuordnen. Basierend auf bedeutenden Intelligenzmodellen wie von Thurstone oder Guilfords Strukturmodell auf der einen Seite und empirisch überzeugenden Belegen auf der anderen entwickelte er ein eigenes Modell. In dem Berliner Intelligenzstrukturmodell unterteilte Jäger (1973) Intelligenzleistungen nach Operations- und Inhaltsklassen. Weiterhin fand er in seiner Analyse sechs Hauptfaktoren, welche er hierarchisch nach *Anschauungsgebendes Denken, Einfallsreichtum und Produktivität, Konzentrationskraft und Tempomotivation, Verarbeitungskapazität, formallogisches Denken und Urteilsfähigkeit, zahlengebundenen Denken und sprachgebundenes Denken* anordnete (Jäger, 1973). Dabei ging Jäger (1973) anders als Guilford von der Existenz eines Generalfaktors aus und meinte, dass ein hierarchisches Modell, welches sowohl General- als auch Primärfaktoren integriert, am adäquatesten Intelligenz beschreibt. Die Klassifikation der Intelligenz nach Inhalt und Operation erwies sich als gut geeignet, um die verschiedenen Bereiche von Intelligenz zu beschreiben. Sowohl Horn, wie bereits beschrieben, nutzte diese Beschreibung der Inhaltsdimension für das Leistungsprüfsystem (1962/1983), als auch Weiß (2003) ordnet die Subtests des erwähnten Grundintelligenztests CFT 20-R dem Faktor der Verarbeitungsgeschwindigkeit nach Jäger zu.

1.2.7 Das Drei-Ebenen-Modell nach Carroll

Ausgangspunkt von Carrolls (1993) Modell war ebenfalls ein hierarchisches Konzept, wobei er drei Ebenen der Intelligenz beschrieb. Carrolls „monumental contribution“, wie Snow auf der Rückseite von Carrolls (1993) Bucheinband zitiert wird, liegt in einer Reanalyse der „world's literature on individual differences in cognitive abilities“. In dieser fasste er bis dahin berichtete Korrelationsmatrizen zusammen und identifizierte durch neu berechnete Faktoranalysen und Vergleiche an 461 Datensätzen ein breites Spektrum an verschiedenen kognitiven Fähigkeiten des Menschen (McGrew, 2005). Die Annahme einer „three-stratum structure of the mayor cognitive abilities“ konnte durch die Ergebnisse der Reanalyse bestätigt werden (Carroll, 1993, S. 625). Einer-

seits fand er Beweise für die Existenz eines alles dominierenden Faktors, den Spearman postulierte, und der die höchste Ebene in seinem Modell einnahm (Carroll, 1993). Andererseits ergaben sich verschiedene Fähigkeitsfaktoren ähnlich den Primärfaktoren Thurstones. Auf der zweiten Ebene oder auch Stratum II genannt, formulierte er acht Faktoren höherer Ebene (siehe Kapitel 3.3.3.1, Abbildung 3.6). Wie stark diese mit dem Generalfaktor korrelierten, ist unterschiedlich. Der kristallinen und fluiden Intelligenz kam basierend auf Cattells (1963) Einteilung allerdings eine besondere Bedeutung zu, da diese die höchsten Korrelationen zu *g* aufweisen. Zusätzlich identifizierte Carroll (1993) auf der Stratum I Ebene stärker spezialisierte Faktoren wie *Induktion, Wortschatz, Sprach- und Schreibfähigkeiten, die Gedächtnisspanne, numerische Fähigkeiten, Reaktionsgeschwindigkeit* oder *das Erkennen von räumlichen Beziehungen*. Diese listete er jeweils unter einem Faktor der zweiten Ebene auf, bemerkte aber, dass diese mehrere Stratum II Faktoren beeinflussen können. Die Faktoren der Stratum I Ebene werden direkt durch die Tests erhoben und repräsentieren die Effekte des Lernens und Erfahrung bzw. kognitiver Strategien. Generell dient das Modell vor allem der Einordnung verschiedener Fähigkeiten in ein System gemäß dem Aufgabentyp, Material und der Generalität des Faktors. Es bildet einen Rahmen, in den auch kognitive Prozesse und die Informationsverarbeitung involviert sind und in die Interpretation einbezogen werden können (Carroll, 1993). Carroll proklamierte: „The model is regarded as a close approximation to the way which abilities actually operate in the ‚real world‘.“ (Carroll, 1993, S. 632). Das LPS-neu, zu dessen Validierung die vorliegende Arbeit beiträgt, lehnt Kreuzpointer (2010) an das Drei-Ebenen Modell an (siehe Kapitel 3.3.3) und kritisiert Horns (1962/1983) Einordnung des LPS in die Modelle von Guilford und Jäger.

1.2.8 Intelligenz – ein komplexes Phänomen

Die beschriebenen Intelligenzstrukturmodelle, welche mit zu den bedeutendsten Vorschlägen zählen, stellen nur einen Auszug der bisher entwickelten dar. Weiterhin sind auch Vernons Einteilung in educational und practical intelligence, Jesens Erweiterung des hierarchischen Konzepts auf das Verhalten und neurophysiologische Prozesse oder Horns Erweiterung von Cattells fluiden und kristallinen Intelligenz erwähnenswert (Carroll, 1993). Interessant ist ein neueres multifaktorielles Modell – die Cattell-Horn-Carroll Theory (CHC), welches die bekanntesten psychometrischen Vorschläge von Carroll und Cattell bzw. Horn integriert und sehr gut empirisch belegt ist (McGrew, 2005).

Sternbergs Meinung zu psychometrischen Modellen ist: „The time perhaps has come to expand our notion and everyone’s notion of what is intelligent.“ (Sternberg,

2003, S. 151). Während er 1985 in der *triarchic theory* neben analytischer Intelligenz die praktische und kreative Intelligenz hervorhob, unterschied er 2003 Intelligenz und *successful intelligence*. Die *successful intelligence* umfasst die Fähigkeiten, welche wichtig für das Anpassen, Formen und Wählen der Umwelt sind (Sternberg, 2003). Dabei ging es ihm vor allem um die Erklärung intelligenten Verhaltens, da eine intelligente Person mit geringer *successful intelligence* im Leben, gemessen an den eigenen Standards, wenig erfolgreich sein kann (Sternberg, 2003).

Kann durch diese postulierten Intelligenzstrukturmodelle eine Antwort auf die Frage, was Intelligenz ist, gegeben werden? Roth (1998b, S. 28) fasste die Ergebnisse der Forschung nach einer Struktur menschlicher Intelligenz zusammen und zog verschiedene Schlüsse: Zwar ließ sich kein einheitliches Intelligenzmodell entwickeln, dennoch stimmen die Forscher weitestgehend darin überein, dass Intelligenz keine Funktionseinheit ist, sondern „ein komplexes Phänomen“ (Roth, 1998b, S. 28) darstellt, welches nur verschiedener Anordnungen und unterschiedlicher Generalität unterliegt. Weiterhin meinte er, dass die jeweilige postulierte Modellstruktur abhängig von den untersuchten Stichproben und Tests ist und betont die Wichtigkeit von objektiven Messungen und Faktorenanalysen bei der Entwicklung der Strukturmodelle. Darüber hinaus kritisierte Roth (1998b), dass viele Modelle wichtige Implikationen, die Intelligenzleistungen bedingen können, wie die Situation, Persönlichkeitsmerkmale, die Entwicklung der Intelligenz über die Lebensspanne und der Ablauf intelligenten Verhaltens vernachlässigen.

1.3 Der Einfluss von Misserfolgsstress in Leistungssituationen

In der modernen Leistungsgesellschaft gehört Stress in Form von ständigem Notendruck für Schüler, Jobauswahlverfahren wie Assessment Center und immerwährender und rapider Wandel in der Arbeitswelt zum Alltagsbild. Statistiken von 2005 der European Agency for Safety and Health at Work unterstreichen die negativen Auswirkungen von Stress in Leistungssituationen: 22 % der Arbeiter von 27 Mitgliedstaaten der EU gaben an unter Gesundheitsproblemen durch arbeitsbedingten Stress zu leiden (Milczarek, Schneider & Gonzalez, 2009). Die damit verbundenen mittel- und langfristigen Kosten, wie die Weltgesundheitsorganisation (Ivanov, 2005) berichtet, spiegeln sich in wachsenden Gesundheitskosten, geringerer Produktivität, erhöhter Mitarbeiterfluktuation und reduzierter Profite von Unternehmen wieder. Bereits 1974 äußerte Selye

Stress ist jedoch unser ständiger Begleiter, solange wir leben. [...] dennoch verdanken wir ihm jeden persönlichen Fortschritt [...]. Er ist die Würze unseres Lebens. Nur der Distress ist jedermanns Feind. Er kann rücksichtslos auf uns einschlagen und trifft nicht nur unseren Körper, sondern auch den Geist und kann uns mit vielen Mitteln zermürben und fertigmachen. Vor ihm allein müssen wir uns hüten. (Selye, 1974, S. 23)

In der vorliegenden Untersuchung wird untersucht, in welchem Ausmaß eine Stresssituation die Leistung beeinflussen kann. Um die Auswirkungen von Stress auf Leistungen zu verstehen, ist es wichtig, aus psychologischer Sichtweise zu analysieren, wie unterschiedliche Personen auf unterschiedliche Stressoren reagieren und ob Stress dabei auch positive Folgen haben kann. Dafür soll zunächst eine theoretische Grundlage zum Verständnis des Stressbegriffs geschaffen werden.

1.3.1 Der Begriff Stress

Das Stresskonzept wird in der Literatur unterschiedlich reflektiert, wobei drei Theorie-tendenzen zu unterscheiden sind (Laux & Vossel, 1982). Der Fokus eines Zweigs der Stressforschung liegt darin, Stress stimulus- bzw. umweltbedingt zu betrachten (Heinrich & Spielberger, 1982), wobei die individuellen Unterschiede der Stressreaktion in Intensität, Qualität und Dauer oft vernachlässigt werden. Eine Person kann eine Situation als Gefahr deuten, während eine andere Person dieselbe Situation als Herausforderung sieht oder eine dritte diese als gänzlich unbedeutend betrachtet (Lazarus & Launier, 1978).

Andererseits sind es umgekehrt die individuellen Unterschiede der Stressreaktion wie die Bewertungen oder Vulnerabilität einer Person, die nach einer zweiten Gruppe der Stressforscher unabhängig von den situationalen Gegebenheiten Stress determinieren. Selye (1974), häufig als „Vater der Stressforschung“ bezeichnet, ist ein Vertreter dieser Sichtweise. Er definierte Stress aus biologischer Perspektive als „unspezifische Reaktion des Körpers auf jede Anforderung“ (Selye, 1974, S. 58), welche nur durch Kampf oder Flucht gelöst werden kann. Aus seiner Sichtweise ist der Stresszustand oder die „biologische Belastung“ (Selye, 1974, S. 18) durch die Aktivierung des vegetativ-endokrinen Systems gekennzeichnet. Selye hob den Unterscheid zwischen negativem *Distress* und positiven *Eustress* hervor, d. h. Stress kann durch bedrohliche und erfreuliche Ereignisse entstehen und Unlust- sowie Lustgefühle erzeugen.

In einer dritten, umfassenderen Sichtweise werden individuelle und situationale Faktoren bzw. deren Interaktionen gleichermaßen berücksichtigt. Ein bedeutendes Beispiel der interaktionalen Perspektive ist das *transaktionale Stresskonzept* nach Lazarus (1966). Lazarus und Launier (1978) beschrieben Stress folgendermaßen:

[...] the term 'Stress' is any event in which environmental or internal demands (or both) tax or exceed the adaptive resources of an individual, social system, or tissue system. (Lazarus & Launier, 1978, S. 296)

Neben dem Einbeziehen beider Faktoren erfuhr auch die einseitig biologische Betrachtung nach Selye (1974) in Lazarus Modell eine Erweiterung. Ein weiterer Unterschied von Lazarus Konzept zu Selyes Ausführungen ist das Eingrenzen der Stressoren auf

bedrohliche Ereignisse. Lazarus Modell (1966) ist in der Psychologie wohl das bis heute einflussreichste Stresskonzept und wird in folgendem Kapitel näher erläutert.

1.3.2 Das transaktionale Stresskonzept nach Lazarus

Das transaktionale Stressmodell nach Lazarus (1966) geht von einer komplexen und dynamischen Wechselbeziehung zwischen Person und Umwelt bei der Entstehung von Stress bzw. emotionalen Zuständen aus. Bereits Schachter und Singer (1962) betonten in Bezug auf die Entstehung von Emotionen im Allgemeinen die Rolle der Bewertungen des physiologischen Zustands. Die empirischen Arbeiten von Lazarus und Launier (1978) heben dabei neben der starken Rolle der kognitiven Bewertungen bei dem Prozess der Stressentstehung auch die individuellen Fähigkeiten zur Stressbewältigung – das *Coping* hervor. Lazarus (1966) beschreibt zwei sich gegenseitig beeinflussende Bewertungen, welche sich anders als bei Schachter und Singer (1962) nicht auf die bloße Erklärung des physiologischen Zustands beziehen. Die *primären Bewertungen* identifizieren, inwiefern ein Ereignis schädigend, positiv oder irrelevant für das Wohlbefinden einer Person ist. Die drei stressrelevanten Belastungen sind dabei Schädigung bzw. Verlust, Bedrohung und Herausforderung. Die *sekundäre Bewertung* überprüft, ob das Individuum genug Ressourcen zur Bewältigung bereitstellen kann. Die Stressbewältigung besteht aus den psychischen und physischen Bemühungen, interne und umweltbedingte Anforderungen oder Konflikte, welche eigene Ressourcen strapazieren oder übersteigern, zu steuern. Umweltbedingt bezieht sich auf externe Ereignisse, welche eine Anforderung erforderlich machen und bei Misserfolg negative Folgen mit sich ziehen. Die internen Anforderungen sind eigene Ziele, Werte oder Motive, deren Vereitelung, abhängig von Art und Stärke der Motivation, ebenfalls negative Konsequenzen haben würde (Lazarus, 1966). Er benennt drei Faktoren der Beschaffenheit von Bedrohung, welche die Bewertung beeinflussen: 1.) die Balance von der Stärke der Bedrohung und der Ressourcen der Person, 2.) die antizipierte Schädigung und Nähe der Konfrontation und 3.) die Ungewissheit in einer Gefahrensituation. Abschließend findet nach der versuchten Stressbewältigung eine *Neubewertung* statt. Die Neubewertung kann wiederum die primären und sekundären Bewertungen, basierend auf der Effektivität und Konsequenzen der vorhergegangenen, eigenen Reaktion, verändern (Lazarus & Launier, 1978). Die Rückmeldungen verdeutlichen, dass sich das Individuum in Lazarus Konzept in ständiger Auseinandersetzung mit der Umwelt befindet.

Die spezifische Art der individuellen Stressbewältigung scheint neben der Intensität und Stärke von Stressepisoden ausschlaggebend für die Gesundheit zu sein (Lazarus, 1966). Coping kann sich in der direkten Handlung, der Hemmung der Handlung, in der

Informationssuche oder anderen psychischen Prozessen äußern (Lazarus & Launier, 1978). Diese dienen einerseits dazu, die Folgen der Stresssituation zu lösen, neu zu bewerten oder zu minimieren. Andererseits kann Coping darauf abzielen, Stresseemotionen wie Angst, Ärger, Frustration oder Hilflosigkeit zu regulieren (Lazarus & Launier, 1978). Intensive Emotionen können momentan ablaufende Prozesse stören, aber eine erfolgreiche Bewältigung kann die negativen Folgen von Stress vereiteln.

Lazarus benennt vier mögliche Kategorien von Stressreaktionen (Lazarus, 1966, S. 6). Eine Kategorie der Stressreaktionen sind die störenden Affekte wie Angst, Ärger und Depressionen (Sarason, Mandler & Craighill, 1952; Taylor & Rechtschaffen, 1959). Tremor, erhöhter Muskeltonus und Sprechschwierigkeiten fasst er unter der Gruppe der motorischen Stressreaktionen zusammen, welche beispielsweise Lantz (1945) untersuchte. Physiologische Reaktionen wie Veränderungen der Atmungs- oder der Herzaktivität, der Transpiration oder andere hormonale Veränderungen bilden die dritte Gruppe, welche eingehend von Selye (1974) erörtert wurden. Für diese Studie ist vor allem die vierte Kategorie, der sich verändernden kognitiven Funktionen, interessant, welche Performance und kognitive Aktivität wie Denken, Wahrnehmung, Bewertungen oder Probleme lösen fördern und hemmen können (Lazarus, Deese & Osler, 1952).

Später bezeichnet Lazarus (1991) das Modell aufgrund der schwierigen Abgrenzung zwischen Stress und Emotionen als Emotionstheorie. Hier kristallisiert sich eine unterschiedliche Sichtweise zur Beziehung von Emotion und Motivation im Vergleich zu Heckhausen (1963) und Atkinson (1964/1975) heraus. Während Atkinson Emotionen als Motivation antizipierend ansieht, kommen nach Lazarus Emotionen eine stärkere Bedeutung zu und Motivation wird eher mit Kognitionen in Verbindung gebracht.

Ziel dieses Abschnitts war es vor allem, ein Verständnis zur Entstehung und den Folgen von Stress zu entwickeln, um die Auswirkungen von Stress auf Leistungen analysieren zu können. Nachfolgend soll nun auf unterschiedliche Studien der psychologischen Forschung eingegangen werden, welche versuchten, Stress durch Instruktionen zu erzeugen, und dabei Leistungsveränderungen erfassten.

1.3.3 Techniken der Stresserzeugung

Die offenkundige Tatsache, dass Menschen in der heutigen westlichen Gesellschaft oft fortwährend Leistungen unter stressigen Bedingungen erbringen müssen, erklärt das große Forschungsinteresse bezüglich der Auswirkungen von Stress. Der Einfluss von Stress auf die Leistung von Individuen ist daher von großer theoretischer und praktischer Bedeutung. Wie konnten aber Stresssituationen experimentell erzeugt werden, um die Effekte von psychologischem Stress adäquat zu untersuchen?

Die Manipulation der *Instruktion*, welche in der experimentellen Psychologie eine wesentliche Variable darstellt, diente auch in der Stressforschung häufig als situative Variable. Unter dem Begriff Instruktion versteht man eine Anweisung, die zur „Übermittlung der Aufgabenstellung an die Versuchsperson bei psychologischen Untersuchungen“ (Häcker & Stapf, 2004, S. 445) dient. Sie kann in vielfältiger Weise, ob verbal, nonverbal, schriftlich oder durch räumliche Gegebenheiten, eingesetzt werden und kann laut Häcker und Stapf (2004) sachliche und motivierende Hinweise liefern. Rotter (1955) erörterte externe Einflüsse der *psychologischen Situation* auf das menschliche Verhalten und räumte ein, dass für eine systematische Untersuchung neben der Instruktion die Komplexität der Situation bedeutend ist, d. h. der Testungsort, der Versuchsleiter und die Zeit müssen beachtet werden. Weiterhin spielen die Persönlichkeit und die Werte und Erfahrungen einer Person eine wichtige Rolle (Rotter, 1955).

Das prinzipielle Problem der experimentellen Stressforschung, geeignete realistische Stresssituationen durch die Instruktion zu schaffen, brachte eine Vielzahl von Techniken der Stresserzeugung hervor. Diese unterschieden sich bezüglich der Stressart, der Instruktionen, der untersuchten subjektiven Stressreaktionen, der physiologischen Reaktionen, der Stichprobe und schließlich auch in den Effekten auf die Leistungen der Vpn. Rotters (1955) geforderte systematische Herangehensweise bei der Untersuchung der psychologischen Situation fand dabei wenig Beachtung. Mit dem Ziel, eine Ordnung in diese Vielfalt hineinzubringen, klassifizierten Lazarus et al. (1952) die Studien nach der verwendeten Technik in durch die Aufgabe verursachten Stress und durch Misserfolg erzeugten Stress. In der Literatur finden sich auch komplexere Aufteilungen von Stressexperimenten. Beispielsweise präsentierte McGrath (1982) eine Einteilung von etwa 200 Studien nach der Bedrohung des Selbstwerts, physikalischer und interpersoneller Bedrohung, welche sich wiederum in eine Vielzahl von Untergruppen aufgliedern. Aufgrund der besseren Übersichtlichkeit sind die folgenden Ausführungen an der Klassifizierung nach Lazarus et al. (1952) orientiert.

In der ersten Gruppe der stresserzeugenden Studien nach Lazarus et al. (1952) wurde versucht, Druck oder Ablenkung durch die Rahmenbedingungen der Aufgabe zu schaffen. Dabei wurde die Situation durch starke sensorische Einflüsse wie laute Geräusche oder Elektroschocks, Ablenkungen oder durch die Schwierigkeit, Komplexität und die Art der Aufgabe manipuliert. Mit steigender Lautstärke (Schulz & Schönplflug, 1982) oder durch angekündigte Elektroschocks und Warnleuchten (Williams, 1947) erhöhte sich die Unsicherheit und die Fehlerzahl der Vpn und es werden weniger korrekte Antworten gegeben. In einer weiteren Studie konnten Schulz und Schönplflug (1982) zeigen, dass unter Zeitdruck oder erhöhtem Schwierigkeitsgrad die Vpn schneller arbeiten, aber die Qualität der Leistung darunter leidet. Bartmann (1963) ermittelte,

dass durch Zeitdruck die Bearbeitungszeit und Genauigkeit bei komplexen Denkprozessen wie Problemlöseaufgaben, aber nicht bei einfachen Ordnungsaufgaben, litten. Er interpretierte dies so, dass unter Zeitdruck die Zielgerichtetheit des Denkhandelns verloren geht und Personen unsicherer bei Lösungsschritten werden. Zuschauer, implizite Konkurrenten oder Geldpreise führten in einer Studie von Baumeister (1984) zu einem Leistungsabfall. Später nannten Baumeister und Showers (1986) vier Möglichkeiten Druck auszulösen: durch eine Belohnung oder Bestrafung, die Anwesenheit von Bewertern oder einer Vergleichsgruppe, die Relevanz der Leistung für das eigene Ego und die Wahrscheinlichkeit keine zweite Chance zu bekommen.

Die meisten Studien verwendeten die zweite Technik, eine aus Misserfolg bzw. drohenden Misserfolg bestehende psychologische Instruktion. Lazarus et al. (1952) nannten dabei drei übliche Arten, um Misserfolg zu erzeugen. Einerseits werden in Untersuchungen unlösbare Aufgaben benutzt (Diggory, 1949), während in anderen Studien die Technik zum Einsatz kam, die Vp vor der Beendigung der Aufgabe zu unterbrechen, welches bei Kendler (1949) in Verbindung mit einer Intelligenztest-Instruktion zu schlechteren Erinnern führte. In der letzten Gruppe wurde Stress durch suggerierten Misserfolg anhand von angegebenen Normen relevanter Vergleichsgruppen (Lantz, 1945) oder eigener vergangener Leistungen (Thorndike & Woodyard, 1934) induziert. Indem die Aufgabe als Intelligenztest angekündigt wurde und die Vpn die Normergebnisse einer fiktiven Studie erhielten, die Erfolg bzw. Misserfolg vermittelten, wurde bei Covin, Donovan und Macintyre (2003) Stress erzeugt.

Da die Misserfolgs-Techniken in der Forschung am häufigsten angewendet wurden und scheinbar am effektivsten waren, fiel auch in dieser Studie die Entscheidung auf eine solche Instruktion. Der nächste Abschnitt verschafft daher einen ausführlicheren Überblick zu ausgewählten Stressexperimenten mit Misserfolgsinstruktion und deren Ergebnisse auf die Leistungen von Individuen.

1.3.4 Die Auswirkungen einer Misserfolgsinstruktion auf die Leistung

In der Stressforschung nennt Selye (1974) Misserfolgsbeurteilungen bzw. den Stress der Enttäuschung im Arbeitsleben als stärksten Stressor der Schaffensfreude beim Arbeiten. Misserfolg kann viele Facetten haben. Einerseits kann er lediglich nicht mit dem erhofften Leistungsergebnis übereinstimmen, andererseits kann er das Herabsetzen des eigenen Status oder Selbstwerts bis hin zum beruflichen Scheitern bedeuten.

Einige der ersten Forscher, welche durch Misserfolg Stress induzierten, waren beispielsweise Sullivan (1927), Thorndike und Woodyard (1934) oder Sears (1937). Thorndike und Woodyard (1934) prognostizierten für die Leistungsergebnisse unter Stress einen eindeutigen Effekt:

It is a matter of common knowledge that a mind which for any reason becomes engaged in an activity and finds itself repeatedly and persistently failing therein, is impelled to intermit or abandon it. (Thorndike & Woodyard, 1934, S. 241)

Bei schwierigen Mathematik- und Anagrammaufgaben mit ständigen Misserfolgsresultaten im Vergleich zu einer vorangegangenen einfacheren Serie konnten sie diese Voraussage bestätigen (Thorndike & Woodyard, 1934). Die Vpn bearbeiteten weniger Items, arbeiteten langsamer und zeigten eine geringe Lernleistung im Vergleich zum Pre-Test sowie im Vergleich zu erfolgreichen Vpn. Die Studie von Sears (1937) ergab, dass Misserfolg im Vergleich zu einer Gruppe mit Erfolgs-Feedback die Leistungseffizienz, d. h. die Schnelligkeit in einer Kartensortierungsaufgabe, und den erwarteten Lernerfolg in einer wiederholten Worterinnerungsaufgabe beeinträchtigte. Der Versuchsleiter manipulierte die Ergebnisse im Vergleich zu vorhergehender eigener Leistung und der anderer Teilnehmer. Laut Sears (1937) hemmt die durch den Misserfolg entstehende Frustration künftige Lernleistungen, während Erfolg zu stetigen Verbesserungen führte. Sears benannte vier Arten von Leistungsvergleichen, um das Gefühl von Misserfolg zu erreichen: Der Vergleich der aktuellen Leistung mit der direkt vorherigen sowie mit dem Durchschnitt aller vorhergegangenen Erfahrungen, mit eigenen Erwartungen und mit anderen Gruppenteilnehmern. In Sullivans (1927) Studienreihe führte bei einer Lernaufgabe und bei einem Memoryspiel die Instruktion, zu den besten Teilnehmern zu gehören, zum schnellsten Lernen und vollständigem Wiedererinnern. Nach der Misserfolgsinstruktion zu den Schlechtesten zu zählen, benötigten Vpn die längste Zeit und erinnerten am wenigsten, auch im Vergleich zur Kontrollgruppe. Interessant in der weiteren Analyse war, dass die Ergebnisse in der separaten Betrachtung zweier Intelligenzgruppen gleich ausfielen, wobei Misserfolg die hoch intelligenten Vpn und Erfolg die wenig intelligenten am stärksten beeinflusste (Sullivan, 1927).

Die große Mehrheit der empirischen Untersuchungen berichten ebenfalls von Verschlechterungen der Leistungen unter Stressinstruktionen. Alper (1946) fand heraus, dass nach einer unlösbaren Aufgabe und einer ego-involving Instruktion, d. h. nachdem den Vpn mitgeteilt wurde, es handle sich um einen wichtigen Eignungstest, die Vpn weniger korrekte Lösungen in einer Satzformierungsaufgabe fanden. Bei einem Versuch von McKinney (1933) wurde den Personen mitgeteilt, dass Vpn mit durchschnittlicher Intelligenz die bevorstehende Aufgabe in einer bestimmten Zeit lösten. Diese Gruppe kam in einer Labyrinth- und Multiplikationsaufgabe seltener zu den richtigen Lösungen (McKinney, 1933). Er erklärte, dass durch die entstandenen Emotionen aktiveres Verhalten zu beobachten sei und diese in geringem Ausmaß stimulieren und motivieren, aber bei Leistungstests durch eine Verringerung der intellektuellen Kontrolle nicht von Vorteil seien. Teilnehmer in dem Versuch von McClelland und Apicella (1947) benötigten nach falsch mitgeteilten Misserfolgsresultaten eine längere Bearbei-

tungszeit in einer Kartensortierungsaufgabe. Die durch Misserfolg entstehende Spannung hemmte laut McClelland und Apicella (1947) anschließende Leistungen, da nach weiteren, erfolgreichen Aufgaben deren Leistungen sogar über die der Kontrollgruppe anstiegen. Nach Misserfolg sank ebenfalls das Lernen und Erinnern von sinnfreien Silben, wobei diese Defizite sich nach einem erneuten, erfolgreichen Versuch wieder ausglich (Zeller, 1950). Eine alternative Erklärung der Reduktion von Misserfolgseffekten bot Zeller (1950), welcher die Leistungsverminderung durch die Verdrängung des Erlernten nach Misserfolg begründete, da sich die Defizite wieder ausgleichen ließen. Einige Forscher untersuchten auch den für diese Studie besonders relevanten Effekt von Stress auf die Ergebnisse in einem Intelligenztest. Bei einem Versuch mit dem Stanford-Binet Intelligenztest wurde Misserfolg durch eine zwischengeschaltete Ballaufgabe erzeugt, wodurch eine höhere Variabilität der Ergebnisse nach dem Feedback erreicht wurde (Lantz, 1945). Misserfolg führte zu langsamerem Arbeiten und weniger korrekten Antworten. Unterschiede zwischen den einzelnen Subtests zeigten, dass Stress einige Fähigkeiten wie verbale Aufgaben stärker hemmte und visuelle Erinnerungen relativ unbeeinflusst blieben. Dass Misserfolg auch zu einer primitiveren selektiven Wahrnehmung von Sätzen, d. h. das Wahrnehmen von mehr bedeutungslosen Sätzen und keinen Verbesserungen im Re-Test führte, wiesen Postman und Bruner (1948) nach. Ein Faktor, welcher ebenfalls durch Misserfolg beeinflusst wird, wie in Kapitel 1.1.2 beschrieben, ist das typische Herabsetzen des Anspruchsniveaus, welches wiederum die Leistung beeinflussen kann (Hoppe, 1930). In einem Versuch, in dem dies gezeigt wurde, informierte der Versuchsleiter die Teilnehmer vorab, wie viele andere Studenten bzw. Hauptschüler die Aufgabe durchschnittlich schafften, brach aber vor dem Vollenden der Aufgabe den Test ab (Schleich, 2007).

Seltener fanden Forscher gegenteilige Belege bzw. keine Unterschiede, nachdem sie Stress erzeugten. McKinney (1933) konnte keine signifikanten Unterschiede der Fehlerzahl oder der Bearbeitungszeit der Stress- und Kontrollgruppe in einer Multiplikationsaufgabe finden. Auch Schleich (2007) fand keine Veränderungen nach Misserfolg in der Arbeitsgeschwindigkeit bei einer motorischen Aufgabe. In den Studien von Thorndike und Woodyard (1934) konnte der Leistungsabfall nach Stress nicht immer bestätigt werden, sondern es fanden sich auch leistungskonstante Gruppen. Wie bereits erläutert, hemmte Stress in einem Experiment von Lantz (1945) einige Fähigkeiten eines Intelligenztest stärker als andere. Keine signifikanten Unterschiede nach Erfolg bzw. Misserfolg berichteten auch Vogel, Raymond und Lazarus (1959) sowie Covin et al. (2003). Einen Leistungsanstieg nach Stress verzeichneten vor allem mit Persönlichkeitskorrelaten kombinierte Studien, welche die Vpn beispielsweise nach unter-

schiedlicher Ängstlichkeit oder Leistungsmotiven untersuchten, auf die im nächsten Kapitel näher eingegangen werden soll.

Weiterhin wurden in den Studien auch andere Stressreaktionen wie Verhaltenskorrelate, Emotionen oder physiologische Reaktionen gemessen. Da in dieser Arbeit das Wohlbefinden und die Bewertungen der Vpn untersucht wurden, seien hier einige dieser Resultate präsentiert. Nach Misserfolg sanken in Fragebögen eines Experiments von Lantz (1945) die angegebene Aufmerksamkeit, Selbstbewusstsein, Wohlfühlen, Ausdauer, Anstrengung und Freundlichkeit. Hingegen waren Impulsivität, physische Aktivität, Haltungsspannung und Kommentare zu entfliehen erhöht (Lantz, 1945). In einer anderen Studie entnahmen McKinney (1933) oder Zander (1944) aus den Berichten und Verhalten der Vpn nach Frustration bei Erinnerungsaufgaben mit Zahlen oder Silben, dass diese weniger aufmerksam, ängstlicher, selbstunsicherer und motorisch aktiver waren. In dem Versuchen von Lantz (1945), Hoppe (1930) oder Postman und Brunner (1948) äußerten die Vpn häufiger Kommentare zu entfliehen oder mit aggressiven bzw. das Experiment oder den Versuchsleiter herabsetzenden Inhalt.

In einer abschließenden Betrachtung wird deutlich, dass die meisten Studien, obwohl sie unterschiedliche Techniken, Arten und Stärken von Stress verwendeten, eine Abnahme der Leistungseffizienz unter Stress bestätigten. Einige Studien berichteten von keinen Unterschieden und größerer Variabilität der Ergebnisse nach Misserfolg. Selten diskutierten die Autoren den speziellen Einfluss der Art der Aufgabe, den Zeitpunkt der Stressinstruktion, die Stressbedingungen oder die beeinflussten Leistungskomponenten. Schneider (1973) räumte ein, um die Resultate genauer interpretieren zu können, sei es wichtig neben der Menge auch die Leistungsgüte, d. h. Fehler, Bearbeitungszeit usw., gezielt zu analysieren, welches in dieser Arbeit Beachtung fand.

Ein Problem, auf das Lazarus et al. (1952) bezüglich der Misserfolgsstudien hinweist, ist, dass oft Persönlichkeitsmerkmale nicht berücksichtigt wurden. Viele Studien erwähnten Selbstwertunterschiede (Alper, 1946), Stabilität der Persönlichkeit (Williams, 1947), Extrovertiertheit (Zander, 1944), das Selbstbewusstsein oder die Intelligenz (Sullivan, 1927) als Indikator für Leistungsunterschiede am Rande, kontrollierten diese aber nicht explizit. Lazarus (1966) diskutierte, dass nicht Misserfolg selbst, sondern was ein Individuum damit verknüpft, ausschlaggebend ist. Demnach können „kurzfristige Stresseinwirkungen zu einem Gewinn oder zu einem Verlust führen“ (Selye, 1974, S. 110). Personen unterscheiden sich darin, ob und wie stark sie von Misserfolg des untersuchten Aufgabentyps beeinflusst werden, wobei ihre Motivation, Erfolg zu haben oder Misserfolg zu meiden, sowie intellektuelle Fähigkeiten, der Selbstwert und die Attribuierung eine große Rolle spielen (Lazarus et al., 1952). Die Leistung kann daher nicht als alleiniger, direkter Indikator für die Richtung und Stärke des Stresseinflusses

herangezogen werden. Ausgehend von diesem Standpunkt wird offensichtlich, dass es sinnvoll ist, Stress in Bezug zu individuellen Persönlichkeitsvariablen wie die motivationalen Merkmale von Individuen zu analysieren.

1.4 Empirische Befunde zum Einfluss der Leistungsmotive und Stress auf die Leistung

Stress ist nur ein sekundäres Konzept, welches basierend auf dem primären Konzept der Motivation und der Situation, in welcher motiviertes Verhalten auftritt, aufgebaut ist (Lazarus et al., 1952). Ist das Erreichen eines Zieles oder Motivs in Gefahr, sollte Stress entstehen. In der vorliegenden Arbeit wird untersucht, in welchem Ausmaß die Testsituation die Leistung in einem Intelligenztest je nach zugrunde liegendem Leistungsmotiv beeinflusst. Im Anschluss werden daher ausgewählte Studien, die sich mit Leistungsmotivation und Stress in Leistungssituationen befassen, vorgestellt.

1.4.1 Empirische Befunde zu Stressinstruktionen und Leistungsmotivation

Das motivationale Verhalten einer Person kann durch externe Einflüsse beeinflusst werden, diese bezeichnet Rotter (1955) als psychologische Situation. Wie in Kapitel 1.1.6 erläutert, verwendeten bereits McClelland et al. (1953) in der Leistungsmotivationsforschung bei der Überarbeitung des TAT verschiedene Settings, um einen situativ angeregten Zustand der Leistungsmotivation zu erzeugen. Während in der neutralen Bedingung keinerlei spezielle Instruktion gegeben wurde, gab der Versuchsleiter in der entspannten Bedingung den Vpn vor, Informationen für seine Abschlussarbeit zu sammeln. In der leistungsorientierten Bedingung informierte er sie wie folgt:

Thus, in addition to general intelligence, [...] these tests demonstrate whether or not a person is situated to be a leader. The present research is being conducted under the auspices of the Office of Naval Research to determine just which individuals possess the leadership qualifications shown by superior performance on the test. (McClelland et al., 1953, S. 102)

Als weitere Instruktion in der leistungsorientierten Bedingung wurden den Probanden Durchschnittsnormen zum Vergleich ihrer Ergebnisse vorgelegt, dabei wurde einer Gruppe Erfolg und einer andere Misserfolg mitgeteilt. Eine letzte Gruppe erhielt abwechselnd Erfolgs- sowie Misserfolgsresultate. Die Ergebnisse der folgenden TAT-Messung zeigten, dass die Leistungsmotivation von entspannt zu leistungsorientiert, wie auch von neutral zu leistungsorientiert anstieg (vgl. Tabelle 1.2). Ebenso ergab sich nach Misserfolg mehr leistungsthematische Rückmeldung als nach Erfolg, wobei die Kombination von Erfolg und Misserfolg die meisten thematischen Nennungen auf-

wies. Dabei nahmen McClelland et al. (1953) an, dass Erfolg das Bedürfnis nach Leistung befriedigt und Misserfolg es anregt.

Tabelle 1.2: Häufigkeit leistungsthematischer Vorstellungen in TAT-Geschichten in Abhängigkeit unterschiedlicher Anregungsbedingungen (nach McClelland et al., 1953, S. 184)

Anregungsbedingung	n	M	SD
Entspannt	39	1.95	4.30
Neutral	39	7.33	5.49
Leistungsorientiert	39	8.77	5.31
Erfolg	39	7.92	6.76
Misserfolg	39	10.10	6.17
Erfolg-Misserfolg	39	10.36	5.67

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung

Bezüglich der Leistungen bei Denk- und Problemlöseaufgaben stellten McClelland et al. (1953) eine Leistungssteigerung bei niedrig- und hochleistungsmotivierten Vpn in einer leistungsorientierten Situation fest, wobei Hochmotivierte besser abschnitten. In einer ähnlichen leistungsorientierten Situation bestätigten Atkinson und Reitman (1956) dieses Muster, wobei sich kein Unterschied in einer Multi-Anreizsituation ergab. Eine Studie von Petty und Harrell (1977) konnte zeigen, dass eine solche leistungsmotivierende Instruktion die aktuelle Leistungsmotivation erhöhte und bessere Leistungen und weniger Fehler in einem Intelligenztest hervorbrachte.

Andere Studien konzentrierten sich auf den Richtungsaspekt der Leistungsmotive: Erfolgsmotivierte wiesen geringe signifikant bessere Leistungen unter Zeitdruck auf, während für Misserfolgsmotivierte keine Unterschiede in den Situationen deutlich wurden, diese aber signifikant erhöhte Probiertätigkeit und verfrühte Fertigmeldungen aufwiesen (Bartmann, 1963). Das Verbessern durch Misserfolgrückmeldungen und das Gleichbleiben der Leistungen nach Erfolg von erfolgsorientierten Vpn sowie das gegenteilige Bild der misserfolgsorientierten konnten Stapf, Fischer und Degner (1986) nachweisen. Dabei hat die Leistungsmotivation einen ähnlichen hohen Einfluss auf Schulleistungen wie die Intelligenz (Steinmayr & Spinath, 2009). Das Konzept der Gesamtleistung beeinflusste die Schulleistung dabei doppelt so stark wie das HE-Konstrukt, wobei auch Ziele, Valenzen und Fähigkeitsselbstkonzepte einen Teil der Varianz in Schulleistungen erklärten (Steinmayr & Spinath, 2009). Allerdings konnte die Leistungsmotivstruktur in anderen Arbeiten nur einen Bruchteil der Varianz von Leistungsunterschieden aufklären. Bei Schneider (1973) verbesserten sich sowohl Personen mit HE als auch Personen mit FM nach einer Misserfolgsmittelung.

Ein anderer Zweig der psychologischen Forschung wendete sich den individuellen Leistungsunterschieden in Stresssituationen zu und beeinflusste die Studien zur Leistungsmotivation. In der Angstforschung zeigten Taylor und Rechtschaffen (1959), ausgehend von Taylors Manifest Anxiety Scala (MAS) zur generellen Ängstlichkeit, dass ängstlichere Vpn durch eine erhöhte Ängstlichkeitserregung – dem so genannten *drive* – schlechtere Leistung erbrachten und langsamer lernten. Empirische Studien von Sarason und Mitarbeitern gingen mit dem selbstentwickelten Test Anxiety Questionnaire (TAQ) speziell auf die Angst in Leistungssituationen ein. Testängstliche erzielten in dem Wechsler-Bellevue Intelligenztest (Mandler & Sarason, 1952) und auch bei einem mathematischen sowie Schulleistungstest (Sarason & Mandler, 1952) schlechtere Ergebnisse, schienen sich aber nach einiger Zeit an die Situation zu gewöhnen und steigerten ihre Leistungen. Zusätzliche experimentelle Sets, wie Erfolgs- und Misserfolgssituationen (Mandler & Sarason, 1952) oder eine ego-involving Instruktion (Sarason et al., 1952), d. h. den Vpn wurde mitgeteilt, dass es sich um einen Intelligenztest handle, führten ebenfalls zu schlechteren Leistungen der Testängstlichen, jedoch zu den besten Ergebnissen der wenig Ängstlichen, wobei in einer nonego-involving und einer neutralen Instruktion die Unterschiede nicht signifikant wurden. Schulz und Schönplugh (1982) berichteten, dass bei einfachen Aufgaben ängstliche Vpn den nicht ängstlichen überlegen waren. Weiterhin ergaben sich häufig negative Korrelationen zwischen Testängstlichkeit und Intelligenz, wobei diese meist nur gering ausfielen (Mahone, 1960; I. G. Sarason, 1959; Sarason & Mandler, 1952; Taylor & Rechtschaffen, 1959).

Einer der ersten Forscher, der die Resultate des Forschungsgebiets zur Ängstlichkeit und Leistungsmotivation verband, war Raphelson (1957), daher war diese kurze Erörterung bedeutsamer experimenteller Befunde der Angstforschung sinnvoll. Wie in Kapitel 1.1.6 beschrieben, gab es in der Leistungsmotivforschung lange kein separates Instrument um die Misserfolgsschreck zu messen, sondern diese wurde mit niedrigen Leistungsmotivwerten beispielsweise im TAT in Verbindung gebracht. Mandler und Sarasons (1952) Fragebogen zur Testängstlichkeit setzten viele an Stelle als Indikator für FM ein (Atkinson & Litwin, 1960; Mahone, 1960; Raphelson, 1957; Wegner, 1973), da laut Atkinson und Litwin (1960) Testängstlichkeit ein ähnlich definiertes Konstrukt wie FM sei. Personen mit niedrigem Leistungsmotiv und hoher Angst, welche der Gruppe mit FM entsprachen, wiesen ein Absinken der Leistung und Erhöhen der Hautleitfähigkeit im Verlauf der Aufgabenbearbeitung auf, während sich für Personen mit hoher Leistungsmotivation und geringer Angst das Gegenteil ergab, ihre Leistungen erhöhten sich und die Hautleitfähigkeit nahm ab (Raphelson, 1957). Die höchsten Hautleitfähigkeiten wurden bei Erfolgsoversichtlichen vor den Tests ermittelt, was laut Raphelson (1957) darauf hinweist, dass neben Angst auch die Aktivierung und das

Interesse eine Rolle spielen. In einer leistungsorientierten Situation werden beide Konstrukte angeregt und weisen zueinander eine hohe negative Korrelation auf (Raphelson, 1957), nicht jedoch in einem neutralen Setting (Atkinson & Litwin, 1960). Bezüglich der Leistung ergab sich bei Wegner (1973) das Bild, je höher die Leistungsmotivation, desto höher waren die Ergebnisse und je niedriger die Ängstlichkeit, desto niedriger fielen die resultierenden Ergebnisse aus.

Die Reaktion der Personen mit FM oder der Testängstlichen scheint abhängig von Komplexität, Schwierigkeit, Aufgabenart, der Situation und der Intelligenz zu sein. Schneider und Gallitz (1973) wiesen nach, dass unter bestimmten Umständen wie einer neutralen Situation oder bei einfachen Aufgaben die Unterschiede beider Gruppen nicht signifikant wurden, sondern nur bei schweren Tests. Misserfolgsmotivierte, klassifiziert mit dem TAQ und TAT, hatten in Feathers (1965) oder Heckhausens (1963) Experimenten bessere Leistungen als Erfolgsmotivierte bei einfachen Additionsaufgaben, während sich bei komplexen das Gegenteil herausstellte. Eine Studie zur Testängstlichkeit ermittelte, dass auf einem niedrigerem Intelligenzlevel Testängstliche schneller die Lösungswege fanden, während in der Gruppe der höher Intelligenzen die weniger Ängstlichen besser abschnitten (Ruebush, 1960). Ruebush räumte ein, dass Testängstliche mit stärker erlernter defensiver Reaktion bei einfachen Aufgaben, wo Vorsichtigkeit wie häufiges Nachschauen sich positiv auswirkte, einen Vorteil hatten. Der Unterschied zwischen testängstlichen und weniger testängstlichen Vpn konnte in einem fachgemäßen Intelligenztest, jedoch nicht in einem spielähnlichen Intelligenztest nachgewiesen werden (Lighthall, Ruebush, Sarason & Zweibelson, 1959).

Zusammenfassend neigen Personen mit FM bei schwierigen und komplexen Aufgaben, wo Konzentration und Problemlösestrategien gefragt sind, sowie bei verbalen Instruktionen, welche die Selbstaufmerksamkeit erhöhen, Stress bzw. Misserfolg erzeugen, zu schlechteren Leistungen im Vergleich zu Personen mit HE. Ob und in welchem Maße Leistungsunterschiede zwischen den Leistungsmotivkomponenten zum tragen kommen, scheint von diesen entscheidenden Randbedingungen abzuhängen.

1.4.2 Integration der Erklärungsversuche der empirischen Befunde

Aufbauend auf den beschriebenen Studien, wird nun erörtert, wie einzelne Autoren die Hintergründe der unterschiedlichen Leistungsveränderungen abhängig vom Leistungsmotiv einer Person erklären. Die Interpretationen von Emotionen und Motivationen vermischen sich dabei oft, da wie bereits in Kapitel 1.1.2 und 1.1.3 erörtert, starke Motivation häufig mit starken Emotionen einhergehen. Angst kann somit nicht nur als Antrieb, sondern auch als motivationale Komponente der Persönlichkeit in Form von FM interpretiert werden. Dies erschwert die Betrachtung der Rolle des Leistungsmotivs

auf die Leistungen unter Stress, weshalb die Erklärungen zu Testängstlichkeit hinzugenommen wurden.

Forscher wie Atkinson beschäftigen sich mit dem Richtungseinfluss von HE und FM in Leistungssituationen. In seinem Leistungsmotivmodell vermutet Atkinson (1957), dass eine Verringerung der subjektiven Wahrscheinlichkeit bei schweren Aufgaben nach Misserfolg für Erfolgsmotivierte zu einem Absinken des Anreizes und bei einfachen Aufgaben zu einer Anreizsteigerung führt. Bei Misserfolgsmotivierten ergibt das Sinken der subjektiven Wahrscheinlichkeit bei schweren Aufgaben nach Misserfolg eine Verringerung der Misserfolgsmotivation und die resultierende Motivation steigt folglich an; während dies bei einfachen Aufgaben zu einer Erhöhung des Misserfolgsmotivs führt (Atkinson, 1957). Wie bereits in den Kritiken zu Atkinsons Modell aufgezeigt, entsprechen diese Voraussagen nicht immer den empirischen Ergebnissen, weshalb auch andere Theorien der Leistungsmotivforscher herangezogen werden sollen.

Laut Lowell (1952) führt die größere Anstrengung Hochleistungsmotivierter im Vergleich zu weniger Motivierten im Aufgabenverlauf zu einer Leistungssteigerung, wobei beide Gruppen anfänglich gleiche Leistungen erbringen, was zeigt, dass diese ihre Möglichkeit, sich zu verbessern, mehr ausschöpfen. Er argumentierte, dass dies nicht durch unterschiedlich hohe Intelligenz zustande kam, da kein signifikanter Unterschied bei Intelligenztest zu ermitteln war und nur einige Subtest signifikant wurden (Lowell, 1952). Außerdem waren leistungsmotivierte Vpn im Vergleich zu niedrigmotivierten zwar in einem verbalen Intelligenztest besser, beide erbrachten bei einer unbekanntem Wortaufgabe jedoch zu Beginn gleiche Leistungen. Bei gewöhnlichen Additionsaufgaben waren leistungsmotivierte Vpn von Beginn an besser, obwohl keine Unterschiede der numerischen Intelligenz gefunden wurden (Lowell, 1952).

Neben den Leistungsmotivtheorien treten oft aus der Angstforschung entspringende Erklärungsversuche bezüglich der Leistungsunterschiede nach Stress in Kombination mit den Leistungsmotiven auf. Die häufig auf Stress folgende Angst interpretiert Taylor (1958) als *drive*, welcher motivierend und energetisierender wirkt und zu Leistungsveränderungen führt. Der Schlüssel einer Leistung ist laut den Drive- bzw. Antriebtheorien „the performer's level of arousal“ (Baumeister & Showers, 1986, S. 363). Eine solche Theorie beschreibt, dass Leistung und Erregungsniveau in einem umgekehrten U-förmigen Zusammenhang stehen (Yerkes & Dodson, 1908). Das so genannte Yerkes-Dodson-Gesetz besagt, dass bei einem optimalen Erregungsniveau auch optimale Leistung folgt, wobei Unterforderung oder Überforderung, d. h. zu hohe oder zu niedrige Erregung, mit einer Leistungsverschlechterung einhergehen. Bei einfachen Aufgaben tritt weitestgehend eine Verbesserung durch Stress auf, je schwerer die Aufgaben werden, desto schneller ist jedoch das optimale Erregungsniveau durch

gaben werden, desto schneller ist jedoch das optimale Erregungsniveau durch stärkere Angst und Motivation erreicht (Yerkes & Dodson, 1908). Lazarus und Eriksen (1952) erklären den häufigen Leistungsabfall durch extrem hohe Motivation, welche zu schnellerem Arbeiten und folglich zu mehr Fehlern führt. Während nach dem Yerkes-Dodson-Gesetz bzw. dem Gesetz der umgekehrten U-Funktion hohe Angst und hohe LM zu schlechten Leistungen führen, sind es laut Atkinson und Litwin (1960) hohe Angst und niedrige LM.

Negative Korrelationen zwischen Testängstlichkeit oder FM und Intelligenz, Schulnoten oder schulrelevanten Fähigkeitstests (Mahone, 1960; I. G. Sarason, 1959; Sarason & Mandler, 1952) werfen die Frage auf, ob misserfolgsfürchtige Vpn grundsätzlich weniger intelligent sind oder die zugrunde liegende Intelligenz durch die Misserfolgsschuld nicht adäquat gemessen werden kann. Es besteht die Annahme (Heckhausen, 1963; Weiner, 1972/1976b), dass durch schlechte, erfolglose Erfahrungen eher FM entwickelt wird bzw. bei guter Begabung und Leistungsbewährung das HE-Motiv verstärkt wird und folglich die Intelligenz die Entwicklung motivationaler Strukturen bedingt. Andererseits kann laut Heckhausen (1963) und Weiner (1972/1976b) durch eine stark ausgeprägte Misserfolgsschuld leistungsorientiertes Handeln gemieden und somit die intellektuelle Entwicklung gehemmt werden, während die Intelligenzentwicklung und somit gute Begabung durch eine eher aufgaben-aufsuchende Erfolgsmotivation gefördert werden kann. Mandler und Sarason (1952) oder Mahone (1960) glauben, dass Angst in der Kindheit durch Misserfolg in Testsituationen erlernt ist und bei stark ängstlichen Vpn durch ausgelöste, aufgabenirrelevante Reaktionen Testleistungen hemmt. FM hängt somit nicht mit intellektuellen Fähigkeiten an sich zusammen, wird aber durch die Intelligenztestsituation geweckt und kann somit die momentane Leistung behindern (Heckhausen, 1963).

Die meisten Forscher stützen diese letzte These. Leistungsvorteile der Gruppe mit FM oder der Testängstlichen unter entspannten Bedingungen oder bei einfachen Aufgaben führen Forscher (Heckhausen, 1980; Mandler & Sarason, 1952; Stapf et al., 1986) zu dem Schluss, dass mehr die motivationalen als die kognitiven Voraussetzungen zu den Unterschieden zwischen HE und FM oder zwischen Testängstlichen führen. Die Befunde von Bartmann (1963) unterstützen die These, dass Misserfolgsmotivierte bereits durch eine schwierige Aufgabe angespannt reagieren. Da er nur Korrelationen zwischen Schulbegabung und Leistung bei weniger komplexen Aufgaben ohne Zeitdruck nachweisen konnte, schlussfolgerte er, dass nur bei Aufgaben dieser Art intellektuelle Fähigkeiten am besten zur Geltung kommen würden. French und Thomas (1958) fanden die Korrelation zwischen Leistungsmotivation und Intelligenz für hochleistungsmotivierte Vpn, aber nicht für niedrigmotivierte. Sie interpretieren, dass Hoch-

leistungsmotivierte sich so sehr anstrengen, dass Intelligenz die Unterschiede dieser Gruppe definiert, während die Niedrigmotivierten sich nicht stark anstrengen und somit nicht ihr volles Fähigkeitspotential nutzen. I. G. Sarason (1959) meint, dass intellektuelle Fähigkeiten am meisten die Varianz in kognitiven Leistungen erklären, da sich meist nur geringe Korrelationen zwischen FM und Intelligenz ergeben und sich für generelle Angst keine signifikanten Ergebnisse zeigen (Raphelson, 1957). Schulz und Schönplflug (1982) interpretieren die längeren Orientierungszeiten bei Aufgabenbeginn und das längere, wiederholte Verharren mit Aufgabeninformationen ängstlicher Vpn im Vergleich zu neutralen Situationen und im Vergleich zu weniger ängstlichen Vpn mit erhöhter Unsicherheit dieser unter Stress. Dass dies nicht durch höhere Erregung, welche bei einfachen Aufgaben fördernd und bei schweren hinderlich ist, zu Stande kommt, können sie durch niedrigere Herzfrequenzen der Ängstlichen in beiden Settings belegen. Schulz und Schönplflug (1982) meinen, dass Ängstliche mehr Kontrollmechanismen entwickelt haben, durch die sie in einfachen Situationen gute Leistungen erzielen können, aber welche die Anforderung schwieriger Aufgaben übersteigt.

I. G. Sarason (1960) kommt nach einer Diskussion verschiedener Studien zu dieser Thematik und Mandler und Sarason (1952) basierend auf ihren experimentellen Studien zu dem Schluss, dass Leistungen bei Testangst oder FM durch eine Ablenkung von relevanten Informationen und eine Zuwendung zu dominierenden, aufgabenirrelevanten und auf sich selbst bezogenen Informationen behindert werden. Lantz (1945) Teilnehmer waren bei einem Intelligenztest unter Stress häufiger geistig abwesend als die Erfolgsgruppe und beschäftigten sich mit Gedanken bezüglich des Aufrechterhaltens des Selbstwerts. Die Vpn von Lantz (1945) sowie von Postmann und Brunner (1948) äußerten häufiger, dass sie gern der Situation entfliehen oder die Aufgabe abbrechen würden. Auch Baumeister (1984) unterstützte mit seinen Studien zur Selbstaufmerksamkeit diese Hypothese und erklärt, dass durch die Aufmerksamkeit auf eigene Leistungsprozesse automatisch ablaufende Prozesse und aufgabenrelevante Gedanken gestört werden. Irrelevante Kognitionen sind nach Mandler und Sarason (1952) Gefühle der Hilflosigkeit, der eigenen Inkompetenz, antizipierte Misserfolgfolgen, Gedanken der Situation zu entfliehen sowie zum Status- und Selbstwertverlust.

Differenzierter erörterte Heckhausen (1982) den Inhalt der irrelevanten Gedanken bezüglich der Leistungsmotive. Erfolgshoffende gaben an, eher selbstvergessener zu sein, schienen weniger stark beeinflusst durch irrelevante Gedanken und hatten bessere Noten. Misserfolgsängstliche fühlten sich durch ihre Kognitionen wie die Folgen der Prüfungsergebnisse, Ursachen des Misserfolgs, das erwartete Ergebnis und emotionale und sich selbst betreffende Gedanken in ihrer Leistung behindert. Da Selbstzweifel negativ mit der Prüfungsnote zusammenhängen, meint Heckhausen (1982), dass der

Grund für schlechtere Noten die unpassenden Gedanken und nicht die geringere Intelligenz sei. Daraus könnte die Schlussfolgerung gezogen werden, dass Personen mit FM schlechter im Coping von irrelevanten Gedanken sind. Eine Untersuchung von Orttner (2006) wies nach, dass Personen mit HE seltener uneffektive Copingstrategien wie Grübeln, Resignieren oder soziale Isolation im Vergleich zu Vpn mit überwiegend FM anwendeten, während beide Gruppen gleich häufig positive Strategien nutzten. Personen mit FM (Orttner, 2006) oder auch Ängstliche (Houston, 1982) waren oft stärker von Stress beeinflusst, da das häufige Einsetzen von unpassendem Coping das Ausmaß der Stressbelastung erhöhte. Durch adäquates Coping nahm die Testangst in Leistungssituationen bei erfolgsorientierten Vpn eine umgekehrte U-Form an, bei misserfolgsorientierten wurde jedoch ein stetiger Anstieg verzeichnet (Becker, 1982).

Leistungsmotivgruppen unterscheiden sich demnach in ihrer Reaktion auf Erfolg und Misserfolg darin, dass die Bewertung der eigenen Fähigkeiten in Leistungssituationen bei misserfolgsfürchtigen Vpn selbstwertgeladene Kognitionen auslöst. Lazarus und Eriksen (1952) heben bei der Entstehung von schlechten Noten und Leistungen neben der Intelligenz den Einfluss von Stresssituationen hervor, da sich bei Konstanthaltung der Intelligenz Vpn mit gutem Notendurchschnitt nach Stress verbesserten, während sich Vpn mit weniger guten Noten verschlechterten.

Ein Zitat von Lowell (1952) soll die Diskussion um die motivationalen und kognitiven Erklärungen von Leistungsunterschieden in Stresssituationen zusammenfassen:

It is just reasonable to argue on the basis of these data alone that the differences in achievement motivation produced differences on an achievement test and on the experimental task, as to argue that intelligence ‚produced‘ the differences in achievement motivation and in output on the experimental task. (Lowell, 1952, S. 39)

Mentale Blockaden, irrelevante Gedanken, durch selbstwertschützende Maßnahmen oder schlechtes Coping werden als alternative Erklärung für die schlechtere Leistung von misserfolgsfürchtigen Vpn in Stresssituationen diskutiert. Dabei ist es schwer möglich, wie auch Lazarus et al. (1952) meinten, die Effekte von Emotion, Intelligenz und Motivation in Stresssituationen zu trennen. Beispielsweise können einige aus Selbstschutz motiviert, andere emotional abgelenkt sein und bei wieder anderen können sich beide Effekte ausgleichen. Es besteht eine breite Variabilität in den Reaktionen sowie in den Ursachen der Stressreaktionen und eine entscheidende Rolle spielen dabei die Leistungsmotivkomponenten.

Ausgehend von den erörterten theoretischen Grundlagen und der empirischen Ergebnisse der drei Themengebiete folgt im nächsten Kapitel die Formulierung der forschungsleitenden Fragen und der zu prüfenden Hypothesen dieser Arbeit.

2. Hypothesen und Fragestellungen

In der vorliegenden empirischen Arbeit werden zwei Hauptziele verfolgt. Einerseits gilt es darzustellen, wie sich psychologischer Stress auf die Ergebnisse eines Intelligenztests auswirkt und andererseits welche Rolle dabei die Ausprägung des Leistungsmotivs der Personen spielt. Zusätzlich werden Zusammenhänge des Wohlbefindens, der Bewertung der Intelligenztests und des subjektiv eingeschätzten Stresslevels zwischen den experimentellen Bedingungen sowie der Leistungsmotive untersucht.

2.1 Einfluss der Stressinstruktion

In dieser Studie wurde, um Stress zu erzeugen, basierend auf den in Kapitel 1.3.4 erörterten Studien, eine Misserfolgsinstruktion verwendet, wobei angenommen wird, dass diese Manipulation erfolgreich verlief. Die Daten der Stressstudien hinsichtlich charakteristischer Züge des Befindens der Vpn weisen durchgehend darauf hin, dass sich Vpn nach Stressinstruktionen unwohl fühlen, weniger aufmerksam und kooperativ sowie stärker angespannt und nervöser sind (Lantz, 1945; McKinney, 1933; Zander, 1944). Wie bereits erwähnt, wurde in dieser Arbeit neben den Leistungsergebnissen gleichzeitig das Wohlbefinden mittels einzelner Gefühlsreaktionen erfasst. Unter der Voraussetzung einer erfolgreichen Instruktion wird entsprechend den genannten Studien die folgende Hypothese angenommen.

H1: Vpn in der neutralen Bedingung fühlen sich im Vergleich zu Vpn der Stressbedingung nach der Stressinstruktion und nach dem LPS-neu Intelligenztest wohler.

2.2 Leistungsmotivation – Wohlbefinden, gefühltes Stresslevel und Bewertung der Intelligenztests

Eine klassische Annahme der Leistungsmotivationsforschung von Atkinson (1957) ist, dass bei starkem Misserfolgsmotiv im Vergleich zum Erfolgsmotiv die resultierende Motivationstendenz negativ ist. Leistungssituationen haben einen geringen Anreiz und werden daher eher gemieden, während sie für Erfolgshoffende einen großen Anreiz bilden. Personen mit starker „Hoffnung auf Erfolg“ (HE) weisen zusätzlich in Leistungssituationen einen selbstwertdienlichen Attribuierungsstil auf, während Testsituationen für Personen mit starker „Furcht vor Misserfolg“ (FM) weniger attraktiv sind, da sie Erfolg mit externalen Ursachen und Misserfolg mit geringer Begabung verbinden (Weiner, 1972/1976b). Zudem stellen Personen mit FM zu hohe Forderungen an sich selbst. Da sie ihre Fähigkeiten unrealistischer einschätzen und sich unrealistischere Ziele setzen (Atkinson & Litwin, 1960; Mahone, 1960), entstehen eine starke Leistungsdiskrepanz

und stärkere Misserfolgsaffekte (Atkinson, 1964/1975). Befunde von Bartmann (1963) verdeutlichen, dass die bloße Ankündigung von Intelligenztests misserfolgsfürchtige Vpn unter Spannung bringt. Zusätzliche Misserfolgs- bzw. Stressinstruktionen beeinflussen Testängstliche bzw. Personen mit FM stärker als erfolgshoffende Vpn (Raphelson, 1957; Sarason et al., 1952) und laut Feather (1965) berichten diese von stärkeren Enttäuschungen über ihre Misserfolge. Gemäß diesen Ergebnissen sollten Personen mit FM Leistungs- und vor allem Stresssituationen mit negativen Affekten verbinden, demnach werden folgende Hypothesen formuliert.

H2: Vpn mit überwiegend HE fühlen sich im Vergleich zu Vpn mit überwiegend FM während der gesamten Testung wohler, wobei größere Unterschiede nach der Stressinstruktion erwartet werden.

H3: Vpn mit überwiegend FM fühlen sich im Vergleich zu Vpn mit überwiegend HE nach der Stressinstruktion gestresster.

Hinsichtlich der Bewertung der getätigten Aufgaben und Tests der Studien wurden meist keine speziellen Fragebögen, sondern nur Verhaltensbeobachtungen durchgeführt. Die Tendenz, nach Misserfolg die Schuld von sich selbst auf eine externe Sache abzuschieben, beobachten Hoppe (1930) und Lantz (1945); Postman und Bruner (1948) bemerken häufiger negative bzw. aggressive Bemerkungen bei diesen Vpn. Die Teilnehmer geben in ihren Versuchen nach Misserfolg oft an, dass das Gerät nicht in Ordnung oder die Aufgabe zu schwer sei, sie setzen den Wert der Aufgabe herab oder beschäftigen sich nur spielerisch mit der Aufgabe. Hoppe erklärt, dass diese Aufgabenbewertungen als eine Art Selbstschutz, um „das ‚Ich‘ möglichst hoch zu halten“, verstanden werden können (Hoppe, 1930, S. 35). Aufgrund dessen und da Personen mit FM Intelligenztests oft mit negativen Emotionen assoziieren, sollen in dieser Untersuchung die folgenden Fragestellungen überprüft werden.

F1: Bewerten Vpn mit überwiegend HE im Vergleich zu Vpn mit überwiegend FM die Intelligenztests positiver?

F2: Bewerten Vpn in der neutralen Bedingung im Vergleich zu Vpn der Stressbedingung den LPS-neu positiver?

2.3 Intelligenztestleistungen und Leistungsmotivation

Grundsätzlich geht aus der erläuterten Theorie hervor, dass Personen mit HE oder hoher Leistungsmotivation den Personen mit Misserfolgsschreck bzw. niedriger Leistungsmotivation in Leistungssituationen oft überlegen sind. Wie die Empirie und auch ein Überblick von Rheinberg (1976) über eine Vielzahl von Studien zwischen Schul- und Studienleistung und Leistungsmotivation zeigen konnten, sind diese jedoch nicht immer im Vorteil.

2.3.1 Intelligenztestleistung und Leistungsmotivation unter neutraler Bedingung

McClelland et al. (1953) kommen zu dem Schluss, dass ein geringerer Anreiz bzw. eine neutrale Situation bei Erfolgsmotivierten nur geringfügig zu einer Anregung des Leistungsmotivs führt und bei Misserfolgsmotivierten nicht die leistungshemmende Furcht auslöst. Entsprechend dieser These finden einige Untersuchungen keine Leistungsunterschiede zwischen den Leistungsmotiven bzw. zwischen hoher und niedriger Testängstlichkeit (Lighthall et al., 1959; Sarason et al., 1952; Schneider, 1973) bzw. einen Vorteil der Misserfolgsschrecklichen in neutralen Situationen oder bei einfachen Aufgaben (Feather, 1965; Heckhausen, 1963). Mandler und Sarason (1952) machen allerdings deutlich, dass ängstliche Vpn Intelligenztests nicht als neutrale Situationen empfinden, da diese im Vergleich zu Kontrollgruppen und den nicht ängstlichen am schlechtesten in Intelligenztests abschneiden. I. G. Sarason (1960) kommt nach einer Zusammenfassung von Studien zur Testängstlichkeit zu dem Schluss, dass schon mündliche Mitteilungen, dass es sich um einen Intelligenztest handle, Stress auslösen können. Diese Betonung führt bei McClelland et al. (1953) zur Anregung des Leistungsmotivs und wird als leistungsorientierte (Atkinson & Reitmann, 1956; Petty & Harrell, 1977; Raphaelson, 1957) oder ego-involving Bedingungen (Alper, 1946; Sarason et al., 1952) in verschiedenen Studien eingesetzt. Diese regen die Leistungsmotivation an, lösen Misserfolgsschreck und aufgaben-irrelevante Kognitionen bei Testängstlichen aus. Außerdem führen schwierige und komplexe Aufgaben, wozu ein Intelligenztest gewiss zählen kann, bei Schneider und Gallitz (1973) oder Feather (1965) zu schlechteren Leistungsergebnissen bei Misserfolgsmotivierten im Vergleich zu Erfolgsmotivierten. Auch Atkinson kommt mit Hilfe seines Risiko-Wahl-Modells zu dem Schluss, dass Misserfolgsschreckliche komplexe Aufgaben von Anfang an als sehr schwer ansehen. Durch eine geringe Erfolgswahrscheinlichkeit sinkt der Anreiz der Aufgabe und die negative Motivationstendenz, welche während der Ausführung einer Handlung hemmend wirkt, erhöht sich (Atkinson, 1964/1975). Da selbst eine neutrale Situation, wenn es sich um einen Intelligenztest oder andere komplexe Aufgaben handelt, gemäß den

Studien als bedrohlich wahrgenommen wird und Stress auslösen kann, werden bereits in neutraler Situation Unterschiede zwischen HE und FM angenommen.

H4: Vpn mit überwiegend HE erreichen im Vergleich zu Vpn mit überwiegend FM bessere Leistungen in den Intelligenztests in der neutraler Situation.

H5: Vpn mit überwiegend HE erreichen im Vergleich zu Vpn mit überwiegend FM bessere Leistungen in den Intelligenztests.

2.3.2 Intelligenztestleistung und Leistungsmotivation unter Stress

Gemäß den in Kapitel 1.3.4 ausführlich dargelegten Studien zu Misserfolgsinstruktionen wird zunächst postuliert, dass Stress durch Misserfolg zu einem Leistungsabfall führt.

H6: Vpn in der neutralen Bedingung erreichen im LPS-neu im Vergleich zu Vpn der Stressbedingung bessere Leistungen.

Hinsichtlich der Personen mit HE und FM löst nicht nur die Ankündigung eines Intelligenztests bzw. dessen Komplexität Leistungsunterschiede aus, sondern die Verbindung mit Stress bzw. Misserfolg sollte die Unterschiede noch verstärken. Durch den niedrigen Intelligenzwert ist die Wahrscheinlichkeit von Misserfolg noch größer, was bei Misserfolgsmotivierten zu stärkerer Furcht, bei Erfolgsmotivierten jedoch zu einem stärkeren Erfolgsanreiz führt (Atkinson, 1964/1975; McClelland et al., 1953). Atkinson spricht sich zusätzlich für eine weitere Erklärungsmöglichkeit aus (1964/1975): Misserfolg ruft Trägheitstendenzen wach, die bei Erfolgsmotivierten die Motivation erhöhen und diese bei Misserfolgsmotivierten verringern. Sarason et al. (1952) und Heckhausen (1982) betonen den Einfluss irrelevanter Gedanken, welche bei Misserfolgsängstlichen bzw. Testängstlichen zu schlechteren Leistungen in schwierigen Testsituationen führen. Sowohl Schneider und Gallitz (1973) in der Leistungsmotivforschung, als auch Testangstforscher wie Sarason et al. (1952) können diese Annahmen belegen. Bei schweren Aufgaben zeigt sich die Wechselwirkung von Leistungsmotivation und der situativen Bedingung, d. h. dass bei Misserfolgsmotivierten die Leistung von der Erfolgs- zur Neutral- und zur Misserfolgsbedingung sinkt, während sich Erfolgsmotivierte umgekehrt verbessern (Schneider & Gallitz, 1973). Nach Misserfolg verschlechtert sich auch die Leistung testängstlicher Vpn im Vergleich zur neutralen Situation, während nicht ängstliche sich verbessern (Sarason et al., 1952). Schlussfolgernd werden aus

den theoretischen Annahmen und empirischen Bestätigungen interindividuelle und intraindividuelle Unterschiede bezüglich der Leistungsmotive unter Stress erwartet.

H7: Vpn mit überwiegend HE erreichen im Vergleich zu Vpn mit überwiegend FM bessere Leistungen in dem Intelligenztest in der Stressbedingung.

H8: Vpn mit überwiegend HE erreichen in der Stressbedingung im Vergleich zur neutralen Bedingung in einem Intelligenztest bessere Leistungen.

H9: Vpn mit überwiegend FM erreichen in der neutralen Bedingung im Vergleich zur Stressbedingung in einem Intelligenztest bessere Leistungen.

2.3.3 Intelligenztestleistung und die Höhe des Leistungsmotivs

Analog der Hypothesen bezüglich der Richtung der Leistungsmotivausprägung werden für die Höhe der Leistungsmotivation ähnliche Einflüsse auf die Leistungsergebnisse angenommen. Gemäß der Yerkes-Dodson-Regel sollte eine zu hohe emotionale Aktiviertheit bzw. in Bezug zur Leistungsmotivation eine zu hohe Motivation, welche durch Stresssituation und durch bloße Intelligenztestsituationen gegeben ist, zu einer Fehlerzunahme, oft durch schnellere Bearbeitungszeiten, führen. Die Befunde von McClelland et al. (1953), Atkinson (1958), French (1958) sowie French und Thomas (1958) ergeben ein gegenteiliges Bild zwischen Hoch- und Niedrigleistungsmotivierten: Hochleistungsmotivierte erbringen im Vergleich zu Niedrigmotivierten bei komplexen Problemlöseaufgaben bessere Leistungen und arbeiten länger an einem Problem. Lowell (1952) belegt, dass hochleistungsmotivierte Vpn bei einer verbalen, jedoch nicht bei einer numerischen Intelligenztestaufgabe besser im Vergleich zu niedriger motivierten abschneiden. French und Thomas (1958) gehen davon aus, dass die Zielgerichtetheit und der energetisierende Effekt Hochleistungsmotivierte dazu führt, die wichtigen Informationen zu erkennen und härter zu arbeiten. Aufgrund dieser empirischen Arbeiten werden die Unterschiede zwischen Hoch- und Niedrigleistungsmotivierter wie folgt vermutet.

H10: Vpn mit hoher GM erreichen im Vergleich zu Vpn mit niedriger GM bessere Leistungen in einem Intelligenztest.

Zwischen den experimentellen Settings in Atkinsons und Reitmans (1956) Studie sind Personen mit hohem Leistungsmotiv in leistungsorientierten Situationen besser als in neutralen bzw. multiplen Anreizsituationen. Auch McClelland et al. (1953) oder Schnei-

der (1973) belegen dies unter leistungsorientierten- und Misserfolgssituationen. Atkinson (1964/1975) führt die Tendenz, bei hochleistungsmotivierten Personen nach Misserfolg bessere Leistung zu erbringen, auf anfängliche hohe Erfolgswahrscheinlichkeit zurück, welche mit dem Misserfolg gesenkt wird und so in einem stärkerem Anreiz resultiert. Auch diese These soll überprüft werden.

H11: Vpn mit hoher GM erreichen in der Stressbedingung im Vergleich zur neutralen Bedingung in einem Intelligenztest bessere Leistungen.

Entgegen den Forschungsbeiträgen zu den Reaktionen Hochmotivierter in unterschiedlichen Untersuchungssituationen, finden die Reaktionen Niedrigmotivierter weniger Beachtung. Zwar ermittelt Becker (1982) den Zusammenhang, dass je niedriger die Leistungsmotivation, desto geringer ist die persönliche Wichtigkeit von Misserfolg; dennoch können McClelland et al. (1953) eine Steigerung bei niedrigmotivierten Vpn in einer Leistungs- bzw. Misserfolgssituation finden. Auch nach der Yerkes-Dodson-Regel, welche einen energetisierenden Effekt von Stress- und Intelligenztestsituationen vermuten lässt, wird folgende These formuliert:

H12: Vpn mit niedriger GM erreichen in der Stressbedingung im Vergleich zur neutralen Bedingung in einem Intelligenztest bessere Leistungen.

3. Methode

Um die in Kapitel 2 dargestellten Fragestellungen zu überprüfen, wurde ein Zwei-Gruppen-Design gewählt, wobei eine Gruppe ein Treatment erhielt. Es handelt sich somit um einen einfaktoriellen Versuchsplan mit einer unabhängigen, zweigestuften Variable (Stress vs. Neutralbedingung). Die Testleistungen nach dem überarbeiteten Leistungsprüfsystem (LPS-neu) dienten für diese Studie als abhängige Variablen. Weiterhin wurden folgende Tests und Fragebögen als Kontrollvariablen erfasst:

- Fragebogen zu den demographischen Daten
- Regensburger Leistungs-Motiv-Inventar für Erwachsene (RLMI-E)
- Grundintelligenztest Skala 2 (CFT 20-R)
- Fragebögen zum aktuellen Wohlbefinden
- Fragebogen zur Testevaluation des CFT 20-R
- Fragebogen zur Testevaluation des LPS-neu
- Fragebogen zur subjektiven Beurteilung der Stressreaktion

Die Beschreibung der Stichprobe, der Verlauf des Experiments und die Erklärung der Untersuchungsmethode werden in folgenden Abschnitten dargestellt.

3.1 Beschreibung der Stichprobe

Insgesamt nahmen 72 Personen an dieser Studie teil, die mit Ausnahme von zwei Teilnehmern alle deutsch als Muttersprache angaben.

Zu einem Teil wurden die Vpn mittels Listen an der Universität Regensburg und zum anderen über Bekanntschaften gewonnen. Die Aufteilung in Experimental- und Kontrollgruppe zu jeweils 36 Teilnehmer erfolgte zufällig. Um eine bessere Vergleichbarkeit der Gruppen in den Untersuchungswerten der Intelligenztests zu erreichen, lag der Fokus der Untersuchung hauptsächlich auf Studenten im Altersbereich junger Erwachsener. Die Vpn nahmen freiwillig an den Testungen teil und die Psychologiestudenten erhielten fünf Vpn-Stunden.

3.1.1 Geschlecht

In der Studie waren 49 Frauen und 23 Männer beteiligt, was einem prozentualen Anteil von 68.1 % Frauen und 31.9 % Männer entspricht (vgl. Tabelle 3.1). Die Experimentalgruppe bestand aus 22 Frauen (61.1 %) und 14 Männern (38.9 %) und in der Kontrollgruppe waren 27 der untersuchten Personen weiblich (75.0 %) und 9 männlich (25.0 %).

Tabelle 3.1: Verteilung nach Geschlecht

Gruppe		Geschlecht		Gesamt
		weiblich	männlich	
Kontrollgruppe	Häufigkeit	27	9	36
	Prozent	75.0	25.0	100.0
Experimentalgruppe	Häufigkeit	22	14	36
	Prozent	61.1	38.9	100.0
Gesamt	Häufigkeit	49	23	72
	Prozent	68.1	31.9	100.0

3.1.2 Alter

Der Altersdurchschnitt der Vpn zum Zeitpunkt der Testung liegt insgesamt bei 23.26 Jahren ($SD = 2.76$), wobei die Spannweite 12 Jahre umfasst. Die Altersspanne reicht in der untersuchten Stichprobe von 19 bis 31 Jahren, was dem Altersfokus der Studie auf junge Erwachsene gerecht wird. Deutlich wird der Fokus ebenfalls durch den Fakt, dass 50 % der Teilnehmer unter bzw. gleich 23 Jahre, und 75 % unter bzw. gleich 25 Jahre alt waren. Die Altersverteilung wird durch die Abbildung 3.1 veranschaulicht.

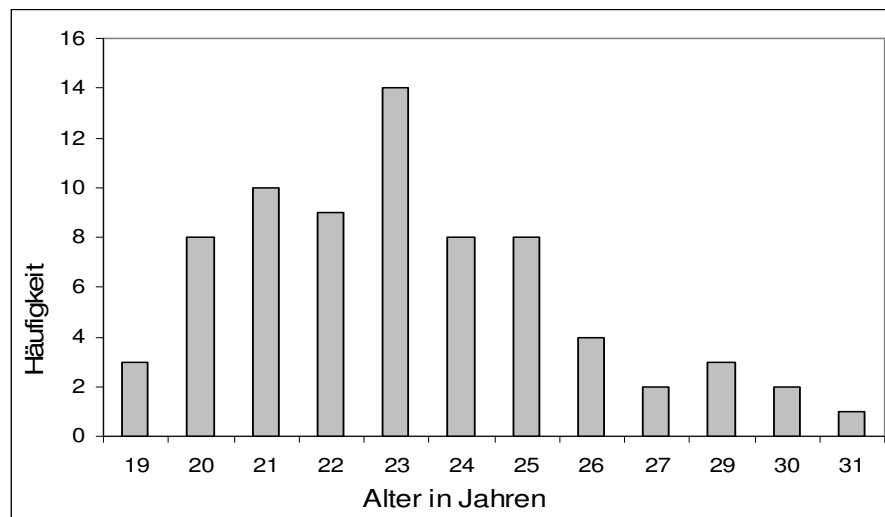


Abbildung 3.1: Verteilung nach Alter

3.1.3 Händigkeit

An der Untersuchung nahmen 67 Rechtshänder (93.1 %) und 4 Linkshänder (5.6 %) teil und eine Person gab an beidhändig zu sein (1.4 %). Somit entspricht der prozentu-

ale Anteil der Linkshänder etwa dem Vorkommen von 5 % in der Gesamtbevölkerung (Psychrembel, 2007, S. 1112).

3.1.4 Studienfach und Semesteranzahl

Mit 61.1 % (44 Personen) sind über die Hälfte der Teilnehmer Psychologiestudenten. Fünf Studenten (6.9 %) entstammen der Fachrichtung Betriebswirtschaftslehre (BWL) und vier (5.6 %) studieren Maschinenbau. Jeweils drei Teilnehmer (4.2 %) sind Studenten der Fachrichtung Medizin und Vergleichende Sprachwissenschaften. Weiterhin nahmen je zwei Personen (2.8 %) aus dem Studiengängen Medientechnik und Lehramt und jeweils eine Person (1.4 %) der Fachrichtungen Volkswirtschaftslehre, Politikwissenschaft, Zahnmedizin, Informatik, Produktion und Management, Jura, Elektrotechnik und Mathematik an der Untersuchung teil. Eine Vp ist bereits berufstätig und gab daher kein Studienfach an. Für weitere Analysen bezüglich des Studienfachs (siehe Kapitel 3.4.2) wurden die Fächer wie folgt gruppiert und es ergeben sich die in Tabelle 3.2 dargestellten Häufigkeiten der einzelnen Studienfachrichtungen:

- Rechts- und Wirtschaftswissenschaften (Betriebswirtschaftslehre, Produktion und Management, Volkswirtschaftslehre, Jura)
- Ingenieurwissenschaften (Medientechnik, Maschinenbau, Elektrotechnik)
- Gesellschafts- und Sozialwissenschaften (Psychologie, Politikwissenschaft, Lehramt)
- Sprach- und Literaturwissenschaften (Vergleichende Sprachwissenschaft)
- Medizin und Gesundheitswesen (Medizin, Zahnmedizin)
- Mathematik und Naturwissenschaften (Mathematik, Informatik)

Tabelle 3.2: Verteilung nach Studienfachgruppen

Studienfachrichtung	Häufigkeit	Prozent
Rechts- und Wirtschaftswissenschaften	8	11.1
Ingenieurwissenschaften	7	9.7
Gesellschafts- und Sozialwissenschaften	47	65.3
Sprach- und Literaturwissenschaften	3	4.2
Medizin und Gesundheitswesen	4	5.5
Mathematik und Naturwissenschaften	2	2.8
Kein Studienfach	1	1.4
Gesamt	72	100

Durchschnittlich befinden sich die Vpn im 6. Semester ($M = 5.86$, $SD = 2.94$), wobei Studenten vom ersten bis zum 12. Semester an der Untersuchung teilnahmen. Die bereits berufstätige Person machte folglich auch zur Semesteranzahl keine Angabe.

3.1.5 Schulabschluss und Abiturnote

Alle teilnehmenden Personen besitzen eine Allgemeine Hochschulreife, wobei 63 (87.5 %) eine allgemeine und 9 Personen (12.5 %) eine fachgebundene Hochschulreife vorweisen können. Die Spanne der dabei erreichten Abiturnoten reicht von 1.0 bis 3.5 mit einem Mittelwert von 2.0 ($SD = 0.6$). Auffällig ist bei der Verteilung der Abiturnoten (vgl. Abbildung 3.2), dass der Median bei 1.8 liegt. Somit gab die Hälfte der Vpn eine Note besser oder gleich 1.8 an und zusätzlich liegen 75 % besser bzw. gleich 2.4 in ihrem Abiturabschluss.

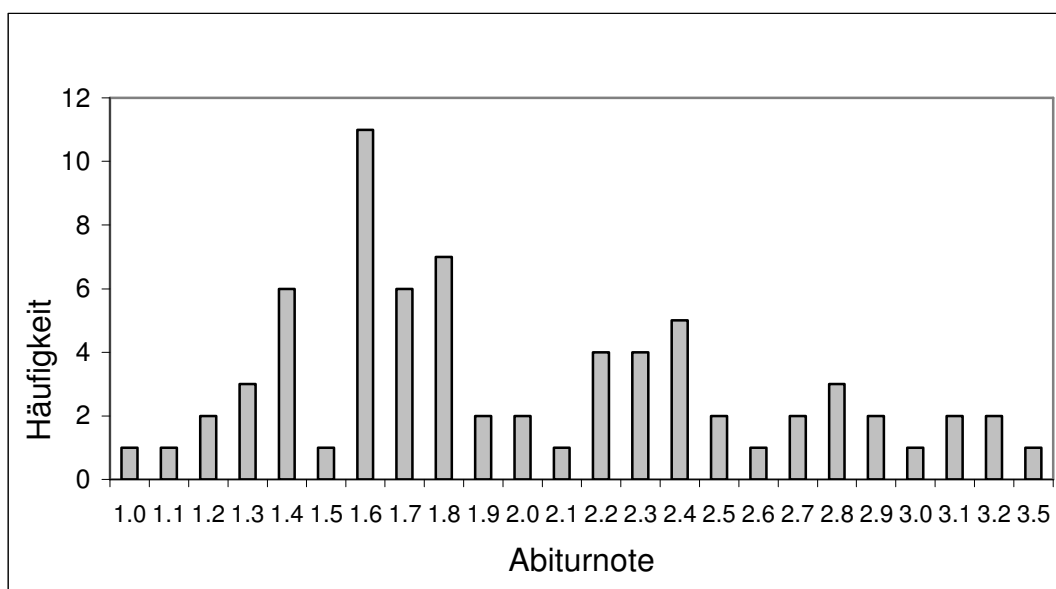


Abbildung 3.2: Häufigkeitsverteilung nach Abiturnote

3.1.6 Berufliche Tätigkeit und berufliche Ausbildung

Alle Teilnehmer gaben zum Zeitpunkt der Testung an, Studenten zu sein, mit Ausnahme einer Person, welche beruflich als Industriekauffrau tätig ist. Neben dem Studium arbeitet eine Person als Mediengestalter und eine weitere als Flugbegleiterin. Zwei Teilnehmer sind momentan nicht berufstätig, erlernten aber bereits die Berufe Kfz-Elektriker und Mediengestalter. Eine Vp gab ebenfalls an, nicht berufstätig zu sein, aber als Praktikantin zu arbeiten. Die restlichen 66 Personen (91.7 %) haben folglich

noch keinen Beruf erlernt und üben momentan keine berufliche Tätigkeit mit Bezug zu ihrer Studienfachrichtung aus.

3.2 Untersuchungsablauf

Der Ablauf des Experiments gliedert sich in drei Schritte: Einen Onlinetest zur Leistungsmotivation und zwei Intelligenztests, welche an zwei unterschiedlichen Tagen stattfanden. Die Vpn wurden zunächst per E-Mail kontaktiert, um sie über den Ablauf und Inhalt der Testung zu informieren. Den Teilnehmern wurde mitgeteilt, dass es sich um ein Experiment zum Thema Intelligenzforschung handelt. Mit dieser E-Mail erhielten sie zusätzlich einen direkten Link, Seriennummer und Passport zu der Online-Version des Regensburger Leistungs-Motiv-Inventars für Erwachsene (RLMI-E), welcher den ersten Teil der Studie darstellte. Diese Computerversion stellte das Hogrefe TestSystem (HTS), ein System zur computerunterstützten Psychodiagnostik, zur Verfügung. Die Teilnehmer konnten den Test an einem beliebigen Ort ausfüllen, wobei sie die Aufforderung bekamen, sich selbstständig einen ruhigen Raum zu wählen, wo sie den Fragebogen ungestört beantworten konnten. Sie erhielten die Bitte, den Motivfragebogen sorgfältig und ehrlich auszufüllen und die vorgegebene Instruktion zum Fragebogen genau zu befolgen. Das Beantworten der Fragen dauerte etwa 30 Minuten und die Versuchsleiter konnten die Ergebnisse jeweils aus dem Internet beziehen.

Der nächste Schritt waren die IQ-Messungen, die an zwei verschiedenen Tagen erfolgten. Die Termine für die Testungen wurden zum Teil telefonisch, zum Teil mit dem www.doodle.com Terminkoordinierungssystem abgestimmt. Pro Termin fanden Testungen mit je zwei bis zehn Probanden statt, da diese Testungen einerseits ökonomischer sind, andererseits in der Stressbedingung die Wirkung der sozialen Bezugsnormen verstärken sollten. Zwischen den Terminen musste aufgrund der Instruktion zur experimentellen Bedingung etwas Zeit liegen. In dieser Stressbedingung erhielten die Vpn die Instruktion, dass der erste abgelegte Test bereits ausgewertet vorliegen würde, worauf später ausführlicher eingegangen wird.

Die Vpn wurden nach der Begrüßung zum zweiten Teil des Experiments, d. h. zum ersten Intelligenztest, in einen ruhigen Raum geführt, um jegliche Art von Ablenkung zu vermeiden. Sie sollten so in einer Sitzordnung Platz nehmen, dass eine Abschreibefahr zum Großteil nicht gegeben war. Nach kurzen Vorbemerkungen zum Thema, zur Wahrung der Identität und zum organisatorischen Ablauf erhielten sie einen Fragebogen zur Person. Danach sollten die Teilnehmer einen Fragebogen zum aktuellen Wohlbefinden ausfüllen, wobei sie die Einschätzungen möglichst spontan vornehmen sollten. Nun konnte mit der ersten Messung zur Intelligenz begonnen werden. Es handelte sich hier um den Grundintelligenztest CFT 20-R, welcher nach Originalversion

instruiert wurde. Die Teilnehmer erhielten den Hinweis, dass wahrscheinlich niemand alle Aufgaben richtig lösen kann, sie jedoch versuchen sollten, möglichst viele zu lösen. Zwei bis drei Beispielaufgaben, welche mündlich vor jeder Aufgabe besprochen wurden, machte die Vpn mit der Aufgabenstellung vertraut. Erst nach eindeutigem Verständnis durften die Teilnehmer mit der jeweiligen Aufgabe beginnen. Die zur Verfügung stehende Zeit pro Untertest entstammte ebenfalls der Originalversion, d. h. je fünf Minuten für die ersten zwei und vier für die letzten zwei Tests. Insgesamt betrug die Bearbeitungszeit für den CFT 20-R etwa 40 Minuten. Abschließend sollten die Vpn erneut einen Fragebogen zum aktuellen Wohlbefinden und einen Fragebogen zur Testevaluation des CFT 20-R ausfüllen. Nach dem Einsammeln der Tests wurden die Termine zur zweiten Testung vereinbart.

Die zweite Testung verlief weitestgehend analog zur ersten und fand auch im gleichen Testraum statt. Der Beginn der Untersuchung war hier ebenfalls ein Fragebogen zum aktuellen Wohlbefinden. Die Kontrollgruppe erhielt danach bei diesem Intelligenztest, einer Revision des Leistungsprüfsystems (LPS-neu), die originale Instruktion. Die Experimentalgruppe bekam zusätzlich eine Stressinstruktion, in der ihnen, wie die Beispiele in Kapitel 1.3.4 beschreiben, ein Misserfolg, d. h. ein für ihr Niveau schlechtes Ergebnis des CFT 20-R suggeriert wurde (vgl. Anhang A-1). Die Vpn erhielten die Information, dass die CFT 20-R Testungen bereits ausgewertet vorlagen und jedem wurde ein Ergebnis 12 % unter der wirklich erreichten Leistung schriftlich mitgeteilt. Somit bekam jeder einen individuell angepassten, fiktiven Wert, anstatt einer standardisierten Norm. Da kein Gütemaßstab der Personen erhoben wurde, wurde ein ipsativer Wert, d. h. auf die vorhergegangenen Daten der Person bezogener Wert, als Norm verwendet (Corsini, 1999). Es besteht die Annahme, dass durch einen ipsativen Wert individuelle Bezugsnormen stärker als durch einen normativen Wert, angeregt werden (siehe Kapitel 1.1.2). Dieser unter dem eigenen Selbstkonzept liegende Wert, sollte bei den Personen Stress erzeugen, „[...] da Personen grundlegend dazu neigen, ihr Selbstbewusstsein auf einem hohen Niveau zu halten.“ (Hoppe, 1930, S. 35). Weiterhin wird davon ausgegangen, da Intelligenz bei Studenten für einen erfolgreichen Studienabschluss äußerst relevant ist (Petty & Harrel, 1977), dass die Valenz eines solchen Tests hoch ist und folglich der Misserfolg eine starke Stresswirkung erzielen. Zusätzlich enthielt die Instruktion eine Abbildung einer gefälschten Häufigkeitsverteilung der Intelligenzwerte von einer angeblichen Studie der Universität Regensburg an 534 Studenten (vgl. Abbildung 3.3), wie es Covin et al. (2003) in ähnlicher Form anwendeten.

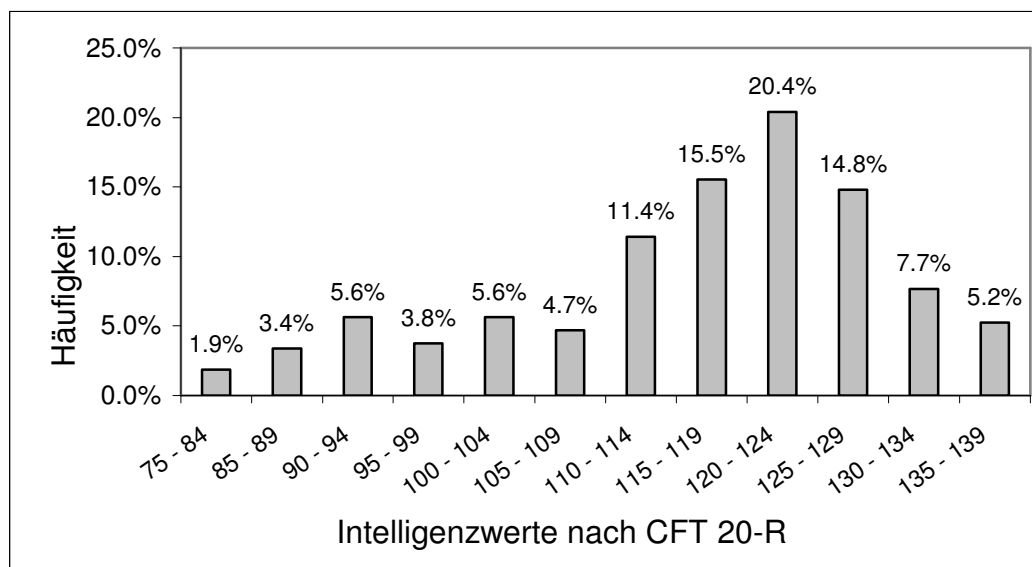


Abbildung 3.3: Erfundene Häufigkeitsverteilung der Intelligenzwerte des CFT 20-R zur Stressinduktion

Diese erfundene Verteilung ist rechtssteil mit einer Spannweite von 75 bis 139, einem Mittelwert von 117.62 ($SD = 13.99$) und einem Median von 118. Somit suggeriert die Grafik ein hohes Niveau der Verteilung der Intelligenzwerte, da ein durchschnittlicher, „normaler“ Intelligenzquotient laut Häcker und Stapf (2009) bei 90 bis 109, ein guter zwischen 110 und 119 und ein sehr guter bis hervorragender ab 120 angesetzt wird. Der Eindruck eines relativ hohen Intelligenzniveaus der untersuchten Studenten der erfundenen Studie sollte dadurch entstehen. Jeder Teilnehmer erhielt zusätzlich eine Gruppenzuordnung nach dem Intelligenzwert und die Prozentwerte, wie viele Personen über bzw. unter dem Wert dieser Gruppe lagen (vgl. Abbildung 3.4). Es besteht die Annahme, dass diese zusätzlichen sozialen Bezugsnormen das Misserfolgserlebnis verstärken und ein zusätzliches Stressgefühl auslösen. Die Teilnehmer hatten Zeit, sich die Abbildung und die gegebenen Werte genau anzuschauen und gegebenenfalls Fragen zu stellen.

Seit unserem ersten Treffen habe ich den ersten Intelligenztest bereits ausgewertet. Dein Wert beträgt 104 und somit fällst Du in die Gruppe 100 - 104. Es hatten 14.7 % der 534 Studenten einen Intelligenzwert unter und 79.7 % einen Wert über Deiner Intelligenzleistung.

Abbildung 3.4: Instruktionsbeispiel der Stressbedingung

Darauf folgte erneut die Erfassung des momentanen Wohlbefindens für die Experimentalgruppe und danach die originale Instruktion des LPS-neu Intelligenztests. Auch in der Stressbedingung wurde darauf hingewiesen, dass es meist nicht möglich ist alle Aufgaben zu bearbeiten, es jedoch wichtig ist, sich Mühe zu geben. Bevor die Vpn mit der Bearbeitung der jeweiligen Aufgabe beginnen konnten, wurden gemeinsam jeweils zwei Beispielaufgaben besprochen. Die Testung dauerte etwa 60 Minuten. Abschließend sollten beide Gruppen den Fragebogen zum aktuellen Wohlbefinden und die Bewertung des LPS-neu mittels eines Evaluationsfragebogens ausfüllen.

Zum Ende der zweiten Intelligenztestung erhielten die Vpn eine Aufklärung über die Untersuchung, der Zugehörigkeit zur Experimental- oder Kontrollgruppe und zu dem zugehörigen Treatment. Sie wurden anhand eines schriftlichen Debriefings über Hintergründe und Zielsetzungen des Experiments informiert und in der Stressbedingung erhielten die Vpn das wahre Ergebnis der CFT 20-R Testung (vgl. Anhang A-2 und A-3). Daraufhin sollten sie die Stärke der Reaktion auf den fiktiven Intelligenzwert subjektiv anhand eines Fragebogens einschätzen. Im Gespräch wurden abschließend offene Fragen beantwortet.

3.3 Untersuchungsmethode

Die Testungen erfolgten durch die handelsüblichen, aktuellen Versionen des RLMI-E und des CFT 20-R sowie der Evaluationsform des LPS-neu, eine Revision des LPS. Der RLMI-E wurde am Lehrstuhl von Prof. Lukesch an der Universität Regensburg entwickelt. Auch das LPS-neu wird an diesem Lehrstuhl überarbeitet, wobei die Daten dieser Studie zur Untersuchung der Validität des LPS-neu beitragen.

3.3.1 Das Regensburger Leistungs-Motiv-Inventar (RLMI-E)

Mit dem Ziel, die Komponenten des Leistungsmotivs „Hoffnung auf Erfolg“ (HE) und „Furcht vor Misserfolg“ (FM) zu messen, wurde das Regensburger Leistungs-Motiv-Inventar für Erwachsene (RLMI-E; Lukesch & Peters-Häderle, 2007) in dieser Untersuchung verwendet.

3.3.1.1 Beschreibung des RLMI-E

Dieses Verfahren arbeitet mit der Szenariotechnik, d. h. es werden verschiedene Situationen, in denen potentiell Erfolgs- oder Misserfolgskognitionen auftreten können, beschrieben. Für jede der 24 Szenarien liegen vier mögliche Reaktionsweisen vor, wobei jede speziell auf die Anregung eines der Motivkomponenten HE, FM sowie „Furcht vor Erfolg“ (FE) und „Hoffnung auf Misserfolg“ (HM) abzielt. Die Vpn mussten jede dieser

Reaktionsmöglichkeiten mit Hilfe einer fünfstufigen Likert-Skala bewerten, wie stark diese Aussage auf sie zutrifft. Dabei entspricht der Wert 1 „trifft gar nicht zu“ und der Wert 5 „trifft sehr zu“. Da Motive bereichsspezifisch sein können, beziehen sich die Situationen speziell auf Inhaltsbereiche. Der Erstentwurf, „das Gitter zur Erfassung von Hoffnung und Furcht bezüglich Erfolg und Misserfolg“ (GEHFEM) von Peters (2001), aus dem der RLMI-E hervorging, umfasste noch sechs Bereiche. Eine ausführliche Beschreibung zur unterschiedlichen konzeptionellen und inhaltlichen Weiterentwicklung des RLMI findet sich bei Lukesch und Peters-Häderle (2007). Eine wichtige Veränderung des RLMI, nach wiederholten Itemtestanalysen, ist der Verzicht auf die zwei Bereiche Sport und Partnerschaft, da sich zu geringe Reliabilitäten ergaben. Somit umfasst die hier verwendete Version die vier Inhaltsbereiche Ausbildung (A), Beruf (B), Freizeit (F) und Aussehen (Optik = O). Schlussfolgernd resultieren die in Tabelle 3.3 dargestellten 16 bereichsspezifischen Skalen und vier Gesamtskalen.

Tabelle 3.3: Die Skalen des Regensburger Leistungs-Motiv-Inventar

		HE	FM	FE	HM
Bereich	Ausbildung	HE in der Ausbildung	FM in der Ausbildung	FE in der Ausbildung	HM in der Ausbildung
	Beruf	HE im Beruf	FM im Beruf	FE im Beruf	HM im Beruf
	Freizeit	HE in der Freizeit	FM in der Freizeit	FE in der Freizeit	HM in der Freizeit
	Aussehen	HE beim Aussehen	FM beim Aussehen	FE beim Aussehen	HM beim Aussehen
Gesamt		HE gesamt	FM gesamt	FE gesamt	HM gesamt

Für jeden Bereich werden je sechs passende Situationen dargestellt. Die Abfolge der Bereiche sowie die motiv-bezogenen Reaktionsweisen variieren zufällig. Es existieren zwei Formen des RLMI-E – eine für Frauen und eine für Männer, wobei sich in den Szenarien nur das Geschlecht des Protagonisten ändert und die sozialen Konstellationen gleich bleiben. Außerdem wurde das RLMI für zwei verschiedene Altersgruppen ausgearbeitet: Neben dem RLMI für Erwachsene (RLMI-E) entstand aufgrund einer Studie von Peters-Häderle (2006) zusätzlich das RLMI für Kinder und Jugendliche zwischen 12 und 18 Jahren (RLMI-K/J).

Die für diese Studie verwendete Onlineversion des RLMI-E ist nur in männlicher Form erhältlich, was aber auf die Testwerte der Frauen keinen bedeutsamen Einfluss haben sollte. Tabelle 3.4 zeigt eine Beispielsituation des RLMI-E aus dem Bereich Freizeit mit den dazugehörigen Antwortvorgaben.

Tabelle 3.4: Beispielsituation aus dem Bereich Freizeit des RLMI-E

Sie nehmen an einer Quizshow im Fernsehen teil, die live übertragen wird. Zuhause sitzen Ihre Freunde und Verwandte vor dem Fernseher und drücken Ihnen die Daumen. Bei dem Quiz spielen immer mehrere Kandidaten gegeneinander.					
Kurz vor Spielbeginn denken Sie:	1 = trifft gar nicht zu		5 = trifft sehr zu		
1. Wenn ich gewinne, erwarten meine Freunde bestimmt, dass ich mich spendabel zeige. FE*	1	2	3	4	5
2. Wenn ich jetzt nichts weiß, mache ich mich zum Gespött der Leute. FM*	1	2	3	4	5
3. Hoffentlich gewinne ich nicht den ersten Preis, der darauf folgende Rummel wäre mir zu viel. HM*	1	2	3	4	5
4. Dieses Spiel gewinne ich und dann bin ich der Held des Tages. HE*	1	2	3	4	5

Anmerkungen: * Skalenzugehörigkeit der jeweiligen Antwortmöglichkeit (im Original nicht angezeigt)

3.3.1.2 Testanalyse des RLMI-E

Da das Hogrefe Testsystem die Itemkennwerte der einzelnen Situationen nicht zur Verfügung stellt, muss auf eine ausführliche Testanalyse des RLMI-E verzichtet werden. Verschiedene Untersuchungen prüften das RLMI-E bereits durchgehend bezüglich der Testgüte und für diese konnten stets zufrieden stellende Ergebnisse ermittelt werden (Eder, 2006; Kornprobst, 2005; Peters, 2001). Daher sind an dieser Stelle die Ergebnisse der internen Konsistenzen der Normierungsstichprobe von Lukesch und Peters-Häderle (2007) über 463 Vpn dargestellt (vgl. Tabelle 3.5). Die Reliabilitäten weisen mit Werten zwischen .85 und .86 auf ein hohes Niveau hin.

Tabelle 3.5: Überblick über die Maße der internen Konsistenz (Cronbachs α) der einzelnen Bereiche

	HE	FM	FE	HM
Ausbildung	.64	.69	.72	.75
Beruf	.69	.68	.67	.66
Freizeit	.55	.53	.57	.53
Aussehen	.69	.67	.63	.62
Gesamt	.85	.86	.85	.86

Für diese Untersuchung sind die Konstrukte HE und FM sowie die von Heckhausen (1980) verwendeten Konstrukte der GM und NH interessant. Daher ist die genaue Ska-

lenanalyse der Konstrukte FE, HM und der einzelnen Bereiche für diese Arbeit nicht direkt von Bedeutung. Ein Überblick der Ergebnisse für alle Leistungsmotivausprägungen sowie der einzelnen Bereiche ist in Tabelle C-1 im Anhang aufgeführt. Die Verteilungen zu den für diese Arbeit bedeutsamen Skalen, gesamt und getrennt für die zu untersuchenden Gruppen, sind in Kapitel 4.2.1 angegeben.

Wie die Gesamtskalenwerte miteinander zusammenhängen, ist in Tabelle 3.6 dargestellt. Die Berechnung der Korrelationen nach Bravais Pearson dieser Studie ähneln den Ergebnissen von Lukesch und Peters-Häderle (2007). Die drei Skalen, welche Misserfolgskognitionen beinhalten, FM, FE und HM, stehen in einem mittleren positiven Zusammenhang zwischen .48 und .63. Die Skala HE weist schwächere und signifikant negative Zusammenhänge zu den anderen drei Skalen zwischen -.26 und -.37 auf.

Tabelle 3.6: Interkorrelation der Skalen des RLMI-E (n = 72)

	HE	FM	FE	HM
HE	-	-.34**	-.26*	-.37**
FM		-	.60**	.48**
FE			-	.63**
HM				-

Anmerkungen: **) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

3.3.2 Der Grundintelligenztest Skala 2 (CFT 20-R)

Als Kontrollverfahren vor dem LPS-neu wurde als zweiter Intelligenztest der Grundintelligenztest CFT 20-R von Rudolf H. Weiß (2006) genutzt.

3.3.2.1 Beschreibung des CFT 20-R

Bei dem CFT 20-R handelt es sich um eine Revision des CFT 20, welcher erstmals 1978 erschien. Ziel der Revision war eine verbesserte Differenzierung im oberen Leistungsbereich, dabei wurden verschiedene Aufgaben verbessert, entfernt oder hinzugefügt. Ergebnis war, dass für den CFT 20-R eine Steigerung der Korrelationskoeffizienten zwischen Subtest und Gesamtwert für alle Subtests erreicht werden konnte. Beide Versionen messen das allgemeine intellektuelle Niveau bzw. den *g*-Faktor (Spearman, 1904) im nicht-verbalen Bereich im Sinne der „General Fluid Ability“ nach Cattell (1963). Die Testaufgaben erfassen das Problemlöseverhalten mit figuralen Beziehungen und formal-logischen Denkproblemen. Durch anschauliche, figurale Darstellung und sprachfreie Aufgaben kann der Test relativ unabhängig von Kultur und Sprache

angewendet werden. Er ist für Kinder und Jugendliche von 8.5 bis 19 Jahren sowie für Erwachsene von 20 bis 60 Jahren einsetzbar. Der Autor legte außerdem besonderen Wert auf die Entwicklung eines ökonomischen Verfahrens. So kann der CFT 20-R auch in der Kurzform (Teil 1) verwendet werden, welcher auch für diese Studie ausreichte.

Wie auch die Originalversion, besteht der CFT 20-R aus zwei gleichartig aufgebauten Testteilen. Der erste Testteil umfasst 56 Items und der zweite 45 Items. Beide Teile setzen sich aus vier unterschiedlichen Subtests zusammen: Reihenfortsetzen, Klassifikationen, Matrizen und Topologien. Die vier Aufgabengruppen sind dadurch gekennzeichnet, dass sie jeweils ein unterschiedliches Design der figuralen Aufgabendarstellung aufweisen. Bei dem Subtest Reihenfortsetzen sollen die Vpn aus fünf möglichen Figuren die auswählen, welche eine Reihe von vier Figuren vervollständigt. Bei den Klassifikationen geht es darum, aus fünf Figuren diejenige herauszufinden, welche nicht in die Reihe passt. Ein leeres Feld einer Matrix mittels fünf möglicher Figuren zu ergänzen, ist die Aufgabe in Subtest 3. Im letzten Subtest soll aus fünf möglichen Varianten, diejenige herausgesucht werden, bei der die Lage eines Punktes genauso möglich ist, wie bei der vorgegebenen Figur.

In Testteil 1 enthalten die ersten drei Subtests 15 Items, während Subtest 4 aus 11 Items besteht. Die 56 Einzelaufgaben sind für jede Aufgabengruppe nach aufsteigender Schwierigkeit geordnet. Für die Antworten muss der Proband in Multiple-Choice-Form aus fünf Alternativen (a bis e), wie in Abbildung 3.5 dargestellt, auswählen. Die richtige Lösung wird dann gesondert auf einen Antwortbogen eingetragen. Beispielaufgaben sind auf einer vorangestellten Seite enthalten und in Testteil 1 mit zusätzlichen Erläuterungen versehen.

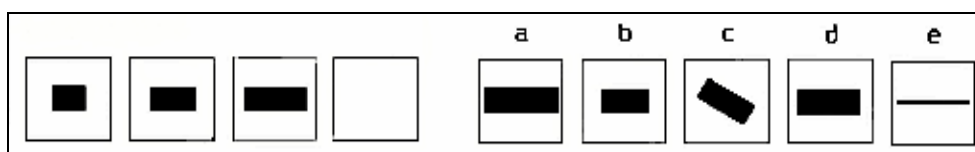


Abbildung 3.5: Beispielaufgaben des CFT 20-R, Testteil 1, 1. Subtest

3.3.2.2 Testanalyse des CFT 20-R

Die Testanalysen über 144 Vpn zu dem CFT 20-R ergaben für den hier verwendeten Test 1 nach Weiß (2006) eine sehr hohe, nach Spearman-Brown-Formel korrigierte, Split-Half-Reliabilität von .92 und für den Gesamttest von .95. Gemäß seinen Faktorenanalysen teilt er den Test in die drei Intelligenzdimensionen „Reasoning“ (Subtest 1 und 3), „Schlussfolgerndes Denken“ (Subtest 4) und „Beziehungsstiftendes Denken“

(Subtest 2). Der Faktor „Reasoning“ kann laut Weiß als Zentralfaktor gelten, da dieser 50 % der Varianz erklärt.

Über die gesamte Stichprobe ergab sich ein Intelligenzmittelwert von 115.69 IQ-Punkten ($SD = 13.72$). Die Ergebnisse der einzelnen Subtests, der Gesamtscore und die IQ-Werte der gesamten Stichprobe sowie der untersuchungsrelevanten Gruppen sind in Tabelle C-2 im Anhang angegeben. Die Mittelwerte und Standardabweichungen der Intelligenzwerte für die gesamte Stichprobe und der unterschiedlichen Gruppen sind in Kapitel 4.3.1 dargestellt.

Die Korrelationen der einzelnen Untertests dieser Arbeit ähneln den Ergebnissen von Weiß (2006) und machen ebenfalls deutlich, dass die einzelnen Subtests und das Gesamtergebnis sowie der Intelligenzwert stark zusammenhängen (vgl. Tabelle 3.7). Alle Korrelationen der Subtests weisen signifikante Zusammenhänge mittlerer Höhe zwischen .30 und .55 auf. Nach den Annahmen von Weiß deuten die signifikanten, positiven Korrelationen darauf hin, dass die Subtests als gemeinsames Testprofil unter dem Faktor „General Fluid Ability“ als Gesamtskala zusammengefasst werden können. Darauf weisen auch die mittleren bis starken Korrelationen der Gesamtergebnisse mit den Subtests hin, welche ebenfalls alle im gewünschten Signifikanzbereich liegen.

Tabelle 3.7: Interkorrelation der Subtests des CFT 20-R (n = 72)

Subtests	CFT 1	CFT 2	CFT 3	CFT 4	CFT R	CFT IQ
CFT 1	-	.38**	.45**	.55**	.82**	.76**
CFT 2		-	.31**	.24*	.68**	.67**
CFT 3			-	.30*	.69**	.60**
CFT 4				-	.72**	.74**
CFT R					-	.96**
CFT IQ						-

Anmerkungen: CFT R = CFT 20-R Gesamt Rohwert, CFT IQ = CFT 20-R IQ-Wert

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

3.3.3 Das neue Leistungsprüfsystem (LPS-neu)

Das überarbeitete Leistungsprüfsystem (LPS-neu) wurde zur Ermittlung der Intelligenzleistung in der Experimentalbedingung eingesetzt.

3.3.3.1 Beschreibung des LPS-neu

Die Originalversion des LPS von Wolfgang Horn wurde bereits 1962 herausgegeben (1962/1983) und basiert auf den Ansichten und Veröffentlichungen der Primary Mental

Abilities nach Thurstone (1938). Wie bereits oben erwähnt, stellt Thurstone das Modell auf, dass die allgemeine Intelligenz eine komplexe Matrix aus vielen voneinander unabhängigen Faktoren ist. Die einzelnen Faktoren korrelieren zum Teil nur gering miteinander und können auch bei Personen gleichen Berufs unterschiedlich ausgeprägt sein (Horn, 1962/1983). Folglich wurde ein Test konstruiert, welcher die wichtigsten Primärfaktoren der Begabung differenziert messen sollte. Die Erfassung der Gesamtingelligenz ist mittels des LPS ebenfalls möglich, aber eher als zweitrangig anzusehen. Die entstandene Version enthielt 15 Untertests, wobei mindestens je zwei Tests gemeinsam einen der von Thurstone postulierten Primärfaktoren *verbal factor*, *reasoning*, *word fluency*, *space*, *closure*, *perceptual speed* und *accuracy* erfassen sollten. Horns Ziel war es, Einblick über die Leistungsstruktur vom „schwachsinnigen Erwachsenen“ bis zum Hochbegabten zu geben und so ist der Test gerade für die Berufs- und Bildungsberatung bedeutsam (Horn, 1962/1982, S. 4). Einsetzbar ist es aber darüber hinaus für Kinder und Erwachsenen von 9 bis 85 Jahren.

In dieser Studie wurde eine von Kreuzpointner (2010) erarbeitete Revision des LPS verwendet, die momentan in der Überprüfungsphase ist. Zu der Validierung des LPS-neu trägt auch diese Studie bei. Ziel von Kreuzpointner ist es ein praktikableres und ökonomischeres Verfahren zu entwickeln, welches die Grundkonzeption des LPS nach Horn beibehalten sollte (Kreuzpointner, 2010). Zusätzlich wird die theoretische Grundlage der Primary Mental Abilities des LPS überdacht und das Drei-Ebenen-Modell von Carroll (1993) als geeignetere Einordnung vorgeschlagen. Dabei überprüft Kreuzpointner, inwieweit sich die einzelnen Subtests des LPS den Faktoren der Stratum II Ebene von Carrolls Modell in Beziehung setzen lassen.

Der Aufbau des LPS nach Horn bleibt in der neuen Version weitestgehend erhalten. Der Test umfasst ebenso vier DIN A 4 Seiten, welche auf einen DIN A 3 Bogen zusammengestellt sind. Durch verbesserten, größeren Druck wurde die häufige Kritik, einer gedrängten und undeutlichen Darstellung der alten Version, beachtet. Eine Änderung der Antwortmodalitäten und das Zusammenfassen der Instruktion auf die erste Seite anstatt vor jeder Aufgabe separat sollen zu einer besseren Übersichtlichkeit beitragen. Die Instruktion des Versuchsleiters wurde der alten Version entnommen und angepasst. Weiterhin wurden durch Kreuzpointner (2010) verschiedene faktoranalytische Studien zum LPS, Interkorrelationsberechnungen und Korrelationen mit anderen Intelligenztests zusammenfassend überprüft. Nach der Auswertung wurden einige Subtests für die neue Version entsprechend verändert, wobei beispielsweise die Itemanzahl von einzelnen Subtests erhöht oder Items geändert wurden. Weiterhin wurden Subtests zusammengefasst oder ganz weggelassen. Der LPS-neu setzt sich aus 11 Subtests zusammen, wodurch die kritisierte neunzigminütige Durchführungsdauer auf

etwa eine Stunde verkürzt wurde. Anschließend werden die Subtest des LPS-neu und die vorgenommenen Zuordnungen der Untertests nach Kreuzpointner (2010) zu den Faktoren der Ebenen des Intelligenzmodells nach Carroll (1993) vorgestellt (vgl. Abbildung 3.6).

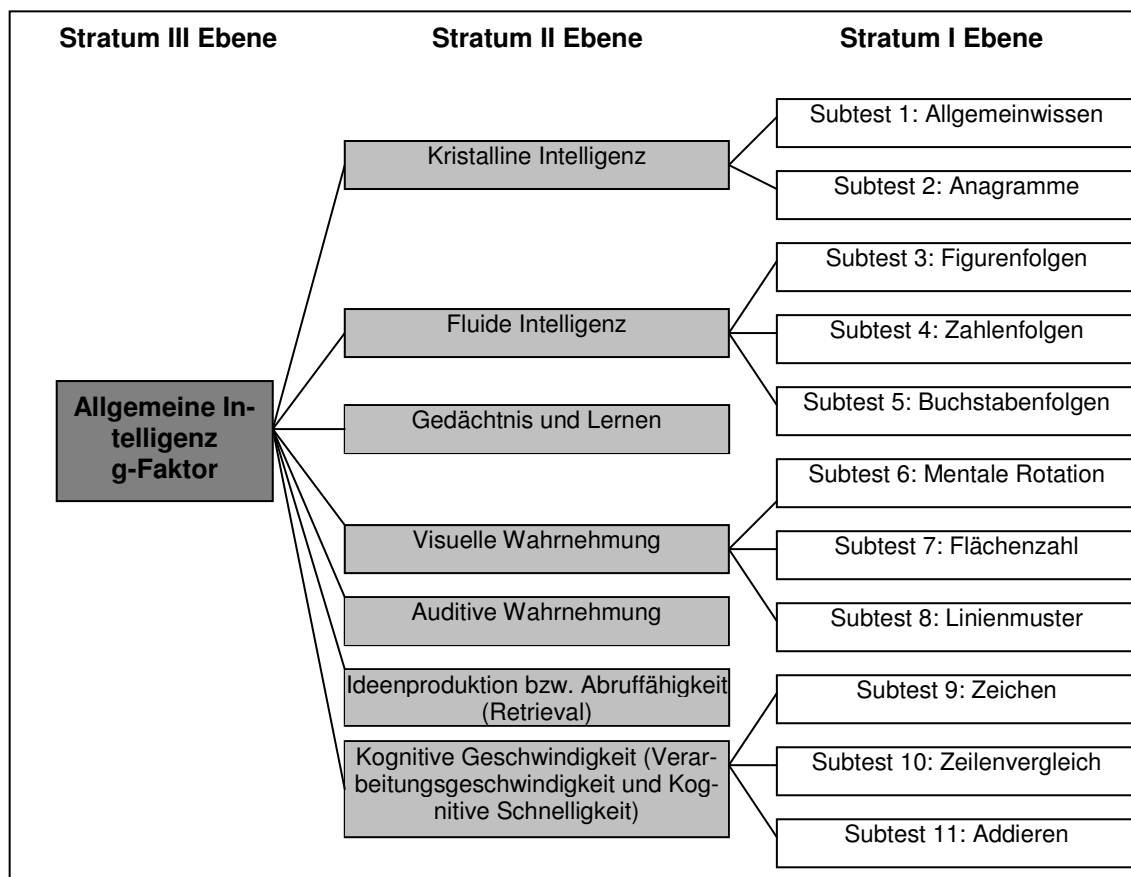


Abbildung 3.6: Einordnung des LPS-neu in das Drei-Ebenen-Modell nach Carroll (1993)

Im ersten Subtest soll weiterhin die Allgemeinbildung über die sprachliche Kompetenz gemessen werden. Die Vpn müssen 60 Wörter erkennen, in denen jeweils ein Buchstabe falsch abgebildet ist und welchen sie markieren sollen. Der Test ist mit einer Durchführungsdauer von drei Minuten angesetzt.

Subtest 2 soll ebenfalls die Allgemeinbildung erfassen. Innerhalb von drei Minuten sollen 40 schwer erkennbare Wörter, die in ihrer Buchstabenreihenfolge durcheinander gebracht sind, herausgefunden werden. Dabei soll der Anfangsbuchstabe eines jeden Wortes für die Lösung gekennzeichnet werden. Laut Kreuzpointner (2010) würde der Subtest 1 den Stratum I Faktoren Verbal Ability und Lexical Knowledge entsprechen,

während in Subtest 2 zusätzlich Reading Decoding erforderlich ist. Durch die Kombination von Test 1 und 2 kann der Stratum II Faktor kristalline Intelligenz erfasst werden.

Subtest 3 ermittelt gemeinsam mit den Subtests 4 und 5 zu je 40 Items den Stratum I Faktor Reasoning, d. h. das schlussfolgernde Denken. Hier soll der Proband das Symbol aus neun dargestellten herausstreichen, welches nicht zu der vorgegebenen Systematik passt. Während in Subtest 3 Formen markiert werden müssen, sind in den Subtests 4 Zahlen und in 5 Buchstaben zu überprüfen. Daher werden nach Kreuzpointner Test 3 und 5 dem Induktive Reasoning und Test 4 dem Sequential oder Quantitative Reasoning beigeordnet. Die drei Tests werden demnach zur fluiden Intelligenz in der zweiten Ebene des Modells nach Carroll zugeordnet. Während für Test 3 drei Minuten zur Bearbeitung zur Verfügung stehen, sind es in Test 4 und 5 fünf Minuten.

Mit Subtest 6 wird der Faktor Space durch visuelle Rotation gemessen, welches zu dem Stratum I Faktor räumliche Beziehung passt. Anhand des Erkennens eines spiegelverkehrten Elements aus fünf möglichen gedrehten Zahlen oder Buchstaben innerhalb von zwei Minuten wird der Faktor erfasst. Die Items des Tests sowie die Tests 7 bis 10 wurden für die neue Version beibehalten, nur die Reihenfolge der Aufgaben wurde leicht verändert.

Die Probanden müssen in Subtest 7 die Anzahl von Flächen eines dreidimensionalen Körpers zählen. Dabei ist die konstante Vorgabe, zwischen zwei bis zehn Flächen auszuwählen, neu.

In der Aufgabenstellung des achten Subtests gilt es, einen von fünf Umrissen in einem Muster zu finden. Die Umrisse sind nebenstehend kleiner abgebildet, aber in ihrer Ausrichtung gleich. Test 7 und 8 sollen in je drei Minuten gelöst werden und während Test 7 Carrolls Faktoren Visualisation und Imagination der ersten Ebene gleichkommt, entspricht Test 8 der Wahrnehmungsgeschwindigkeit. Schlussfolgernd werden in den Aufgaben 6 bis 8 der Stratum II Faktor der visuellen Wahrnehmung getestet, wobei jeder Test aus 40 Items zusammengesetzt ist.

Subtest 9 wird erst nach dem darauf folgenden Test bearbeitet und hier haben die Teilnehmer zwei Minuten Zeit. Aufgabe ist hier zuerst, jede achte „0“ und anschließend jede achte „1“, „2“, „3“ usw. der 40 Items zu markieren. Neben der Schnelligkeit ist hier die Genauigkeit der Bearbeitung bedeutsam.

Der Test 10 wird vor Nummer 9 durchgeführt und besteht aus 60 Items, für die eine fünfminütige Bearbeitungsdauer zur Verfügung steht. Hier sind die Unterschiede der sich sonst gleichenden Zeilen von Spalte 10 und 9 herauszufinden und zu markieren.

In dem letzten, sechsminütigen Subtest 11 soll aus einer Reihe von zehn Ziffern die Summe gebildet werden. Hier geht es darum, abschließend den Stratum Faktor II der kognitiven Schnelligkeit zu erheben, in dem die Konzentrations- und Leistungsfähigkeit

gefordert wird. Die Tests 9, 10 und 11 werden nach Kreuzpointners (2010) Vorschlag unter Kognitive Geschwindigkeit zusammengefasst, welcher Carrolls Stratum II Faktoren Verarbeitungsgeschwindigkeit und Kognitive Schnelligkeit vereint.

Für die Stratum II Faktoren auditive Wahrnehmung, allgemeine Gedächtnisfähigkeit und Ideenproduktion bzw. Abruffähigkeit enthält der LPS-neu keine Entsprechungen.

Abbildung 3.7 verdeutlicht den Aufbau des LPS-neu anhand der ersten vier Subtests mit den Aufgabenbeispielen der Instruktionsseite.

		1	2	3				4					
KRAIDE	G Z W E R	○	○	○	○	○	○	○	○	○	○	○	○
TELLOR	C K E R A		+		+		+						
KRAIDE	G Z W E R	○	○	○	○	○	○	○	○	○	○	○	○
TELLOR	G Z W E R	○	○	○	○	○	○	○	○	○	○	○	○
KRAIDE	G Z W E R	○	○	○	○	○	○	○	○	○	○	○	○
		2	2	2	2	2	2	2	2	3	2	2	2
		4	5	5	4	4	5	4	4	5	4	4	5
		2	2	2	2	2	2	3	2	2	2	2	2

Abbildung 3.7: Beispielaufgaben des LPS-neu

3.3.3.2 Testanalyse des LPS-neu

Mit den Ergebnissen der 72 Vpn dieser Studie wurde eine Testanalyse durchgeführt. Tabelle 3.8 stellt die Ergebnisse der Analyse mittels der Mittelwerte, Standardabweichungen, der Schiefe, Exzess sowie Cronbachs Alpha für die einzelnen Subtests dar. Weiterhin wurden zur Reliabilitätüberprüfung die Split-Half-Koeffizienten berechnet. Jeder Test wurde dafür in zwei gleichgroße Teile anhand der „ungeraden“ und „geraden“ Aufgaben geteilt, deren deskriptiven Werte ebenfalls in der Tabelle 3.8 enthalten sind. Da in dieser Untersuchung zusätzlich die Güte der Leistungen betrachtet wird, sind in Kapitel 4.3.1 Mittelwert, Standardabweichung, Schiefe und Exzess der korrekten, falschen, nicht bearbeiteten, berichtigten und ausgelassenen Aufgaben angegeben. Die Mittelwerte der Subtests zeigen, dass meist mehr als die Hälfte richtig beantwortet wurde, nur für Test 4, 9 und 11 traf dies nicht zu.

Durch die Schiefe der Skalen kann überprüft werden, inwieweit die Form einer Normalverteilung ähnelt. Dies ist vor allem für die Anwendung weiterer statistischer Tests von Bedeutung. Werte kleiner Null werden als rechtsteil, größer Null als linksteil und Werte um die Null als symmetrisch bezeichnet (Bortz, 1993). Um die Stärke der Abweichung von der Normalverteilung für Schiefe zu bewerten, werden Ausführungen von Hopkins und Weeks (1990) herangezogen. Anhand von Verteilungsbeispielen bewerten sie Ergebnisse von -.33 als leicht verschoben, .58 oder -.71 als moderat, 1.15 als substantiell und 1.76 als extrem bzw. 1.88 als sehr extrem verschoben. Kreuzpointner (2010) schlägt folglich auf der Basis dieser Einschätzung von Effektstärken vor,

Grenzen für geringe Schiefe unter .3, für mittlere zwischen .3 und .8 sowie für starke Verschiebung über .8 zusetzen. Für die Tests 1, 2, 6 bis 10 lässt sich eine leichte Verschiebung nach rechts feststellen. Die Tests 3, 4, 5 sowie Test 11 fallen positiv und daher eher linksschief aus. Insgesamt sind außer Test 5 und 8 alle annähernd symmetrisch und ähneln in etwa einer normalverteilten Form, wobei auch diese nur geringe Abweichungen aufweisen. Für Test 8 beobachtete bereits Kreuzpointner (2010) eine stärkere rechtsschiefe Verteilung und schlägt aufgrund dessen eine Verkürzung der Testzeit auf zwei Minuten vor.

Der Exzess oder Kurtosis, welcher die Wölbung einer Verteilung angibt, ist ein weiteres Maß zur Prüfung der Abweichungen zu einer Normalverteilung. Liegen die Werte unter Null wird die Verteilung als breitgipflig bezeichnet, über Null als schmalgipflig und bei gleich Null ähnelt die Form einer Normalverteilung (Bortz, 1993). Die Stärke des Exzesses soll auch hier nach den vorgeschlagenen Grenzen von Kreuzpointner bewertet werden. Die im positiven Bereich liegenden Subtests 3, 4, 5, 7, 8, 10 und 11 weisen eher schmalgipflige Verteilungsformen auf. Die Subtests 1, 2, 6 und 9 nehmen eher eine breite Verteilungsform an. Insgesamt sind die Formen, mit Ausnahme der Subtests 5 und 10, einer Normalverteilung ähnlich. Um die Verteilungen zu Schiefe und Exzess genauer zu überprüfen, sind im Anhang die Histogramme angefügt (vgl. Abbildung E-1). Anhand dieser ist zu erkennen, dass die Abweichungen der Tests 5 und 10 höchstwahrscheinlich durch die Konzentration der Ergebnisse um den Mittelwert zu begründen sind.

Die Ergebnisse zu Cronbachs Alpha geben an, wie stark die Items eines jeden Subtests das Gleiche messen. Diese internen Konsistenzen liegen für die Tests zwischen .71 und .93 auf einem hohen Niveau.

Zur weiteren Prüfung der Reliabilität sind in Tabelle 3.8 die Split-Half-Koeffizienten aufgeführt. Aus den Mittelwerten und Standardabweichungen der aufgeteilten Untertests geht hervor, dass einander ebenbürtige Testpaare gebildet wurden. Die nach Spearman-Brown-Formel korrigierten Halbierungsreliabilitäten sind für die Subtests zwischen .74 und .98 zufrieden stellend.

Tabelle 3.8: Deskriptive Statistik, Reliabilitätsanalyse nach Cronbachs Alpha und Test-halbierungsmethode der LPS-neu Subtests

Subtest	Items	<i>M</i>	<i>SD</i>	Schiefe	Exzess	α	<i>M</i> ¹⁾	<i>SD</i> ¹⁾	<i>r</i> _{tt} ¹⁾
Subtest 1	60	41.92	9.97	-0.39	-0.42	.92	18.06 21.87	4.42 4.8	.91 (.83)
Subtest 2	40	26.14	7.37	-0.32	-0.45	.91	13.43 12.71	3.80 3.79	.94 (.88)
Subtest 3	40	26.19	3.48	0.01	0.36	.76	8.33 8.86	1.90 1.92	.80 (.66)
Subtest 4	40	17.51	2.83	0.14	0.29	.71	6.03 6.49	2.51 2.54	.74 (.59)
Subtest 5	40	20.49	3.63	0.80	1.35	.80	7.65 5.83	1.81 2.07	.86 (.75)
Subtest 6	40	25.82	7.40	-0.01	-0.72	.93	12.26 11.56	3.70 3.80	.98 (.95)
Subtest 7	40	31.22	3.90	-0.47	0.53	.78	15.43 15.79	2.05 2.17	.83 (.71)
Subtest 8	40	33.21	5.25	-0.97	0.76	.89	12.83 12.38	2.73 2.76	.90 (.82)
Subtest 9	40	13.78	5.30	-0.08	-0.08	.87	7.36 6.42	2.76 2.72	.93 (.87)
Subtest 10	60	35.31	5.37	-0.72	2.51	.90	12.92 12.39	2.67 2.76	.98 (.96)
Subtest 11	80	19.89	7.51	0.21	0.51	.92	10.08 9.81	3.77 3.86	.97 (.94)

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, α = Cronbachs Alpha, *r*_{tt} = korrigierter Split-Half-Koeffizient nach Spearman-Brown-Formel

¹⁾ Teilung in zwei gleichgroße Subtest (ungerade vs. gerade), Angabe in Klammern = unkorrigierte Split-Half-Koeffizient

Tabelle 3.9 gibt einen Überblick der Interkorrelationen der einzelnen Subtests des LPS-neu sowie des Gesamtscores. Für die meisten Korrelationen kann die gewünschte bzw. eine tendenzielle Signifikanz bezüglich der postulierten Stratum II Faktoren festgestellt werden, d. h. die postulierten Zusammenhänge zwischen den Subtests zur kristallinen Intelligenz (1 und 2), zur fluiden Intelligenz (3, 4 und 5), zur visuellen Wahrnehmung (6, 7 und 8) sowie zur kognitive Geschwindigkeit (9, 10 und 11) sind signifikant. Diese Korrelationen sind positiv bei etwa mittlerer Stärke, wobei nur der Subtest 6 mit 7 und 8 geringere Zusammenhänge aufweist. Die signifikanten Korrelationen der Subtests zum LPS-neu Gesamtscore sind durchweg zufrieden stellend zwischen .47 und .75.

Tabelle 3.9: Interkorrelation der Subtests des LPS-neu (n = 72)

Subtests LPS-neu	1	2	3	4	5	6	7	8	9	10	11	Ges
1	-	.52**	.29*	.12	.25*	.32**	.14	.19°	.28*	.26*	.35**	.66**
2		-	.38**	.41**	.48**	.37**	.39**	.32**	.31**	.20°	.50**	.75**
3			-	.37**	.47**	.45**	.36**	.40**	.17	.28*	.40**	.62**
4				-	.59**	.20°	.36**	.44**	.30*	.21°	.44**	.55**
5					-	.36**	.43**	.50**	.33**	.23°	.41**	.67**
6						-	.19°	.18°	-.15	.26*	.40**	.57**
7							-	.48**	.34**	.43**	.42**	.60**
8								-	.22°	.38**	.32**	.59**
9									-	.30*	.31**	.47**
10										-	.34**	.56**
11											-	.74**

Anmerkungen: Ges = Korrekte Antworten insgesamt

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

3.3.4 Fragebogen zum aktuellen Wohlbefinden

3.3.4.1 Beschreibung des Fragebogens zum aktuellen Wohlbefinden

Wie stark die Stressinstruktion das momentane Befinden der Vpn beeinflusst hat, wurde mittels eines Fragebogens zum aktuellen Wohlbefinden erfasst (vgl. Anhang B-2). Damit eine Veränderung im gefühlten Stresslevel festgestellt werden konnte, mussten die Teilnehmer den Fragebogen insgesamt fünf Mal bearbeiten. Einmal kamen vor und nach der Intelligenztestung des CFT 20-R die Fragebögen zum Wohlbefinden 1 und 2 zum Einsatz. Fragebogen 3 wurde vor dem LPS-neu ausgehändigt, die Experimentalgruppe erhielt zusätzlich Fragebogen 4 nach der Stressinstruktion. Abschließend beantworteten alle Vpn nach dem LPS-neu den fünften Fragebogen zum aktuellen Wohlbefinden.

Angelehnt an einen Zusatzfragebogen der Beanspruchungsmessskalen (BMS – Ermüdung, Monotonie, Sättigung, Stress) von Plath und Richter (1984) wurde der Fragebogen zum aktuellen Wohlbefinden speziell für diese Untersuchung erstellt. Das BMS ist ein Verfahren, welches erlebte Ansprüche, besonders in Bezug auf den Arbeitsplatz, skaliert erfasst. Pro Item stehen sich zwei Adjektive gegenüber, wobei ein Adjektiv das positive und das andere das negative entgegengesetzte Gefühl bezeichnet. Wie in Abbildung 3.8 dargestellt, muss die Vp mittels einer Skala zwischen dem Adjektivpaar einschätzen, welcher Pol ihr aktuelles Befinden am besten beschreibt.

Der Fragebogen zum aktuellen Wohlbefinden enthält insgesamt acht bipolare Items. Vier entsprechen dabei dem Zusatzfragebogen des BMS (Entspannung, Interesse, Frische, Aufmerksamkeit) und 3 wurden neu hinzugefügt (Unbekümmertheit, Vertrauen, Behaglichkeit) bzw. eines umformuliert (Gelassenheit). Bei der Auswertung der Skala werden die gesetzten Striche jeweils in die Richtung des geforderten Gefühls gemessen. Bei dem in Abbildung 3.8 dargestellten ersten Item, wird von links nach rechts gemessen, um das angegebene Entspannungslevel zu erfassen. Für Interesse verhält es sich genau umgekehrt.

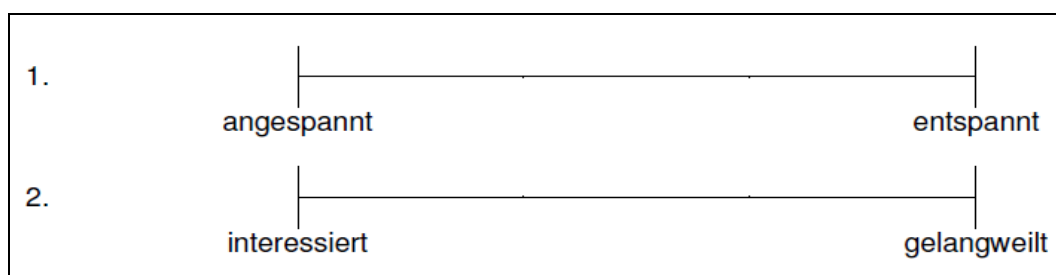


Abbildung 3.8: Zwei Itembeispiele des Fragebogens zum aktuellen Wohlbefinden

3.3.4.2 Itemanalyse der Fragebögen zum aktuellen Wohlbefinden

Die Analyse der beschriebenen acht Items zur Beurteilung des momentanen Befindens für alle fünf Fragebögen ist in der Tabelle 3.10 dargestellt. Aus der Tabelle sind die Mittelwerte, Standardabweichungen und Trennschärfen der Items zu entnehmen. Die Zustimmung zu dem jeweiligen Gefühl ist umso höher, je höher der jeweilige Wert des Mittelwerts der Tabelle ist. Die Trennschärfen geben durch die Korrelation des Items mit dem Gesamtergebnis des Tests repräsentiert.

Tabelle 3.10: Itemkennwerte des Fragebogens zum aktuellen Wohlbefinden 1 (n = 72)

	W1	W1	W2	W2	W3	W3	W4¹⁾	W4¹⁾	W5	W5
	M (SD)	r_{it}	M (SD)	r_{it}	M (SD)	r_{it}	M (SD)	r_{it}	M (SD)	r_{it}
E	60.06 (20.29)	.50	52.13 (21.64)	.56	62.90 (18.45)	.68	54.11 (22.95)	.78	46.46 (22.08)	.62
I	74.19 (18.20)	.52	71.06 (17.80)	.38	65.74 (20.44)	.54	64.25 (21.81)	.40	61.81 (22.08)	.53
U	63.61 (23.24)	.62	58.60 (19.72)	.62	64.26 (20.85)	.74	52.28 (24.58)	.82	55.50 (22.24)	.73
F	53.10 (24.59)	.42	51.36 (20.71)	.56	50.92 (24.35)	.53	45.11 (22.26)	.32	38.28 (23.78)	.63
G	61.06 (21.39)	.55	59.72 (19.03)	.59	59.61 (22.00)	.57	52.17 (23.15)	.78	57.44 (21.90)	.63
V	66.19 (19.67)	.40	61.17 (23.31)	.58	66.13 (20.10)	.64	55.36 (23.62)	.72	55.26 (23.10)	.69
B	66.19 (18.61)	.71	60.86 (19.22)	.74	66.00 (18.27)	.82	54.14 (24.19)	.75	55.17 (23.05)	.84
A	70.74 (18.99)	.51	66.78 (18.63)	.51	63.64 (19.82)	.61	58.67 (22.61)	.45	53.57 (26.45)	.59

Anmerkungen: M = Mittelwert, SD = Standardabweichung, r_{it} = Trennschärfe, W = Wohlbefinden, E = Entspannung, I = Interesse, U = Unbekümmertheit, F = Frische, G = Gelassenheit, V = Vertrauen, B = Behaglichkeit, A = Aufmerksamkeit
¹⁾ für Wohlbefinden 4 ist n = 16

Aus den Mittelwerten zum Fragebogen 1 ergibt sich, dass der Komponente Interesse am meisten und Frische am wenigsten zugestimmt wurde (vgl. Tabelle 3.10). Insgesamt zeigt sich, dass die Vpn den mittleren bis hohen Bereich für ihre Bewertungen bevorzugten. Die Trennschärfen liegen durchgehend in einem guten Bereich und variieren zwischen .50 und .71, somit muss auf keines der Items verzichtet werden.

Die Itemkennwerte zu Wohlbefinden 2 zeigen, dass auch hier die Mittelwerte im mittleren bis hohen Bereich liegen. Interesse weist die höchste und Frische die niedrigste Zustimmung auf. Die Trennschärfen der Skalen liegen in einem gewünschten Bereich zwischen .51 und .74, nur Interesse weist eine Trennschärfe von .38 auf.

Bei dem Fragebogen 3 variieren die Itemmittelwerte zwischen 50.92 (SD = 24.35) für Frische und 66.13 (SD = 20.10) für Vertrauen. Alle Items liegen im mittleren Bereich und somit etwas niedriger im Vergleich zu den ersten zwei Fragebögen. Sie weisen auch hier auf eine eher zögerliche Zustimmung hin. Alle Trennschärfen liegen wiederum in dem gewünschten Bereich zwischen .54 und .82.

Für die Mittelwerte des vierten Fragebogens ergibt sich ebenfalls für Frische die niedrigste und Interesse die höchste Zustimmung. Im Vergleich zu den anderen Bögen wurde den Aussagen hier geringer zugestimmt, da die Mittelwerte niedriger ausfallen.

Es ergaben sich hohe Trennschärfen zwischen .72 und .78 für Entspannung, Unbekümmertheit, Gelassenheit, Vertrauen und Behaglichkeit. Die Items Interesse, Frische und Aufmerksamkeit weisen nur mittlere Trennschärfen zwischen .32 und .45 auf.

Die Mittelwerte des letzten Fragebogens zum aktuellen Wohlbefinden sind ebenfalls eher niedrig. Auch in diesem Fragebogen ist die gefühlte Frische am niedrigsten und Interesse am höchsten ausgeprägt. Die ermittelten Trennschärfen für den fünften Fragebogen sind durchgehend in einem gewünschten Bereich zwischen .53 und .84.

3.3.4.3 Skalenanalyse der Fragebögen zum aktuellen Wohlbefinden

Die Verteilung der Gesamtskalen zu den fünf Fragebögen des aktuellen Wohlbefindens ist in Tabelle 3.11 verdeutlicht. Die Mittelwerte liegen zwischen 64.40 ($SD = 13.56$) bei dem Fragebogen 1 und 52.94 ($SD = 17.22$) für Fragebogen 4. Das Wohlbefinden liegt somit allgemein im mittleren Bereich bzw. deutet tendenziell auf ein durchschnittliches Wohlfühlgefühl der Personen hin, wobei dies von der ersten zur letzten Messung leicht abnimmt und nur zu Beginn der zweiten Testung nochmals ansteigt. Insgesamt sind die Verteilungen relativ symmetrisch, nur die Fragebögen 2 und 4 zeigen eine leicht linkssteile Verteilung. Während die Fragebögen 1, 3 und 5 eher breitgipflig sind, weisen 2 und 4 eine schmalgipflige Form auf. Auch die Abweichungen der Gipfformen befinden sich nur im geringen bis mittleren Bereich. Die Ergebnisse zu Schiefe und Exzess erlauben folglich die Annahme, dass die Verteilungen einer Normalverteilung ähneln, was auch aus den Histogrammen im Anhang (vgl. Abbildung E-3) deutlich wird.

Die internen Konsistenzen der Fragebögen zum Wohlbefinden sind in Tabelle 3.11 angegeben. Es konnten Werte für Cronbachs Alpha zwischen .81 bis zu .89 ermittelt werden, was auf mittlere bis hohe Reliabilitäten schließen lässt. Die internen Konsistenzen liegen relativ hoch und somit scheinen die einzelnen 8 Items mit der jeweiligen Skala zum Allgemeinen Wohlbefinden 1 bis 5 relativ hoch zu korrelieren.

Tabelle 3.11: Deskriptive Statistik der Fragebögen zum aktuellen Wohlbefinden 1 bis 5 (jeweils 8 Items)

	<i>M</i>	<i>SD</i>	Schiefe	Exzess	α
Wohlbefinden Gesamt 1	64.40	13.56	0.09	-0.26	.81
Wohlbefinden Gesamt 2	60.21	13.72	0.36	0.60	.84
Wohlbefinden Gesamt 3	62.40	15.05	0.05	-0.66	.88
Wohlbefinden Gesamt 4	54.51	16.78	0.28	0.39	.87
Wohlbefinden Gesamt 5	52.94	17.22	0.01	-0.18	.89

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, α = Cronbachs Alpha

3.3.4.4 Faktorenanalyse der Fragebögen zum aktuellen Wohlbefinden

Um zu überprüfen, ob sich die einzelnen Komponenten des Fragebogens zum momentanen Befinden unter verschiedenen Faktoren zusammenfassen lassen, wurde eine Faktorenanalyse durchgeführt. Zunächst konnte mit den Berechnungen des Kaiser-Meyer-Olkin-Koeffizienten, welcher sich für die fünf Fragebögen zwischen .70 und .84 bewegte, gezeigt werden, dass die Daten für eine Faktorenanalyse gut geeignet sind.

Die Hauptkomponentenanalysen mit Varimax-Rotation mittels der Eigenwertmethode (Eigenwert > 1) liefern keine eindeutige Ergebnisse (vgl. Anhang D, Tabelle D-1). Es konnten zwei statistisch unterscheidbare Faktoren mit Eigenwert größer 1 für Fragebogen 2 bis 5 und für Fragebogen 1 drei Faktoren ermittelt werden. Für den ersten Fragebogen wurde erneut eine Hauptkomponentenanalyse, mit Begrenzung auf zwei Faktoren, durchgeführt. Bei den Faktorladungen zeigt sich tendenziell, dass Interesse, Frische und Aufmerksamkeit auf dem zweiten Faktor laden. Interesse und Vertrauen können jeweils in einer und Behaglichkeit in drei Analysen beiden Faktoren zugeordnet werden. Durch diese uneindeutigen Ergebnisse der Faktorenanalyse ist es schwierig, eine sinnvolle Einteilung zu finden. Mit Faktor 1 kann aber bei allen Fragebögen über die Hälfte der Varianz erklärt werden. Dadurch und aufgrund hoher interner Konsistenzen und somit hoher Korrelationen aller Items auf der Gesamtskala „Wohlbefinden“ werden die einzelnen Faktoren nicht weiter unterteilt.

3.3.5 Evaluationsfragebogen des CFT 20-R und des LPS-neu

3.3.5.1 Beschreibung des Evaluationsfragebogens

Um zu ermitteln, wie die Teilnehmer die beiden Intelligenztests bewerteten, wurde eigens für diese Studie ein Fragebogen zur Testevaluation für den LPS-neu und den CFT 20-R zusammengestellt (vgl. Anhang B-3). Hauptsächlich sollte diese Evaluation die Entspannungsinstruktion der Parallelstudie verstärken, in der den Vpn die ausschließliche Überprüfung einer Neuentwicklung des Intelligenztests LPS-neu suggeriert wurde (Stadler, 2010). Die Fragebögen zum CFT 20-R und dem LPS-neu sind analog aufgebaut und beinhalten acht Aussagen, welche die Vpn anhand einer Likert-Skala bewerten sollen, wie stark diese auf sie zutreffen. Dabei entspricht 1 „trifft gar nicht zu“ und 5 „trifft sehr zu“ (vgl. Tabelle 3.12).

Tabelle 3.12: Zwei Beispielitems zur Testevaluation des CFT 20-R und des LPS-neu

	Antwortmuster: 1 = trifft gar nicht zu 2 = trifft eher nicht zu 3 = teils teils 4 = trifft eher zu 5 = trifft sehr zu				
1. Die Anweisungen zu den Aufgaben sind verständlich.	1	2	3	4	5
2. Die Schwierigkeit der Aufgaben ist zu hoch.	1	2	3	4	5

3.3.5.2 Itemanalyse der Evaluationsfragebögen

In Tabellen 3.13 sind Mittelwerte, Standardabweichungen und Trennschärfen für die acht Items des Evaluationsfragebogens zum CFT 20-R und LPS-neu angegeben. Für die Items „Schwierigkeit ist zu hoch“, „Schwierigkeit ist zu niedrig“ und „Testzeit ist insgesamt zu lang“ wurde für weitere Analysen eine Umkodierung vorgenommen. Somit können die Mittelwerte aller acht Aussagen als positive bzw. negative Testbewertung interpretiert werden. Dabei gilt, je höher das arithmetische Mittel ist, desto positiver wird der Intelligenztest bewertet.

Tabelle 3.13: Itemkennwerte der Evaluationsfragebögen (n = 72)

	Eval. CFT 20-R			Eval. LPS-neu		
	<i>M</i> (<i>SD</i>)	r_{it}	$r_{it}^{2)}$	<i>M</i> (<i>SD</i>)	r_{it}	$r_{it}^{2)}$
Die Anweisungen sind verständlich	4.71 (.59)	.38	.39	4.18 (.94)	.30	.28
Schwierigkeit ist zu hoch ¹⁾	3.47 (.92)	.06	.13	2.71 (.83)	.24	.33
Schwierigkeit ist zu niedrig ¹⁾	3.61 (.83)	.03	-	4.10 (.79)	-.02	-
Verfügbare Zeit ist angemessen	3.57 (.95)	.34	.39	2.15 (.88)	.19	.23
Aufgabenanordnung ist übersichtlich	4.50 (.90)	.38	.37	3.75 (1.24)	.51	.54
Antwortbogen ist übersichtlich	4.15 (1.12)	.32	.33	3.69 (1.11)	.59	.58
Testzeit insgesamt zu lang	4.22 (.77)	.31	.28	3.78 (1.22)	.32	.47
Test ist interessant	3.89 (.88)	.29	.24	3.58 (1.02)	.45	.26

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, r_{it} = Trennschärfe

¹⁾ = Umkodierte Items bei denen 1 = trifft sehr zu und 5 = trifft gar nicht zu

²⁾ = Trennschärfe ohne dem Item „Schwierigkeit ist zu niedrig“

Bei der Bewertung des CFT 20-R wurde der Aussage „Test ist interessant“ am geringsten ($M = 3.89$, $SD = 0.88$), während der Aussage „Die Anweisungen sind verständlich“ ($M = 4.71$, $SD = .59$) am meisten zugestimmt wurde (vgl. Tabelle 3.13). Die Schwierigkeit wurde tendenziell als weder zu hoch noch zu niedrig und die Testlänge als nicht zu lang bewertet wurde. Die Mittelwerte über alle acht Items liegen über 3, dem Mittelpunkt der Skala, und folglich wurden die Tests eher positiv bewertet. Für vier Komponenten konnten gute Trennschärfen $>.30$ ermittelt werden. Die Trennschärfen der einzelnen Komponenten der zu niedrigen und zu hohen Schwierigkeit gehen gegen Null. Das Item „Schwierigkeit zu niedrig“ wird aufgrund der geringsten Trennschärfe aus weiteren Analysen ausgeschlossen. Dadurch kann nun auch für „Schwierigkeit ist zu hoch“ eine bessere Trennschärfe ermittelt werden, während die Trennschärfen der übrigen Items sich weiterhin zwischen Werten von $.24$ und $.39$ bewegen.

Die Ergebnisse der Itemanalyse für den zweiten Evaluations-Fragebogen zum LPS-neu sind ebenfalls in Tabelle 3.13 dargestellt. Die Mittelwerte der nicht umkodierten Items dieses Fragebogens variieren hier zwischen 2.15 ($SD = .88$) und 4.18 ($SD = .94$). Dem Item „Verfügbare Zeit ist angemessen“ wurde demnach am wenigsten zugestimmt und der Mittelwert für „Anweisungen sind verständlich“ ist auch in diesem Fragebogen am höchsten. Es ergaben sich annehmbare Trennschärfen zwischen $.24$ und $.59$ für sechs Items. Die Trennschärfe für „Verfügbare Zeit ist angemessen“ liegt bei $.19$ und für „Schwierigkeit ist zu niedrig“ ergab sich sogar eine negative Trennschärfe. Deshalb wird auch für diesen Fragebogen das Item „Schwierigkeit ist zu niedrig“ für weitere Analysen nicht beibehalten. Aus den ermittelten Trennschärfen nach der Selektion ist bereits zu erkennen, dass die Güte des Fragebogens steigt, da sich nun durchgängig Trennschärfen $>.20$ ermitteln lassen.

3.3.5.3 Skalenanalyse der Evaluationsfragebögen

Ein Überblick zu den Skalenwerten der Bewertungsfragebögen ist in Tabelle 3.14 zu sehen, wobei das Item „Schwierigkeit ist zu niedrig“ in diese Analyse nicht mit einfließt. Insgesamt betrachtet, bewerten die Vpn den CFT 20-R positiver als den LPS-neu, da für diesen Evaluationsfragebogen sich ein höherer Mittelwert von 4.07 mit einer Standardabweichung von $.47$ ergibt (vgl. Tabelle 3.14). Die durchschnittliche Zustimmung zum LPS-neu beträgt 3.41 ($SD = .61$). Bei beiden Evaluationsfragebögen zeigen sich rechtsschiefe Verteilungen, welche beim CFT 20-R eine schmalgipflige und beim LPS-neu tendenziell eine breitgipflige Form annehmen. Deutlich wird aus den Werten und den im Anhang angegebenen Histogrammen (vgl. Tabelle E-4), dass die Evaluation des LPS ähnlich einer Normalverteilung ist, die des CFT 20-R jedoch etwas höher

abweicht. Die Werte für Cronbachs Alpha liegen im Bereich von .57 und .67, somit ist eine noch gute interne Konsistenz gewährleistet.

Tabelle 3.14: Deskriptive Statistik der Evaluationsfragebögen des CFT 20-R und des LPS-neu (jeweils 7 Items)

	<i>M</i>	<i>SD</i>	<i>Schiefe</i>	<i>Exzess</i>	<i>α</i>
Evaluation Gesamt CFT	4.07	0.47	-0.97	1.23	.57
Evaluation Gesamt LPS-neu	3.41	0.61	-0.36	-0.26	.67

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, *α* = Cronbachs Alpha

3.3.5.4 Faktorenanalyse der Evaluationsfragebögen

Eine Faktorenanalyse, bestehend aus der Hauptkomponentenanalyse mit Varimax-Rotation, wurde für beide Bewertungsfragebögen durchgeführt (vgl. Anhang D, Tabelle D-2). Vorab wurde das Kaiser-Meyer-Olkin-Kriterium zur Stichprobeneignung für eine Faktorenanalyse ermittelt, welches mäßig gute Werte von .62 und .60 annahm.

Bei dem ersten Fragebogen ergaben sich drei unterscheidbare Faktoren und beim zweiten wurden zwei Faktoren ermittelt. Daher wurde erneut für den Fragebogen zum CFT 20-R eine faktoranalytische Überprüfung mit der Begrenzung auf zwei Faktoren durchgeführt. Die Varianz, die durch diese Faktoren erklärt wird, liegt bei dem ersten Faktor zwischen 26.1 % und 30.6 % und bei dem zweiten zwischen 18.7 % und 21.2 %. Es zeigen allerdings nur die zwei Items „Verständlichkeit der Anweisungen“ und „Übersichtlichkeit des Antwortbogens“ bei beiden Fragebögen ähnliche Ladungen. Die übrigen laden entweder auf entgegengesetzten Faktoren bzw. auf beiden. Aus diesem Grund kann keine sinnvolle Trennung der Items auf unterschiedliche Faktoren vorgenommen werden. Die Gesamtskala, welche auch eine gute Reliabilität zeigt, wird für weitere Analysen beibehalten.

3.3.6 Fragebogen zur subjektiven Beurteilung der Stressreaktion

3.3.6.1 Beschreibung des Stressreaktions-Fragebogens

Nach der Aufklärung der Experimentalgruppe über die Zielstellungen der Untersuchung sowie über die erfundenen Ergebnisse und Verteilungen, erhielten die Vpn einen Fragebogen, um ihr subjektiv erlebtes Stressniveau einzuschätzen. Die Vpn sollten in einem für diese Studie entwickelten Stressreaktions-Fragebogen beurteilen, wie sie das schlechte Ergebnis aus dem ersten Intelligenztest empfunden haben und wie sie ihre darauf folgende Reaktion einschätzen. Der Fragebogen war in dem schriftlichen

Debriefing enthalten und bestand aus drei Aussagen (vgl. Anhang A-3). Analog zum Evaluationsfragebogen sollten die Vpn die Aussagen anhand einer Likert-Skala zwischen 1 „trifft gar nicht zu“ bis 5 „trifft sehr zu“ bewerten (vgl. Tabelle 3.15). Die drei Aussagen umfassen Ärger, Frustration und Druck, welche bereits Selye (1974) unter anderem zu wichtigen stressrelevanten Themen zählt.

Tabelle 3.15: Die Items zur Beurteilung der Reaktion auf den fiktiven Intelligenzwert

	Antwortmuster: 1 = trifft gar nicht zu 2 = trifft eher nicht zu 3 = teils teils 4 = trifft eher zu 5 = trifft sehr zu				
1. Das Ergebnis meines Intelligenzwerts hat mich frustriert.	1	2	3	4	5
2. Ich habe mich über das Ergebnis meiner Leistung geärgert	1	2	3	4	5
3. Ich habe mich durch den Wert unter Druck gesetzt gefühlt.	1	2	3	4	5

3.3.6.2 Testanalyse des Stressreaktions-Fragebogens

Für den Fragebogen zur subjektiven Beurteilung der Stressreaktion wurden Mittelwerte, Standardabweichungen und Trennschärfen für die drei Items ermittelt (vgl. Tabelle 3.16). Auch hier gilt nach der 5-stufigen Likert-Skala, je höher die Zustimmung, desto höher der Wert. Die Mittelwerte sind relativ gleich hoch ausgeprägt und liegen zwischen 3.22 ($SD = 1.44$) für den gefühlten Druck und 3.78 ($SD = 1.22$) für Frustration. Mit Werten über .73 liegen die Trennschärfen in einem optimalen Bereich.

Tabelle 3.16: Itemkennwerte des Stressreaktions-Fragebogens

	<i>M</i>	<i>SD</i>	<i>r_{it}</i>
Frustration	3.78	1.22	.80
Ärger	3.42	1.34	.75
Druck	3.22	1.44	.73

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung *r_{it}* = Trennschärfe

Insgesamt konnte ein Mittel von 3.47 mit einer Standardabweichung von 0.54 für den Fragebogen zur Stressreaktion berechnet werden (vgl. Tabelle 3.17). Die Verteilung ist

leicht rechtssteil und die Form eher breitgipflig ausgeprägt. Somit weicht die Verteilung nicht stark von der Normalverteilung ab, was auch die grafische Verteilung im Anhang offensichtlich macht (vgl. Tabelle E-5). Die Reliabilität der Gesamtskala liegt bei .87 und ist somit optimal, d. h. dass die drei abgefragten Items mit der Gesamtskala Stress in einem hohen Zusammenhang stehen.

Tabelle 3.17: Deskriptive Statistik des Stressreaktions-Fragebogens

	<i>M</i>	<i>SD</i>	Schiefe	Exzess	α
Stressreaktion Gesamt	3.47	0.54	-0.50	-0.65	.87

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, α = Cronbachs Alpha

Bei der Berechnung der Interkorrelationen des Fragebogens zum gefühlten Stresslevel ergeben sich drei hoch signifikante, positive Zusammenhänge (vgl. Tabelle 3.18). Die höchste Korrelation liegt für Frustration und Ärger mit .74 vor.

Tabelle 3.18: Interkorrelationen der 3 Komponenten des Stressreaktions-Fragebogens

	Frustration	Ärger	Druck
Frustration	-	.74**	.71**
Ärger			.65**
Druck			-

Anmerkungen: **) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

3.3.6.3 Faktorenanalyse des Stressreaktions-Fragebogens

Der ermittelte Kaiser-Meyer-Olkin-Koeffizient von .72 zeigt, dass die Daten für eine weitere Faktorenanalyse geeignet sind. Nach der Faktorenanalyse der drei Items des Fragebogens zur eingeschätzten Stressreaktion ergibt sich ein statistisch unterscheidbarer Faktor mit Eigenwert größer 1. Alle drei Items können folglich, wie beabsichtigt, unter einem Faktor zusammengefasst werden und somit ist keine weitere Analyse mit Varimax-Methode nötig. Die Items werden dem Oberbegriff „gefühltes Stresslevel“ zugeordnet und dieser Faktor erklärt 80.0 % der Gesamtvarianz.

3.3.7 Demografischer Fragebogen

Mit dem Fragebogen zur Person wurden die Daten zum Geschlecht, Geburtsdatum, Alter, Studienfach und Semesteranzahl sowie Berufstätigkeit und erlernter Beruf, Händigkeit, Muttersprache, Schulabschluss und Abiturnote erfasst (vgl. Anhang B-1). Zusätzlich konnten die Teilnehmer im Debriefing abschließend angeben, wer über die Ergebnisse der Intelligenztests informiert werden möchte und wer nicht.

3.4 Statistische Datenauswertung

Die Auswertung der Daten wird mittels der Version 16.0 des „Statistical Package for the Social Sciences“ (SPSS) und „Microsoft Office Excel XP“ vorgenommen.

3.4.1 Verwendete statistische Verfahren

Zur Auswertung der Daten der untersuchten Stichproben wurden vor Beginn der Berechnungen alle Variablen auf Skalenniveau geprüft, um die postulierten Zusammenhänge der Variablen mittels geeigneter Korrelationsmaße zu prüfen.

Einzelne Vergleiche der Merkmale werden mittels des Welch-Tests anstelle des üblichen *t*-Tests für unabhängige Stichproben berechnet. Bei einem *t*-Test sollten die Voraussetzungen der Normalverteilung und der gleichen Varianzen gegeben sein, während der *W*-Test auch ohne diese aufgrund stärkerer Robustheit des Tests einsetzbar ist (Kubinger, Rasch & Moder, 2009). Bei bestehender Homogenität liefert er annähernd die gleichen Ergebnisse wie ein *t*-Test. Außerdem steigt durch die Voraussetzungsüberprüfung für den *t*-Test mittels mehrerer statistischer Tests die Wahrscheinlichkeit des α -Fehlers (Kubinger et al., 2009). Daher wurde der *W*-Test durchgängig dem *t*-Test vorgezogen.

Weiterhin wurden für einzelne Gegenüberstellungen der unterschiedlichen Messzeitpunkte des Wohlbefindens Differenzvergleiche mittels *t*-Tests für abhängige Stichproben durchgeführt. Die Effektstärken der einzelnen Mittelwertvergleiche wurden nach Hedges *g* berechnet. Dabei wird die mittlere Differenz der beiden Mittelwerte durch die gemittelte Standardabweichung geteilt. Bortz (1993) gibt einen Vorschlag zur Einteilung der Stärke der Effektgrößen an. Werte um .20 wertet er als schwach, um .50 als mittlerer Effekt und als starker Effekt werden Werte mit .80 betrachtet.

Um mehrere Mittelwertsunterschiede zu überprüfen, sind für diese Studie multivariate Varianzanalysen und für die Post-hoc Vergleiche univariate Varianzanalysen geeignet. Für die Anwendung der multivariaten Varianzanalyse wurde stets der Box-Test auf Gleichheit der Kovarianzmatrizen durchgeführt. Um die Stärke des Einflusses der einzelnen Variablen anzugeben, wird das Effektstärkemaß Eta-Quadrat (η^2) bzw.

partielltes Eta-Quadrat ($_{\text{partial}}\eta^2$) angegeben (Cohen, 1973). Es gibt an, wie viel der Variation der abhängigen Variable auf die unabhängige zurückgeführt werden kann. Während für Eta-Quadrat die Quadratsumme der unabhängigen Variable durch die gesamte geteilt wird, steht bei der Berechnung partielltes Eta-Quadrat die Summe aus der Quadratsumme der Variable und der Fehlerquadratsumme im Nenner (Cohen, 1973). Für diese Studie ist das partielle Eta-Quadrat interessant, da es die Stärke des Einflusses der einzelnen Variable auf das Modell, unabhängig von anderen einbezogenen Variablen, ermittelt. Für die Bewertung der Stärke der Effektgröße schlägt Cohen (1988) folgende Einteilung in .01 als schwach, .06 als mittel und .14 als starker Effekt vor. Oft dient der Intelligenzwert des CFT 20-R, entsprechend der Homogenitätsanalyse der Gruppen (siehe Kapitel 3.4.2), als Kovariate. So wurde der Einfluss dieses Faktors vorab herauspartialisiert. Die Interpretationen der Effektstärken der multivariaten Kovarianzanalysen wurden aufgrund möglicher Fehlinterpretationen weggelassen, da es schwierig ist, die Kovariate bei der Interpretation der Effekte zu berücksichtigen.

Um die genannten Verfahren anzuwenden, müssen gewisse Voraussetzungen gegeben sein. Einerseits sollten die zu vergleichenden Gruppen generell relativ homogen sein, dafür folgt in Abschnitt 3.4.2 die genaue Analyse der Gruppen auf ihre Vergleichbarkeit. Eine weitere wichtige Voraussetzung ist neben der Homogenität der Gruppen die Normalverteilung der relevanten Variablen. Da die in Kapitel 3.3 beschriebene Schiefe der interessierenden Variablen tendenzielle Symmetrie aufweisen, kann auf weitere Analysen bezüglich der Verteilung verzichtet werden. Im Fall einer symmetrischen Verteilung sind die genannten statistischen Verfahren nach Ausführungen von Diehl und Arbinger (2001) relativ robust. Diehl und Arbinger (2001) weisen allerdings darauf hin, beim Auftreten einer schiefen Verteilung ein strengeres Signifikanzniveau zu wählen. Die Anwendung von Verfahren zur Prüfung von Mittelwertsunterschieden bei gegebener Symmetrie ist daher unbedenklich und außerdem würden, wie bei Kubinger, Rasch und Moder (2009) diskutiert, weitere Tests eine Vergrößerung des α -Fehlers zur Folge haben. Ähnliches gilt, wie bereits beschrieben, für die dritte Voraussetzung der Varianzhomogenität bzw. der Varianzhomogenität der Fehlervarianzen für die Varianzanalysen. Um die weiteren Tests zu vermeiden, wird für diese Studie der Welch-Test als robusteres Verfahren gewählt. Bezüglich der Varianzanalyse ist nach Diehl und Arbinger (2001) im Falle heterogener Varianzen bei gegebener, tendenziell symmetrischer Verteilung und bei gleichgroßen Stichproben (Bortz, 1993) das Verfahren robust.

Für die Prüfung der Signifikanz der Ergebnisse werden Werte unter .01 als sehr signifikant, unter .05 als signifikant bewertet. Alles über .05 gilt als nicht signifikant, wobei Werte unter .10 als tendenziell signifikante Ergebnisse gekennzeichnet sind.

3.4.2 Prüfung der Homogenität der zu vergleichenden Gruppen

Um in der Auswertung die Daten der einzelnen Subgruppen der Gesamtstichprobe auf die verschiedenen Variablen miteinander zu vergleichen, soll zunächst die Vergleichbarkeit der einzelnen Gruppen analysiert werden. Einen Überblick der Homogenitätsprüfung anhand der angegebenen demographischen Daten liefert Tabelle 3.19. Es wurde untersucht, ob sich die Experimental- und Kontrollgruppe sowie die Motivationsgruppen (HE und FM) in den Variablen signifikant voneinander unterscheiden. Bezüglich der ausführlichen Erläuterung der Aufteilung der Stichprobe in Personen mit überwiegend HE und überwiegend FM wird auf Kapitel 4.2.1 verwiesen. Anhand der Kriterien Geschlecht, Alter, Studienfach, Semesteranzahl, Abiturnote, HE, FM, GM und NH, der zugehörigen Motivationsgruppe und der Intelligenzwerte des CFT 20-R wurden die Gruppen mittels Welch-Test oder χ^2 -Test analysiert.

Tabelle 3.19: Überblick der Homogenitätsprüfung der zu vergleichenden Gruppen anhand der demographischen Daten

	Gruppen (KG vs. EG)			LM-Gruppen (HE vs. FM)		
	df	χ^2/t	p	df	χ^2/t	p
Geschlecht (χ^2 -Test)	1	1.60	.21	1	6.77	<.01**
Alter (W-Test)	69.82	-3.82	.70	59.23	1.70	.10
Studienfach ¹⁾ (χ^2 -Test)	5	6.82	.23	5	6.01	.31
Semesterzahl (W-Test)	69.21	-0.72	.47	64.02	1.08	.28
Abiturnote (W-Test)	68.25	2.15	.04*	57.56	1.27	.21
HE (W-Test)	65.94	0.78	.44	-	-	-
FM (W-Test)	67.13	0.51	.61	-	-	-
GM (W-Test)	69.37	1.05	.30	-	-	-
NH (W-Test)	69.89	0.02	.98	-	-	-
LM- Gruppe (χ^2 -Test)	1	0.51	.48	-	-	-
CFT IQ-Wert (W-Test)	69.84	-3.67	<.01**	60.02	0.15	.89

Anmerkungen: KG = Kontrollgruppe (n = 36), EG = Experimentalgruppe (n = 36), LM = Leistungsmotivgruppe, HE: n = 31, FM: n = 41, df = Anzahl der Freiheitsgrade, t und χ^2 = Testwert, p = Signifikanzwert

¹⁾ Studienfachgruppierung verwendet, eine fehlende Angabe

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

Bei der Überprüfung der Unterschiede zwischen Experimental- und Kontrollgruppe bezüglich der verschiedenen Merkmalsausprägungen zeigen sich zwei signifikante Unterschiede. Die Experimentalgruppe wies eine mit einem Durchschnitt von 1.9 (SD = 0.5) geringfügig bessere Abiturnoten auf. In der Kontrollgruppe liegt der Mittel-

wert der erreichten Abiturnoten bei 2.2 ($SD = 0.6$). In Abbildung 3.9 sind die Boxplotverteilungen der Abiturnoten getrennt für Kontroll- und Experimentalgruppe dargestellt. Der Modus der Experimentalgruppe beträgt 1.8 und die Spannweite der Noten reicht von 1.0 bis 3.2. Bei der Kontrollgruppe liegt der Modus bei 2.2 und die Spannweite umfasst die Noten zwischen 1.1 bis 3.5.

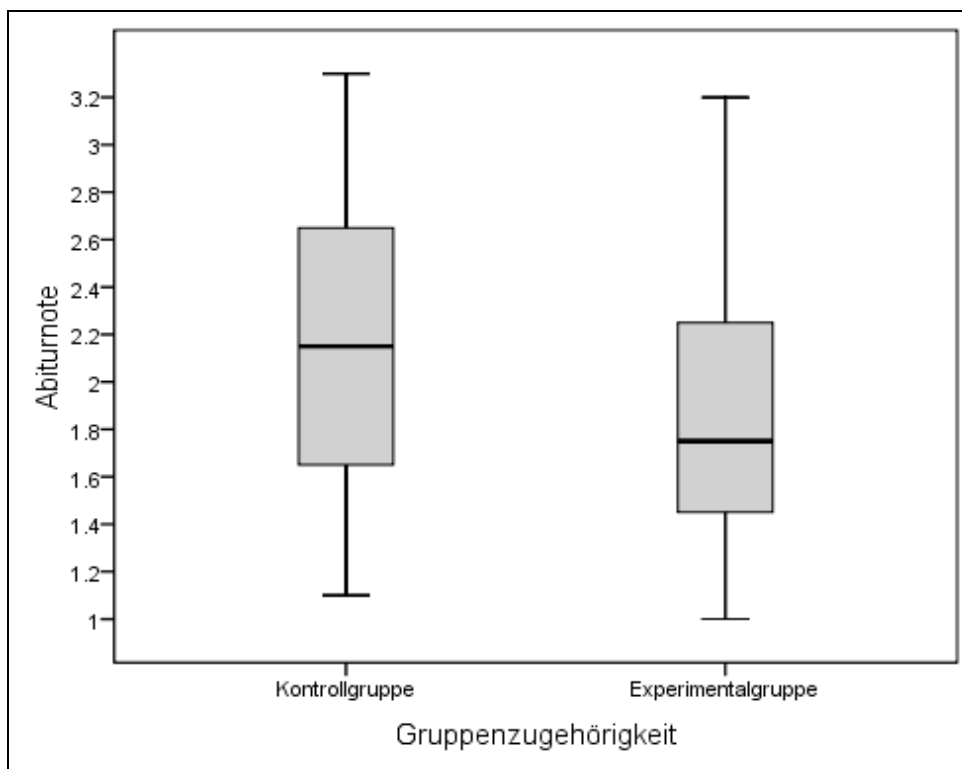


Abbildung 3.9: Boxplot der Abiturnoten getrennt für Experimental- und Kontrollgruppe

Weiterhin wurden signifikante Unterschiede bezüglich des ermittelten CFT 20-R IQ-Werts zwischen der Gruppe mit und ohne Stressinstruktion festgestellt. Die Experimentalgruppe weist hier signifikant höhere Werte ($M = 121.17$, $SD = 12.34$) im Vergleich zur Kontrollgruppe auf ($M = 110.22$, $SD = 12.95$). Der Vergleich der übrigen Kategorien bezüglich der Gruppenzugehörigkeit liefert keine weiteren signifikanten Unterschiede.

Beim Vergleich der Gruppen mit überwiegend HE zu der Gruppe mit überwiegend FM ergab sich in der Homogenitätsprüfung ein signifikanter Unterschied bezüglich des Geschlechts. Zwar wurden etwa gleich viele Männer und Frauen auf die Motivgruppe HE zugeteilt, aber bei der Ausprägung FM werden die Unterschiede deutlich. Frauen weisen seltener die Motivkomponente HE im Vergleich zu FM auf (HE = 32.7 %, FM = 67.3 %), während bei Männern prozentual HE häufiger vorkommt (HE = 65.2 %, FM = 34.8 %) (vgl. Abbildung 3.10).

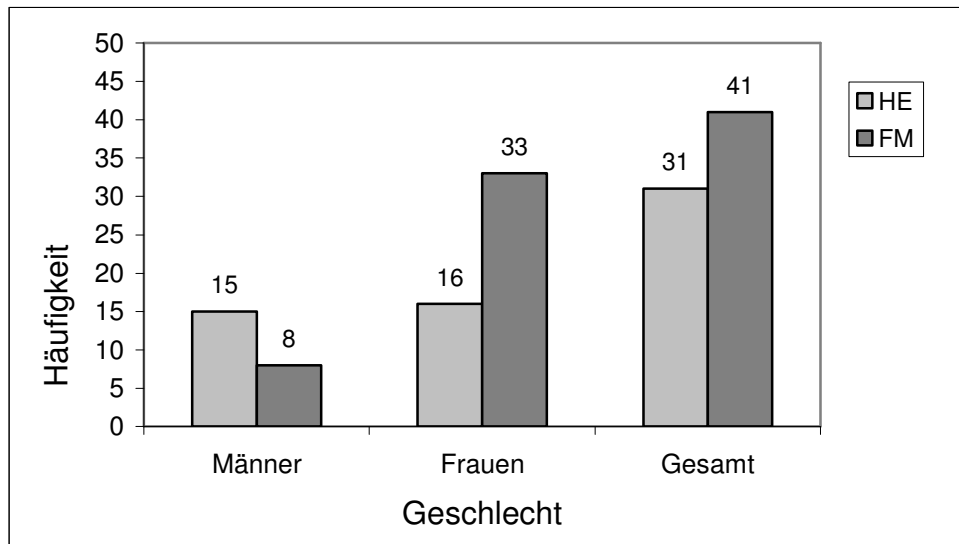


Abbildung 3.10: Verteilung von Männern und Frauen bezüglich der Zuordnung zu den Leistungsmotivgruppen

Zusammenfassend ist zu sagen, dass die Gruppenvergleiche drei signifikante Unterschiede ergeben. Experimental- und Kontrollgruppe unterscheiden sich bezüglich der Abiturnote und des CFT 20-R Intelligenzwerts. Die Leistungsmotivgruppen weisen einen signifikanten Unterschied bezüglich der Verteilung von Männern und Frauen auf.

4. Ergebnisse

Im folgenden Kapitel werden die Ergebnisse der Untersuchung gemäß den Hypothesen aus Kapitel 2 vorgestellt. Darüber hinaus sind Analysen bezüglich des Wohlbefindens, dem gefühltem Stresslevel, den Bewertungen, den Ergebnissen des RLMI-E und der Intelligenztests sowie Zusammenhangsanalysen der Testresultate in dem Kapitel enthalten.

4.1 Einfluss der Stressinstruktion

Im Bezug auf die Misserfolgsinstruktion, eine angeblich schlechtere Leistung eines vorhergehenden Intelligenztests, wird angenommen, dass die Manipulation erfolgreich verlief und bei den Vpn Stress auslöste. Zur Prüfung des Einflusses der Stressinstruktion sind in diesem Kapitel die Ergebnisse der Berechnungen zum Wohlbefinden, dem gefühlten Stresslevel und den Bewertungen der Intelligenztests aufgeführt.

4.1.1 Stressinstruktion und Wohlbefinden

Die Ergebnisse des Fragebogens zum Allgemeinen Wohlbefinden sind in Tabelle 4.1 mittels t -Tests auf Signifikanz der Differenzen abhängiger Stichproben dargestellt. Vorab ist zu bemerken, dass sich die fünf Mittelwerte, wie in Abbildung 4.1 dargestellt, des Wohlbefindens vom ersten Messzeitpunkt ($M = 64.39$, $SD = 13.56$) bis zum letzten Zeitpunkt ($M = 52.94$, $SD = 17.22$) verringern, wobei es bei Wohlbefinden 3 zu Beginn der zweiten Testung ($M = 62.40$, $SD = 15.05$) nochmals ansteigt.

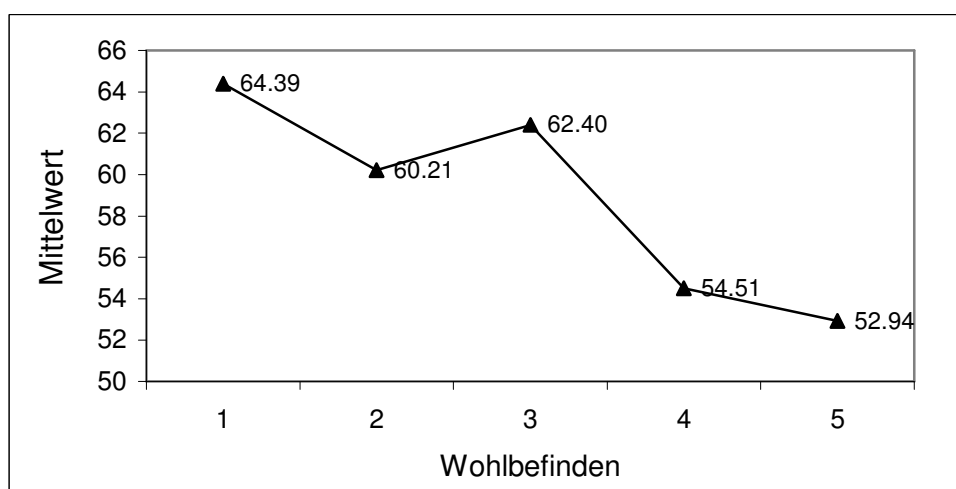


Abbildung 4.1: Vergleich der einzelnen Messzeitpunkte des Wohlbefindens 1 – 5 ($n = 72$, für Wohlbefinden 4 ist $n = 36$)

Die Messwertpaare der t -Tests vor im Vergleich zu nach den Intelligenztests (CFT 20-R: W1 – W2, LPS-neu: W3 – W5) ergeben signifikante Unterschiede, wobei die Effektstärken zeigen, dass die Abnahme des Wohlbefindens für den LPS-neu stärker als im Vergleich zum CFT 20-R ausfällt (vgl. Tabelle 4.1). Bei der Betrachtung des Vergleichs des Wohlbefindens 3 und 5 getrennt für die Kontroll- und Experimentalgruppe wird ersichtlich, dass sich ein stärkerer Effekt der Abnahme des Wohlbefindens in der Experimentalgruppe ($g = .82$) im Vergleich zur Kontrollgruppe ($g = .52$) ergibt. Die Vergleiche in der Experimentalgruppe vor im Vergleich zu nach der Stressinstruktion (W3 – W4) sowie nach der Stressinstruktion und nach dem zweiten Intelligenztest (W4 – W5) ergeben ebenfalls signifikante Werte. Hier sind mittlere Effekte zu beobachten. Kein signifikanter Unterschied besteht zwischen den zwei Fragebögen vor den Intelligenztests.

Tabelle 4.1: Mittelwertsvergleiche der Differenz einzelner Messzeitpunkte des Wohlbefindensfragebogen ($n = 72$)

	Differenz (<i>SD</i>)	t (df)	p	g
CFT: W1 – W2	4.18 (11.63)	3.05 (71)	<.01**	0.36
LPS: W3 – W5	9.46 (14.21)	5.65 (71)	<.01**	0.68
KG W3 – W5	7.54 (14.48)	3.13 (35)	<.01**	0.52
EG W3 – W5	11.38 (13.86)	4.93 (35)	<.01**	0.82
EG W3 – W4	6.06 (8.86)	4.11 (35)	<.01**	0.67
EG W4 – W5	5.32 (10.53)	3.00 (35)	<.01**	0.50
W1 – W3	1.99 (10.52)	1.61 (71)	.11	0.19

Anmerkungen: KG = Kontrollgruppe ($n = 36$), EG = Experimentalgruppe ($n = 36$), W = Wohlbefinden, Differenz = Differenz der Mittelwerte, *SD* = Standardabweichung, t = Testwert, df = Anzahl der Freiheitsgrade, p = Signifikanzwert, g = Effektstärke nach Hedges
**) sehr signifikant auf dem Niveau von .01 (einseitig)

Mittels multivariater Varianzanalyse wird der Einfluss der experimentellen Bedingung auf das Wohlbefinden untersucht. In Abbildung 4.2 sind die Vergleiche dazu grafisch verdeutlicht.

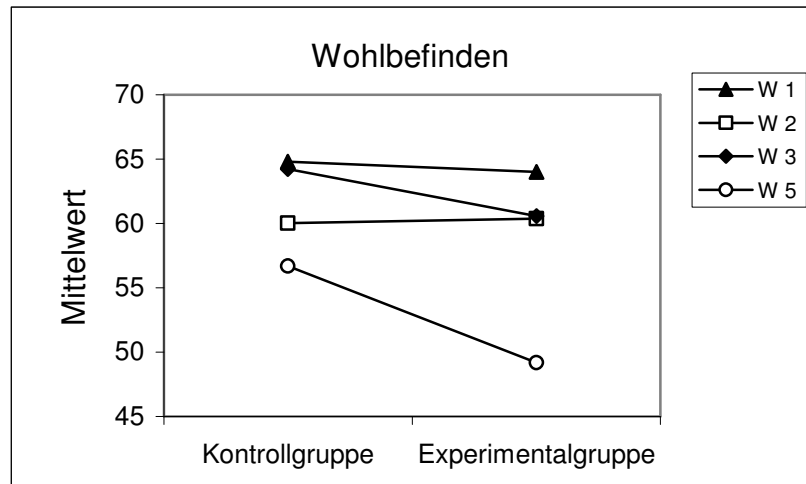


Abbildung 4.2: Vergleich des Wohlbefindens (W 1 – 5) der Kontroll- und Experimentalgruppe

Die Voraussetzung der Homogenität der Kovarianzen der abhängigen Variablen ist gegeben, da sich mittels Box-Test kein signifikanter Wert ermitteln lässt. Nach dieser Überprüfung ergibt sich für die Gruppenzugehörigkeit nur kein signifikanter Einfluss ($F(4, 67) = 1.92, p = .27, \text{partial } \eta^2 = .10$). Die einzelnen Post-hoc Vergleiche zeigen, dass sich die Gruppen bezüglich Wohlbefinden 1, 2 und 3 nicht signifikant unterscheiden und die Effekte gering ausfallen (vgl. Tabelle 4.2). In den Kontrollbedingungen vor und nach dem CFT 20-R und vor dem LPS-neu weisen beide Gruppen ein ähnliches Wohlbefinden auf.

H1: Vpn in der neutralen Bedingung fühlen sich im Vergleich zu Vpn der Stressbedingung nach der Stressinstruktion und nach dem LPS-neu Intelligenztest wohler.

Bezüglich Wohlbefinden 5 ergibt sich ein tendenziell signifikanter Wert mit mittlerem Effekt. Die Kontrollgruppe fühlt sich im Vergleich zur Gruppe mit Stressinstruktion nach dem LPS-neu tendenziell weniger unwohl (vgl. Tabelle 4.2). An dieser Stelle wird darauf hingewiesen, dass sich die Ergebnisse zu den mittleren Quadratsummen der berechneten Varianzanalysen aufgrund einer besseren Übersichtlichkeit im Anhang (vgl. Tabelle F-1 bis F-13) befinden.

Tabelle 4.2: Univariate Varianzanalysen der Kontroll- und Experimentalgruppe bezüglich des Wohlbefindens

	KG, n = 36 M (SD)	EG, n = 36 M (SD)	F	p	$\rho\eta^2$
Wohlbefinden 1	64.78 (13.95)	64.00 (13.34)	0.06	.81	.00
Wohlbefinden 2	60.04 (14.28)	60.38 (13.33)	0.01	.92	.00
Wohlbefinden 3	64.23 (14.82)	60.57 (15.27)	1.06	.31	.02
Wohlbefinden 5	56.68 (15.68)	49.19 (18.08)	3.53	.07°	.05

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 70
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.1.2 Stressinstruktion und gefühltes Stresslevel

Weiterhin soll mittels der subjektiven Einschätzung des gefühlten Stresslevels der Einfluss der Stressinstruktion auf die Experimentalgruppe untersucht werden. Abbildung 4.3 verdeutlicht die Häufigkeitsverteilung des subjektiven Stressempfindens.

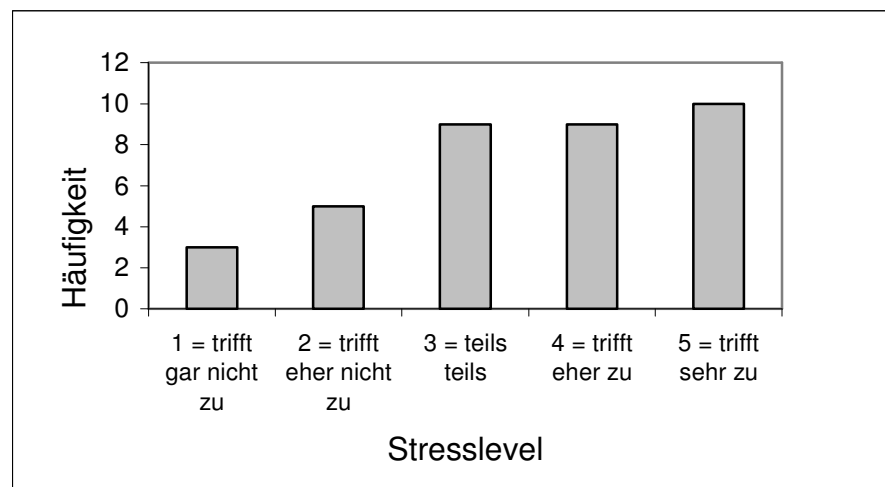


Abbildung 4.3: Verteilung des Stresslevels

Ein Viertel der Vpn gab an, teilweise von der Stressinstruktion beeinflusst zu sein, da sie sich für das mittlere Stresslevel (3 = teils, teils) entschieden (vgl. Tabelle 4.3). Über die Hälfte (52.8 %) wählten in der Skala 4 bzw. 5 und fühlten sich somit gestresst.

Tabelle 4.3: Häufigkeitsverteilung des gefühlten Stresslevels (n = 36)

	Häufigkeit	Prozent	Kumulierte Prozent
1 = trifft gar nicht zu	3	8.3	8.3
2 = trifft eher nicht zu	5	13.9	22.2
3 = teils teils	9	25.0	47.2
4 = trifft eher zu	9	25.0	72.2
5 = trifft sehr zu	10	27.8	100.0
Gesamt	36	100.0	

4.1.3 Zusammenhänge zwischen Wohlbefinden, gefühltem Stresslevel und Bewertung

In diesem Abschnitt werden die Zusammenhänge zwischen dem Wohlbefinden, dem eingeschätzten Stresslevel und den Bewertungen untersucht. Für diese Analyse werden die Produkt-Moment-Korrelationen nach Braivais Pearson verwendet.

In Tabelle 4.4 ist zu sehen, dass das Wohlbefinden der Experimentalgruppe, mit Ausnahme des Zusammenhangs zwischen Ärger und Wohlbefinden 3, signifikant negativ mit dem empfundenen Stress korreliert. Die Zusammenhänge liegen in einem mittleren bis hohen Bereich, wobei für Wohlbefinden 4 und 5 stets höhere Korrelationen vorliegen.

Tabelle 4.4: Korrelationen zwischen dem Stresslevel und Wohlbefinden in der Experimentalgruppe (n = 36)

	Frustration	Ärger	Druck	Stress Gesamt
Wohlbefinden 3	-.45**	-.20	-.51**	-.43**
Wohlbefinden 4	-.63**	-.38*	-.70**	-.64**
Wohlbefinden 5	-.64**	-.40*	-.66**	-.63**

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

In Tabelle 4.5 sind die Korrelationen des Wohlbefindens mit den Bewertungen des CFT 20-R und des LPS-neu angegeben. Fast alle Zusammenhänge ergeben positive und hoch signifikante mittlere Werte. Die zwei tendenziell signifikanten Werte zwischen dem Wohlbefinden 1 mit der CFT 20-R Bewertung und dem Wohlbefinden 3 mit der Bewertung des LPS-neu in der Experimentalgruppe fallen dabei aus der Reihe. Das Wohlbefinden nach den Intelligenztestungen und nach der Stressinstruktion korreliert jedoch signifikant mit den Bewertungen.

Tabelle 4.5: Korrelationen zwischen den Bewertungen und Wohlbefinden

	Eval. CFT (n = 72)	Eval. LPS-neu (n = 72)	KG (n = 36) Eval. LPS-neu	EG (n = 36) Eval. LPS-neu
Wohlbefinden 1	.21°	-	-	-
Wohlbefinden 2	.31**	-	-	-
Wohlbefinden 3	-	.44**	.68**	.31°
Wohlbefinden 4	-	-	-	.44**
Wohlbefinden 5	-	.55**	.55**	.52**

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson
n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe
**) signifikant auf dem Niveau von .01 (zweiseitig)
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Aus Tabelle 4.6 wird deutlich, dass ein negativer Zusammenhang zwischen der Bewertung des LPS-neu und dem empfundenen Stresslevel besteht. Die Korrelationen sind niedrig und für Frustration und Druck nur tendenziell signifikant, jedoch für Ärger und die interessierende Gesamtskala ist der Zusammenhang signifikant.

Tabelle 4.6: Korrelation zwischen dem Stresslevel und der Bewertung des LPS-neu in der Experimentalgruppe (n = 36)

	Frustration ¹⁾	Ärger	Druck	Stress Gesamt
Evaluation LPS-neu	-.32°	-.34*	-.28°	-.35*

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson
*) signifikant auf dem Niveau von .05 (zweiseitig)
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.2 Leistungsmotivation

4.2.1 Verteilung der Leistungsmotivation in der Stichprobe

Einen differenzierten Überblick der einzelnen Leistungsmotivausprägungen, welche in die Analyse der Studie mit einfließen, liefert Tabelle 4.7. Da für diese Untersuchung die Konstrukte „Hoffnung auf Erfolg“ (HE) und „Furcht vor Misserfolg“ (FM) sowie die sich dadurch ergebende Gesamtmotivation (GM) und Nettohoffnung (NH), welche bereits Heckhausen (1980) gebrauchte, interessant sind, wird auf die Darstellung und die Auswertung der Komponenten „Furcht vor Erfolg“ (FE) und „Hoffnung auf Misserfolg“ (HM) und der einzelnen Bereiche verzichtet.

Für die gesamte Stichprobe ergibt sich ein durchschnittlicher Wert für HE von 101.85 ($SD = 7.98$) und für FM von 104.00 ($SD = 11.45$). Die Berechnungen zur GM, summiert aus den Mittelwerten von HE und FM, liefern einen Mittelwert von 205.85 mit

einer Standardabweichung von 11.54. Die durchschnittliche NH, welche sich aus der Differenz von HE und FM ergibt, beträgt -2.15 ($SD = 16.02$). Für die Verteilungen von HE und der NH ergeben sich relativ symmetrische bzw. leicht linkssteile Verteilungen. Hingegen sind die Verteilungen für FM und der GM geringfügig rechtssteil. Für HE und der GM zeigen sich leicht breitgipflige und für die restlichen zwei Skalen eher schmalgipflige Verteilungen. Die grafischen Verteilungen zu Schiefe und Exzess der RLMI-E Skalen sind im Anhang angefügt (vgl. Abbildung E-6).

Bei der getrennten Betrachtung der Gruppen und im Vergleich zur Verteilung über alle Vpn ergeben sich für die einzelnen Skalen ähnliche Mittelwerte und Standardabweichungen. Beispielsweise ist der Mittelwert für HE in der Kontrollgruppe 101.11 ($SD = 6.94$) und in der Experimentalgruppe 102.58 ($SD = 8.94$). Die Verteilungen der Skalen beider Gruppen sind leicht linkssteil und breitgipflig. Geringfügige Unterschiede zeigt die Verteilung der GM, welche in der Kontrollgruppe eher rechtssteil und breitgipflig und für die Experimentalgruppe linkssteil und schmalgipflig ausfällt. Im Wesentlichen sind die Verteilungen relativ symmetrisch und die Formen der einzelnen Skalen weichen gering von der Normalverteilung ab.

Tabelle 4.7: Verteilung der Leistungsmotivausprägung in den unterschiedlichen Gruppen und für die Gesamtstichprobe

	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Exzess
Kontrollgruppe (n = 36)						
HE	101.11	6.94	88	113	0.11	-0.94
FM	103.31	12.65	71	129	-0.20	0.20
GM	204.42	12.07	180	226	-0.27	-0.82
NH	-2.19	16.45	-35	41	0.22	0.41
Experimentalgruppe (n = 36)						
HE	102.58	8.94	88	124	0.38	-0.44
FM	104.69	10.25	82	129	0.13	-0.06
GM	207.28	10.97	187	234	0.31	0.32
NH	-2.11	15.80	-33	31	0.30	-0.49
Gesamt (n = 72)						
HE	101.85	7.98	88	124	0.36	-0.37
FM	104.00	11.45	71	129	-0.13	0.21
GM	205.85	11.54	180	234	-0.05	-0.24
NH	-2.15	16.02	-35	41	0.25	-0.09

Anmerkungen: n = Anzahl der Vpn, *M* = Mittelwert, *SD* = Standardabweichung, Min = Minimum, Max = Maximum

Nach der Berechnung der Leistungsmotivkennwerte des RLMI-E fand eine Halbierung der Teilnehmer in Personen mit überwiegend HE oder FM statt. Die Vpn mit einer NH über bzw. gleich Null wurden der Gruppe mit HE und mit einer NH unter Null der Gruppe FM zugeordnet. Insgesamt nahmen 41 Personen mit FM (56.9 %) und 31 (43.1 %) mit HE an der Studie teil (vgl. Tabelle 4.8). In den einzelnen Gruppen zeigte sich ein ähnliches Bild der Prozentanteile, d. h. dass in beiden Gruppen Personen mit FM häufiger im Vergleich zu Personen mit HE vertreten waren.

Tabelle 4.8: Verteilung nach der Zuordnung zu einer Motivationsgruppe

Gruppe		LM – Gruppe		Gesamt
		HE	FM	
Kontrollgruppe	Häufigkeit	14	22	36
	Prozent	38.9	61.1	100.0
Experimentalgruppe	Häufigkeit	17	19	36
	Prozent	47.2	52.8	100.0
Gesamt	Häufigkeit	31	41	72
	Prozent	43.1	56.9	100.0

Anmerkungen: LM = Leistungsmotiv

Interessant für die Analyse ist ferner die Betrachtung der Stärke des Leistungsmotivs. Dafür wurden die Teilnehmer an der Mediangrenze der GM von 206.50 in zwei Gruppen aufgeteilt. Bei einer GM unter dem Median wurden die Vpn der Gruppe mit niedriger Leistungsmotivation und über dem Median der Gruppe mit hoher Leistungsmotivation zugeordnet. Die untersuchte Stichprobe enthielt 36 niedrig und 36 hoch leistungsmotivierte Vpn. Die Häufigkeiten sind in der Experimental- und Kontrollgruppe ähnlich verteilt, obwohl sie sich genau entgegengesetzt verhalten (vgl. Tabelle 4.9).

Tabelle 4.9: Verteilung nach der Höhe der Leistungsmotivation

Gruppe		LM – Gruppe		Gesamt
		Niedrig LM	Hoch LM	
Kontrollgruppe	Häufigkeit	17	19	36
	Prozent	47.2	52.8	100.0
Experimentalgruppe	Häufigkeit	19	17	36
	Prozent	52.8	47.2	100.0
Gesamt	Häufigkeit	36	36	72
	Prozent	50.0	50.0	100.0

Anmerkungen: LM = Leistungsmotiv

4.2.2 Leistungsmotivation und Wohlbefinden

H2: Vpn mit überwiegend HE fühlen sich im Vergleich zu Vpn mit überwiegend FM während der gesamten Testung wohler, wobei größere Unterschiede nach der Stressinstruktion erwartet werden.

Hinsichtlich Hypothese 2 wird eine multivariate Varianzanalyse durchgeführt, wobei die Gleichheit der Kovarianzenmatrizen der abhängigen Variablen als gegeben gelten kann. Die einzelnen Mittelwerte in Abbildung 4.4 verdeutlichen, dass die Gruppe mit HE stets höhere Werte im Wohlbefinden angab bzw. beim Fragebogen 3 etwa ähnliche Wohlbefindensangaben im Vergleich zur Gruppe mit FM machte.

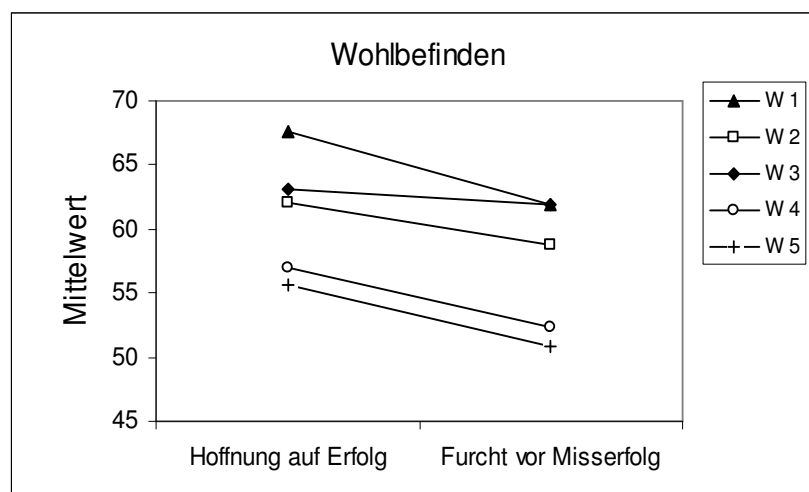


Abbildung 4.4: Vergleich des Wohlbefindens (W 1 – 5) der Leistungsmotivgruppen

Für den Vergleich zwischen den Leistungsmotivgruppen bezüglich des Wohlbefindens wird der Testwert $F(5, 30) = 0.41$ im Modell nicht signifikant ($p = .84$, $\text{partial } \eta^2 = .06$). Für keinen der fünf Post-hoc Tests der Wohlbefindensmessungen lassen sich signifikante Unterschiede bezüglich der Leistungsmotive feststellen (vgl. Tabelle 4.10).

Tabelle 4.10: Univariate Varianzanalysen zur Prüfung der Leistungsmotivgruppe bezüglich des Wohlbefindens

	HE, n = 31 M (SD)	FM, n = 41 M (SD)	F	p	$p\eta^2$
Wohlbefinden 1	67.62 (14.65)	61.95 (12.29)	0.64	.43	.02
Wohlbefinden 2	62.06 (16.92)	58.80 (10.71)	0.27	.61	.01
Wohlbefinden 3	63.08 (16.72)	61.88 (13.85)	0.06	.81	.00
Wohlbefinden 4	56.98 (19.41)	52.30 (14.21)	0.68	.41	.02
Wohlbefinden 5	55.70 (19.54)	50.85 (15.16)	1.20	.28	.03

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $p\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 34

Werden die beiden Leistungsmotivgruppen getrennt auf die Veränderung der einzelnen Zeitpunkte des Wohlbefindens geprüft, sind die meisten Unterschiede signifikant (vgl. Tabelle 4.11). Die Differenzen der HE-Gruppe sind alle signifikant, mit jeweils mittleren Effektstärken zwischen .43 und .61. Für die Gruppe mit überwiegend FM liegen fast alle Vergleiche im annehmbaren Signifikanzbereich mit hohen Effektstärken. In der Kontrollbedingung (CFT 20-R: W1 – W2) ergeben sich nur tendenzielle Signifikanzen mit geringeren Effekten. Im Vergleich zu HE sinken die Mittelwerte vor im Vergleich zu nach der Stressinstruktion in der Experimentalgruppe der misserfolgsängstlichen Personen stärker ab (LPS-neu: W3 – W4, W3 – W5); hier sind insgesamt die stärkeren Effekte zwischen .75 und .86 zu verzeichnen.

Zusätzlich wurde eine Analyse bezüglich der Änderung von Wohlbefinden 3 und 5 getrennt für die Kontroll- und Experimentalgruppe vorgenommen. Es zeigt sich, dass Personen mit HE keine signifikante Veränderung in der neutralen Bedingung, wohl aber in der Stressbedingung aufweisen. Für misserfolgsfürchtige Vpn wird der Unterschied in beiden Bedingungen mit starkem Effekt zwischen .86 und .94 signifikant.

Tabelle 4.11: Mittelwertsvergleiche der Differenz einzelner Messzeitpunkte des Wohlbefindensfragebogen der Leistungsmotivgruppen (n = 72)

	HE (n = 31)				FM (n = 41)			
	<i>d</i> (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i>	<i>g</i>	<i>d</i> (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i>	<i>g</i>
CFT: W1 – W2	5.56 (12.53)	2.47 (30)	.02*	0.44	3.15 (10.95)	1.84 (40)	.07°	0.29
LPS: W3 – W5	7.38 (15.89)	2.59 (30)	.02*	0.46	11.04 (12.77)	5.54 (40)	<.01**	0.86
KG W3 – W5	5.96 (20.66)	1.08 (13)	.30	0.29	8.57 (9.10)	4.41 (21)	<.01**	0.94
EG W3 – W5	8.56 (11.11)	3.18 (16)	<.01**	0.77	13.91 (15.79)	3.84 (18)	<.01**	0.86
EG W3 – W4	4.26 (6.93)	2.53 (16)	.02*	0.61	7.68 (10.21)	3.28 (18)	<.01**	0.75
EG W4 – W5	4.30 (7.73)	2.29 (16)	.04*	0.56	6.23 (12.87)	2.11 (18)	.05*	0.48
W1 – W3	4.54 (10.51)	2.41 (30)	.02*	0.43	0.07 (10.24)	0.04 (40)	.97	0.01

Anmerkungen: n = Anzahl der Vpn, W = Wohlbefinden, KG = Kontrollgruppe (HE n = 14, FM n = 22), EG = Experimentalgruppe (HE n = 17, FM n = 19), *d* = Differenz der Mittelwerte, *SD* = Standardabweichung, *t* = Testwert, *df* = Anzahl der Freiheitsgrade, *p* = Signifikanzwert, *g* = Effektstärke nach Hedges

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Weiterhin werden die fünf Wohlbefindensangaben auf ihren Zusammenhang mit den RLMI-Skalen HE und FM sowie für die berechnete GM und NH untersucht. Es ergeben sich signifikant positive Zusammenhänge zwischen HE mit dem Wohlbefinden vor beiden Intelligenztests und nach dem Intelligenztest 2 (vgl. Tabelle 4.12). Negative Korrelationen wurden für FM in der Kontrollbedingung und nach dem zweiten Intelligenztest ermittelt. Bezüglich der GM sind keine signifikanten Zusammenhänge ermittelbar, während die NH signifikant positiv mit den Wohlbefindensmessungen 1, 2, 3 und 5 sowie mit 4 tendenziell, korreliert. Auffällig ist, dass in der Stressbedingung nach der Instruktion die Werte nicht signifikant werden.

Tabelle 4.12: Korrelationen zwischen Leistungsmotivausprägungen und Wohlbefindens 1 – 5 (n = 72)

	Wohlbef. 1	Wohlbef. 2	Wohlbef. 3	Wohlbef. 4 ¹⁾	Wohlbef. 5
HE	.32**	.18	.26*	.15	.24*
FM	-.35**	-.30*	-.16	-.28 ^o	-.28*
GM	-.13	-.17	.02	-.21	-.11
NH	.41**	.30*	.24*	.28 ^o	.32**

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson

¹⁾ für Wohlbefinden 4 ist n = 36

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

^{o)} tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.2.3 Leistungsmotivation und gefühltes Stresslevel

H3: Vpn mit überwiegend FM fühlen sich im Vergleich zu Vpn mit überwiegend HE nach der Stressinstruktion gestresster.

Für die Prüfung von Hypothese 3 wird bezüglich des gesamten Stresslevels eine univariate Varianzanalyse berechnet, welche ein sehr signifikantes Ergebnis ergibt.

Eine multivariate Varianzanalyse wurde für die drei Stressitems berechnet. Der Unterschied zwischen FM und HE ist tendenziell signifikant ($F(3, 32) = 2.54, p = .07$) mit einem starkem Effekt ($\text{partial } \eta^2 = .19$). Die Post-hoc Vergleiche sind ebenfalls signifikant bzw. Druck nur tendenziell. Die Effektstärken liegen in einem hohen Bereich. Die Mittelwerte in Tabelle 4.13 und der Abbildung 4.5 verdeutlichen, dass sich stets die FM-Gruppe stärker von der Stressinstruktion beeinflusst fühlte.

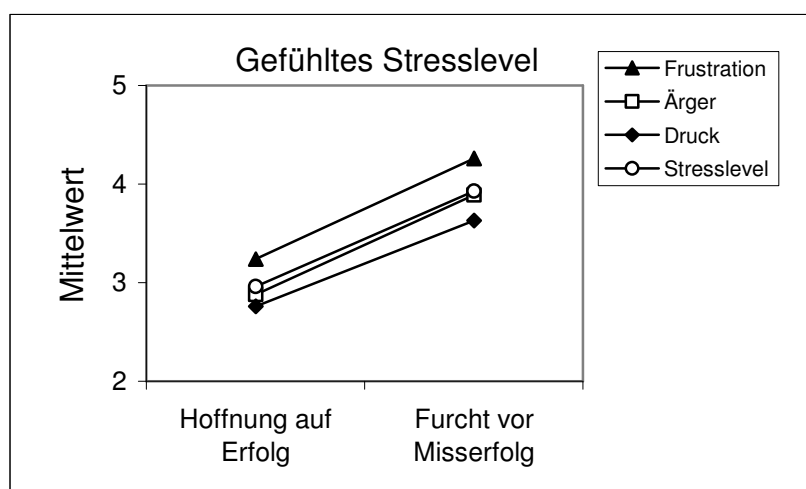


Abbildung 4.5: Vergleich des Stresslevels der Leistungsmotivgruppen

Tabelle 4.13: Univariate Varianzanalysen der Leistungsmotivgruppen bezüglich des gefühlten Stresslevels

	HE, n = 17 M (SD)	FM, n = 19 M (SD)	F	p	ρ^2
Frustration	3.24 (1.20)	4.26 (1.05)	7.54	.01**	.18
Ärger	2.88 (1.36)	3.89 (1.15)	5.84	.02*	.15
Druck	2.76 (1.52)	3.63 (1.26)	3.50	.07°	.09
Stress Gesamt	2.96 (1.18)	3.93 (1.03)	6.95	.01**	.17

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, ρ^2 = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 34

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Durch die Zusammenhangsanalyse in Tabelle 4.14 wird deutlich, dass FM positiv mit dem empfundenen Stress korreliert, während HE bei Frustration und dem gesamten Stresslevel tendenziell negative Korrelationen aufweist. Signifikant positive Zusammenhänge lassen sich für die GM feststellen, wobei Druck nur tendenzielle Signifikanz erzielt. Für den Zusammenhang zwischen den Stresslevelskalen und der NH lassen sich signifikante negative Korrelationen ermitteln. Die signifikanten Zusammenhänge sind alle von mittlerer Stärke zwischen .34 und .61.

Tabelle 4.14: Korrelationen zwischen Leistungsmotivausprägungen und des Stresslevels (n = 36)

	Frustration	Ärger	Druck	Stress Gesamt
HE	-.28°	-.24	-.26	-.29°
FM	.61**	.45**	.44**	.56**
GM	.48**	.34*	.31°	.42*
NH	-.59**	-.45**	-.45**	-.55**

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.2.4 Leistungsmotivation und Bewertung

Zur Unterschiedsprüfung der experimentellen Gruppen und der Leistungsmotivgruppen bezüglich der Bewertungen der Intelligenztests sind in den Tabellen 4.15 und 4.16 die Ergebnisse der zweifaktoriellen Kovarianzanalysen zu finden. Der IQ-Wert nach dem CFT 20-R wurde als Kovarianzfaktor hinzugenommen.

F1: Bewerten Vpn mit überwiegend HE im Vergleich zu Vpn mit überwiegend FM die Intelligenztests positiver?

Für die Evaluation des CFT 20-R ist in Abbildung 4.6 bereits zu erkennen, dass die Gruppen sehr ähnliche Bewertungen abgeben. Der Kovarianzfaktor zeigt keinen signifikanten Einfluss auf das Modell ($F(1, 61) = 0.30$, $p = .59$, $\text{partial}\eta^2 = .00$). Aus Tabelle 4.15 wird deutlich, dass weder die Interaktion, noch die beiden Haupteffekte für die CFT 20-R Evaluationen unter dem gewünschten Signifikanzniveau liegen.

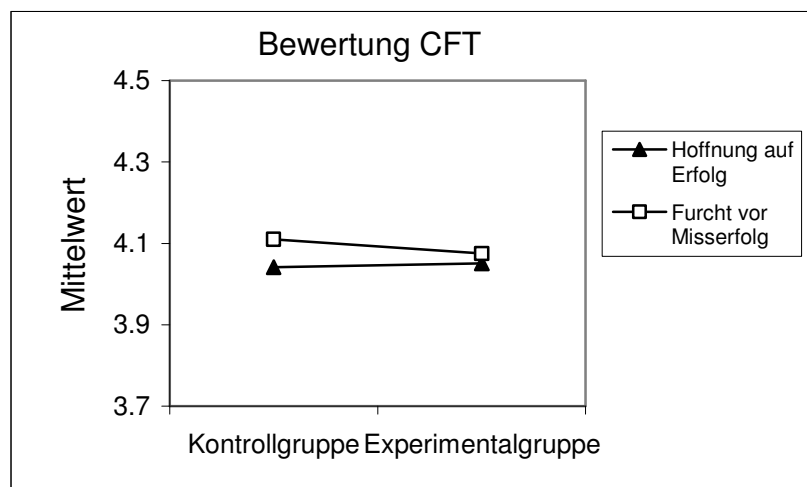


Abbildung 4.6: Vergleich der Bewertung des CFT 20-R von Versuchsbedingungen und der Leistungsmotivgruppen

Tabelle 4.15: Zweifaktorielle Kovarianzanalyse der Kontroll- und Experimentalgruppe und der Leistungsmotivgruppen bezüglich der Bewertung des CFT 20-R

		n	M	SD	F	p	$\text{partial}\eta^2$
Gruppe	KG	36	4.08	0.49	0.10	.75	.00
	EG	36	4.06	0.46			
LM-Gruppe	HE	41	4.08	0.49	0.16	.69	.00
	FM	31	4.09	0.40			
Interaktion					0.05	.82	.00

Anmerkungen: n = Anzahl der Vpn, LM-Gruppe = Leistungsmotivgruppe, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, p = Signifikanzwert, $\text{partial}\eta^2$ = Effektstärke partielles Eta-Quadrat Kovariate = CFT 20-R IQ-Wert

F2: Bewerten Vpn in der neutralen Bedingung im Vergleich zu Vpn der Stressbedingung den LPS-neu positiver?

Die Ergebnisse zur Bewertung des LPS-neu zeigen ein ähnliches Bild (vgl. Abbildung 4.7, Tabelle 4.16). Die verwendete Kovariate CFT IQ-Wert ($F(1, 67) = 0.44$, $p = .51$, $\text{partial } \eta^2 = .01$) und die Interaktion ergeben keinen signifikanten Werte. Der Effekt der Leistungsmotivgruppen hat keinen signifikanten Einfluss. Der Unterschied zwischen den experimentellen Gruppen ist ebenfalls nicht signifikant, obwohl sich hier ein geringer Effekt ermitteln lässt.

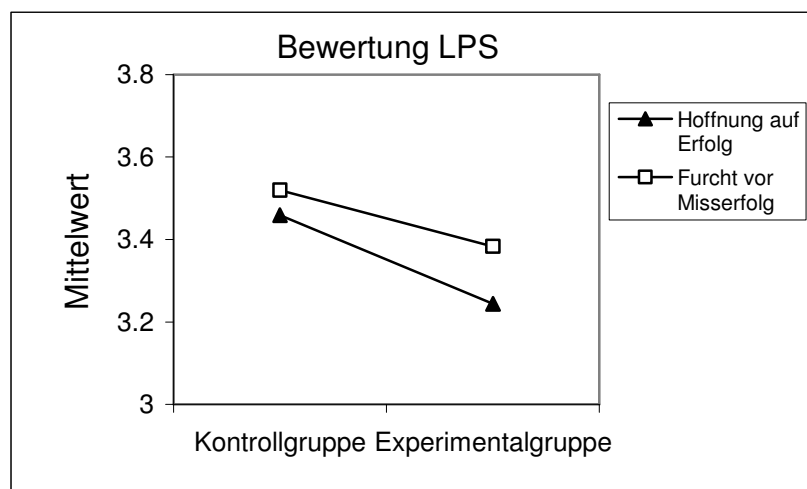


Abbildung 4.7: Vergleich der LPS-neu Bewertungen

Tabelle 4.16: Zweifaktorielle Kovarianzanalyse der Kontroll- und Experimentalgruppe und der Leistungsmotivgruppen bezüglich der Bewertung des LPS-neu

		n	M	SD	F (df1, df2)	p	$\text{partial } \eta^2$
Gruppe	KG	36	3.50	0.54	1.86	.18	.03
	EG	36	3.32	0.66			
LM-Gruppe	HE	41	3.34	0.77	0.45	.50	.01
	FM	31	3.46	0.45			
Interaktion					0.05	.82	.00

Anmerkungen: n = Anzahl der Vpn, LM-Gruppe = Leistungsmotivgruppe, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, p = Signifikanzwert, $\text{partial } \eta^2$ = Effektstärke partielles Eta-Quadrat, Kovariate = CFT 20-R IQ-Wert
 **) sehr signifikant auf dem Niveau von .01 (zweiseitig)

Bei der Überprüfung des Zusammenhangs der Bewertungen der zwei Intelligenztests mit den Motivausprägungen des RLMI ergeben sich nur geringe und nicht signifikante Werte (vgl. Tabelle 4.17). Nur für die Korrelation HE und der Evaluation des LPS-neu in der Kontrollgruppe ist das Ergebnis tendenziell signifikant.

Tabelle 4.17: Korrelation zwischen den Leistungsmotivausprägungen und den Bewertungen von CFT 20-R und LPS-neu

	Eval. CFT 20-R (n = 72)	Eval. LPS-neu (n = 72)	Eval. LPS-neu KG (n = 36)	Eval. LPS-neu EG (n = 36)
HE	.10	.05	.31 [°]	-.25
FM	-.01	-.03	-.03	-.05
GM	.05	.00	.23	-.20
NH	.06	.05	.19	-.06

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson

n = Anzahl der Vpn, Eval. = Evaluation, KG = Kontrollgruppe, EG = Experimentalgruppe

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.3 Die Intelligenztestleistungen

4.3.1 Die Ergebnisse der Intelligenztests in der Stichprobe

Nachdem in den Kapiteln 3.3.2.2 und 3.3.3.2 die Ergebnisse der Gesamtleistungen und der einzelnen Subtests aufgeführt sind, wird in diesem Abschnitt auf weitere für die folgenden Analysen relevante Mittelwerte und Standardabweichungen eingegangen (vgl. Tabelle 4.18).

Zwischen beiden experimentellen Gruppen erbringt die Gruppe mit Stressinstruktion im CFT 20-R und im LPS-neu insgesamt die besseren Leistungen. In dieser Gruppe werden Aufgaben seltener falsch beantwortet, ausgelassen und nicht bearbeitet und mehr Aufgaben berichtet.

Betreffend der Leistungsmotivgruppen erzielen beide Gruppen im CFT 20-R durchschnittlich die gleichen Ergebnisse. Im LPS-neu sind es Personen mit HE, die im Vergleich zur FM-Gruppe häufiger korrekt antworten, mehr Fehler machen, häufiger Aufgaben auslassen und stärker berichtigen, wobei die FM-Gruppe seltener Aufgaben nicht bearbeitet.

Tabelle 4.18: Ergebnisse des CFT 20-R und LPS-neu nach den Experimental- und Leistungsmotivgruppen

	<i>M</i> Gesamt (<i>SD</i>), n = 72	<i>M</i> KG (<i>SD</i>) n = 36	<i>M</i> EG (<i>SD</i>) n = 36	<i>M</i> HE (<i>SD</i>) n = 31	<i>M</i> FM (<i>SD</i>) n = 41
CFT IQ-Wert	115.69 (13.72)	110.22 (12.95)	121.17 (12.34)	115.97 (14.80)	115.49 (13.03)
LPS-neu:					
Korrekt	291.46 (38.83)	284.72 (41.44)	298.19 (35.34)	295.90 (43.37)	288.10 (35.21)
Fehler	51.17 (16.97)	51.75 (15.95)	50.58 (18.14)	52.32 (19.54)	50.29 (14.93)
Ausgelassen	16.17 (10.92)	16.92 (10.00)	15.42 (11.87)	17.52 (12.03)	15.15 (10.04)
Nicht Bearbeitet	161.01 (34.96)	166.22 (36.54)	155.81 (32.98)	154.26 (37.11)	166.12 (32.77)
Berichtigt	1.88 (1.99)	1.61 (1.52)	2.14 (2.36)	2.10 (2.23)	1.71 (1.79)

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung

Die Ergebnisse der Tests, welche in die folgenden Berechnungen einfließen, werden auf ihre Abweichung zu einer Normalverteilung mittels Schiefe und Exzess geprüft (vgl. Tabelle 4.19). Die jeweiligen grafischen Darstellungen sind im Anhang veranschaulicht (vgl. Tabelle E-2 und E-7). Für die Intelligenzwerte nach dem CFT 20-R ergeben sich annähernd symmetrische Verteilungen. Auch der Exzess der CFT 20-R Intelligenzwerte weist auf Gipfformen ähnlich zu einer Normalverteilung hin.

Die LPS Ergebnisse der korrekten Antworten und der nicht bearbeiteten Aufgaben weichen nach den Werten für Schiefe und Exzess nicht stark von einer Normalverteilung ab. Moderate bis stärkere Abweichungen in der Verteilung sind für die ausgelassenen und berichtigten Antworten zu verzeichnen. Hier ergeben sich mit Ausnahme der berichtigten Antworten der Kontrollgruppe linksschiefe und schmalgipflige Verteilungen. Fehler sind in der Regel nicht normalverteilt, die Abweichungen von einer symmetrischen Verteilung sind hier allerdings nicht zu hoch, so dass auch die Ergebnisse der Fehler mit in die Analysen einfließen können.

Tabelle 4.19: Schiefe und Exzess des CFT 20-R und LPS-neu nach den Experimental- und Leistungsmotivgruppen

	Gesamt, n = 72		KG, n = 36		EG, n = 36		HE, n = 31		FM, n = 41	
	Schiefe	Exzess	Schiefe	Exzess	Schiefe	Exzess	Schiefe	Exzess	Schiefe	Exzess
IQ-Wert	-0.28	0.07	-0.44	-0.14	-0.15	-0.03	-0.40	0.24	-0.17	0.00
Korrekt	-0.03	0.30	-0.19	0.18	0.45	-0.02	0.09	-0.54	-0.36	1.40
Fehler	1.59	4.17	1.57	4.50	1.67	4.56	1.41	3.07	1.78	6.11
Ausgel.	0.60	0.25	0.68	0.78	0.63	0.07	0.88	0.72	0.15	-1.01
N. Bearb.	0.23	-0.24	0.16	-0.41	0.24	0.09	0.30	-0.50	0.33	0.14
Berichtigt	1.57	2.47	0.76	-0.40	1.54	1.68	1.32	1.60	1.83	3.95

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, Ausgel. = Ausgelassen, N. Bearb. = Nicht Bearbeitet

In Tabelle 4.20 sind die Mittelwerte und Standardabweichungen der Leistungsmotivgruppen einzeln für die Kontroll- und Experimentalgruppe dargestellt. In der Kontrollgruppe konnten für die HE-Gruppe höhere Werte im CFT 20-R und den korrekten Antworten beim LPS-neu ermittelt werden. Geringfügig machen diese Personen im Vergleich zu der FM-Gruppe mehr Fehler, lassen mehr Aufgaben aus, berichtigen häufiger, aber sie schaffen mehr Aufgaben. Die FM-Gruppe erzielt nach der Stressinstruktion im Vergleich zur HE-Gruppe bessere Leistungen im CFT 20-R. Sie haben eine geringere Fehleranzahl im LPS-neu, lassen weniger Aufgaben aus und berichtigen seltener. Die Gruppe mit überwiegend HE antwortet hingegen im LPS-neu häufiger korrekt und lässt seltener Aufgaben aus.

Die varianzanalytischen Berechnungen bezüglich der Ergebnisse der Intelligenztests sind in Kapitel 4.4 dargestellt.

Tabelle 4.20: Ergebnisse des CFT 20-R und LPS-neu der Leistungsmotivgruppen getrennt für Kontroll- und Experimentalgruppe

	<i>M</i> KG – HE (<i>SD</i>), n = 14	<i>M</i> KG – FM (<i>SD</i>), n = 22	<i>M</i> EG – HE (<i>SD</i>), n = 17	<i>M</i> EG – FM (<i>SD</i>), n = 19
CFT IQ-Wert	111.07 (14.77)	109.68 (11.99)	120.00 (13.96)	122.21 (10.98)
Korrekt	288.57 (42.46)	282.27 (41.59)	301.94 (44.45)	294.84 (25.43)
Fehler	51.86 (13.04)	51.68 (17.85)	52.71 (24.02)	48.68 (10.90)
Ausgelassen	18.21 (10.96)	16.09 (9.51)	16.94 (13.15)	14.05 (10.78)
Nicht Bearbeitet	161.36 (33.20)	169.32 (38.95)	148.41 (40.08)	162.42 (24.28)
Berichtigt	1.79 (1.76)	1.50 (1.37)	2.35 (2.57)	1.95 (2.20)

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung

4.3.2 Konstruktvalidierung

Der CFT 20-R kam in der Kontrollbedingung zum Einsatz, um die Intelligenz der Gruppen vor dem Treatment zu vergleichen. Weiterhin sollte gesichert werden, dass Personen, die im LPS-neu hohe Leistungen erbringen, auch vorher im CFT 20-R gute IQ-Werte erreichten. Diese theoretische Annahme der Untersuchung, dass beide Intelligenztests ähnliche Ergebnisse über die Intelligenzstruktur der Teilnehmer erbringen, d. h. dass ein positiver Zusammenhang des CFT 20-R und der korrekten Antworten des LPS-neu besteht, wird an dieser Stelle geprüft. Durch die Bestimmung der Korrelationen wird eine empirische Konstruktvalidierung vorgenommen.

In der Tabelle 4.21 sind die Korrelationen für die einzelnen Untertests sowie für die Gesamtergebnisse der Kontrollgruppe aufgeführt. Für diese Studie ist besonders der Zusammenhang des gesamten CFT 20-R Intelligenzwerts mit der Gesamtzahl der richtigen Antworten des LPS-neu interessant. Für die zwei Korrelationen ergeben sich zwei signifikante, positive Zusammenhänge von etwa mittlerer Stärke (.54). Eine Person, welche im CFT 20-R gute Leistungen erbringt, schneidet auch im LPS-neu gut ab. Der Einsatz des CFT 20-R als Kontrollvariable ist durch die Bestätigung der Annahme, dass CFT 20-R und LPS-neu miteinander hoch korrelieren, begründet.

Tabelle 4.21: Korrelation der Skalen des LPS-neu und des CFT 20-R für die Kontrollgruppe (n = 36)

	CFT 1	CFT 2	CFT 3	CFT 4	CFT R	CFT IQ
LPS-neu 1	.43**	.08	-.05	.23	.24	.24
LPS-neu 2	.44**	.37*	.35*	.07	.41*	.39*
LPS-neu 3	.36*	.22	.28°	.25	.37*	.35*
LPS-neu 4	.57**	.28°	.53**	.33°	.57**	.53**
LPS-neu 5	.42*	.19	.29°	.18	.36*	.37*
LPS-neu 6	.35*	.17	.16	.14	.27°	.24
LPS-neu 7	.50**	.37*	.53**	.34*	.58**	.57**
LPS-neu 8	.48**	.10	.35*	.18	.37*	.41*
LPS-neu 9	.28°	.19	.32°	.40*	.39*	.40*
LPS-neu 10	.18	.32°	.15	.10	.25	.24
LPS-neu 11	.40*	.25	.35*	.40*	.46**	.44**
Korrekt	.59**	.33*	.39*	.34*	.55**	.54**
Fehler	-.37*	-.09	-.19	-.26	-.30°	-.30°
Ausgelassen	-.09	-.15	.19	.20	.04	.02
Nicht Bearbeitet	-.49**	-.28°	-.42**	-.32°	-.50**	-.48**
Berichtigt	.09	.12	.28°	-.14	.12	.09

Anmerkungen: CFT R = CFT 20-R Gesamt Rohwert, CFT IQ = CFT 20-R IQ-Wert

***) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.3.3 Zusammenhänge der Intelligenztests und der Stressinstruktion

Die ermittelten Korrelationen zwischen dem Intelligenzwert des CFT 20-R zum Wohlbefinden, dem gefühltem Stresslevel und der Evaluationen sind in Tabelle 4.22 gesamt und getrennt für die experimentellen Gruppen aufgeführt. Es ergaben sich sehr geringe, nicht signifikante Zusammenhänge. Das gefühlte Stresslevel im Zusammenhang mit dem Intelligenzwert ergab in der Experimentalgruppe einen tendenziell signifikanten Wert von $-.27$. Zusätzlich fand eine Überprüfung statt, ob Personen mit höherem IQ den LPS-neu positiver bewerten. Auch für diese Korrelation konnten sich keine signifikanten Werte ermitteln lassen.

Tabelle 4.22: Korrelationen zwischen der CFT 20-R IQ-Werte und Wohlbefinden, gefühltem Stresslevel und Bewertung

	CFT IQ-Wert (n = 72)	KG – CFT IQ-Wert (n = 36)	EG – CFT IQ-Wert (n = 36)
Wohlbefinden 1	.02	.12	-.06
Wohlbefinden 2	.12	.04	.23
Wohlbefinden 3	.02	.07	.09
Wohlbefinden 4	-	-	.14
Wohlbefinden 5	-.09	-.16	.13
Stresslevel Gesamt	-	-	-.27°
Evaluation CFT 20-R	.05	-.06	.20
Evaluation LPS-neu	.02	.00	.16

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson
 n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe
 °) tendenziell signifikant auf dem Niveau $p = .10$ (zweiseitig)

Nun werden die Zusammenhänge der Leistungen im LPS-neu, d. h. die korrekten Antworten insgesamt mit dem Wohlbefinden, dem Stresslevel und der Evaluation des LPS-neu in Zusammenhang gebracht (vgl. Tabelle 4.23). Die Korrelationen zwischen den fünf Wohlbefindensmessungen und dem LPS-neu sind niedrig und nicht signifikant. Das gefühlte Stresslevel der Experimentalgruppe weist einen leichten, negativen Zusammenhang mit den korrekten Antworten des LPS-neu auf, der auf dem Niveau $p = .05$ signifikant wird. Personen, die sich durch den fiktiven IQ-Wert gestresster fühlten, erbrachten tendenziell schlechtere Leistungen. Die Zusammenhänge zwischen Leistungen des LPS-neu und der Bewertung dieses Intelligenztests sind nicht bzw. für die Experimentalgruppe tendenziell signifikant.

Tabelle 4.23: Korrelationen zwischen der Leistung im LPS und Wohlbefinden, gefühltem Stresslevel und Bewertung

	LPS-neu Korrekt Gesamt (n = 72)	LPS-neu Korrekt KG (n = 36)	LPS-neu Korrekt EG (n = 36)
Wohlbefinden 3	.07	.22	-.06
Wohlbefinden 4	-	-	.07
Wohlbefinden 5	-.05	-.12	.09
Stresslevel Gesamt	-	-	-.32*
Evaluation LPS-neu	.14	.07	.28°

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson

n = Anzahl der Vpn, Korrekt = korrekten Antworten über alle Aufgaben, KG = Kontrollgruppe, EG = Experimentalgruppe

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.4 Intelligenztestleistung und Leistungsmotivation

4.4.1 Intelligenztestleistung und Leistungsmotivation unter neutraler Bedingung

Aus den Ergebnissen der Homogenitätsüberprüfung in Kapitel 3.4.2 geht bereits hervor, dass sich die Experimental- und Kontrollgruppe bezüglich der CFT 20-R Intelligenzwerte signifikant unterscheiden ($t(69.84) = -3.67, p = <.01$). Die Ergebnisse dieser Welch-Test Berechnung soll in weiteren Analysen Beachtung finden.

H4: Vpn mit überwiegend HE erreichen im Vergleich zu Vpn mit überwiegend FM bessere Leistungen in den Intelligenztests in der neutraler Situation.

Auch für Hypothese 4 wurde in der Homogenitätsüberprüfung der Unterschied der Leistungsmotivgruppen bezüglich des CFT 20-R mittels Welch-Test überprüft. Die Behauptung, dass beide hier unterschiedliche Werte erzielen, ergibt kein signifikantes Ergebnis ($t(60.02) = 0.15, p = .89$). Beide Gruppen erzielen ähnliche IQ-Werte (HE: $M = 115.97, SD = 14.80$; FM: $M = 115.49, SD = 13.03$).

Nun soll für Hypothese 4 die Gruppe ohne Stressinstruktion bezüglich der Leistungen des LPS-neu getrennt betrachtet werden. Dabei wird der Einwand von Schneider (1973) neben der Menge auch die Güte der Leistung, d. h. Fehler, Bearbeitungszeit usw. gezielt zu analysieren, beachtet. Für diesen Zweck wird eine multivariate Varianzanalyse mit den Leistungsmotivgruppen als unabhängige Variable berechnet. Der Faktor der Leistungsmotivgruppen ist dabei nicht signifikant ($F(5, 30) = 0.34, p = .88, \text{partial}\eta^2 = .05$). Die in Tabelle 4.24 demonstrierten Post-hoc Vergleiche der einzelnen

LPS-neu Leistungen weisen keine Signifikanz und nur geringe Effektstärken auf, obwohl die Gruppe mit überwiegend HE mehr korrekte Antworten erzielt.

Tabelle 4.24: Univariate Varianzanalysen der Gruppen HE und FM bezüglich der LPS-neu Leistung in neutraler Situation

	HE, n = 14 M (SD)	FM, n = 22 M (SD)	F	p	$p\eta^2$
Korrekt	288.57 (42.46)	282.27 (41.59)	0.19	.66	.01
Fehler	51.86 (13.04)	51.68 (17.85)	0.00	.97	.00
Ausgelassen	18.21 (10.96)	16.09 (9.51)	0.38	.54	.01
Nicht Bearbeitet	161.36 (33.20)	169.32 (38.95)	0.40	.53	.01
Berichtigt	1.79 (1.76)	1.50 (1.37)	0.30	.59	.01

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung F = Testwert, p = Signifikanzwert, $p\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 34

Nach Carrolls (1993) Standpunkt ist es wichtig, neben dem Gesamtergebnis auch die einzelnen Fähigkeiten, welche durch die unterschiedlichen Subtests präsentiert sind, zu beachten. Daher wird für die Ergebnisse der Subtests ebenfalls eine multivariate Varianzanalyse bezüglich Hypothese 4 durchgeführt. Der Einfluss der Motivgruppen ergibt im Gesamtmodell kein signifikantes Ergebnis, aber einen hohen Effekt ($F(11, 24) = 1.21$, $p = .34$, $partial\eta^2 = .36$). Die Post-hoc Analysen aus Tabelle 4.25 weisen zwei tendenziell signifikante Werte für Subtest 2 und 3 auf. Beide Effekte liegen in einem mittleren Bereich. Obwohl nach den Mittelwerten die HE-Gruppe bei 11 Vergleichen und auch bei Subtest 3 häufiger korrekt antwortet, ist bei Subtest 2 sowie bei Subtest 1 und 4 die Misserfolgsfürchtige die bessere Gruppe.

Tabelle 4.25: Univariate Varianzanalysen der Gruppen HE und FM bezüglich LPS-neu Subtests in neutraler Situation

	HE, n = 14 M (SD)	FM, n = 22 M (SD)	F	p	$\rho\eta^2$
Subtest 1	39.50 (8.27)	39.73 (9.90)	0.01	.94	.00
Subtest 2	21.86 (7.53)	26.27 (7.45)	2.98	.09°	.08
Subtest 3	27.79 (3.77)	25.36 (3.55)	3.80	.06°	.10
Subtest 4	16.86 (3.13)	17.00 (2.33)	0.02	.88	.00
Subtest 5	20.93 (3.27)	20.18 (3.28)	0.45	.51	.01
Subtest 6	26.29 (7.54)	25.36 (7.64)	0.13	.72	.00
Subtest 7	32.14 (4.50)	31.27 (3.12)	0.47	.50	.01
Subtest 8	33.71 (7.16)	32.05 (4.88)	0.69	.41	.02
Subtest 9	14.07 (3.85)	12.68 (6.37)	0.54	.47	.02
Subtest10	36.14 (5.28)	33.82 (6.68)	1.21	.28	.03
Subtest 11	19.29 (4.94)	18.55 (8.75)	0.08	.78	.00

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 34
 °) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.4.2 Intelligenztestleistungen der Leistungsmotiv- und experimentellen Gruppen

In diesem Abschnitt soll geklärt werden, ob sich die Kontroll- und Experimentalgruppe und die Motivgruppen in ihren Leistungen im LPS-neu unterscheiden.

H5: Vpn mit überwiegend HE erreichen im Vergleich zu Vpn mit überwiegend FM bessere Leistungen in den Intelligenztests.

H6: Vpn in der neutralen Bedingung erreichen im LPS-neu im Vergleich zu Vpn der Stressbedingung bessere Leistungen.

Das gewählte multivariate Modell umfasst die unabhängigen Variablen der Testsituation, der Leistungsmotivgruppen sowie die Kovariate CFT 20-R IQ-Wert. Als abhängige Variablen dienen die LPS-neu Leistungen und die einzelnen Subtests. Die Kovariate beeinflusst das Modell signifikant ($F(5, 63) = 6.72$, $p = <.01$, $\text{partial}\eta^2 = .39$) und auch in den Post-hoc Vergleichen für korrekte, falsche und nicht bearbeitete Aufgaben ergibt sich ein signifikanter Einfluss der Kovariaten (vgl. Anhang F, Tabelle F-6). Daher sind in Abbildung 4.8 die Vergleiche in Abhängigkeit der Kovariate dargestellt. Die Vergleiche ohne den Einfluss der Kovariate sind im Anhang zu finden (vgl. Abbildung G-1).

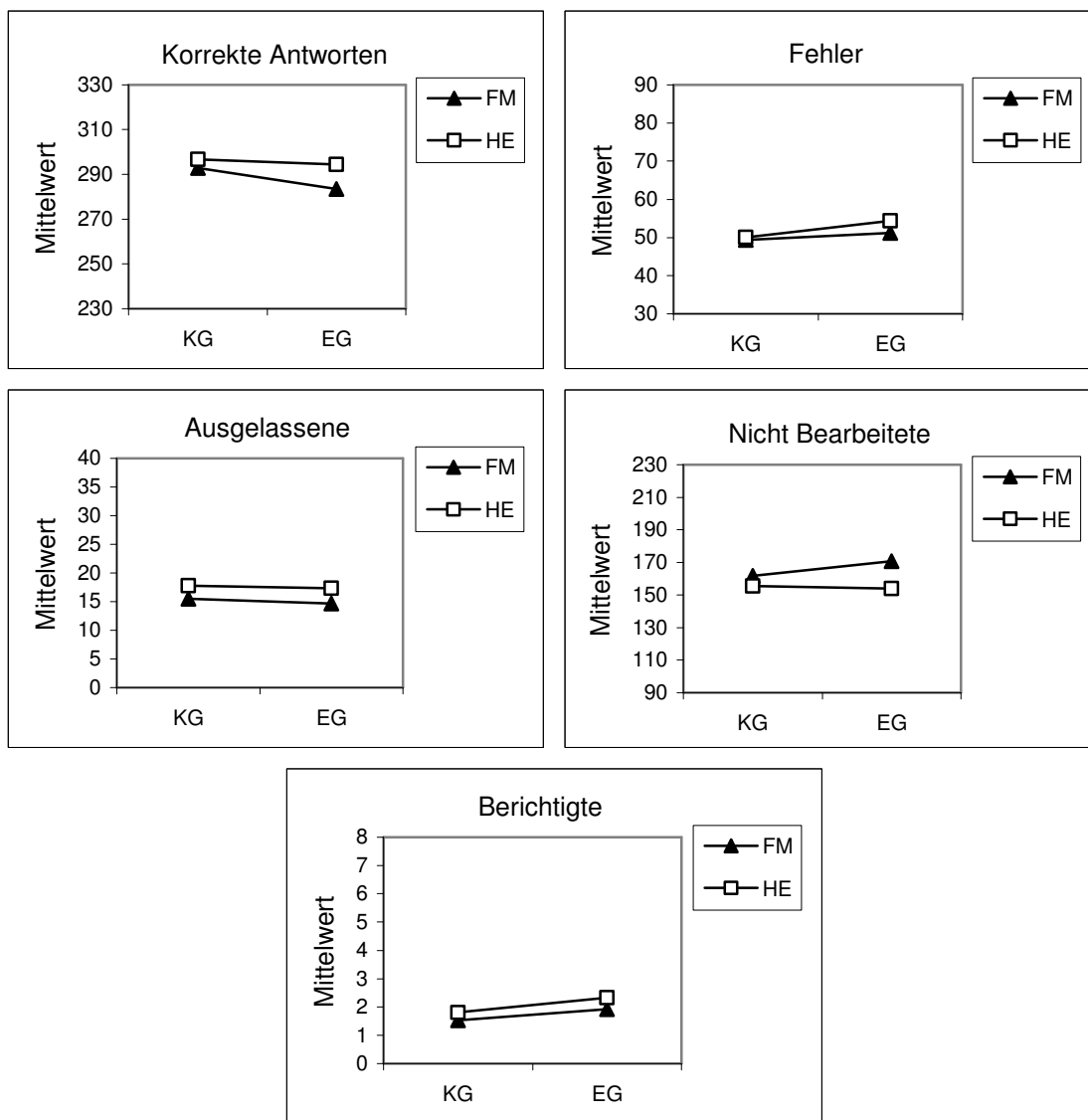


Abbildung 4.8: Vergleiche der Leistungen des LPS-neu der Experimentellen- und der Leistungsmotivgruppen unter dem Einfluss der Kovariaten CFT IQ-Wert

Die unabhängige Variable „Gruppe“ liefert keinen signifikanten Wert ($F(5, 63) = 0.30$, $p = .91$, $\text{partial}\eta^2 = .02$) und die einzelnen Post-hoc Vergleiche sind ebenfalls nicht signifikant. Der untersuchte Effekt betrifft einen Unterschied der Motivgruppen ist nicht signifikant ($F(5, 63) = 0.99$, $p = .43$, $\text{partial}\eta^2 = .07$). Die getrennten Analysen ergeben Werte, welche nicht im gewünschten Signifikanzbereich liegen. Für die nicht bearbeiteten Aufgaben ergibt sich allerdings ein mittlerer Effekt.

Bezüglich der Interaktion der Durchführungsbedingung und der beiden Leistungsmotivgruppen konnte kein signifikanter Effekt nachgewiesen werden ($F(5, 63) = 0.16$, $p = .98$, $\text{partial}\eta^2 = .01$). Für die Post-hoc Analysen der Interaktion ergeben sich keine signifikanten Werte (vgl. Tabelle 4.26).

Tabelle 4.26: Univariate Kovarianzanalysen der experimentellen Gruppen und der Motivgruppen bezüglich der Leistung im LPS-neu

	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>F</i>	<i>p</i>	$\rho\eta^2$
Gruppe	KG (n = 36)	EG (n = 36)			
Korrekt	284.72 (41.44)	298.19 (35.34)	0.49	.49	.01
Fehler	51.75 (15.95)	50.58 (18.14)	0.51	.48	.01
Ausgelassen	16.92 (10.00)	15.42 (11.87)	0.05	.83	.00
Nicht Bearbeitet	166.22 (36.54)	155.81 (32.98)	0.20	.65	.00
Berichtigt	1.61 (1.52)	2.14 (2.36)	0.77	.38	.01
Motivgruppe	HE (n = 31)	FM (n = 41)			
Korrekt	295.90 (43.37)	288.10 (35.21)	0.94	.34	.01
Fehler	52.32 (19.54)	50.29 (14.93)	0.24	.63	.00
Ausgelassen	17.52 (12.03)	15.15 (10.04)	0.87	.36	.01
Nicht Bearbeitet	154.26 (37.11)	166.12 (32.77)	2.40	.13	.04
Berichtigt	2.10 (2.23)	1.71 (1.79)	0.52	.47	.01
Interaktion¹⁾					
Korrekt			0.21	.65	.00
Fehler			0.09	.76	.00
Ausgelassen			0.01	.94	.00
Nicht Bearbeitet			0.51	.48	.01
Berichtigt			0.02	.89	.00

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Testwert, *p* = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, Kovariate = CFT 20-R IQ-Wert

¹⁾ für *M* und *SD* der Interaktion vgl. Tabelle 4.20

Eine multivariate Kovarianzanalyse soll Aufschluss über die postulierten Unterschiede zwischen Kontroll- und Experimentalgruppe und der Leistungsmotivgruppen bezüglich der LPS-neu Subtests geben. Als Kovariate diente der CFT 20-R IQ-Wert, welcher einen signifikanten Einfluss ($F(11, 57) = 4.87$, $p < .01$, $\text{partial}\eta^2 = .49$) hat. Bei jedem Subtest besitzt die Kovariate einen signifikanten Einfluss (vgl. Anhang F, Tabelle F-7).

Der Einfluss der experimentellen Gruppe liefert kein signifikantes Ergebnis für das Gesamtmodell ($F(11, 57) = 1.53$, $p = .15$, $\text{partial}\eta^2 = .23$). Der Einfluss der Leistungsmotivgruppen ist nicht signifikant ($F(11, 57) = 1.08$, $p = .39$, $\text{partial}\eta^2 = .17$). Bezüglich der Interaktion zwischen experimenteller Bedingung und Leistungsmotiv erreicht der Wert nicht die gewünschte Signifikanzgrenze ($F(11, 57) = 1.05$, $p = .41$, $\text{partial}\eta^2 = .17$).

In den univariaten Kovarianzanalysen der experimentellen Gruppen konnte für Subtests 7, 5 und 3 ein signifikantes Ergebnis ermittelt werden, wobei die Kontrollgruppe

bei Test 3 und 7 tendenziell besser abschneidet (vgl. Tabelle 4.27). Bei den anderen Mittelwerten liegt die Experimentalgruppe stets höher.

Die Post-hoc Analysen der Leistungsmotivgruppen liefern für Subtest 3 und 10 eine tendenzielle Signifikanz (vgl. Tabelle 4.27). Mit Ausnahme von Subtest 2 und 8 liegen die Misserfolgsfürchtigen stets niedriger in ihren Leistungen.

Tabelle 4.27: Univariate Kovarianzanalysen der experimentellen Gruppen und der Motivgruppen bezüglich der LPS-neu Tests

	<i>M (SD)</i>	<i>M (SD)</i>	<i>F</i>	<i>p</i>	$\rho\eta^2$
Gruppe	KG (n = 36)	EG (n = 36)			
Subtest 1	39.64 (9.18)	44.12 (10.32)	1.29	.26	.02
Subtest 2	24.56 (7.69)	27.72 (6.78)	0.25	.62	.00
Subtest 3	26.31 (3.78)	26.08 (3.21)	3.36	.07°	.05
Subtest 4	16.94 (2.63)	18.08 (2.95)	0.10	.75	.00
Subtest 5	20.47 (3.25)	20.50 (4.02)	4.01	.05°	.06
Subtest 6	25.72 (7.51)	25.94 (7.35)	0.29	.59	.00
Subtest 7	31.61 (3.68)	30.83 (4.12)	7.74	.01**	.10
Subtest 8	32.69 (5.84)	33.72 (4.61)	0.55	.46	.01
Subtest 9	13.22 (5.51)	14.33 (5.10)	0.62	.44	.01
Subtest 10	34.72 (6.20)	35.89 (4.40)	0.06	.80	.00
Subtest 11	18.83 (7.43)	20.94 (7.54)	0.04	.84	.00
Motivgruppe	HE (n =31)	FM (n = 41)			
Subtest 1	42.03 (10.11)	41.78 (9.96)	0.00	.98	.00
Subtest 2	25.39 (8.08)	26.71 (6.83)	1.05	.31	.02
Subtest 3	26.97 (3.66)	25.61 (3.27)	3.53	.06°	.05
Subtest 4	17.52 (3.33)	17.51 (2.44)	0.01	.94	.00
Subtest 5	20.97 (4.08)	20.12 (3.25)	1.45	.23	.02
Subtest 6	27.16 (7.66)	24.83 (7.08)	1.80	.18	.03
Subtest 7	31.35 (4.67)	31.12 (3.25)	0.21	.65	.00
Subtest 8	33.10 (6.39)	33.29 (4.27)	0.02	.88	.00
Subtest 9	14.10 (4.68)	13.54 (5.77)	0.24	.63	.00
Subtest 10	36.48 (5.20)	34.41 (5.39)	2.73	.10°	.04
Subtest 11	20.84 (6.96)	19.17 (7.91)	0.89	.35	.01

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Testwert, *p* = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, Kovariate = CFT 20-R IQ-Wert

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Für die einzelnen Post-hoc Berechnungen des Interaktionseinflusses konnte für Subtest 2 ein signifikanter Unterschied ermittelt werden (vgl. Tabelle 4.28). Der Effekt liegt

in einem mittleren Bereich. Während die Gruppe mit HE bei Subtest 2 in der Kontrollbedingung schlechtere Leistungen erbringt, ist die FM-Gruppe unter Stress schlechter.

Tabelle 4.28: Univariate Kovarianzanalysen der Interaktion zwischen den experimentellen Gruppen und den Motivgruppen bezüglich des LPS-neu Tests

Interaktion	KG – HE <i>M (SD)</i>	KG – FM <i>M (SD)</i>	EG – HE <i>M (SD)</i>	EG – FM <i>M (SD)</i>	<i>F</i>	<i>p</i>	η^2
Subtest 1	39.50 (8.27)	39.73 (9.90)	44.12 (11.21)	44.16 (9.75)	0.02	.88	.00
Subtest 2	21.86 (7.53)	26.27 (7.45)	28.29 (7.51)	27.21 (6.21)	4.49	.04*	.06
Subtest 3	27.79 (3.77)	25.36 (3.55)	26.29 (3.53)	25.89 (2.98)	1.16	.29	.02
Subtest 4	16.86 (3.13)	17.00 (2.33)	18.06 (3.47)	18.11 (2.49)	0.24	.63	.00
Subtest 5	20.93 (3.27)	20.18 (3.28)	21.00 (4.74)	20.05 (3.31)	0.24	.63	.00
Subtest 6	26.29 (7.54)	25.36 (7.64)	27.88 (7.92)	24.21 (6.52)	0.82	.37	.01
Subtest 7	32.14 (4.50)	31.27 (3.12)	30.71 (4.84)	30.95 (3.47)	0.12	.73	.00
Subtest 8	33.71 (7.16)	32.05 (4.88)	32.59 (5.85)	34.74 (2.94)	1.93	.17	.03
Subtest 9	14.07 (3.85)	12.68 (6.37)	14.12 (5.38)	14.53 (4.97)	0.24	.62	.00
Subtest 10	36.14 (5.27)	33.82 (6.68)	36.76 (5.27)	35.11 (3.40)	0.01	.93	.00
Subtest 11	19.29 (4.94)	18.55 (8.75)	22.12 (8.19)	19.89 (6.97)	0.47	.49	.01

Anmerkungen: *n* = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Testwert, *p* = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, Kovariate = CFT 20-R IQ-Wert

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Zur grafischen Verdeutlichung der signifikanten Unterschiede dient Abbildung 4.9. Da die Kovariate auf das Modell einen signifikanten Einfluss hat, wurden in den Diagrammen die Kovariate CFT 20-R IQ-Wert berücksichtigt, d. h. in den Abbildungen sind die geschätzten Randmittelwerte abgebildet. Alle Diagramme ohne den Einfluss der Kovariate sind im Anhang dargestellt (vgl. Abbildung G-3). Im Anhang sind zusätzlich auch die Diagramme der nicht signifikanten Subtests unter dem Einfluss der Kovariate zu finden (vgl. Abbildung G-2).

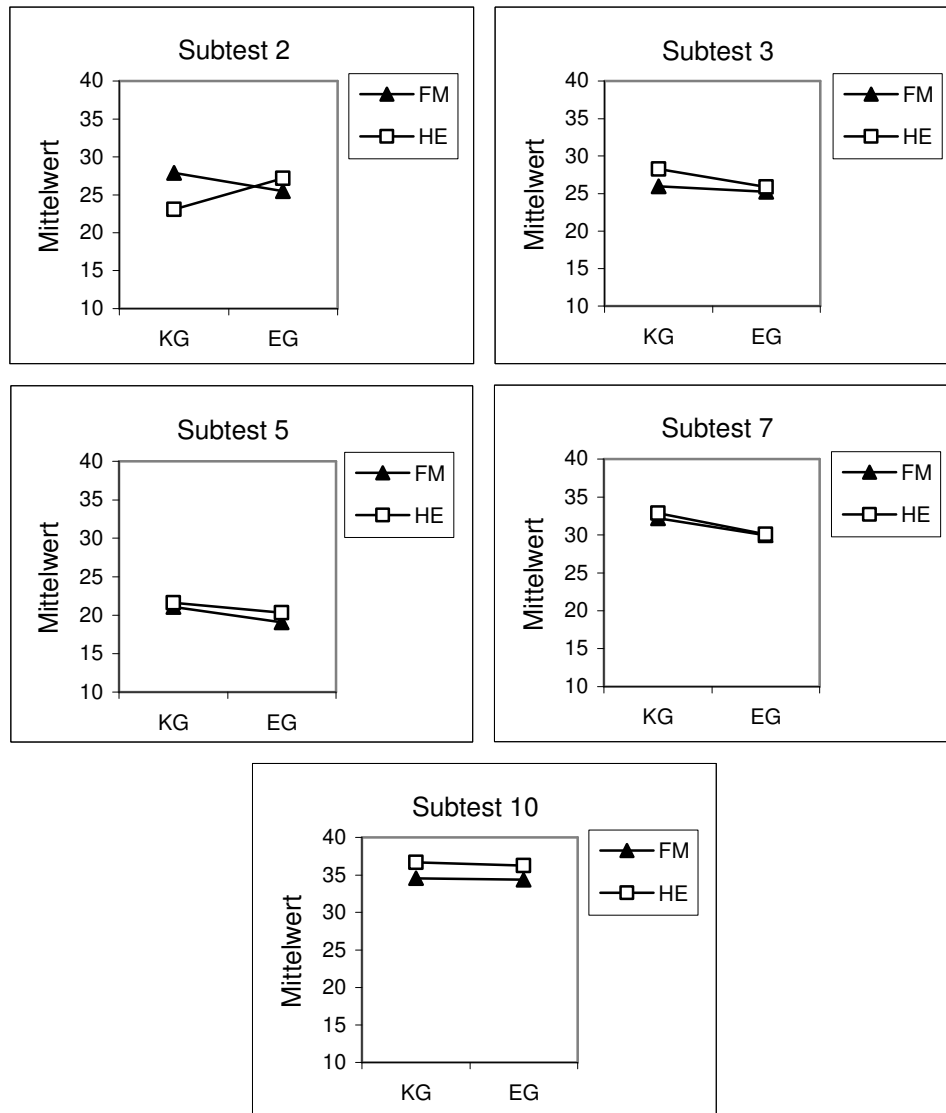


Abbildung 4.9: Vergleich der LPS-neu Subtests der experimentellen Gruppen und der Leistungsmotivgruppen unter dem Einfluss der Kovariate CFT 20-R IQ-Wert

4.4.3 Intelligenztestleistung und Leistungsmotivation unter Stress

H7: Vpn mit überwiegend HE erreichen im Vergleich zu Vpn mit überwiegend FM bessere Leistungen in dem Intelligenztest in der Stressbedingung.

Nun wird die Leistung im LPS-neu bezüglich der Motivgruppen mittels multivariater Kovarianzanalyse noch getrennt für die Gruppe mit Stressinstruktion betrachtet. Die Kovariate CFT 20-R IQ-Wert ergab einen signifikanten Einfluss ($F(4, 30) = 5.18$, $p = <.01$, $partial\eta^2 = .41$). Der Effekt des überwiegenden Leistungsmotivs ist nicht signifikant ($F(4, 30) = .82$, $p = .52$, $partial\eta^2 = .10$). Die berechneten univariaten Kovarianzana-

lysen der LPS-neu Ergebnisse verdeutlichen, dass die Unterschiede zwischen den Motivgruppen nicht signifikant werden (vgl. Tabelle 4.29). Unter der Bedingung eines Signifikanzniveaus von 10 % liefert das Ergebnis der nicht bearbeiteten Aufgaben einen tendenziell signifikanten Unterschied mit mittlerem Effekt. Bei der Betrachtung der Mittelwerte wird deutlich, dass Personen mit FM im Vergleich zu denen mit HE mehr Aufgaben nicht bearbeiten. Für die nicht bearbeiteten Aufgaben und für die korrekten Antworten hat die Kovariate einen signifikanten Einfluss (vgl. Anhang F, Tabelle F-8).

Tabelle 4.29: Univariate Kovarianzanalysen der Leistungsmotivgruppen bezüglich der LPS-neu Leistung in der Stressbedingung

	HE, n = 17 M (SD)	FM, n = 19 M (SD)	F	p	$p\eta^2$
Korrekt	301.94 (44.45)	294.84 (25.43)	1.36	.25	.04
Fehler	52.71 (24.02)	48.68 (10.90)	0.27	.61	.01
Ausgelassen	16.94 (13.15)	14.05 (10.78)	0.37	.55	.01
Nicht Bearbeitet	148.41 (40.08)	162.42 (24.28)	2.79	.10°	.08
Berichtigt	2.35 (2.57)	1.95 (2.20)	0.25	.62	.01

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung F = Testwert, p = Signifikanzwert, $p\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, Kovariate = CFT 20-R IQ-Wert
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Für die Subtests wird eine separate multivariate Kovarianzanalyse für die Stressgruppe vorgenommen. Die Kovariate CFT 20-R IQ-Wert fließt in die Analysen ein und hat keinen signifikanten Einfluss auf das Modell ($F(11, 23) = 0.59$, $p = .82$, $partial\eta^2 = .22$). Nur für die drei Subtests 1, 3 und 10 ergibt sich ein signifikanter Einfluss der Kovariate (vgl. Anhang F, Tabelle F-8). Der Haupteffekt der Leistungsmotivgruppen ist nicht signifikant ($F(11, 23) = 1.33$, $p = .27$, $partial\eta^2 = .39$).

Bei der Betrachtung der univariaten Kovarianzanalysen der Subtests konnten keine signifikanten Unterschiede zwischen den Motivgruppen gefunden werden (vgl. Tabelle 4.30). Beide Gruppen erzielen ähnliche Ergebnisse, wobei in sieben Subtests die Vpn mit überwiegend HE höhere Leistungen erbringen.

Tabelle 4.30: Univariate Kovarianzanalysen der Leistungsmotivgruppen bezüglich der LPS-neu Subtests in der Stressbedingung

	HE, n = 17 M (SD)	FM, n = 19 M (SD)	F	p	$\rho\eta^2$
Subtest 1	44.12 (11.21)	44.16 (9.75)	0.01	.94	.00
Subtest 2	28.29 (7.51)	27.21 (6.21)	0.78	.38	.02
Subtest 3	26.29 (3.53)	25.89 (2.98)	0.36	.55	.01
Subtest 4	18.06 (3.47)	18.11 (2.49)	0.11	.74	.00
Subtest 5	21.00 (4.74)	20.05 (3.31)	1.89	.18	.05
Subtest 6	27.88 (7.92)	24.21 (6.52)	2.57	.12	.07
Subtest 7	30.71 (4.84)	30.95 (3.47)	0.00	.96	.00
Subtest 8	32.59 (5.85)	34.74 (2.94)	1.65	.21	.05
Subtest 9	14.12 (5.38)	14.53 (4.97)	0.00	.99	.00
Subtest 10	36.76 (5.27)	35.11 (3.40)	1.99	.17	.06
Subtest 11	22.12 (8.19)	19.89 (6.97)	1.23	.28	.04

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, Kovariate = CFT 20-R IQ-Wert
 **) sehr signifikant auf dem Niveau von .01 (zweiseitig)
 *) signifikant auf dem Niveau von .05 (zweiseitig)
 °) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

H8: Vpn mit überwiegend HE erreichen in der Stressbedingung im Vergleich zur neutralen Bedingung in einem Intelligenztest bessere Leistungen.

Die Gruppe mit überwiegend HE wurde für Hypothese 8 auf einen Unterschied zwischen den experimentellen Gruppen geprüft. In einem multivariaten Kovarianzmodell ergab die Kovariate CFT 20-R IQ-Wert keinen signifikanten Einfluss ($F(4, 25) = 5.26$, $p = <.01$, $\text{partial}\eta^2 = .46$). Es konnte kein signifikanter Effekt der Durchführungsbedingung gefunden werden ($F(4, 25) = .27$, $p = .90$, $\text{partial}\eta^2 = .04$). Für die einzelnen univariaten Kovarianzanalysen lag keiner der Vergleiche im gewünschten Signifikanzbereich (vgl. Tabelle 4.31).

Tabelle 4.31: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Leistung der HE-Gruppe

	KG, n = 14 M (SD)	EG, n = 17 M (SD)	F	p	$\rho\eta^2$
Korrekt	288.57 (42.46)	301.94 (44.45)	0.14	.71	.01
Fehler	51.86 (13.04)	52.71 (24.02)	0.46	.50	.02
Ausgelassen	18.21 (10.96)	16.94 (13.15)	0.00	.99	.00
Nicht Bearbeitet	161.36 (33.20)	148.41 (40.08)	0.00	.98	.00
Berichtigt	1.79 (1.76)	2.35 (2.57)	0.68	.42	.02

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 28, Kovariante = CFT 20-R IQ-Wert

Die Analyse für Hypothese 8 wird analog für die Subtests des LPS-neu durchgeführt. Die Kovariante CFT 20-R IQ-Wert ergab in der multivariaten Kovarianzanalyse ein tendenziell signifikantes Ergebnis für die Subtests des LPS-neu ($F(11, 18) = 2.27, p = .06, \rho\eta^2 = .58$). Bezüglich des Hauptfaktors der experimentellen Bedingung wird der Effekt nicht signifikant ($F(11, 18) = 1.12, p = .40, \rho\eta^2 = .41$). Bei den univariaten Kovarianzanalysen beeinflusst der Kovarianzfaktor alle Subtests signifikant (vgl. Anhang F, Tabelle F-9). Die Post-hoc Analysen ergeben für Subtest 2, 3 und 7 tendenziell signifikante Ergebnisse (vgl. Tabelle 4.32). Die Richtung der signifikanten Ergebnisse stimmt allerdings nur bei Subtest 2, d. h. die Gruppe mit Stressinstruktion erreichte hier bessere Ergebnisse im Vergleich zur Kontrollgruppe. Insgesamt sind die Ergebnisse beider Gruppen ähnlich, obwohl die Gruppe mit Stressinstruktion in acht der Subtests leicht höhere Ergebnisse erzielt.

Tabelle 4.32: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Subtests der HE-Gruppe

	KG, n = 14 M (SD)	EG, n = 17 M (SD)	F	p	$\rho\eta^2$
Subtest 1	39.50 (8.27)	44.12 (11.21)	0.33	.57	.01
Subtest 2	21.86 (7.53)	28.29 (7.51)	2.91	.10°	.09
Subtest 3	27.79 (3.77)	26.29 (3.53)	4.02	.06°	.13
Subtest 4	16.86 (3.13)	18.06 (3.47)	0.00	.96	.00
Subtest 5	20.93 (3.27)	21.00 (4.74)	1.95	.17	.07
Subtest 6	26.29 (7.54)	27.88 (7.92)	0.00	.96	.00
Subtest 7	32.14 (4.50)	30.71 (4.84)	3.74	.06°	.12
Subtest 8	33.71 (7.16)	32.59 (5.85)	2.00	.17	.07
Subtest 9	14.07 (3.85)	14.12 (5.38)	0.29	.59	.01
Subtest 10	36.14 (5.27)	36.76 (5.27)	0.37	.55	.01
Subtest 11	19.29 (4.94)	22.12 (8.19)	0.24	.63	.01

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 28, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Kovariate = CFT 20-R IQ-Wert

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

H9: Vpn mit überwiegend FM erreichen in der neutralen Bedingung im Vergleich zur Stressbedingung in einem Intelligenztest bessere Leistungen.

Für Vpn mit überwiegend FM ist für Hypothese 9 der Unterschied zwischen den Durchführungsbedingungen mittels multivariater Kovarianzanalyse getestet worden. Der IQ-Wert des CFT 20-R wurde als Kovariate hinzugefügt und beeinflusst das Gesamtmodell signifikant ($F(5, 34) = 2.65$, $p = .04$, $\text{partial}\eta^2 = .28$). Der Haupteffekt zwischen der Neutral- und Stressbedingung ist nicht signifikant ($F(6, 34) = .11$, $p = .99$, $\text{partial}\eta^2 = .02$). Die univariaten Kovarianzanalysen ergeben keine signifikanten Werte (vgl. Tabelle 4.33), wobei die Kovariate die korrekten und nicht bearbeiteten Aufgaben signifikant beeinflusst (vgl. Anhang F, Tabelle F-10).

Tabelle 4.33: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Leistung der FM-Gruppe

	KG, n = 22 M (SD)	EG, n = 19 M (SD)	F	p	$\rho\eta^2$
Korrekt	282.27 (41.59)	294.84 (25.43)	0.24	.62	.01
Fehler	51.68 (17.85)	48.68 (10.90)	0.03	.87	.00
Ausgelassen	16.09 (9.51)	14.05 (10.78)	0.18	.68	.01
Nicht Bearbeitet	169.32 (38.95)	162.42 (24.28)	0.36	.55	.01
Berichtigt	1.50 (1.37)	1.95 (2.20)	0.02	.88	.00

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 38, Kovariate = CFT 20-R IQ-Wert

Für die Subtests wurde für Vpn mit überwiegend FM analog eine multivariate Kovarianzanalyse genutzt, wobei der Wert des CFT 20-R IQ-Wert vorab als Kovariate aus dem Modell herausgerechnet wurde. Der Einfluss der Kovariate ist im Gesamtmodell ($F(11, 28) = 3.19$, $p = .01$, $\text{partial}\eta^2 = .56$) und außer für Subtests 1, 6, 8 und 10 auch in den univariaten Kovarianzanalysen signifikant (vgl. Anhang F, Tabelle F-10). Es konnte kein signifikanter Unterschied zwischen der Stress- und Kontrollgruppe für Vpn mit überwiegend FM gefunden werden ($F(11, 28) = 0.90$, $p = .55$, $\text{partial}\eta^2 = .26$). Die Berechnungen der univariaten Kovarianzanalysen, die sich für die Subtests ergeben, liegen nur für Subtest 7 in einem tendenziell signifikanten Bereich (vgl. Tabelle 4.34). Vpn der Kontrollbedingung sind in Subtest 7 besser im Vergleich zur Stressgruppe. Die restlichen, nicht signifikanten Mittelwerte deuten auf einen Vorteil der Gruppe unter Stress hin, da diese in acht Subtests höhere Werte erzielen.

Tabelle 4.34: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Subtests der FM-Gruppe

	KG, n = 22 M (SD)	EG, n = 19 M (SD)	F	p	$\rho\eta^2$
Subtest 1	39.73 (9.90)	44.16 (9.75)	1.60	.21	.04
Subtest 2	26.27 (7.45)	27.21 (6.21)	2.19	.15	.05
Subtest 3	25.36 (3.55)	25.89 (2.98)	0.17	.68	.00
Subtest 4	17.00 (2.33)	18.11 (2.49)	0.12	.73	.00
Subtest 5	20.18 (3.28)	20.05 (3.31)	1.55	.22	.04
Subtest 6	25.36 (7.64)	24.21 (6.52)	0.33	.57	.01
Subtest 7	31.27 (3.12)	30.95 (3.47)	3.21	.08°	.08
Subtest 8	32.05 (4.88)	34.74 (2.94)	1.14	.29	.03
Subtest 9	12.68 (6.37)	14.53 (4.97)	0.65	.42	.02
Subtest 10	33.82 (6.68)	35.11 (3.40)	0.15	.70	.00
Subtest 11	18.55 (8.75)	19.89 (6.97)	0.69	.41	.02

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 38, Kovariate = CFT 20-R IQ-Wert
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

4.4.4 Intelligenztestleistung und die Höhe des Leistungsmotivs

H10: Vpn mit hoher GM erreichen im Vergleich zu Vpn mit niedriger GM bessere Leistungen in einem Intelligenztest.

Für die Hypothese 10 ist eine getrennte Betrachtung der Ergebnisse des CFT 20-R und LPS-neu nötig. Bezüglich des CFT 20-R ist eine Überprüfung mittels Welch-Test angemessen. Der Unterschied zwischen den Gruppen mit hoher und niedriger GM zeigt keinen signifikanten Einfluss ($t(69.54) = 1.44$, $p = .16$) mit mittlerem Effekt ($g = .34$). Vpn mit niedriger GM erzielten dabei einen höheren IQ-Wert ($M = 118.00$, $SD = 13.05$) im Vergleich zur Gruppe mit hoher GM ($M = 113.39$, $SD = 14.16$).

Für die Hypothese 10 wird eine multivariate Kovarianzanalyse zwischen den hochleistungsmotivierten und den niedrigleistungsmotivierten Vpn bezüglich der Leistungen des LPS-neu berechnet. Die Kovariate CFT 20-R IQ-Wert hat einen signifikanten Einfluss auf das Modell ($F(5, 65) = 7.65$, $p = <.01$, $\text{partial}\eta^2 = .37$) und die korrekten, falschen sowie nicht bearbeiteten Aufgaben (vgl. Anhang F, Tabelle F-11). Der untersuchte Haupteffekt der beiden Gesamtmotivationsgruppen ist nicht signifikant ($F(5, 65) = 0.66$, $p = .66$, $\text{partial}\eta^2 = .05$). Die Ergebnisse der Post-hoc Analysen in Tabelle 4.35 weisen keine Signifikanz und nur geringe Effektstärken auf. Die Gruppen mit

hoher und die mit niedriger GM scheinen sich in ihrer Leistung im LPS-neu nicht zu unterscheiden, wenn der Einfluss des IQ-Werts kontrolliert wird. Nur für Fehler ließ sich ein mittlerer Effekt ermitteln, wobei Vpn mit niedriger Gesamtmotivation mehr Fehler bei der Bearbeitung des LPS-neu begehen, aber insgesamt häufiger korrekt antworten.

Tabelle 4.35: Univariate Kovarianzanalyse der Gruppen mit niedriger GM ($n = 36$) und hoher GM ($n = 36$) bezüglich LPS-neu Leistung

	Niedrige GM <i>M (SD)</i>	Hohe GM <i>M (SD)</i>	<i>F</i>	<i>p</i>	$\rho\eta^2$
Korrekt	293.00 (41.70)	289.92 (36.27)	0.38	.54	.01
Fehler	53.36 (20.00)	48.97 (13.20)	2.52	.12	.04
Ausgelassen	16.17 (11.54)	16.17 (10.43)	0.04	.85	.00
Nicht Bearbeitet	157.47 (33.86)	164.56 (36.15)	0.05	.83	.00
Berichtigt	2.00 (2.24)	1.75 (1.71)	0.18	.67	.00

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, *F* = Testwert, *p* = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 69, Kovariate = CFT 20-R IQ-Wert

Zur Hypothese 10 wird die multivariate Kovarianzanalyse analog für die Subtests berechnet. Der CFT 20-R IQ-Wert geht dabei als Kovariate in das Modell ein und ergibt keinen signifikanten Einfluss für die Subtests ($F(11, 59) = 5.42$, $p = <.01$, $\rho\eta^2 = .50$). Die unabhängige Variable der Gesamtmotivationsgruppen hat keinen signifikanten Einfluss auf das Modell ($F(11, 59) = 0.57$, $p = .85$, $\rho\eta^2 = .10$). Die univariaten Kovarianzanalysen weisen keinen signifikanten Wert für die Subtests bzw. nur einen tendenziell signifikanten Wert für Subtest 9 auf. Dabei übt die Kovariate auf alle Subtests einen signifikanten Einfluss aus (vgl. Anhang F, Tabelle F-11). In Subtest 9 ist es die Gruppe mit hoher GM, welche in dem Test besser abschneidet. Die nicht signifikanten Ergebnisse bestärken die vorherigen Ergebnisse, dass kein Unterschied zwischen beiden Gruppen besteht (vgl. Tabelle 4.36).

Tabelle 4.36: Univariate Varianzanalysen der Gruppen mit niedriger GM (n = 36) und hoher GM (n = 36) bezüglich LPS-neu Subtests

	Niedrige GM M (SD)	Hohe GM M (SD)	F	p	$p\eta^2$
Subtest 1	42.67 (10.11)	41.11 (9.88)	0.09	.77	.00
Subtest 2	26.39 (8.21)	25.89 (6.54)	0.22	.64	.00
Subtest 3	26.17 (3.44)	26.22 (3.57)	0.30	.58	.00
Subtest 4	17.67 (3.00)	17.36 (2.68)	0.22	.64	.00
Subtest 5	20.92 (4.14)	20.06 (3.02)	0.14	.71	.00
Subtest 6	26.47 (6.91)	25.19 (7.86)	0.24	.63	.00
Subtest 7	31.31 (4.37)	31.14 (3.42)	0.19	.66	.00
Subtest 8	33.00 (5.85)	33.42 (4.64)	1.00	.32	.01
Subtest 9	13.25 (5.49)	14.31 (5.13)	2.77	.10°	.04
Subtest 10	35.06 (5.74)	35.56 (5.05)	0.75	.39	.01
Subtest 11	20.11 (7.89)	19.67 (7.21)	0.12	.73	.00

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $p\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 69, Kovariate = CFT 20-R IQ-Wert

H11: Vpn mit hoher GM erreichen in der Stressbedingung im Vergleich zur neutralen Bedingung in einem Intelligenztest bessere Leistungen.

Um Hypothese 11 für Vpn mit hoher GM zu überprüfen, ist ebenfalls eine multivariate Kovarianzanalyse geeignet. Der CFT 20-R IQ-Wert ist als Kovariate vorab herauspartialisieren worden, da dessen Einfluss signifikant ist ($F(5, 29) = 2.86, p = .03, \text{partial}\eta^2 = .33$). Die Kovariate hat auch auf die einzelnen Leistungen, außer für die falschen Antworten, einen signifikanten Einfluss (vgl. Anhang F, Tabelle F-12). Es konnte kein signifikanter Effekt für die unabhängige Variable der Durchführungsbedingung festgestellt werden ($F(5, 29) = 0.14, p = .98, \text{partial}\eta^2 = .02$). Aus Tabelle 4.37 geht hervor, dass die Berechnungen der univariaten Kovarianzanalysen nicht signifikant sind und sich nur geringe Effekte ermitteln lassen. Vpn der Stressbedingung antworten bei Betrachtung der Mittelwerte häufiger korrekt.

Tabelle 4.37: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Leistung der Gruppe mit hoher GM

	KG, n = 19 M (SD)	EG, n = 17 M (SD)	F	p	$\rho\eta^2$
Korrekt	283.74 (37.02)	296.82 (35.21)	0.20	.66	.01
Fehler	48.89 (11.35)	49.06 (15.37)	0.24	.63	.01
Ausgelassen	17.32 (11.01)	14.88 (9.93)	0.03	.87	.00
Nicht Bearbeitet	169.32 (37.93)	159.24 (34.39)	0.05	.82	.00
Berichtigt	1.53 (1.26)	2.00 (2.12)	0.01	.92	.00

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, Kovariate = CFT 20-R IQ-Wert

Für die Analyse der LPS-neu Subtests der HE-Gruppe diene ebenfalls eine multivariate Kovarianzanalyse. Der Einfluss der Kovariate CFT 20-R IQ-Wert ist signifikant ($F(11, 23) = 4.28, p = <.01, \text{partial}\eta^2 = .67$). Der Einfluss der experimentellen Bedingung ist nicht signifikant ($F(11, 23) = 1.05, p = .44, \text{partial}\eta^2 = .33$). Nur für Subtest 7 lässt sich ein signifikanter Unterschied zwischen den Gruppen feststellen, wobei die Kontrollgruppe die bessere Gruppe ist (vgl. Tabelle 4.38). Subtest 7 und auch sechs weitere Tests werden von der Kovariate signifikant beeinflusst (vgl. Anhang F, Tabelle F-12).

Tabelle 4.38: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Subtests der Gruppe mit hoher GM

	KG, n = 19 M (SD)	EG, n = 17 M (SD)	F	p	$\rho\eta^2$
Subtest 1	39.47 (9.13)	42.94 (10.63)	0.57	.46	.02
Subtest 2	24.89 (5.47)	27.00 (7.58)	0.37	.55	.01
Subtest 3	26.37 (3.76)	26.06 (3.46)	1.88	.18	.05
Subtest 4	16.68 (2.69)	18.12 (2.55)	0.02	.88	.00
Subtest 5	19.79 (2.76)	20.35 (3.35)	0.04	.84	.00
Subtest 6	24.95 (8.40)	25.47 (7.47)	0.01	.90	.00
Subtest 7	31.32 (3.53)	30.94 (3.40)	6.00	.02*	.15
Subtest 8	32.47 (5.28)	34.47 (3.68)	0.07	.79	.00
Subtest 9	13.68 (5.55)	15.00 (4.68)	0.04	.83	.00
Subtest 10	35.37 (5.80)	35.76 (4.21)	0.10	.75	.00
Subtest 11	18.74 (5.65)	20.71 (8.70)	0.03	.86	.00

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, Kovariate = CFT 20-R IQ-Wert

*) signifikant auf dem Niveau von .05 (zweiseitig)

H12: Vpn mit niedriger GM erreichen in der Stressbedingung im Vergleich zur neutralen Bedingung in einem Intelligenztest bessere Leistungen.

Für die Gruppe mit niedriger GM wird bezüglich der LPS-neu Leistungen eine multivariate Kovarianzanalyse berechnet. Als Kovariate dient der IQ-Wert CFT 20-R, welcher das Modell signifikant beeinflusst ($F(4, 30) = 6.30$, $p = <.01$, $\text{partial}\eta^2 = .46$). Bezüglich der experimentellen Situation konnte kein signifikanter Einfluss für das Modell ermittelt werden ($F(4, 30) = 0.31$, $p = .87$, $\text{partial}\eta^2 = .04$). Keiner der Post-hoc Vergleiche weist eine Signifikanz unter 10 % auf (vgl. Tabelle 4.39); die korrekten, falschen sowie nicht bearbeiteten Aufgaben werden dabei ebenfalls signifikant von der Kovariate beeinflusst (vgl. Anhang F, Tabelle F-13).

Tabelle 4.39: Univariate Kovarianzanalyse der experimentellen Gruppen bezüglich der LPS-neu Leistung der Gruppe mit niedriger GM

	KG, n = 17 M (SD)	EG, n = 19 M (SD)	F	p	$\rho\eta^2$
Korrekt	285.82 (47.03)	299.42 (36.37)	0.14	.71	.00
Fehler	54.94 (19.77)	51.95 (20.63)	0.19	.66	.01
Ausgelassen	16.47 (9.06)	15.89 (13.63)	0.04	.84	.00
Nicht Bearbeitet	162.76 (35.76)	152.74 (32.29)	0.05	.83	.00
Berichtigt	1.71 (1.79)	2.26 (2.60)	1.14	.29	.03

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, Kovariate = CFT 20-R IQ-Wert

Die Hypothese H 12 bezüglich der Gruppe mit überwiegend FM wird mittels multivariater Kovarianzanalyse bezüglich der Subtests des LPS-neu untersucht. Der CFT 20-R IQ-Wert wird dabei als Kovariate kontrolliert und hat einen signifikanten Einfluss auf das Gesamtmodell ($F(11, 23) = 3.97$, $p = <.01$, $\text{partial}\eta^2 = .66$) und auf die Subtests in den univariaten Kovarianzanalysen (vgl. Anhang F, Tabelle F-13). Zwischen den experimentellen Gruppen wurde der Unterschied im Gesamtmodell nicht signifikant ($F(11, 23) = 1.20$, $p = .34$, $\text{partial}\eta^2 = .37$). Ein signifikanter Unterschied konnte in den Post-hoc Vergleichen bei Subtest 5 gefunden werden (vgl. Tabelle 4.40), wobei die Kontrollgruppe bessere Werte erzielt.

Tabelle 4.40: Univariate Kovarianzanalyse der experimentellen Gruppen bezüglich der LPS-neu Subtests der Gruppe mit niedriger GM

	KG, n = 17 M (SD)	EG, n = 19 M (SD)	F	p	$\rho\eta^2$
Subtest 1	39.82 (9.50)	45.21 (10.20)	0.96	.33	.03
Subtest 2	24.18 (9.77)	28.37 (6.12)	0.54	.47	.02
Subtest 3	26.24 (3.91)	26.11 (3.07)	0.65	.43	.02
Subtest 4	17.24 (2.61)	18.05 (3.34)	0.11	.75	.00
Subtest 5	21.24 (3.65)	20.63 (4.62)	5.50	.03*	.14
Subtest 6	26.59 (6.51)	26.37 (7.42)	0.32	.57	.01
Subtest 7	31.94 (3.93)	30.74 (4.76)	2.71	.11	.08
Subtest 8	32.94 (6.55)	33.05 (5.33)	0.35	.56	.01
Subtest 9	12.71 (5.59)	13.74 (5.51)	0.49	.49	.01
Subtest 10	34.00 (6.72)	36.00 (4.68)	0.07	.79	.00
Subtest 11	18.94 (9.21)	21.16 (6.59)	0.00	.98	.00

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, Kovariate = CFT 20-R IQ-Wert

*) signifikant auf dem Niveau von .05 (zweiseitig)

4.4.5 Zusammenhänge zwischen Leistungsmotiven und Intelligenztestleistungen

Die Zusammenhänge zwischen der Leistungsmotivation und den Intelligenzwerten sind in Tabelle 4.41 aufgeführt. Nur die Korrelation zwischen dem CFT 20-R IQ-Wert und FM weist einen tendenziell signifikanten Wert auf.

Tabelle 4.41: Korrelationen zwischen IQ-Wert des CFT 20-R, der Leistung im LPS und der Leistungsmotivausprägungen (n = 72)

	CFT IQ-Wert Gesamt	LPS-neu Korrekt Gesamt	LPS-neu Korrekt KG(n = 36)	LPS-neu Korrekt EG (n = 36)
HE	.00	.07	.10	.07
FM	-.19°	-.15	-.11	-.18
GM	-.19	-.10	-.02	-.15
NH	.14	.14	.13	.17

Anmerkungen: Produkt-Moment-Korrelation nach Braivais Pearson

n = Anzahl der Vpn, Korrekt = korrekten Antworten über alle Aufgaben, KG = Kontrollgruppe, EG = Experimentalgruppe

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

5. Diskussion

Die vorgestellten Ergebnisse der statistischen Hypothesenprüfung werden in den folgenden Abschnitten unter Bezugnahme auf die theoretischen Ausführungen aus Kapitel 1 diskutiert. Anschließend werden die Durchführungsbedingungen und verwendeten Messmethoden kritisch analysiert und Anregungen für künftige Forschungsarbeiten abgeleitet.

5.1 Diskussion der Ergebnisse

5.1.1 Einfluss der Stressreaktion

Bei der Betrachtung der Angaben zum Wohlbefinden der Vpn wird deutlich, dass sowohl die Stressinstruktion als auch die Intelligenztests eine Abnahme bewirken. Dabei ergeben die Analysen, dass die Stressinstruktion das Wohlbefinden stärker beeinflusst als die alleinige Intelligenztestung. Einerseits sinkt das Wohlbefinden der Experimentalgruppe nach der Stressinstruktion stärker als nach der Intelligenztestung, andererseits ist der Effekt der Abnahme des Wohlbefindens vor im Vergleich zu nach der LPS-neu Testung in der Experimentalgruppe im Vergleich zur Kontrollgruppe größer. Dass die Abnahme des Wohlbefindens bei gesamter und getrennter Betrachtung für das LPS-neu stärker als beim CFT 20-R ausfällt und das LPS-neu auch allgemein schlechtere Bewertungen bekommt, lässt vermuten, dass die Vpn dieses als schwieriger und weniger verständlich empfinden, wodurch auch ihr Wohlbefinden stärker leidet. Möglich ist auch, dass sich die Vpn in der zweiten Intelligenztestung von vornherein weniger wohl fühlten, woraufhin ihre Bewertungen des LPS-neu geringer ausfielen.

Die Unterschiede bezüglich des Wohlbefindens zwischen den Durchführungsbedingungen zu den fünf Messzeitpunkten entsprechen, wenn die Vpn korrekte Angaben machten, der Hypothese 1. Kontroll- und Experimentalgruppe weisen während der gesamten Testungen mit dem CFT 20-R sowie vor dem LPS-neu und der Stressinstruktion ein ähnliches Befindlichkeitsmuster auf. Indes lässt sich aufgrund der Ergebnisse vermuten, dass sich die Experimentalgruppe nach dem Misserfolg durchschnittlich weniger wohl fühlt, wie dies bereits McKinney (1933), Zander (1944) und Lantz (1945) berichten. Die subjektive Einschätzung des gefühlten Stresslevels bestätigt, dass sich etwa dreiviertel der Teilnehmer (77.8 %) durch die Stresssituation beeinflusst fühlen. Beide Fragebögen dienen der Kontrolle des Einflusses der Stressinstruktion und gemäß den Ergebnissen können beide Anhaltspunkte zur Einschätzung der Emotionen der Teilnehmer eine geeignete Manipulation bestätigen.

Ob sich die Emotionen in den Bewertungen der Intelligenztests niederschlagen, wird anhand von Vergleichen der Evaluationen zwischen den Durchführungsbedingungen geprüft. Der Unterschied der Bewertung des LPS-neu zwischen den Gruppen mit und ohne Stressinstruktion, der in der Fragestellung 2 überprüft wird, wird in dieser Arbeit statistisch nicht bedeutsam. Die Bewertungen sind folglich weniger abhängig von Stress. Die Beobachtungen in den Untersuchungen von Hoppe (1930), Lantz (1945) sowie Postman und Brunner (1948), dass Vpn einen Misserfolg zum Selbstschutz oft mit externen Ursachen rechtfertigen wie eine negative Bemerkung zur Aufgabe oder speziell zur Aufgabenschwierigkeit, kann in dieser Arbeit nicht empirisch belegt werden. Während die Ergebnisse der drei genannten Studien auf Verhaltensbeobachtungen ausgewählter Teilnehmer beruhen, wird in der vorliegenden Studie versucht, die Beobachtungen mittels Fragebögen zu replizieren. Es kann vermutet werden, dass die negativen Bewertungen relativ individuell und eher spontan im Affekt geäußert werden. Bei einer direkten Befragung mittels Fragebogen, welche durchaus mit wichtigen Testbewertungen in Verbindung gebracht werden können, könnte sich eine eher vernünftige als gefühlsbetonte Evaluation durchsetzen und die Vpn könnten zu einer möglichst objektiven Bewertung „gezwungen“ werden. Tendenziell ist tatsächlich ein gewisser Effekt von $_{partial}I^2 = .03$ in der Bewertung des LPS-neu, jedoch nicht beim CFT 20-R, vorhanden, der eine leicht schlechtere Bewertung durch die Experimentalgruppe vermuten lässt. Interessant wäre es zu analysieren, ob dieser Unterschied signifikant wird, wenn die Teilnehmer aufgefordert werden, möglichst subjektive Bewertungen abzugeben.

Die mittleren Zusammenhänge zwischen den Testbewertungen, dem Wohlbefinden und dem gefühlten Stresslevel korrelieren größtenteils signifikant miteinander. Zwar sind keine Rückschlüsse auf Kausalzusammenhänge möglich, es kann dennoch vermutet werden, dass mit steigendem Wohlbefinden auch die Bewertung beider Gruppen positiver ausfällt und das gefühlte Stresslevel der Teilnehmer der Experimentalgruppe abnimmt. Zusätzlich gilt, je gestresster eine Vpn durch einen Misserfolg ist, desto schlechter bewertet sie den LPS-neu. Diese Zusammenhänge beider Gruppen gelten in geringerem Maß bereits schon vor den Intelligenztests und vor der Stressinstruktion, d. h. wie gestresst eine Vp reagiert und die Tests bewertet, hängt auch mit dem Wohlbefinden beim Testungsbeginn zusammen. Gehen Vpn schon mit einem geringen Wohlbefinden in die Testung, scheinen diese auch eine erhöhte Vulnerabilität oder gemäß Lazarus (1966) weniger Ressourcen zur Bewältigung der Stressinstruktion zur Verfügung zu haben. Je schlechter ihr Wohlbefinden am Anfang ist, desto schwieriger wird es, den erlebten Misserfolg zu kompensieren. Dies spiegelt sich auch darin wieder, dass Personen, die sich bereits vor der Testung unwohl fühlen, auch nachher die

Tests schlechter bewerten. Dieser Effekt lässt vermuten, dass einige Vpn der Testung grundsätzlich mit einer negativen Einstellung gegenüberstehen und mit einem zusätzlichen Misserfolg sich deren negative Erwartungen und der Zusammenhang verstärken.

Die Korrelationen dieser drei Fragebögen mit den Ergebnissen der zwei verwendeten Intelligenztests CFT 20-R und LPS-neu ergeben geringe Zusammenhänge. Obwohl das Wohlbefinden durch Stressinstruktion beeinflusst wird, hat dies keine weiteren Auswirkungen auf die Intelligenzleistungen. Es besteht allerdings eine geringe Tendenz, dass je höher die Stressbelastung empfunden wird, desto schlechter fallen die Ergebnisse im CFT 20-R ($r = -.27$) und im LPS-neu ($r = -.32$) aus. Obwohl der CFT 20-R die Kontrollbedingung darstellt, scheint das Stresserleben auch mit diesem Wert zusammenzuhängen. Es könnte vermutet werden, dass weniger intelligente Personen sich in Intelligenztestsituationen allgemein gestresster fühlen oder umgekehrt, dass schlechtere Resultate durch Angst vor Intelligenztests zustande kommen. Bei der separaten Betrachtung der Intelligenztests könnte auch gemutmaßt werden, dass die mitgeteilten fiktiven Misserfolge der Personen, die im CFT 20-R schlecht abschneiden, im Vergleich zu der fiktiv präsentierten Studie die höchsten Differenzen aufweisen und tendenziell höheren Stress auslösen. Für das LPS-neu entspricht die Korrelation den empirischen Überlegungen, d. h. je gestresster eine Person sich fühlt, desto geringere Leistungsresultate erzielt sie. Auf der anderen Seite wäre aber auch denkbar, dass Personen, die Schwierigkeiten bei der Bearbeitung mit dem LPS-neu haben und schlechte Leistungen erbringen, sich im Nachhinein gestresster fühlen. Ähnliches gilt für die Bewertung: Je schlechter eine Person im LPS-neu ist, desto negativer wird der Test tendenziell bewertet.

5.1.2 Leistungsmotivation – Wohlbefinden, gefühltes Stresslevel und Bewertung der Intelligenztests

Bezüglich der drei Einschätzungsfragebögen wird weiterhin ein Unterschied zwischen den Ausprägungen des Leistungsmotivs überprüft. Ursprünglich wird von Atkinsons (1964/1975) Definition des Erfolgs- und Misserfolgsmotiv ausgegangen, dass Personen mit starker Ausprägung von „Hoffnung auf Erfolg“ (HE) Leistungssituationen positive Affekte und Personen mit „Furcht vor Misserfolg“ (FM) negative Affekte bzw. Valenzen zuordnen. Die erhöhte Spannung von Misserfolgsmotivierten bei Intelligenztests (Bartmann, 1963) kann bezüglich der Items zum Wohlbefinden dieser Studie wie Unruhe, Unwohlsein und Anspannung nicht repliziert werden, was zur Ablehnung von Hypothese 2 führt. Die Mittelwerte weisen jedoch in die Richtung der Hypothese, da Erfolgshoffende in allen Messzeitpunkten ein höheres Wohlbefinden angeben und bei der ersten Messung vor der unbekanntem Testsituation und nach der Stressbedingung sich

stärkere Effekte abzeichnen. Die Korrelationen zwischen den Leistungsmotivausprägungen und dem Wohlbefinden bestätigen diesen Trend, dass je mehr FM eine Person ausmacht, desto geringer ist das Wohlbefinden in den Testsituationen und je höher HE, desto höhere Wohlbefindenswerte werden angegeben. Tatsächlich sind die Zusammenhänge teilweise eher gering und nicht bzw. nur tendenziell signifikant. Vor Beginn der ersten Testung scheinen sich die in Leistungssituationen entgegengesetzten Antizipationen der Leistungsmotivausprägungen am stärksten in höheren Zusammenhängen widerzuspiegeln. Im Verlauf der Testung lässt der Einfluss des Leistungsmotivs auf das Wohlbefinden nach. Die Betrachtung der Veränderung des Wohlbefindens über den Testzeitraum weist wiederum sowohl für Personen mit FM, als auch mit HE eine Senkung des Wohlbefindens auf, wobei der Effekt für die FM Gruppe nach der Stressinstruktion stärker ausfällt. Die Anregung des Leistungsmotivs lässt scheinbar im Verlauf der Testung nach, was an einer Steigerung weiterer äußerer Einflüsse wie der Anregung anderer Motive (French, 1958), Kompensationsmöglichkeiten, die Vertrautheit der Testsituation (Petty & Harrell, 1977) oder einfach an der Müdigkeit liegen kann. Vergleichsweise eindeutig ist, dass sich bei einem stärker ausgeprägtem HE-Motiv im Vergleich zu FM die höchsten und durchgehend signifikanten Zusammenhänge zum Wohlbefinden herausbilden. Bei der Betrachtung einer der Leistungsmotivskalen wird allerdings vernachlässigt, dass jeder das Erfolgs- und Misserfolgsmotiv als stabile Personenvariablen in unterschiedlichem Ausmaß besitzt und erst das Zusammenwirken beider ein umfassendes Bild ergibt.

Ein eindeutigeres Bild über die momentane Verfassung der Teilnehmer ergibt sich, wenn sie direkt nach dem empfundenen Stress befragt werden: Vpn mit überwiegend FM scheinen nach Misserfolg Stress intensiver zu erleben, was die Erwartungen von Hypothese 3 bestätigt. Die Ausführungen von Weiner (1972/1976b), dass Personen mit FM Misserfolg mit geringer Begabung und Personen mit HE diesen mit externalen Ursachen verbinden, kann als Erklärung für den stärkeren Einfluss bei Misserfolgfüchtigen herangezogen werden. Daraus kann geschlussfolgert werden, dass der Misserfolg als zusätzliche Belastung ihre generelle Furcht in Leistungssituationen verstärkt. Die positiven Korrelationen zwischen dem eingeschätzten Stresslevel und FM sowie der „Gesamtmotivation“ (GM) und die negativen zu HE sowie der „Nettohoffnung“ (NH) unterstützen diese Annahme. Besonders bedeutsam scheint hier, dass mit steigender Misserfolgsschreck das Scheitern am gravierendsten erlebt wird. Es kann erneut nicht klar entschieden werden, ob ein Kausalzusammenhang besteht oder eine mögliche Drittvariable die Korrelation beeinflusst. Die obigen Ausführungen von Weiner sind jedoch einleuchtend, um die Zusammenhänge zwischen Stresslevel und FM, HE sowie der NH zu erklären. Bei genaueren Überlegungen ist auch nicht überraschend, dass

Personen mit einer stark ausgeprägten GM bei unerwartetem Misserfolg deutlich mehr Stress empfinden. Hochleistungsmotivierte haben in Leistungssituationen grundsätzlich eine hohe Erfolgserwartung (Atkinson, 1964/1975) und da sich mit steigender Leistungsmotivation die persönliche Wichtigkeit von Misserfolg erhöht (Becker, 1982), übt die Stressinstruktion intensiveren psychischen Druck aus.

Insgesamt finden sich die aufgrund früherer Erfahrungen entstandenen negativen Valenzen gegenüber Leistungssituationen bei Vpn mit FM und vergleichsweise positiven Valenzen Erfolgsmotivierter (Weiner, 1972/1976b; Selbstbegründungsmodell nach Heckhausen, 1963, 1972) im generellen Wohlbefinden nicht eindeutig wieder. Auch bezüglich der Bewertungen kommt die in Fragestellung 1 erwartete negative Einstellung der Misserfolgsfürchtigen gegenüber Intelligenztests weder in den Unterschiedsprüfungen noch bezüglich der Korrelationen zum Ausdruck. Selbst bei einem unerwarteten Misserfolg bleiben die Bewertungen und das Wohlbefinden unbeeinflusst vom Leistungsmotiv, jedoch löst Misserfolg bei der Gruppe mit FM ein stärkeres Stressempfinden aus. Die Antizipation negativer Folgen aufgrund von vergangenem Misserfolg scheint vermutlich erst bei der Konfrontation mit neuem Misserfolg wach zu werden und sich in dem Stresserleben der Personen mit FM zu zeigen. Es kann davon ausgegangen werden, dass Misserfolgsfürchtige generell ein niedrigeres Selbstkonzept der eigenen Begabung aufweisen (Sowa, 2009); sie erreichen ihr eigenes Leistungspotential bzw. die individuellen Ziele gemäß dem Anspruchsniveau (Atkinson, 1964/1975) nicht und empfinden mehr Stress als Vpn mit HE. Dass sich dies nicht in einer Veränderung des Wohlbefindens zeigt, ist dennoch überraschend. Ein Versuch dies zu erklären, könnte von den in Kapitel 3.3 berichteten Faktoranalysen ausgehen. Während die Items des Fragebogens zum Stresslevel eindeutig auf einem Faktor laden, sind die Analysen zum Wohlbefindensfragebogen uneindeutig, worauf im Kapitel 5.2.1.1 näher eingegangen wird. Als eine weitere Erklärung könnte auch eine Implikation des Risiko-Wahl-Modells dienen, die voraussagt, dass das Gefühl von Scham über Misserfolg und Stolz über Erfolg am größten bei Aufgaben mittlerer Schwierigkeit ist. Personen mit Misserfolgsschuld sollten diese eher meiden und schwierige oder leichte Aufgaben bevorzugen, wie dies auch Studien von Schneider und Meise (1973) oder von Atkinson und Litwin (1960) bestätigen. Ein Intelligenztest kann so gesehen als schwere Aufgabe gelten. Da die subjektive Erfolgswahrscheinlichkeit bei schweren Aufgaben gering ist, vermutet Atkinson (1957), dass Misserfolgsfürchtige aufgrund geringerer Selbstschuld bei Misserfolg diese wählen. Es sollte jedoch bei dieser Erörterung zum Wohlbefinden beachtet werden, dass die Angaben des Befindens insgesamt nicht sehr extrem sind, sondern durchschnittlich im mittleren Bereich liegen, d. h. keine Person sich in der Studie extrem wohl oder unwohl fühlte.

Diese Schlussfolgerungen könnten zusammenfassend dazu führen, dass das Wohlbefinden der Vpn mit FM nicht geringer als das der Erfolgsmotivierten ausfällt, wohl aber bergen Instruktionen, die den Selbstwert herabsetzen, ein höheres Stresspotential für Misserfolgsfürchtige. Die Leistungsmotive scheinen bezüglich des Wohlbefindens und der Bewertungen keine direkte Rolle zu spielen, wohl aber bei der Höhe des empfundenen Stresses. Wie sich eine Person fühlt, hängt, wie in Kapitel 5.1.1 erläutert, eher mit der Stressbelastung zusammen und daher scheint der Zusammenhang zwischen den Leistungsmotiven und Wohlbefinden eher indirekt zu sein.

5.1.3 Intelligenztestleistungen und Leistungsmotivation

Das Hauptanliegen dieser Arbeit war es, den Einfluss der Durchführungsbedingung in einer Leistungssituation in Abhängigkeit der Leistungsmotivausprägungen zu untersuchen. Die in den Hypothesen 4 bis 12 postulierten Leistungsunterschiede zwischen den verschiedenen Intensitäts- und Richtungsausprägungen der Leistungsmotive ließen sich in der vorliegenden Arbeit nur geringfügig finden. In folgenden Abschnitten sollen die zentralen Ergebnisse erst getrennt für die Einteilungen in HE und FM bezüglich der experimentellen Bedingungen und generell für die gesamte Stichprobe sowie für die Höhe der GM diskutiert werden, um die Intelligenztestresultate abschließend in einer umfassenden Betrachtung zu erörtern.

5.1.3.1 Intelligenztestleistungen und Leistungsmotivation unter neutraler Bedingung

Bereits aus den vorgestellten Studien geht hervor, dass die Befundlage inkonsistent ist, ob und wann Personen mit FM niedrigere Leistungen als die mit überwiegend HE erbringen. Die hemmende FM sollte gemäß dem Risiko-Wahl-Modell nach Atkinson (1957, 1964/1975) generell zu schlechteren Leistungen führen und Personen mit HE sollten in Leistungssituationen eine höhere Anregung und bessere Leistungen verzeichnen (McClelland et al., 1953). Untersuchungen konnten dies einerseits generell (Atkinson & Litwin, 1960), andererseits häufig nur für komplexe und schwere Aufgaben nachweisen (Feather, 1965; Schneider & Gallitz, 1973). Gerade schwierige Aufgaben mit niedriger Erfolgswahrscheinlichkeit führen für Vpn mit HE zu einem höheren Anreiz und stärkerer Motivation, während Misserfolgsmotivierte durch stärkere Furcht (Schneider, 1973) oder aufgaben-irrelevante Gedanken (Sarason et al., 1952) gehemmt sind. Es besteht daher die Annahme, dass Personen mit HE bei Intelligenztestaufgaben generell einen Vorteil haben, was in Hypothese 4 für die neutrale Situation und in Hypothese 6 für die gesamte Stichprobe geprüft wurde.

Was die Intelligenzleistung des CFT 20-R mit seinen eher figuralen Aufgaben und des LPS-neu in der neutralen Situation betrifft, sind beide Leistungsmotivgruppen sich sehr ähnlich. Atkinsons Vermutungen lassen sich somit in der aktuellen Studie nicht bestätigen. Nur tendenziell sind die Vpn mit HE diejenigen, die korrekter antworten, mehr Aufgaben auslassen, weniger nicht bearbeiten und bei acht der elf Subtests bessere Ergebnisse erzielen. Dabei erreicht der Subtest 3 (Figurenfolgen) ein tendenziell signifikantes Ergebnis in Richtung der Hypothese 4. Bezüglich der Anagrammaufgabe, dem Subtest 2, sind es allerdings Misserfolgsfürchtige, die entgegen der Hypothese 4 den höheren Wert erreichen und somit die Hypothese widerlegen. Die erwarteten Unterschiede, die unter anderem Schneider und Gallitz (1973) finden konnten, kommen in den Daten dieser Untersuchung nicht zum Tragen. Als Erklärung könnten die Ergebnisse von McClelland et al. (1953), Sarason et al. (1952) oder von Schneider (1973) dienen, die zeigen, dass in einer neutralen Situation durch eine unzureichende Anregung der Leistungsmotive oder Testängstlichkeit die Unterschiede nicht signifikant werden. Diese Erklärung ist jedoch fragwürdig, da Intelligenztests oft bereits als leistungsorientierte oder den Selbstwert ansprechende Situation verwendet werden (Alper, 1946; Atkinson & Reitman, 1956; Petty & Harrell, 1977; Sarason et al., 1952) und diese zu schlechteren Leistungen und bei Misserfolgsfürchtigen zu hohen Spannungen führen (Bartmann, 1963). Außerdem erklärt diese Vermutung die Ergebnisse der Studie nicht ausreichend, da sich, wie in den folgenden Kapitel 5.1.3.2 und 5.1.3.3 deutlich wird, auch in der Stresssituation keine signifikanten Unterschiede finden lassen.

5.1.3.2 Intelligenztestleistungen der Leistungsmotiv- und experimentellen Gruppen

Die Leistungsunterschiede zwischen HE und FM über die gesamte Stichprobe sind zwar gering und nicht signifikant und führen zur Ablehnung von Hypothese 6, dennoch ist eine gewisse Tendenz erkennbar. Erfolgshoffende geben im LPS-neu häufiger korrekte Antworten und bearbeiten mehr Aufgaben, wobei sich hier sogar ein geringer Effekt von $_{partial}r^2 = .04$ ermitteln lässt. Bezüglich der Subtests ist es der dritte und zehnte Subtest, Zeilenvergleichen, in denen die HE-Gruppe besser abschneidet.

In Hypothese 5 gilt es zu untersuchen, ob der Misserfolgsstress im Vergleich zur neutralen Bedingung die Testleistungen entsprechend den Studien von Thorndike und Woodyard (1934), Lantz (1945) oder anderen in Kapitel 1.3.4 erläuterten Studien reduziert. Zwar sind für diese Unterschiedsprüfungen ebenfalls nur drei Subtests signifikant, diesmal sind allerdings zwei davon, Zahlenfolgen (Subtest 4) und Flächen zählen (Subtest 7), hypothesenkonform. Der Trend der Mittelwerte und des dritten signifikanten Subtests, Buchstabenfolgen (Subtest 5), weist jedoch einen Vorteil der Experimental-

gruppe auf, die häufiger korrekt antwortet, weniger Fehler macht und mehr Aufgaben bearbeitet und berichtet. Für die Hypothese 5 kann folglich nicht gezeigt werden, dass die bei Misserfolg entstehende Frustration die Leistung hemmt, sondern es liegen auch gegenteilige Hinweise vor. Ein Grund für diese überraschenden Ergebnisse könnte das unterschiedliche Intelligenzlevel der beiden Gruppen sein. Gemäß der Homogenitätsprüfung der Subgruppen (siehe Kapitel 3.4.2) sind sowohl die Abiturnoten als auch die CFT 20-R Intelligenzquotienten der Experimentalgruppe signifikant besser als die der Kontrollgruppe. Die Vpn der Experimentalgruppe ($M = 121.17$, $SD = 12.34$) scheinen demnach generell intelligenter als die der Kontrollgruppe ($M = 110.22$, $SD = 12.95$) zu sein und daher ist es durchaus möglich, dass diese trotz Stress nicht unter der Leistung der Kontrollgruppe im LPS-neu liegen. Denkbar wäre auch, dass intelligentere Personen bessere Erfahrungen in Testsituationen machen und daher weniger ängstlich sind oder bessere Copingstrategien entwickelt haben. Diese Überlegungen sind allerdings dahingehend einzuschränken, da bereits vor den Berechnungen der Einfluss der Intelligenz mittels Kovarianzanalysen herauspartialisiert wurde. Es scheint, dass die Stressinstruktion eher Antrieb als Hemmung bedeutet, worauf in der zusammenfassenden Betrachtung der Intelligenztestleistung näher eingegangen wird.

Interessant ist ein signifikanter Wert, der sich bezüglich der Interaktion zwischen den beiden Gruppen in der Anagrammaufgabe (Subtest 2) ermitteln lässt. Die Gruppe mit FM ist in der Kontrollbedingung im Vergleich zur HE-Gruppe im Vorteil, das Gegenteil zeigt sich in der Stressbedingung. Anders als bei Schneider können beide Gruppen nach dem Misserfolg ihre Leistungen erhöhen, die Erfolgshoffenden scheinen sich durch einen stärkeren Anreiz und Anstrengung nur in größerem Maße zu verbessern.

5.1.3.3 Intelligenztestleistungen und Leistungsmotivation unter Stress

Ausgehend von Atkinsons (1964/1975) Erweiterung des Risiko-Wahl-Modells mit der Trägheitstendenz und den Befunden von Schneider (1973) sollte Misserfolg Erfolgsorientierte zu besseren Leistungen motivieren, während dies für Personen mit FM einen Motivationsverlust und somit eine Leistungsverschlechterung bedeutet. Der erwähnte Interaktionseffekt kann die Analysen von Schneider (1973) teilweise replizieren, in der vorliegenden Studie verbessern sich allerdings beide Motivgruppen nach Misserfolgstress. Für die separate Berechnung der Unterschiede in der Stressbedingung liefern die Gesamtleistung des LPS-neu und die Subtestvergleiche auch für diese, in Hypothese 7 formulierte Annahme, keine signifikanten Belege. Geringe Effekte verweisen jedoch tendenziell auf einen Vorteil der Vpn mit HE unter Stress im Vergleich zu Vpn mit FM; dazu kommt, dass Misserfolgssüchtige tendenziell signifikant weniger Aufgaben bearbeiten. Da Intelligenztests bereits stressauslösend für Personen mit FM sind,

sollte ein zusätzlicher Misserfolg, wie Studien von Raphelson (1957) oder Sarason et al. (1952) bestätigen, einer noch höheren Belastung entsprechen. Neben dem Misserfolg führen auch hohe Normwerte einer relevanten Gruppe, wie McKinney (1933) zeigen konnte, zu höherem Stress und schlechteren Leistungen. Der Unterschied zwischen den Leistungsmotivgruppen sollte daher hier dominierender sein, da eine zusätzliche Angstbelastung und ausgelöste aufgabenirrelevante Gedanken Misserfolgsfürchtige stärker beeinflussen (Heckhausen, 1982). Dazu kommt, dass Studien eine unrealistische Fähigkeitseinschätzung und Zielsetzung bei Vpn mit FM im Vergleich zu Vpn mit HE feststellen konnten (Mahone, 1960), was nach Ausführungen von Atkinson (1964/1975) bezüglich des Anspruchsniveaus zu unterschiedlichen Misserfolgseffekten führt: Wenn sich Personen mit HE realistischere Ziele und Misserfolgsfürchtige sich zu hohe Ziele setzten, ist anzunehmen, dass der gleiche niedrige Intelligenzquotient die Erfolgshoffenden weniger stresst und weniger in ihrer Leistung hemmt. Auch die selbstwertdienliche Attribution der Erfolgshoffenden nach Misserfolg auf mangelnde Anstrengung oder Pech, im Gegensatz zu der Misserfolgsattribution mit mangelnder Begabung von Personen mit FM, könnte ein Grund sein, warum Personen mit FM stärker von der Instruktion beeinflusst sind. Die Effekte dieser Arbeit spiegeln dies teilweise wider, da sich tendenziell stärkere Effekte in der Misserfolgsbedingung als in der Kontrollbedingung ergeben.

Entsprechend Hypothese 8 liegen die Gesamtmittelwerte des LPS-neu der Personen mit HE in der Stressbedingung höher als in der neutralen, wie dies bereits Schneider und Gallitz (1973) oder Stapf, Fischer und Degner (1986) zeigen konnten. Unter Stress machen diese mehr Fehler, berichtigen aber häufiger und bearbeiten mehr Aufgaben. Die höhere Leistung spricht nach McClelland et al. (1953) für einen größeren Leistungsanreiz für Vpn mit HE nach Misserfolg. Im Gegensatz zu Schneider und Gallitz (1973) erreicht keiner dieser Werte jedoch statistische Signifikanz, womit Hypothese 8 nicht bestätigt werden kann. Tendenziell signifikant sind drei der Subtests, wobei nur die Anagrammaufgabe (Subtest 2), jedoch nicht die Zahlen- und Flächenzählaufgabe (Subtest 4 und 7) der Hypothesenrichtung entspricht.

Wird den Annahmen intensivierter irrelevanter Gedanken von Vpn mit FM unter Stress von Heckhausen (1982) oder Sarason et al. (1952) nachgekommen, so sollten diese gehemmt unter Stress sein und in ihrer Leistung weiter absinken. Da Untersuchungen wie von Schneider und Gallitz (1973) empirische Belege dafür fanden, wurde für die Gruppe mit FM in Hypothese 9 diese Reaktion erwartet. Hier wurden die Ergebnisse nicht nur nicht signifikant, Personen mit FM sind entsprechend der Mittelwerte sogar unter Stress insgesamt und separat bezüglich der Subtests besser, machen weniger Fehler, bearbeiten mehr Aufgaben, lassen weniger aus und berichtigen häufiger.

Entsprechend Hypothese 9 kann die Aufgabe Flächenzählen tendenziell jedoch bestätigen, dass sich bei Personen mit FM durch Stress die Leistung verschlechtert. Trotzdem muss durch die erläuterten Ergebnisse die Hypothese 9 verworfen werden.

5.1.3.4 Intelligenztestleistungen und die Höhe des Leistungsmotivs

Weiterhin sind die Ergebnisse ebenfalls gemäß der Höhe des Leistungsmotivs analysiert worden. Aus früheren Untersuchungen (Atkinson, 1958; French & Thomas, 1958) besteht die Annahme, dass bei hoher Motivation die Leistungen besser im Vergleich zu niedriger Motivation ausfallen. Den Erwartungen der Hypothese 10 entspricht ein tendenziell signifikanter Wert des Subtests 9, in dem die Schnelligkeit und Genauigkeit eine wichtige Rolle spielen. Auch Lowell (1952) konnte zeigen, dass nur bestimmte Fähigkeitstests die Unterschiede zum Tragen bringen. Allerdings unterscheiden sie sich bei Lowell bezüglich der verbalen Intelligenz, während diese Fähigkeit in dem Subtest 9 der vorliegenden Arbeit nicht gefragt ist. Die restlichen nicht signifikanten Ergebnisse können diese Befunde und folglich die Hypothese 10 nicht unterstützen, da sich tendenziell in den Mittelwerten und für die Subtests sogar ein Vorteil der niedrigmotivierten Personen ergibt. Diese antworten häufiger korrekt, machen zwar mehr Fehler, korrigieren aber auch häufiger und erledigen mehr Aufgaben. Dies entkräftet die These von French und Thomas (1958), dass der energetisierende Effekt einer hohen Leistungsmotivation dazu führt, härter und effektiver zu arbeiten.

Nach einer Misserfolgsinstruktion wurden die Veränderungen der Intelligenztestleistungen der Hochmotivierten und der weniger Motivierten untersucht. Wie frühere Studien von McClelland et al. (1953) belegen, steigern hoch- und niedrigleistungsmotivierte Personen unter Misserfolgssituationen ihre Leistung. Atkinson (1964/1975) führt diese Tendenz auf eine anfänglich hohe Erfolgswahrscheinlichkeit zurück, welche mit Misserfolg sinkt und so in einen höheren Anreiz resultiert. Bekräftigt wird diese Aussage anhand der Mittelwerte, signifikant werden aber lediglich einige der Subtests. Für Personen mit hoher Leistungsmotivation ist nur der Subtest 7 des Flächenzählens signifikant, allerdings entgegen der Hypothese 11 sind Hochmotivierte hier unter neutraler Bedingung besser. Die Daten verweisen tendenziell jedoch darauf, dass diese Gruppe unter Stress mehr korrekte Antworten erzielt, aber auch mehr Fehler macht und mehr Items bearbeitet und berichtigt. Für die Niedrigmotivierten weisen die Mittelwerte ähnliche Tendenzen auf, obwohl diese zusätzlich auch weniger Fehler unter Stress machen. Diese Unterschiede sind allerdings nicht signifikant, daher kann Hypothese 12 ebenfalls nicht bestätigt werden. Nur der Wert des Subtests 5, Buchstabenfolgen, liegt im gewünschten Signifikanzbereich. Da sich hier jedoch höhere Ergebnisse für die

Personen der Kontrollgruppe mit niedriger GM ergeben, konnte Hypothese 12 widerlegt werden.

5.1.3.5 Resümee bezüglich der Intelligenztestleistungen und Leistungsmotivation

Zusammenfassend kann keiner der erwarteten Leistungsunterschiede konkret nachgewiesen werden. Bei dem Einfluss der Instruktion auf die Intelligenzleistung in Abhängigkeit des Leistungsmotivs scheint es sich um keinen einfachen, geradlinigen Zusammenhang, sondern vielmehr um einen komplexen Wechselwirkungsprozess zwischen Person und Situation zu handeln. Vielmehr scheinen für die Interpretation der Ergebnisse differenziertere Aussagen nötig.

Tendenziell ergibt sich für die Einteilung nach der Richtung des Leistungsmotivs in dieser Studie, dass Erfolgshoffende höhere Mittelwerte im Vergleich zu Misserfolgsmotivierten in allen Bedingungen erreichen, sich allerdings beide Gruppen tendenziell unter Stress verbessern. Bezüglich der Höhe des Leistungsmotivs sind es entgegen den Annahmen die Vpn mit geringer GM, die höhere Mittelwerte erreichten, wobei sich auch hier Vpn mit hoher und mit niedriger GM nach der Misserfolgsinstruktion in ihrer Leistung steigern können. Beim Vergleich der Durchführungssituation generell ist es die Experimentalgruppe, die tendenziell höhere Werte im Vergleich zur Kontrollgruppe erzielt. Auf der Basis der nicht signifikanten Ergebnisse ist es schwierig, Aussagen zu den postulierten Unterschieden zu treffen. Wie bereits angesprochen, liegt die Schlussfolgerung nahe, dass die durch die Misserfolgsinstruktion ausgelöste Motivation eher einen Antrieb als eine Hemmung für die Vpn bedeutete; oder in Atkinsons (1964/1975) Worten ausgedrückt, die Motivation eine Lokomotion des Verhaltens hervorrief. Nach McKinneys (1933) Auffassung und Belegen hat Stress zwar einen energetisierenden Effekt, dieser sollte jedoch in Testsituationen aufgrund einer Verringerung der intellektuellen Kontrolle nicht von Vorteil sein. Eine weitere Erklärungsmöglichkeit für den geringen Einfluss bzw. den motivierenden Effekt der Misserfolgsinstruktion wäre nach Hoppe (1930) die Tendenz von Personen, ihr Selbstbewusstsein auf hohem Niveau zu halten und Misserfolg eher zu verleugnen. Zwar geben etwa Dreiviertel der Vpn zu, sich gestresst gefühlt zu haben, allerdings fand die Befragung nach dem eingeschätzten Stresslevel erst nach der Aufklärung statt. So können Vpn im Testverlauf den hemmenden Einfluss des Misserfolgs verdrängen, da sie versuchen, ihren Selbstwert zu schützen. Nachdem im Debriefing offen feststand, dass der Wert unter ihrem Intelligenzquotient lag, konnten sie die Misserfolgsgedanken ohne jegliche Herabsetzung des Selbstwerts zulassen. Anhand des Debriefings kann die Vp auch einen Eindruck darüber gewonnen haben, welche Antworten der Versuchsleiter von ihr erwartet. Gemäß dem Versuchsleiter-Erwartungseffekt ist es möglich, dass einige Personen ange-

ben, sich gestresst zu fühlen, wobei der Misserfolg für sie im Testungsverlauf eher unbedeutend war bzw. sie ihn verdrängen konnten.

Bezüglich der Unterteilung in Erfolgs- und Misserfolgsmotivierte konnte eine Untersuchung von Schneider (1973) ebenfalls keine Unterschiede in einer neutralen Bedingung und bei einer Misserfolgsbedingung feststellen. Er postulierte in einem späteren Experiment, dass es wichtig sei, besonders die Güte der Leistung zu beachten: Nach einer Misserfolgsinduktion sind Erfolgsmotivierte schneller in der Bearbeitung und verbessern sich häufiger, Misserfolgsfürchtige verschlechtern sich, d. h. sie arbeiten langsamer und machen häufiger Fehler. Obwohl keine signifikanten Unterschiede bestehen, ist die Tatsache interessant, dass sich ähnliche Unterschiede wie bei Schneider (1973) in der Bearbeitungsart der Aufgaben niederschlagen. Die Erfolgshoffenden sind tendenziell erfolgreicher als die Misserfolgsfürchtigen, bearbeiten mehr Aufgaben, lassen zwar mehr Aufgaben aus und machen mehr Fehler, aber berichtigen diese auch häufiger. Für Misserfolgsfürchtige stellt sich hingegen heraus, dass sie zwar geringfügig weniger Fehler in beiden Bedingungen machen, aber weniger Aufgaben schaffen und die Leistung folglich insgesamt niedriger ausfällt. Dabei können, wie postuliert, höhere Effekte unter Stress zwischen beiden Gruppen mit einem Vorteil der HE-Personen verzeichnet werden. Die Interpretation liegt nahe, dass Misserfolgsfürchtige durch erhöhte Unsicherheit (Schulz & Schönplug, 1982) sorgfältiger arbeiten oder durch ablenkende Gedanken (Heckhausen, 1982) sich länger bei Aufgaben aufhalten. Erfolgshoffende neigen eher dazu, Aufgaben, die sie nicht lösen können, schneller auszulassen oder zu raten und weisen dadurch ein effektiveres Leistungsmuster auf. Misserfolgsfürchtige gehen insgesamt steifer in der Bearbeitung vor und konzentrieren sich scheinbar zu sehr darauf, einen Misserfolg zu vermeiden und dieses Fokussieren auf den Performanceprozess kann nachweislich die Leistung behindern (Baumeister, 1984). Separat betrachtet scheinen Personen mit HE dieses „Gütemuster“ in der Stressbedingung im Vergleich zur neutralen Situation noch zu steigern. Auch Misserfolgsfürchtige können unter Stress die Leistungsgüte erhöhen, nur dass diese weniger Fehler unter Stress begehen. Die Effekte deuten darauf hin, dass sich Erfolgshoffende jedoch stärker steigern können, was eher die Befunde von Stapf et al. (1986) repliziert, die zwar eine Leistungssteigerung bei Vpn mit HE, jedoch das Gleichbleiben der Leistung der FM-Gruppe nach Misserfolg voraussagen.

Bei der Betrachtung der Höhe des Leistungsmotivs zeichnet sich in der aktuellen Studie das deutliche Bild von Schneiders (1973) gefundenem starken Vorteil der Hochmotivierten, welchen French und Thomas (1958) durch eine stärkere Anstrengung und Lowell (1952) mit einem effektiveren Ausschöpfen des eigenen Leistungspotentials erklären, allerdings nicht ab. Der Trend der vorliegenden Daten offenbart sogar oft ei-

nen Vorteil der Niedrigmotivierten, die ein ähnliches Schema der Leistungsgüte wie die HE-Gruppe aufweisen. Auch der Vergleich zwischen den Durchführungsbedingungen separat für Niedrig- und Hochmotivierte ergibt dieses Muster, nur begehen Personen mit geringerer GM weniger Fehler. Beide erbringen unter Stress höhere Leistungen und weisen ähnliche Effektstärken der Verbesserung auf, was McClellands et al. (1953) Annahme eines angeregten Bedürfnisses nach Leistung durch Misserfolg bestätigen kann. Die deutliche Verbesserung von Personen mit hoher GM, wie es Schneider (1973) nachwies, kann in dieser Untersuchung nicht bestätigt werden. Die Tatsache, dass Hochmotivierte unter Stress mehr Fehler machen und generell schlechter als die weniger Motivierten sind, kann gemäß dem Yerkes-Dodsen Effekt (Yerkes & Dodsen, 1908) interpretiert werden. Das hohe Erregungsniveau durch Intelligenztests allgemein und Misserfolg speziell kann bei Hochmotivierten zu Übermotivation führen, da diese generell von einer charakteristisch hohen Motivation geprägt sind. Dies kann zwar zu schnellerem Arbeiten, aber auch zu mehr Fehlern und folglich schlechteren Leistungen führen. Zudem erwarten Hochleistungsmotivierte nachweislich höhere Ergebnisse (McClelland et al., 1953), was die intensivere Enttäuschung erklären würde. Das Erregungsniveau der Personen mit geringerer Leistungsmotivation ist möglicherweise in solchen Situationen optimaler, da Leistungssituationen und Misserfolg, wie Becker (1982) nachweisen konnte, in geringerem Ausmaß persönlich wichtig sind.

Allerdings kann insgesamt von keinen großen Unterschieden gesprochen werden, da diese relativ gering und die Inkonsistenz der Befunde auf eher zufällige als statistisch relevante Unterschiede schließen lassen. Zudem können die Ergebnisse des Befindens- und des Stressfragebogens auch nicht belegen, dass die Instruktion einen unterschiedlichen Einfluss auf beide Motivausprägungen hatte. Eine mögliche Erklärung der nicht signifikanten Unterschiede wäre, wenn im CFT 20-R vorher der Erfolg im Intelligenztest ein Ziel der Vpn mit HE war und sie sich anstrebten, um gemäß dem Selbstbegründungsmodell (Heckhausen, 1963, 1974) ihre eigene Tüchtigkeit möglichst hoch zu halten, sehen sie nach dem eingetretenem Misserfolg die vorherige Anstrengung als sinnlos. Sie senken ihr Anspruchsniveau. Hoppe (1930) bemerkt, dass Personen durch geringen „Mut zur Wirklichkeit“ (Hoppe, 1930, S. 41) auch zu unrealistischen Veränderungen neigen. Wird den Befunden von Mahone (1960) oder Atkinson und Litwin (1960) nachgekommen, so entsprechen unrealistischere Zielsetzungen und Fähigkeitseinschätzungen charakteristischen Zügen der Personen mit FM. Durch die Verdrängung des Misserfolgs und das Setzen eines zu hohen Anspruchsniveaus auch nach Misserfolg, könnte vielleicht bei den Misserfolgsfürchtigen noch ein „Hoffnungsfunkel“ bestehen. Diese Annahme muss allerdings schon laut den Befunden von Schleich (2007) verworfen werden, da sie zwar die postulierten Veränderungen des

Anspruchsniveaus nach Misserfolg nachweisen konnte, jedoch keinen Unterschied der Anspruchsniveausetzung nach Misserfolg zwischen den Leistungsmotiven fand.

Laut Feather (1965) sinkt nach Misserfolg die subjektive Wahrscheinlichkeit auf Erfolg. Für Erfolgsmotivierte führt dies nur bei leichten Aufgaben zu einer Anreizsteigerung, bei schweren Aufgaben, wozu auch die Intelligenztests dieser Untersuchung zählen können, nimmt der Anreiz jedoch ab (Atkinson, 1957). Betrifft weiterhin der Misserfolg eine *stabile* Persönlichkeitseigenschaft wie die Intelligenz, könnte die subjektive Wahrscheinlichkeit auf Erfolg äußerst niedrig ausfallen, was bei Vpn mit FM weniger Angst aufgrund eines geringeren Gefühls von Scham und Selbstverschulden auslösen kann (Atkinson, 1957). Beim Weiterführen dieses Gedankens könnten die Misserfolgsfürchtigen, obwohl sie in der aktuellen Studie gestresster sind, den niedrigen Intelligenzquotient als erleichternd, aufgrund einer Bestätigung ihrer eigenen niedrigen Fähigkeitserwartung (Kornprobst, 2005), empfinden. Möglicherweise agieren sie danach losgelöster von der eigenen Angst, da sie weiteren Misserfolg auf die Intelligenz assoziieren können. Ähnliches könnte bei überwiegend HE zutreffen, wenn ein Misserfolg nicht bei einer Tätigkeit auftritt, sondern das eigene Fähigkeitspotential betrifft. Ausführungen von Atkinson (1957) und Experimente von McClelland (1956) zur subjektiven Erfolgswahrscheinlichkeit oder zur Theorie der Informationsgewinnung (Schneider, 1973; Meyer, 1976) zeigen, dass selbst Personen mit HE zu Aufgaben mit einer Erfolgswahrscheinlichkeit von etwa .50 bzw. .40 tendieren. Bei zu niedrigen Erfolgschancen ist der Anreiz auf Stolz zu gering, um die leistungsfördernde Motivkomponente HE entscheidend anzuregen. Die Vp könnte dies als „an seine Grenzen stoßend“ interpretieren und jede weitere Anstrengung, da ein geringer Intelligenzwert stabiler ist, als sinnlos empfinden. Ein schlechterer Intelligenzquotient kann viel mehr als nur einen Misserfolg bei einer Problemlöse- oder Erinnerungsaufgabe bedeuten. Eine solche Information über die eigene Begabung kann ernster genommen werden, was eine sonst erfolgshoffende Person zu keinen weiteren Leistungssteigerungen ziehen kann.

Hinzu kommt der Einfluss des Drucks durch soziale Faktoren (Baumeister, 1984) wie eine Gruppentestung oder die sozialen Bezugsnormen (Heckhausen, 1974) der Instruktion durch die Verteilung der Intelligenzwerte der fiktiven Studie und die Zuordnung der Vpn zu einer Gruppe gemäß dem Intelligenzquotienten. Diese Faktoren könnten dazu führen, dass die Stresssituation nicht als rein leistungsorientiert wahrgenommen und zusätzlich das Anschlussmotiv angeregt wurde, was gemäß Befunden von Atkinson und Reitman (1956) mit einer Multianreizsituation die Unterschiede zwischen den Leistungsmotivgruppen aufheben kann. Vpn sind also in ihrer Leistung auch abhängig von der Ausprägung des Anschlussmotivs, d. h. wie stark ihr Bedürfnis nach Freundschaft, Kooperation oder nach Angliederung an diese Gruppe ist (Murray,

1938). So kann es dazu kommen, dass auch Personen mit FM oder niedriger Motivation, wenn die Ausprägung des Anschlussmotivs hoch ist, effizienter in einer Gruppe arbeiten (French, 1958). Außerdem kann potentieller Leistungsstress nur bei Vpn mit überwiegendem Leistungsmotiv im Vergleich zu überwiegendem Anschlussmotiv einen Einfluss nehmen (Vogel et al., 1959). Zwar sind beide Intelligenztests für Gruppentestungen prädestiniert, dennoch sollte eine Vergleichsstudie mit Einzeltestungen oder der Messung des Anschlussmotivs Aufschluss über solche Fehlerquellen geben.

Schlussfolgernd können durch diese unrealistische Verschiebung des Anspruchsniveaus, der geringen subjektiven Erfolgswahrscheinlichkeit nach der Mitteilung niedriger Intelligenzwerte und sozialer Faktoren die Unterschiede zwischen den Leistungsmotivkomponenten weniger stark ausfallen. Wenn lediglich die signifikanten Ergebnisse betrachtet werden, ergibt sich ebenfalls kein schlüssiges Muster bezüglich der Leistungsunterschiede. Für Lantz (1945) war klar, dass Stress die Aufgaben mit verbalen Fähigkeiten am stärksten hemmt und visuelle Fähigkeiten weniger betroffen sind. In dieser Arbeit ist die visuelle Wahrnehmung allerdings am stärksten betroffen, da Subtest 7 am häufigsten signifikant wird. Auch andere vereinzelt Tests sind in der aktuellen Untersuchung signifikant, wobei nicht eindeutig abgrenzbar ist, welche Aufgaben und somit Stratum I Faktoren nach Carroll (1993) durch die Leistungsmotive und Stresssituation beeinflusst werden. Zudem widerlegen einige signifikante Subtests die Ergebnisse in gegenteilige Richtung und ermitteln einen Vorteil für Personen mit FM.

Interessant ist das Faktum, dass Misserfolgsfürchtige in einigen Tests signifikant höhere Leistungen im Vergleich zu Personen mit HE oder zur Kontrollgruppe erzielen. Dies widerspricht der Annahme, dass diese in Leistungssituationen und unter Stress in ihrer Leistung gehemmt sind. Verschiedene Autoren gehen davon aus, dass Personen mit FM bei einfachen Aufgaben (Feather, 1965) und neutralen Situationen (Schneider & Gallitz, 1973) überlegen sind, andere hingegen meinen, dass diese stets unterlegen sein sollten (Atkinson, 1964/1973). Die vorliegende Studie kann nicht bestätigen, dass Vpn mit stärker ausgeprägtem Misserfolgsmotiv grundsätzlich schlechtere Leistungen aufweisen bzw. sogar weniger intelligent sind. Dadurch, dass diese teilweise auch höhere Leistungen erbringen, besteht ein Hinweis darauf, dass Misserfolgsfürchtige nicht weniger intelligent, wohl aber durch die stärker ausgeprägte Furchtkomponente häufiger negativ in Testsituationen beeinflusst sind. Dass Misserfolgsmotivierte einerseits im Vorteil und andererseits im Nachteil bei Tests sein können, erinnert an eine Untergliederung von FM durch Schmalt (1976) in FM1 als Faktor zur aktiven Misserfolgsvermeidung und FM2 als passive Misserfolgsschuld. Dieser Vorteil wird beispielsweise in einer Studie mit einfachen Aufgaben von Ruebush (1980) deutlich, in der testängstliche Vpn aktiv häufiger nachschauen und kontrollieren. Es gilt zu beachten, dass, wie in Kapi-

tel 1.1.7 beschrieben, ein Konzept von vier gesonderten Leistungsmotiven besteht und die Leistungsmotivausprägung einer jeden Person durch diese vier Motive ausgemacht wird. Um detaillierte Aussagen bezüglich der Leistungsunterschiede anhand der Leistungsmotive zu treffen, sollte daher versucht werden, auch den Einfluss der Komponenten „Furcht vor Erfolg“ (FE) und „Hoffnung auf Misserfolg“ (HM) und deren Wechselwirkung mit HE und FM in die Analysen mit einzubeziehen. Eine Kategorisierung in auffällige Ausprägungsmuster nach den vier Leistungsmotiven könnte eine Möglichkeit darstellen, um dies in experimentellen Studien umzusetzen.

Grundsätzlich kann festgestellt werden, dass bei der Betrachtung der Mittelwerte zwar postulierte Unterschiede zwischen den Leistungsmotiven – wie von Schneider (1973) nachgewiesen – entdeckt werden können, die meisten statistischen Überprüfungen in dieser Stichprobe diese aber nicht eindeutig unterstützen. Der Annahme von I. G. Sarason (1959), dass kognitive Leistungen am meisten durch die intellektuelle Kompetenz bestimmt werden, kann durch den geringen Einfluss des Leistungsmotivs und da auch der Einfluss der Intelligenz in der Kovarianzanalyse meist sehr signifikant ausfällt, zugestimmt werden. Intelligenztests scheinen wohl eine eigene Stellung in der Erforschung von Misserfolg auf die Leistung einzunehmen. Das Scheitern in einem Intelligenztest, wie oben ausgeführt, ist möglicherweise gravierender und die Auswirkungen auf einen weiteren Test sind sehr komplex. Selbst wenn die Misserfolgsinstruktion wenig signifikante Unterschiede zwischen den Leistungsmotivkomponenten und den Durchführungsbedingungen hervorbringt, weisen stärkere Unterschiedseffekte in der Stressbedingung sowie die Veränderungen im Wohlbefinden und die Angaben des subjektiven Stresslevels auf einen Einfluss der Testsituation hin. Damit unterstreicht die vorliegende Arbeit, dass die Testsituation und das Verhalten des Testleiters zu abweichenden Intelligenztestergebnissen führen können, was häufig außer Acht gelassen wird. Zu einer möglichst objektiven Durchführung von Intelligenztests, ist es daher notwendig, sich stets an die originale Instruktion zu halten und so eine maximale Standardisierung der Testsituation zu gewährleisten. Dass die erwarteten Effekte in dieser Arbeit nur tendenziell und nicht signifikant sind, könnte auch an den Eigenheiten der untersuchten Stichprobe oder den Durchführungsbedingungen liegen. Diese methodischen Probleme sollen im nächsten Kapitel detaillierter beleuchtet werden und Aufschluss über mögliche Ursachen der nicht signifikanten Ergebnisse geben.

5.2 Diskussion der Methode

Des Weiteren ist es wichtig, die verwendeten Messverfahren, die Besonderheiten der untersuchten Stichprobe und Probleme des Untersuchungsablaufs bei der Interpretation zu berücksichtigen; daher sollen diese im Folgenden kritisch beurteilt werden.

5.2.1 Untersuchungsmethode

5.2.1.1 Das Regensburger Leistungs-Motiv-Inventar

Bewährt hat sich in der Untersuchung die Verwendung der Onlineversion des Regensburger Leistungs-Motiv-Inventars für Erwachsene (RLMI-E). Durch dieses Verfahren kann enorm an Aufwand und Zeit sowohl für den Versuchsleiter als auch für die Teilnehmer gespart werden. Da die Testung ohnehin aus zwei Treffen á 45 bis 60 Minuten bestand, erhöhte dies die Wahrscheinlichkeit der Teilnahme der Vpn. Die Testgüte kann für diesen Test aufgrund der fehlenden Itemkennwerte durch die Onlinetestung nicht analysiert werden. Frühere Reliabilitätsberechnungen (Eder, 2006; Kornprobst, 2005; Peters, 2001) belegen jedoch stets die Güte des Verfahrens und die Interkorrelationsmuster der vier Gesamtskalen dieser Studie (siehe Kapitel 3.3.1.2).

5.2.1.2 Die Intelligenztests CFT 20-R und LPS-neu

Bezüglich des CFT 20-R wird ebenfalls aufgrund bisheriger ausreichender Überprüfung der Testgüte auf eine Item- und Skalenanalyse in dieser Arbeit verzichtet. Die Interkorrelationen (siehe Kapitel 3.3.2.2) machen deutlich, dass alle vier Subtests hoch mit dem Gesamtrawwert korrelieren und replizieren somit die Analyse von Weiß (2006).

Diese Untersuchung dient vor allem auch der Validierung des Intelligenztests LPS-neu. Die vorliegende Studie kann eine sehr gute Testgüte des LPS-neu unterstreichen, da die Reliabilitätsanalysen zufriedenstellende Werte für Cronbachs Alpha zwischen .71 und .93 und mittels der Testhalbierungsmethode zwischen .74 und .98 erzielen (siehe Kapitel 3.3.3.2). Die Subtests des LPS-neu können aufgrund dieser Ergebnisse als bewährt gelten. Die signifikanten mittleren bis starken Zusammenhänge der Subtests mit dem Gesamtscore zwischen .47 und .75 deuten daraufhin, dass Carrolls (1993) postulierter, alles determinierender Generalfaktor der Intelligenz zu existieren scheint, der die höchste Ebene in seinem Drei-Ebenen Modell bildet. Untereinander korrelieren die Tests geringer und einige Tests weisen gar keine bzw. nur tendenzielle Signifikanz auf, was die Existenz von verschiedenen Fähigkeitsfaktoren auf der untersten Stufe von Carrolls Modell bestätigt. Die jeweils auf der Stratum II Ebene nach Kreuzpointner (2010) zusammengehörigen Subtests Test 1 und 2 (kristalline Intelligenz), Test 3, 4 und 5 (fluide Intelligenz), Test 9, 10 und 11 (kognitive Geschwindigkeit) korrelieren stets signifikant miteinander. Ausnahme bildet Subtest 6, der gemeinsam mit Test 7 und 8 die Fähigkeit der visuellen Wahrnehmung beschreibt, jedoch mit beiden Tests nur tendenziell zusammenhängt. Dieser Subtest zeigt in der vorliegenden Arbeit auch keine statistisch bedeutsamen Unterschiede zwischen den Subgruppen

auf. Die meisten Subtests weisen allerdings auch ähnlich hohe Korrelationen außerhalb der Stratum II Faktoren mit anderen Subtests auf.

Da der CFT 20-R als Kontrollinstrument in dieser Studie dienen sollte, wurde die Konstruktvalidität beider Tests analysiert. Die Gesamtscores korrelieren mit .58 signifikant in einem mittleren bis starken Bereich miteinander, d. h. Personen, die im ersten Intelligenztest gut abschneiden, erbringen auch im LPS-neu insgesamt gute Leistungen. Um den alleinigen Einfluss der Leistungsmotive unabhängig von der Intelligenz zu untersuchen, musste die Intelligenz kontrolliert werden und daher konnte der CFT 20-R Intelligenzwert berechtigt als Kovariate in die Analysen mit aufgenommen werden. Dass die Intelligenztests nicht höher miteinander zusammenhängen, hängt allein schon mit dem Aufbau und dem zugrunde liegenden Intelligenzstrukturmodellen beider Tests zusammen. Während der CFT 20-R figurale Probleme enthält und im Sinne des g-Faktors interpretiert werden kann, sind die Aufgaben im LPS-neu von sprachlichen Fähigkeiten über numerische, visuelle bis zu Geschwindigkeitsfähigkeiten vielfältiger und damit ein Beispiel für die Existenz mehrerer zum Teil unabhängiger Intelligenzfaktoren. Die Dominanz liegt bei dem CFT 20-R auf dem Faktor der fluiden Intelligenz (siehe Kapitel 3.3.2.1), welchen die Subtests 3, 4 und 5 im LPS-neu erfassen (siehe Kapitel 3.3.3.1). Auffällig ist, dass nur Aufgabe 4 des LPS-neu teilweise signifikant mit dem CFT 20-R korreliert und so anscheinend nicht alle der drei Tests den gleichen postulierten Faktor der fluiden Intelligenz messen. Es sind eher die Subtests 2 (Anagrammaufgaben), 7 (Flächenzählen) und 9 (Zeichenzählen), welche in geringem Maß signifikante Ähnlichkeiten mit den Dimensionen des CFT 20-R aufweisen.

5.2.1.3 Die Einschätzungsfragebögen

Der für diese Studie entwickelte Wohlbefindensfragebogen erzielt hohe Werte bei der Berechnung der internen Konsistenz zwischen .81 und .89 und auch die Trennschärfen variieren zwischen .42 und .84 ebenfalls in einem zufriedenstellenden Bereich (siehe Kapitel 3.3.4). Daher wurde trotz inkonsistenter Faktorladungen nur ein allgemeiner Wohlbefindensfaktor gebildet. Wie bereits angesprochen, ist dies möglicherweise nicht differenziert genug und einige statistisch nicht bedeutsame Unterschiede könnten darin begründet sein. Es wäre zu überlegen, den Fragebogen anhand größerer Stichproben zu überprüfen und zu analysieren, ob sich eine bestimmte Unterteilung herauskristallisiert und die Testgüte dadurch gesteigert werden kann. Eine inhaltlich sinnvolle Unterteilung könnte die Abgrenzung der Items Entspannung, Unbekümmertheit, Gelassenheit, Vertrauen und Behaglichkeit zu den Items Interesse, Frische und Aufmerksamkeit darstellen. Diese Einteilung erinnert an zwei Komponenten von Testangst in Leistungssituationen: einerseits die emotionale Aufregtheit, d. h. die physiologische Erregtheit,

und andererseits die kognitive Besorgnis, d. h. die interferierenden Kognitionen (Morris & Liebert, 1970). Während Prüfungssituationen steigt die Aufgeregtheit an, während die Besorgniskomponente weitestgehend konstant bleibt, obwohl es gerade die letztere ist, die negative Korrelationen mit den nachfolgenden Leistungen aufweist. Diese Unterteilung könnte einen besseren Aufschluss und womöglich höhere Effekte bezüglich des Befindens der Leistungsmotivgruppen in Intelligenztests und unter Stress geben. Denkbar wäre, vorab direkt Items hinzuzunehmen, welche inhaltlich stärker einer der Angstkomponenten zugeordnet werden können. Danach ist es wichtig, Korrelationsanalysen mit anderen Befindensfragebögen wie der Beanspruchungsmessskala (BMS) von Plath und Richter (1984) durchzuführen, um zu überprüfen, wie konstruktvalide der Fragebogen zum aktuellen Wohlbefinden ist.

Eine Zusammenhangsanalyse fand bereits innerhalb dieser Studie mit dem ebenfalls selbst entwickelten Fragebogen zur subjektiven Beurteilung der Stressreaktion statt. Beide Fragebögen korrelieren im mittleren bis hohem Bereich signifikant positiv miteinander und scheinen daher ähnliche Konstrukte zu erfassen. Dass die Fragebögen nicht das Gleiche ermitteln, wird deutlich, da sich Vpn mit Erfolgshoffnung und Misserfolgsschmerz im gefühlten Stresslevel, aber nicht bezüglich des Wohlbefindens unterscheiden. Die Items des Fragebogens laden durchgehend auf einem Faktor und die interne Konsistenz von .87 sowie die Trennschärfen ($r_{it} = .73 - .80$) deuten auf eine gute Inhaltsvalidität hin (siehe Kapitel 3.3.6). Es ist jedoch offensichtlich, dass bei einer Auswahl von drei Items die Möglichkeit eines einheitlichen Faktors im Gegensatz zu dem Befindlichkeitsfragebogen höher ist. Um den Stressreaktionsfragebogen für die Anwendung in verschiedenen Leistungssituationen mit Feedback einzusetzen, sollte allerdings eine Erweiterung der Itemliste in Betracht gezogen und diese in einer erneuten Stichprobe überprüft werden. Zusätzlich erscheint eine Verallgemeinerung der zu bewertenden Aussagen sinnvoll, um ihn außerhalb der Intelligenzforschung in Studien und möglicherweise auch in der Praxis anzuwenden. Wie belastend Stress beispielsweise durch eine schlechte Schulnote auf eine Person wirkt und welche Auswirkungen dies auf weitere Leistungen haben kann, wäre eine praktische Anwendungsmöglichkeit. Nach hinreichender Validitätsprüfung besteht durchaus die Möglichkeit, die beiden Fragebögen gemeinsam oder getrennt im genannten pädagogischen Bereich oder im Bezug auf Arbeitsbelastungssituationen einzusetzen.

Was den neu entwickelten Bewertungsfragebogen betrifft, erzielten Item- und Skalenanalyse gute Werte, diese fallen jedoch teilweise grenzwertig aus (siehe Kapitel 3.3.5). Einzelne Trennschärfen liegen in einem nicht befriedigenden Bereich, woraufhin sogar auf ein Item mit der niedrigsten Trennschärfe „Schwierigkeit ist zu niedrig“ verzichtet wurde. Dadurch verbesserten sich zwar die Trennschärfen der restlichen sieben

Items, trotzdem liegen nach wie vor drei Items unter $<.30$ und die restlichen variieren zwischen $.33$ und $.58$. Der Test ist daher weniger reliabel als die obigen Fragebögen, liegt aber immer noch mit $.57$ und $.67$ in einem annehmbaren Testgütebereich. Weiterhin sind die Faktorladungen inkonsistent, laden einerseits bei den zwei Fragebögen sehr unterschiedlich und erklären nur ein Drittel der Varianz. Die einzelnen Items sind daher schwer in eine Struktur einzuordnen, da sie scheinbar zu unterschiedliche inhaltliche Bereiche der Bewertung ansprechen. Nicht signifikante Unterschiede zwischen Experimental- und Kontrollgruppe sowie zwischen den Leistungsmotivgruppen könnten in diesen Schwächen des Evaluationsfragebogens begründet sein, da sie inhaltlich, wie oben beschrieben, durchaus nachvollziehbar sind. Außerdem bestehen auch signifikant mittlere bis hohe Zusammenhänge, wie in Kapitel 5.1.1 erörtert, zu den zwei weiteren Fragebögen zum Wohlbefinden und dem Stresslevel zwischen $.31$ und $.68$. Es ist denkbar, auch wenn nicht klar beurteilt werden kann, ob es sich hier um Kausalzusammenhänge handelt, dass auch die Bewertungen der Gruppen gerade nach der Stressinstruktion nicht unbeeinflusst blieben. Dies wurde möglicherweise nur in dieser Datenlage nicht statistisch bedeutsam. Daher sollte eine Überarbeitung dieser Messmethode durchgeführt werden, da sie als Kriterium der tatsächlichen Bewertung nicht hinreichend zuverlässig scheint. Insgesamt ist die interne Konsistenz des Tests jedoch annehmbar und somit wäre vor der Entwicklung neuer Items eine Umformulierung und Verfeinerung der bestehenden Items sinnvoll. Zusätzlich sollte wohl auch die Itemmenge erweitert und durch weitere Faktoranalysen die Struktur des Tests untersucht werden, um ein Verfahren für die Bewertung von Intelligenztests von unterschiedlichen Gruppen und in unterschiedlichen Situationen zu erhalten.

Die drei Fragebögen sind subjektive Selbstberichte und daher trotz Betonung der Anonymisierung durch die Tendenz der sozialen Erwünschtheit möglicherweise verzerrt. So könnte sich beispielsweise ein Teilnehmer nicht eingestehen wollen, dass ihn das fiktive Ergebnis beunruhigt hat, aus Angst sich zu offenbaren, die Situation nicht durchschaut zu haben oder die persönliche Wichtigkeit der eigenen Intelligenz vor sich selbst bzw. Bekannten zugeben zu müssen. Ein Anderer hingegen gab möglicherweise, wie bereits erwähnt, gemäß dem Versuchsleiter-Erwartungseffekt an, sich gestresst zu fühlen.

5.2.2 Stichprobe

In diesem Abschnitt werden die Ergebnisse zusätzlich bezüglich der Eigenheiten der untersuchten Stichprobe, die in dieser Arbeit fast nur aus Studenten bestand, erörtert. Atkinson und Litwin (1960) formulieren bezüglich Studenten, dass es unwahrscheinlich ist, dass FM bei diesen höher als HE ausgeprägt ist. Diese Vermutung scheint nach

der Betrachtung der Leistungsmotivverteilung dieser Stichprobe zutreffend, jedoch etwas pauschal gefasst. Vielmehr fallen HE sowie FM beide relativ hoch aus. Weiterhin verteilen sich Personen mit überwiegend HE und welche mit überwiegend FM in der Stichprobe relativ homogen, wobei die Studie tendenziell sogar mehr Personen mit überwiegend FM enthält. Für dieses für Studenten untypische Muster ist eine Erklärung nahe liegend. Die Studie enthält prozentual mehr Frauen (68.1 %) und da bei Frauen in dieser Arbeit (siehe Kapitel 3.4.2) sowie in anderen Studien (Kornprobst, 2005; Peters, 2001) FM stärker im Vergleich zu HE ausgeprägt ist (HE = 32.7 %, FE = 67.3 %), sind sie in der FM-Gruppe gemäß der Homogenitätsprüfung überrepräsentiert. Es wäre daher interessant zu überprüfen, ob Frauen und Männer mit deren typischen Leistungsmotivmustern sich in ihren Reaktionen unter Stress unterscheiden. Aufgrund der Größe der Stichprobe und der geringen Anzahl von Männern musste auf den Vergleich in dieser Arbeit verzichtet werden.

Die Verteilung der NH generell zeigt weiterhin, dass die durchschnittliche Differenz zwischen HE und FM nicht sehr groß ausfiel und die untersuchte Stichprobe keine Extremgruppen mit hoher und niedriger Richtung der Leistungsmotivausprägung enthielt. Da selten eine Motivkomponente überwiegt und beide relativ stark ausgeprägt sind, sollte der Konflikt zwischen den Motiven Erfolg aufzusuchen und Misserfolg zu meiden relativ stark sein (Atkinson, 1964/1975), was, wie oben bereits angesprochen, den Effekt der Motive ausgleichen und nicht signifikante Unterschiede erklären kann. Heckhausen (1963) meint, dass dieser Konflikt durch höhere Zielsetzungen aufgelöst werden kann, was die Häufigkeit dieses Motivmusters unter Hochschulabsolventen erklären würde. Demnach befanden sich durch die hohe Ausprägung beider Motivkomponenten bezüglich der GM wenig Niedrigmotivierte unter den Studenten, was die besseren Leistungen der Personen mit geringer GM verständlicher erscheinen lässt. Ob es sich hier um eine charakteristische Leistungsmotivausprägung für Studenten handelt oder Besonderheiten der untersuchten Stichprobe sind, bleibt dabei ungeklärt. Gleichzeitig wurden die Gruppen anhand des Medians unterschieden und nicht explizit die Extremgruppen untersucht. Auch Lighthall et al. (1959) führten ihre nicht signifikanten Leistungsunterschiede von testängstlichen Personen zwischen einem stressigen und einem spielähnlichen Test auf die Mediantrennung zurück, da Zweibelson (1956) im gleichen Experiment anhand von Extremgruppen einen signifikanten Unterschied fand.

Da die meisten der Studenten sich bereits in höheren Semestern befanden, durchliefen sie bis dahin eine Vielzahl an Testsituationen und Auswahl- sowie Bewertungsverfahren von Grundschule über Gymnasium bis hin zu zulassungsbeschränkten Studiengängen und verschiedensten Hochschulprüfungen. Es kann vermutet werden, dass eine solche „Selektion“ der Grundgesamtheit ein charakteristisches Muster der

Leistungsmotive bevorzugt. Extreme Leistungsmotivgruppen wie starke FM und geringer HE neigen einerseits eher dazu, Leistungssituation zu vermeiden bzw. sehen in diesen keinen so hohen Anreiz und Scheitern womöglich aufgrund ihrer Furcht und irrelevanten Gedanken häufiger. Ebenso ist denkbar, dass Niedrigermotivierte in einem lernintensiven und selbstständigen Studium wenig Anreiz finden oder dem Einsatz und der Zielstrebigkeit Hochmotivierter (French & Thomas, 1958) unterliegen.

Weiterhin kann die häufige Konfrontation mit Testsituationen auch bedeuten, dass sie eine Art Vertrautheit entwickelten und lernten, Testängstlichkeit bzw. FM zu kompensieren und so ihre Testresultate steigern konnten (Petty & Harrell, 1977; Sarason & Mandler, 1952). Zwar gehen auch Schulz und Schönplug (1982) von besser erlernten Kontrollmechanismen ängstlicher Personen in Leistungssituationen aus, meinen allerdings, dass diese bei schweren Aufgaben überstiegen werden. Für Studenten könnte ein anderes Bild gemutmaßt werden: Während ein Intelligenztest normalerweise eine Gefahr für das Selbstkonzept darstellt, sehen Studenten sie vielleicht gemäß Lazarus und Launier (1978) als Herausforderung. Da es sich in der Stichprobe zu 64.4 % um Psychologiestudenten handelt, kann von einer noch größeren Ähnlichkeit der Stichprobe ausgegangen werden. Außerdem werden während des Psychologiestudiums Intelligenztests und ihre Schwächen behandelt, woraus aufgrund theoretischen Wissens bzw. Erfahrungen ein stärkeres Testwissen und Vertrautheit anzunehmen ist. Dies kann auch bei der Attribuierung von Misserfolg vorteilhaft sein (Weiner, 1972/1976b), da es bei Hintergrundwissen wahrscheinlich ist, auch externe Ursachen wie die Tagesform oder spezielle Testarten anstatt nur die schlechte Begabung bei der Ergebnisbewertung zu berücksichtigen. Schachter und Singer (1962) belegen beispielsweise, dass kognitive Erklärungen Emotionen kompensieren können. Womöglich führen diese Gründe nicht zu weniger Stresseemotionen, aber zu besseren Bewältigungsstrategien des Stresslevels, wie die Regulation dieser Emotionen (Lazarus & Launier, 1978).

Bezüglich der Intelligenz liefert die Homogenitätsprüfung der Subgruppen (siehe Kapitel 3.4.2), wie bereits erörtert, weitere wichtige Hinweise auf eine mögliche Fehlerquelle bezüglich der Unterschiede der Fähigkeiten zwischen Kontroll- und Experimentalgruppe. Zwar ist die Verteilung der Leistungsmotive sowie des Geschlechts in diesen Subgruppen annähernd gleich, jedoch scheint die Experimentalgruppe im Vergleich zur Kontrollgruppe gemäß den Abiturnoten und CFT 20-R Intelligenzquotienten generell intelligenter. Damit durch die ungleichmäßige Verteilung der Intelligenz die Auswirkungen der Leistungsmotive in der statistischen Analyse nicht ausgeschaltet werden, wurden, wie oben beschrieben, die Messergebnisse vor den Berechnungen von dem Einfluss der Intelligenz mittels Kovarianzanalysen bereinigt. In dieser Stichprobe ist die Verteilung der Intelligenzquotienten von Studenten mit einem Mittelwert von 115.69

($SD = 13.72$) relativ hoch. Generell differieren Studenten in ihrer Höhe der Intelligenz geringfügiger und durch die Homogenität der untersuchten Personen ist eine Generalisierung der Ergebnisse fragwürdig. Tatsächlich zeigen verschiedene Studien, dass es gerade auf die Verteilung der Intelligenzwerte ankommt: Bei Gruppen mit hoher Intelligenz nimmt Misserfolg einen stärkeren Einfluss (Sullivan, 1927) und Heckhausen (1963) meint, dass sich nur bei Hochintelligenten das Leistungsmotiv positiv auf die Leistung auswirken kann. Insgesamt scheint es in der aktuellen Studie so zu sein, dass sich die Unterschiede zwischen den Misserfolgsfürchtigen und Erfolgshoffenden zwar andeuten, aber durch die ähnlichen Ausprägungen der Leistungsmotive und der Intelligenz nicht so stark zum Tragen kommen, dass sie signifikant werden.

Die teilweise Vernachlässigung der Untersuchung von Extremgruppen, des Geschlechts und der Homogenität einer studentischen Stichprobe könnten Gründe für die statistisch nicht signifikanten Ergebnisse sein. Wie bereits erwähnt, ist die Stichprobe dieser Studie allerdings zu klein, um die Gruppen weiterhin nach Geschlecht oder Extremgruppen differenziert zu analysieren. Zwar sind die Stichprobengrößen der Hauptgruppen ausreichend, gerade bei den separaten Analysen für beispielsweise HE-Personen, zwischen Stress- und Neutralsituation sind sie allerdings bereits relativ klein, wodurch die statistische Power verringert ist. Es sollte in Erwägung gezogen werden, in nachfolgenden Experimenten größere Stichproben zu untersuchen, um die Effekte der genannten Aspekte explizit in der statistischen Analyse zu beachten.

5.2.3 Untersuchungsablauf

Trotz sorgfältigem Vorgehen bei der Untersuchungsdurchführung sind nicht alle methodischen Schwächen vermeidbar. Durch die Freiwilligkeit der Studenten ist denkbar, dass es sich um einen bestimmten Personenkreis handelt, und zwar laut Atkinson (1975) konkret um Hochleistungsmotivierte, die sich in ein Experiment eintragen, indem Intelligenz gemessen wird. Eine Studentin, die sich scheinbar blind in die Liste eintrug, nahm nicht an der Studie teil, da sie Angst vor Intelligenztests und deren Ergebnissen hatte. Somit könnte die einseitige Verteilung der Leistungsmotive auch darin begründet sein, dass es durchaus stark misserfolgsfürchtige Studenten gibt, diese aber ein solches Leistungsexperiment meiden würden, da dies eine zusätzliche Belastung für sie darstellt. Das spricht aber dafür, dass für diejenigen, die sich eintragen, ein Intelligenztest positive Valenzen hat, da sie dafür Interesse zeigten. Lantz (1945) betont die Wichtigkeit einer hohen Valenz gegenüber der Testsituation und dem folgenden Resultat, damit Misserfolg einen relevanten Einfluss auf die Vpn haben kann. Der Prozess der Motivation kann laut Heckhausen (1980) erst zustande kommen, wenn ein Ziel oder eine Situation bedeutungsvoll für die Vpn ist. Intelligenztests dürften bei Studen-

ten generell einen hohen Stellenwert einnehmen, da deren Intelligenz wichtig für ihren Studienerfolg ist und Intelligenztests häufig in Jobauswahlverfahren zum Einsatz kommen. Nach der Meinung von Baumeister und Showers (1986) gehen vor allem Studenten bei sich von einem hohen Intelligenzniveau aus und stehen dadurch bei Intelligenztests stärker unter Druck. Wenn von einer solchen persönlichen Relevanz ausgegangen wird, sollte auch das Leistungsmotiv dieser Personen durch einen Misserfolg genügend angeregt werden (Schneider, 1973). Gestützt wird diese Annahme, da Dreiviertel der untersuchten Personen sich nach Misserfolg gestresst fühlen. Dabei sollte aber nicht außer Acht gelassen werden, dass verschiedene Studienrichtungen in dieser Stichprobe enthalten sind. Für Personen der technischen Studienrichtungen könnte ein Intelligenztest einen geringeren Stellenwert als ihre technische, praktische Begabung einnehmen. Hinzu kommt, dass über die Hälfte der Vpn Psychologiestudenten sind, welche Versuchspersonenstunden sammeln müssen. Diese nahmen folglich nicht gänzlich freiwillig teil, wodurch die Relevanz der experimentellen Studie ebenfalls reduziert sein könnte.

Bezüglich der Valenzen der Testsituationen soll im Rahmen des Untersuchungsablaufs ein weiterer Kritikpunkt betrachtet werden. Einige Vpn fragten erst nach den Testungen, ob sie die Ergebnisse der Intelligenztests mitgeteilt bekommen könnten, obgleich sie bereits in den ersten E-Mails darüber informiert wurden, dass sie über die Resultate informiert werden. Das Nichtwissen, ob die Testung einen individuellen Informationswert zur Begabungseinschätzung enthalten könnte, würde gemäß der Theorie der Informationsmaximierung zu geringer Motivanregung führen (Meyer, 1976). Möglicherweise verstärkte die Betonung der anonymen Behandlung der Daten diesen Effekt und senkte die persönliche Relevanz der Studienergebnisse. Dies könnte für Erfolgsmotivierte weniger Anreiz und gleichzeitig weniger Furcht für Misserfolgsmotivierte bedeuten, wodurch die erwarteten Unterschiede der Leistungsmotive ausblieben. Generell wurde nicht explizit nach dem Anreiz bzw. der Bedeutung eines Intelligenztests oder des eigenen Intelligenzquotienten gefragt und so können nur Vermutungen diesbezüglich angestellt werden.

Die meisten vorgestellten Misserfolgstudien messen die Leistung nach der Manipulation direkt anhand *einer* Aufgabe (Atkinson, 1958; Bartmann, 1963; McClelland, 1953; Sarason & Mandler, 1952), während der LPS-neu sich jedoch in 11 Tests gliedert und etwa 60 Minuten dauert. Die individuelle Wirkung des Leistungsmotivs oder des Misserfolgs kann daher im Verlauf der Testung variieren. Hochmotivierte können nach Lowell (1952) ihr Potential über die Zeit besser ausschöpfen und ihre Leistung steigern, während Niedrigmotivierte eher konstant bleiben. Nach den Befunden von French und Thomas (1958) beschäftigen sich Hochmotivierte bei einem längeren Test-

verlauf ausdauernder mit den Aufgaben, die Lösung fanden beide Motivgruppen jedoch in ähnlicher Zeit. Doch auch diese Aufgaben dauern nur 10 bis 20 Minuten.

Bezüglich der Misserfolgsinstruktion können unterschiedliche Vermutungen der Art und Richtung der Veränderung des Einflusses über den Zeitverlauf der Testung aufgestellt werden. Baumeister bemerkt (1984, S. 617) eine Art „giving up“-Reaktion bei seinen Vpn, wenn ein weiterer Misserfolg im Testverlauf folgt (beispielsweise in der vorliegenden Studie das Nichtvollenden eines Subtests), welches das experimentelle Setting vereitelt und zu besseren Leistungen führt. Gleichzeitig kann aber auch aufgrund von Kommentaren der Vpn während der Aufgabenbearbeitung ein gegenteiliger Effekt eines weiteren Misserfolgs gemutmaßt werden. Durch die erhöhte Schwierigkeit des LPS-neu im Vergleich zum CFT 20-R, dass Aufgaben seltener vollendet werden konnten oder Items häufiger übersprungen werden, scheint sich laut der verbalen Äußerungen der Vpn die Misserfolgsinstruktion zu intensivieren. Schafft die Vp dann wiederum eine Aufgabe vollständig – ist somit erfolgreich, könnte das den erzeugten Stress der Instruktion aufheben und die Leistungen fördern, was ebenfalls von McClelland und Apicella (1947) und auch von Zeller (1950) nachgewiesen wurde. Bereits Hoppe (1930) belegt, dass das Erleben von Erfolg und Misserfolg mit vorherigen Leistungserfahrungen gemäß der Anspruchsniveausetzung variiert. Durch das erfolgreiche Beenden einer Aufgabe vor der gegebenen Zeit könnte folglich auch „Erfolg“ die Stressinstruktion vereiteln und gemäß Sears (1937) zu weiteren Leistungsverbesserungen führen. Mit diesen Annahmen schließt die aktuelle Arbeit an die Meinung von Feather (1965) an, dass die Motivation nach einer Misserfolgsinstruktion mit den Erfolgen und Misserfolgen während der Aufgabenbearbeitung weiter variieren kann. Wie genau sich im weiteren Testverlauf der Einfluss der Stressinstruktion erhöht oder verringert, ist sehr komplex und individuell abhängig und sollte in weiteren Analysen überprüft werden. In dieser Untersuchung war es jedoch aufgrund der Validierung des LPS-neu wichtig, den Test nach der Stressinstruktion gemäß der Originalinstruktion durchzuführen und Messungen zum Einfluss von weiterem Erfolg und Misserfolg hätten unnötige Unterbrechungen bedeutet.

Eine wichtige Fehlerquelle bei langen Testungen im Untersuchungsverlauf von Leistungstests sind vor allem auch Müdigkeit und Aufmerksamkeit, die die Ergebnisse beeinflussen können. Während des Experimentverlaufs können Müdigkeitserscheinungen auftreten, obwohl die Vpn vielleicht zum Zeitpunkt der Wohlbefindensmessung, d. h. vor und nach den Tests, angibt, aufmerksam zu sein und sich frisch zu fühlen. Da die Testungen an unterschiedlichen Tagen stattfanden und auch aufgrund der Raumkoordination abweichende Tageszeiten nicht vermieden werden konnten, muss die unterschiedliche Erschöpfung des Alltags als weitere Einflussquelle auf die Intelligenz-

leistung berücksichtigt werden. Um die Unterschiede zwischen den Leistungsmotiven unter Stress deutlich zu machen, wird daher die Benutzung einer kürzeren Testvariante vorgeschlagen, die aufgrund der in dieser Arbeit zu überprüfenden neuen LPS Version nicht benutzt werden konnte.

Neben der Aufgabe, mit der die Leistungsergebnisse gemessen werden, scheint weiterhin, wie bereits erwähnt, nicht unbedeutend, in welcher Aufgabe der Teilnehmer seinen Misserfolg erfährt. Hoppe (1930), Lantz (1945) sowie Postman und Bruner (1948) gehen von den Misserfolgen der momentanen Aufgaben aus. Es scheint, dass nicht der Misserfolg eines vorherigen Tests, sondern der in der momentan getätigten Aufgabe ausschlaggebend für diese selbstschützenden Abwertungen einer Aufgabe ist. Weiterhin sprach sich beispielsweise Lantz (1945) dafür aus, dass die Misserfolgssituation, wenn danach die Veränderung bei einem Intelligenztest geprüft wird, einem Intelligenztest nicht ähnlich sein sollte. Es ist durchaus möglich, dass daher unterschiedliche Resultate im Vergleich zu den Studien mit Misserfolg (siehe Kapitel 1.3.4) zustande kommen und Werte nicht im gewünschten Signifikanzbereich liegen. Widersprüchlich dazu sind die Ausführungen zur Trägheitstendenz, in denen behauptet wurde, dass der Effekt von Misserfolg nur bei verwandten Tätigkeiten auftritt (Schneider, 1973). Ziel dieser Arbeit war es außerdem, Stress auszulösen; der Misserfolg wurde dabei vorrangig als Mittel zum Zweck verwendet. Daher scheint die Forderung von Lantz für die vorliegende Untersuchung zu weit gegriffen. Vorstellbar wäre aber in einer Replikation, die Auswirkungen von einer dazwischen geschalteten Stress- bzw. Misserfolgsaufgabe zu überprüfen.

Die eigentliche Schwierigkeit dieser Untersuchung lag vor allem in der Operationalisierung einer geeigneten Stresssituation und der Glaubhaftigkeit dieser für die Vpn. Viele Befunde belegen, dass Misserfolg als Stresssituation wirkungsvoll ist. Das Erleben des Misserfolgs ist allerdings nach Hoppe (1930) von dem Anspruchsniveau, der individuellen Zielsetzung einer Person an die zukünftige eigene Leistung, abhängig. Über das Anspruchsniveau von Studenten kann nur gemutmaßt werden, dass es relativ hoch liegt. Inwiefern die Personen bereits Intelligenztests gemacht haben und sich so ein individuelles Ziel für diesen Test bildeten bzw. wie hoch oder niedrig dieses war, wurde in den Daten dieser Untersuchung nicht festgehalten. Da wenige Effekte der Misserfolgsinstruktion zu verzeichnen sind, ist es möglich, dass diese durch ein niedriges Anspruchsniveau oder noch keine Vorerfahrungen nicht als Misserfolg erlebt wurde. Auf der anderen Seite könnten Studenten eine starke Einschätzung des eigenen Intelligenzniveaus aufgrund der Vielzahl an Testerfahrungen haben und so den unter ihrem eigenen Niveau liegenden Wert anzweifeln und folglich die Glaubhaftigkeit des Experiments in Frage stellen. Es konnte aber mittels des Stress- und Wohlbefin-

densfragebogens und auch einigen Kommentaren der Teilnehmer ein Einfluss der Misserfolgssituation nachgewiesen werden. Inwiefern es sich dabei tatsächlich um Stress im Sinne von Lazarus und Launier (1978) handelt, kann nicht eindeutig bestätigt werden. Zudem kommen bereits Lazarus et al. (1952) zu dem Schluss, dass Stress vielfältige Facetten hat, d. h. die unterschiedlichen experimentellen Umsetzungen einer Stresssituation (siehe Kapitel 1.3.3 und 1.3.4) auch verschiedene Stressarten reflektieren. Daher ist es nicht verwunderlich, dass diese Arbeit und die vorgestellten Studien aufgrund unterschiedlicher Settings, Aufgaben und zudem verschiedener Instrumente zur Messung der Leistungsmotive oder Testängstlichkeit abweichende Befunde hervorbringen. Zu beachten ist daher, dass mit dieser Untersuchung kein Anspruch einer Generalisierbarkeit auf Stress allgemein erhoben werden soll und kann. Das verwendete Setting sollte daher nur speziell als Misserfolgsstress in Leistungssituationen interpretiert werden. Spannend wäre dabei zu beobachten, ob sich bezüglich der Leistungsunterschiede zwischen den Leistungsmotivgruppen bzw. Durchführungssituationen stärkere Effekte bzw. andere Richtungen der Effekte durch verschiedene Stressformen (z. B. Zeitdruck, physikalische Bedrohung usw.) ergeben würden.

Nicht möglich ist es in der untersuchten Stichprobe die Veränderung der Intelligenzleistung zwischen den zwei Messzeitpunkten miteinander zu vergleichen, da beide Intelligenztests sich zu sehr unterscheiden und für den LPS-neu noch keine Normierung erfolgte. Natürlich sollte aufgrund der Lernerfahrung nicht ein Intelligenztest zweimal verwendet werden, jedoch wäre ein Vorschlag, zwei Testunterformen vor und nach dem experimentellen Setting zu verwenden, wie es beispielsweise Lantz (1945) durchführte. Zwei parallele Testversionen des LPS-neu stehen bereits in der Entwicklungsphase und könnten demnach für folgende Experimente genutzt werden.

Bezüglich der Testdurchführung lässt sich festhalten, dass neben der Operationalisierung der Stresssituation die Wirkung der Instruktion von der Bedeutung von Intelligenztests für die Vpn und von der Aufgabe, in der die Person den Misserfolg erfährt sowie von der Art und Länge der darauf folgenden Aufgabe, in der weitere Erfolge und Misserfolge den Einfluss der Misserfolgsinstruktion modellieren, abhängig ist.

5.2.4 Ausblick

Insgesamt lässt sich zwar ein Nachweis des Einflusses auf Wohlbefinden und gefühltes Stresslevel der Vpn allgemein erbringen, Auswirkungen des Leistungsmotivs und weiterhin auf die Leistungen zeichnen sich allerdings nur in geringem Maß ab. Die ausführliche Diskussion verdeutlicht, dass sich zahlreiche Punkte ergeben, an denen weiterführende Untersuchungen aufbauend auf den Ergebnisse dieser Studie anknüpfen

können. Darüber hinaus kristallisierten sich auch Ideen für inhaltlich wichtige Themen, welche nicht unmittelbar dem Inhalt dieser Studie entsprechen, heraus.

Es wurden bereits die Leistungsmotivstruktur von Studenten und deren mögliche Besonderheiten, wie das Fehlen von Extremgruppen oder Niedrigmotivierten, im Vergleich zur Gesamtpopulation erörtert. Weitere Untersuchungen könnten klären, ob es sich dabei um eine Zufälligkeit der Stichprobe dieser Arbeit oder tatsächlich um eine charakteristische Leistungsmotivausprägung von Studenten handelt. Sollte dies der Fall sein, ist es nötig, um generalisierbarere Aussagen über die untersuchte Thematik treffen zu können, diese Studie an stärker heterogenen Stichproben bzw. auch anderen spezifischen Gruppen (z. B. Schüler, Hochbegabte, verschiedene Berufsgruppen usw.) im Vergleich zu überprüfen. Langzeitstudien könnten zudem Aufschluss darüber bringen, ob wirklich eine bestimmte Stärke und Kombination der Leistungsmotivkomponenten Vorteile im Auswahlprozess von Studenten und beim weiteren Studienerfolg hat bzw. sich diese bevorzugt durchsetzt. Oder spielt die Erfahrung mit Testsituationen im Verlaufe des Bildungsprozesses eines Individuums eine größere Rolle, d. h. können auch Personen mit FM lernen, ihre Angst zu überwinden? Umstritten ist diesbezüglich der negative Zusammenhang zwischen FM und Intelligenz, der auch in dieser Arbeit zum CFT 20-R ($r = -.19$) als einziger Zusammenhang zwischen Leistungsmotiv und Testleistung auftaucht und der die Diskussion, ob Personen mit FM generell weniger intelligent und deswegen weniger erfolgreich sind, erneut entfacht. Die vorliegende Arbeit kann neben anderen Studien (siehe Kapitel 1.4.2) eine solch pauschale Argumentation jedoch nicht stützen, da Teilnehmer mit überwiegend FM teilweise auch bessere Ergebnisse erzielen. Somit unterstreicht die aktuelle Studie die Meinung verschiedener Forscher (Heckhausen, 1980; Mandler & Sarason, 1952; Stapf et al., 1986), dass mehr die motivationalen als die kognitiven Voraussetzungen zu den Unterschieden zwischen HE und FM führen. Methodisch sehr aufwändig, jedoch inhaltlich interessant, wäre in einer Langzeitstudie die Richtung des Zusammenhangs zwischen Intelligenz und Leistungsmotiv zu analysieren: Beeinflusst das Leistungsmotiv durch das Bevorzugen oder das Vermeiden von Leistungssituationen die Entwicklung der Intelligenzstrukturen? Trägt die Intelligenz durch erhöhte Leistungserfolge bzw. Misserfolge zur Festigung der Leistungsmotivkomponenten bei oder sind beide Konstrukte relativ unabhängig und durch eine Drittvariable wie Testangst verbunden? Eine allgemein anerkannte Antwort steht bis heute aus; die Ergebnisse der vorliegenden Arbeit weisen allerdings daraufhin, dass Misserfolgssüchtige durch ihre stärker ausgeprägte Furcht in Leistungssituationen gehemmt sind und nicht die Intelligenz die überwiegend schlechteren Leistungsergebnisse dieser Gruppe erklärt, da diese teilweise auch bessere Ergebnisse erreichten. Es scheint daher sinnvoll, erfolgreiche Trainings zur

Kompensation von FM zu entwickeln, um diesen die Möglichkeit zu bieten, in Schule, Studium und Beruf ihr volles Potential zu entfalten. Studien wie von Petty und Harrell (1977) zeigen, dass Leistungen durch Kompensationsmöglichkeiten wie Testwissen und angstreduzierende Instruktionen wie Training von Relaxationsmethoden (Houston, 1982) steigerbar sind. Weiner (1972/1976b) schlug vor, dass unterschiedliche Bekräftigungsraten von Erfolg und Misserfolg für die ungleich Motivierten notwendig sind. Basierend auf der Vorstellung von Motiven als stabilisierendes System nach Heckhausen (1963, 1974) Selbstbewertungsmodell konnten Studien bereits durch Trainings zur Erlernung günstigerer Zielsetzungen, Ursachenerklärungen und Selbstbewertungen positive Effekte auf die Erfolgszuversicht und Intelligenztestergebnisse nachweisen (Krug & Hanel, 1976). Das häufigere negative Coping bei Misserfolgs- im Vergleich zu Erfolgsorientierten (Houston, 1982; Orttensburger, 2006) ist ein weiterer Aspekt, dem sich bevorstehende Trainingsprogramme widmen könnten. Künftige Studien sollten an diesen Forschungsbeiträgen anknüpfen und versuchen, Copingmaßnahmen speziell für Personen mit FM zu entwickeln.

Angesprochene Themenbereiche, die bereits in dieser Untersuchung als Erklärungsmöglichkeit dienten, wie etwa Testängstlichkeit, Selbstkonzept, Kausalattribution, Aufmerksamkeit oder vorhergehende Erfahrungen bzw. das Anspruchsniveau, scheinen untrennbar mit der Leistungsmotivthematik und Performance verbunden zu sein. Studien, die Zusammenhänge zwischen den genannten Themen und den Leistungsmotiven unterstreichen konnten, gibt es bereits zahlreich (Köppl, 2006; Kornprobst, 2005; Schleich, 2007; Wegner, 1973). Auch alternative Erklärungen wie das Anspruchsniveau und unterschiedliche Attributionsmuster sowie die Integration der festgestellten Zusammenhänge zu Testängstlichkeit und Selbstkonzept könnten differenziertere Aussagen über die Ergebnisse liefern. Es wäre wichtig, diese durch Fragebögen in weiterführenden Studien zu erfassen und sie in eine tiefgründigere Analyse bezüglich der Auswirkungen von Stress auf die Intelligenzleistung unterschiedlich motivierter Person mit einzubeziehen.

6. Zusammenfassung

Ziel der vorliegenden Arbeit war es zu untersuchen, in welchem Ausmaß das Testsetting in Form von Misserfolgsstress die Intelligenztestleistung in Abhängigkeit der individuellen Ausprägung der Leistungsmotive „Hoffnung auf Erfolg“ und „Furcht vor Misserfolg“ beeinflusst. Weiterhin trägt diese Studie zur Validierung des überarbeiteten Leistungsprüfsystems nach Horn (1962/1983), dem LPS-neu (Kreuzpointner, 2010), bei. Zunächst wurde die Leistungsmotivausprägung der 72 Teilnehmer mit Hilfe des Regensburger Leistungs-Motiv-Inventar für Erwachsene (RLMI-E; Lukesch & Peters-Häderle, 2007) erfasst. Eine Kontrollgruppe bearbeitete zwei Intelligenztests (CFT 20-R; Weiß, 2006; LPS-neu) in einer neutralen Situation, während die Experimentalgruppe den LPS-neu in einer Stressbedingung absolvieren musste. Um Misserfolgsstress zu erzeugen, bekamen diese schlechte, unter ihrem Intelligenzwert liegende Ergebnisse des zuvor ausgewerteten CFT 20-R zurückgemeldet. Zudem wurde ihnen ein relativ hohes Niveau der Intelligenzquotienten von Studenten der Universität Regensburg anhand einer fiktiven Studie mitgeteilt. Jeweils vor und nach den Intelligenztests sowie nach der Stressinstruktion sollten die Teilnehmer den für diese Studie entwickelten Fragebogen zum aktuellen Wohlbefinden ausfüllen. Weiterhin wurden ihre Meinungen zu den Intelligenztests mittels eines Evaluationsfragebogens und die subjektive Beurteilung ihrer Stressreaktion in der Experimentalgruppe erfragt. Die Testanalysen dieser Fragebögen sowie des LPS-neu liefern insgesamt befriedigende Ergebnisse.

Das signifikant stärkere Absinken des Wohlbefindens nach dem Misserfolgsstress, das tendenziell signifikant niedrigere Wohlbefinden der Experimentalgruppe nach der Instruktion und die Ergebnisse des Stressreaktionsfragebogens, in dem über die Hälfte der Vpn (52,8 %) angaben sich gestresst bzw. sehr gestresst zu fühlen, können einen Einfluss der Manipulation bestätigen.

Ausgehend der theoretischen Ausführungen (Atkinson, 1964/1975; Heckhausen, 1963, 1972; Weiner, 1972/1976b) sowie verschiedener Studien zu Misserfolgsstress wird ein Vorteil der Erfolgshoffenden gegenüber den Misserfolgsfürchtigen, der Hoch- gegenüber den Niedrigmotivierten sowie der Kontroll- gegenüber der Experimentalgruppe bei der Bearbeitung eines Intelligenztests erwartet. Gleichzeitig sollten sich Personen mit überwiegend „Hoffnung auf Erfolg“ sowie Hoch- und Niedrigmotivierte unter Misserfolgsstress durch einen gestiegenen Anreiz verbessern, Personen mit überwiegend „Furcht vor Misserfolg“ sollten sich jedoch durch stärkere Furcht oder irrelevante Gedanken in ihrer Leistung verschlechtern.

Die Mehrheit dieser vermuteten Unterschiede konnte nicht belegt werden. Einige signifikante Subtestvergleiche des LPS-neu ergeben kein schlüssiges Muster, da sie zum Teil Belege in Richtung der Hypothesen und zum Teil auch entgegengesetzte Belege erbringen. So deutet der Trend der Daten auf einen Vorteil der Experimental- und der Gruppe der Niedrigmotivierten gegenüber ihren Vergleichsgruppen und auf eine Leistungssteigerung der genannten Leistungsmotivgruppen unter Misserfolgsstress hin. Tendenziell erzielten die Erfolgsorientierten jedoch höhere Leistungen, wobei sich ein größerer Unterschied in der Stressbedingung abzeichnet. Darüber hinaus konnte ein signifikant negativer Zusammenhang zwischen der Intelligenztestleistung des CFT 20-R und „Furcht vor Misserfolg“ festgestellt werden. Die Leistungsmotive scheinen bezüglich des Wohlbefindens und der Bewertungen keine direkte Rolle zu spielen, wohl aber konnte erwartungsgemäß ein höher gefühltes Stresslevel der Misserfolgsfürchtigen nach der Stressinstruktion festgestellt werden.

Im Hinblick auf die Durchführungsrichtlinien eines Intelligenztests unterstreichen diese Ergebnisse einen Einfluss der Testsituation auf die Intelligenztestleistung und folglich die Wichtigkeit der Verwendung einer standardisierten Instruktion, um die Unterschiede zwischen den Leistungsmotivgruppen möglichst gering zu halten. Auf der Basis dieser Ergebnisse wurde die Relevanz möglicher Trainingsprogramme speziell für Personen mit überwiegend „Furcht vor Misserfolg“ zur besseren Bewältigung von Testsituationen hervorgehoben.

Literaturverzeichnis

- Alper, T. G. (1946). Memory for completed and incompleted tasks as a function of personality: An analysis of group data. *Journal of Abnormal Psychology*, 41, 403-420.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behaviour. *Psychological Review*, 64, 359-372.
- Atkinson, J. W. (1958). Towards experimental analysis of human motivation in terms of motives, expectancies, and incentives. In J. W. Atkinson (Hrsg.), *Motives in fantasy, action and society. A method of assessment and study* (S. 288-305). Princeton: Van Nostrand.
- Atkinson, J. W. (1975). *Einführung in die Motivationsforschung*. Stuttgart: Klett. (Original erschienen 1964: An introduction to motivation)
- Atkinson, J. W. & Litwin, G. H. (1960). Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *Journal of Abnormal and Social Psychology*, 60, 52-63.
- Atkinson, J. W. & Reitman, W. R. (1956). Performance as a function of motive strength and expectancy of Goal-Attainment. In J. W. Atkinson (Hrsg.), *Motives in fantasy, action and society. A method of assessment and study* (S. 278-287). Princeton: Van Nostrand.
- Bartmann, T. (1963). Der Einfluß von Zeitdruck auf die Leistung und das Denkverhalten bei Volksschülern. *Psychologische Forschung*, 27, 1-61.
- Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, 46, 610-620.
- Baumeister, R. J. & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, 16, 361-383.
- Becker, P. (1982). Fear reactions and achievement behavior of students approaching an examination. In H. W. Krohne & L. Laux (Hrsg.), *Achievement, stress, and anxiety* (S. 275-290). Washington: Hemisphere.
- Binet, A. & Simon, T. (1911). *A method of measuring the development of intelligence of young children*. Lincoln, Illinois: Courier company.
- Bortz, J. (1993). *Statistik für Sozialwissenschaftler* (3. Auflage). Berlin: Springer.
- Canavan-Gumpert, D., Garner, K. & Gumpert, P. (1978). *The success-fearing personality*. Lexington, Massachusetts: D.C. Heath and Company.
- Caroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: University Press.

- Cattell, J. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22.
- Cohen, J. (1973). Eta-Squared and partial eta-squared in fixed factor anova designs. *Educational and Psychological Measurement, 33*, 107-112.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2. Ausgabe). Hillsdale, New York: Erlbaum Associates.
- Corsini, J. R. (1999). *The dictionary of psychology*. London, Philadelphia, Levittown: Brunner/Mazel.
- Covin, R., Donovan, L. A. & Macintyre, P. D. (2003). The relationship between self-esteem and performance when information regarding others performance is available. *Journal of Social Psychology, 143* (3), 541-544.
- Diehl, J. M. & Arbinger, R. (2001). *Einführung in die Inferenzstatistik*. Eschborn: Klotz.
- Diggory, J. C. (1949). Responses to experimentally induced failure. *The American Journal of Psychology, 62*, 48-61.
- Eder, M. (2006). *Der Einfluss von Depressivität auf die Leistungsmotivation*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Eysenck, H. J. (2004). *Die IQ-Bibel. Intelligenz verstehen und messen*. Stuttgart: Klett-Cotta.
- Feather, N. T. (1965). Performance at a difficult task in relation to initial expectation of success, test anxiety, and need achievement. *Journal of Psychology, 33*, 200-217.
- French, E. G. (1958). Effects of the interaction of motivation and feedback on task performance. In J. W. Atkinson (Hrsg.), *Motives in fantasy, action and society. A method of assessment and study* (S. 400-408). Princeton: Van Nostrand.
- French, G. E. & Thomas, H. T. (1958). The relation of achievement motivation to problem solving effectiveness. *Journal of Abnormal and Social Psychology, 56*, 45-48.
- Griffore, R. (1977). Fear of success and task difficulty: Effects on graduate students final exam performance. *Journal of Educational Psychology, 69* (5), 630-635.
- Guilford, J. P. (1963). Intellectual resources and their values as seen by scientists. In C. W. Taylor & F. Barron (Hrsg.), *Scientific Creativity. Its recognition and development* (S. 101-160). New York: Willey.
- Guthke, J. (1974). *Zur Diagnostik der intellektuellen Lernfähigkeit* (2. Auflage). Berlin: VEB.
- Häcker, H. & Stapf, K. H. (2004). *Psychologisches Wörterbuch* (14. Auflage). Huber: Göttingen.
- Häcker, H. & Stapf, K. H. (2009). *Psychologisches Wörterbuch* (15. Auflage). Huber: Göttingen.

- Heckhausen, H. (1963). *Hoffnung und Furcht in der Leistungsmotivation*. Meisenheim am Glan: Anton Hain.
- Heckhausen, H. (1972). Die Interaktion der Sozialisationsvariablen in der Genese des Leistungsmotivs. In C. F. Graumann (Hrsg.), *Handbuch der Psychologie*. Band 7/2 (S. 955-1019). Göttingen: Hogrefe.
- Heckhausen, H. (1974). *Leistung und Chancengleichheit*. Göttingen: Hogrefe.
- Heckhausen, H. (1975). Fear of failure as a self-reinforcing motive system. In I. G. Sarason & C. Spielberger (Hrsg.), *Stress and anxiety*, Band 2 (S. 117-128). Washington, D. C.: Hemisphere.
- Heckhausen, H. (1980). *Motivation und Handeln. Lehrbuch der Motivationspsychologie*. Berlin: Springer.
- Heckhausen, H. (1982). Task-irrelevant cognitions during exam. In H. W. Krohne & L. Laux (Hrsg.), *Achievement, stress, and anxiety* (S. 247-274). Washington: Hemisphere.
- Heckhausen, H. & Heckhausen, J. (2006). *Motivation und Handeln. Lehrbuch der Motivationspsychologie* (3. Auflage). Berlin: Springer.
- Heinrich, D. L. & Spielberger, C. D. (1982). Anxiety and complex learning. In H. W. Krohne & L. Laux (Hrsg.), *Achievement, stress, and anxiety* (S. 145-165). Washington: Hemisphere.
- Hesse, H. (1906/1970). *Unterm Rad* (11. Auflage). Frankfurt am Main: Suhrkamp.
- Hopkins, K. D. & Weeks, D. L. (1990). Tests for normality and measures of skewness and kurtosis: Their place in research reporting. *Educational and Psychological Measurement*, 50, 717-729.
- Hoppe, F. (1930). Erfolg und Misserfolg. *Psychologische Forschung*, 4, 1-62.
- Horn, W. (1962/1983). *Leistungsprüfsystem LPS. Handanweisung für die Durchführung, Auswertung und Interpretation*. Göttingen: Hogrefe.
- Horner, M. S. (1968). *Sex differences in achievement motivation and performance in competitive and non-competitive situations*. University of Michigan: Unveröffentlichte Dissertation.
- Houstan, B. K. (1982). Trait anxiety and cognitive coping behavior. In H. W. Krohne & L. Laux (Hrsg.), *Achievement, stress, and anxiety* (S. 195-206). Washington: Hemisphere.
- Ivanov, I. (2005). *Psychische Gesundheit und Arbeitsleben*. Helsinki: Weltgesundheitsorganisation Europa. Download unter: <http://www.euro.who.int/document/mnh/qbrief06.pdf> [Stand: 10.02.2010]
- Jäger, A. O. (1973). *Dimensionen der Intelligenz* (3. Auflage). Göttingen: Hogrefe.

- Kendler, T. S. (1949). The effect of success and failure on the recall of tasks. *The Journal of General Psychology*, 41, 79-87.
- Köppl, J. (2006). *Leistungsmotive, Kausalattribution, Anspruchsniveau und Arbeitszufriedenheit – Eine Validierungsstudie zum RLMI*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Kornprobst, N. (2005). *Erfolgsfurcht und Selbstkonzept*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Krug, S. & Hanel, J. (1976). Motivänderung: Erprobung eines theoriegeleiteten Trainingsprogrammes. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 8, 274-287.
- Kreuzpointner, L. (2010). *Bedingung für die Äquivalenz von Papier-Bleistift-Version und Computerversion bei Leistungstests*. Universität Regensburg: Dissertation.
- Kubinger, K. D., Rasch, D. & Moder, K. (2009). Zur Legende der Voraussetzungen des *t*-Tests für unabhängige Stichproben. *Psychologische Rundschau*, 60, 26-27.
- Lantz, B. (1945). Some dynamics aspects of success and failure. *Psychological Monographs*, 59, 1-40.
- Laux, L. & Vossel, G. (1982). Theoretical and methodological issues in achievement-related Stress and anxiety research. In H. W. Krohne & L. Laux (Hrsg.), *Achievement, stress, and anxiety* (S. 3-18). Washington: Hemisphere.
- Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.
- Lazraus, R. S. (1991). *Emotion and adaption*. New York: Oxford University Press.
- Lazarus, R. S. & Eriksen, C. W. (1952). Psychological stress and ist personality correlates: Part I. The effects of failure stress upon skilled performance. *Journal of Experimental Psychology*, 43, 100-105.
- Lazarus, R. S. & Launier, R. (1978). Stress related transactions between person and environment. In L. A. Pervin & M.Lewis (Hrsg.), *Perspectives interactional psychology* (S. 287-327). New York: Plenum.
- Lazarus, R S., Deese, J. & Osler, S. F. (1952). The effects of psychological stress upon performance. *Psychological Bulletin*, 49, 293-317.
- Lighthall, F., Ruebush, B., Sarason, S. & Zweibelson, I. (1959). Change in mental ability as a function of test anxiety and type of mental test. *Journal of Consulting Psychology*, 23 (3), 34-38.
- Lowell, E. L. (1952). The effect of need for achievement on learning and speed of performance. *Journal of Psychology*, 33, 31-40.
- Lukesch, H. (1998). *Einführung in die pädagogisch-psychologische Diagnostik* (2. Auflage). Regensburg: Roderer.

- Lukesch, H. (2006). *Einführung in die Pädagogische Psychologie* (4. Auflage). Regensburg: Roderer.
- Lukesch, H., Kornprobst, N., Köppl, J. & Peters-Häderle, K. (2008). *Entwicklung und Validerung des Regensburger Leistungs-Motiv-Inventar für Erwachsene (RLMI-E)*. Universität Regensburg: Institut für Psychologie, Lehrstuhl für Pädagogische Psychologie und Medienpsychologie. Download unter: <http://epub.uni-regensburg.de/3679/> [Stand: 03.08.2009]
- Lukesch, H. & Peters-Häderle, K. (2007). *Das Regensburger Leistungs-Motiv-Inventar für Erwachsene (RLMI-E)*. Göttingen: Hogrefe.
- Mahone, C. H. (1960). Fear of failure and unrealistic vocational aspiration. *Journal of Abnormal and Social Psychology*, 60 (2), 253-261.
- Mandler, G. & Sarason, S. B. (1952). A study of anxiety and learning. *Journal of Abnormal Social Psychology*, 47, 166-173.
- McClelland, D. C. (1956). Risk taking in children with high and low need for achievement. In J. W. Atkinson (Hrsg.), *Motives in fantasy, action and society. A method of assessment and study* (S. 306-320). Princeton: Van Nostrand.
- McClelland, D. C. & Apicella, F. S. (1947). Reminiscence following experimental induced failure. *Journal of Experimental Psychology*, 37, 159-169.
- McClelland, D. C., Atkinson, J. W., Clark, R. A. & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.
- McKinney, F. (1933). Certain emotional factors in learning and efficiency. *Journal of General Psychology*, 9, 101-116.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present and future. In D. P. Flanagan & P. L. Harison (Hrsg.) *Contemporary intellectual assessment: Theorie, tests, and issues* (S. 136-202). New York: Guilford Press.
- Meer, E. van der (1998). Intelligenz als Informationsverarbeitung. In E. Roth (Hrsg.), *Intelligenz. Grundlagen und neuere Forschung* (S. 161-184). Berlin: Kohlhammer.
- Meyer, W. U. (1976). Leistungsorientiertes Verhalten als Funktion von wahrgenommener eigener Begabung und wahrgenommener Aufgabenschwierigkeit. In H. D. Schmalt & W. U. Meyer (Hrsg.), *Leistungsmotivation und Verhalten* (S. 101-136). Stuttgart: Klett.
- Milczarek, M., Schneider, E. & Gonzalez, E. R. (2009). Stress at work – facts and figures. European risk observation report. Luxemburg: European Agency for Safety and Health at Work, European Communities. Download unter: http://osha.europa.eu/en/publications/reports/TE-81-08-478-EN-C_OSH_in_figures_stress_at_work [Stand: 10.02.2010]
- Morris, L. W. & Liebert, R. M. (1970). Relationship of cognitive and emotional components of test anxiety to physiological arousal and academic performance. *Journal of Consulting and Clinical Psychology*, 35, 332-337.

- Murray, H. A. (1938). *Explorations in personality. A clinical and experimental study of fifty men of college age*. New York: Oxford Univ. Press.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J. & Urbina, S. (1996). Intelligence: knowns and unknowns. *American Psychologist*, 51 (2), 77-101.
- Nemzek, C. L. (1933). The constancy of the I.Q.. *Psychological Bulletin*, 30, 143-153.
- Orttenburger, D. (2006). *Leistungsmotivation, Stress und Coping bei Salespromotern*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Peters, K. (2001). *Erfolgsfurcht bei Frauen*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Peters-Häderle, K. (2006). *Erfolgsfurcht und Leistungsangst bei Kindern*. Universität Regensburg: Unveröffentlichte Dissertation.
- Petty, N. E. & Harrell, E. H. (1977). Effect of programmed instruction related to motivation, anxiety, and test wiseness on group IQ test performance. *Journal of Educational Psychology*, 69 (5), 630-635.
- Piaget, J. (1967). *Psychologie der Intelligenz*. Zürich: Rascher.
- Plath, H.-E. & Richter, P. (1984). *Ermüdung, Monotonie, Sättigung, Stress. (BMS). Verfahren zur skalierten Erfassung erlebter Beanspruchungsfolgen*. Göttingen: Hogrefe.
- Postman, L. & Bruner, J. S. (1948). Perception under stress. *Psychological Review*, 55, 314-323.
- Psychrempel (2007). *Klinisches Wörterbuch* (261. Auflage). Berlin: De Gruyter.
- Raphelson, A. C. (1957). The relationship among imaginative, direct verbal, and physiological measures of anxiety in an achievement situation. In J. W. Atkinson (Hrsg.), *Motives in fantasy, action and society. A method of assessment and study* (S. 340-349). Princeton: Van Nostrand.
- Rheinberg, F. (1976). Situative Determinanten der Beziehung zwischen Leistungsmotiv und Schul- und Studienleistung. In H. D. Schmalt & W. U. Meyer (Hrsg.), *Leistungsmotivation und Verhalten* (S. 249-246). Stuttgart: Klett.
- Rheinberg, F. (2000). *Motivation* (3. Auflage). Stuttgart: Kohlhammer.
- Rheinberg, F., Schwarz, N. & Singer, G. M. (1987). Symbolische Selbstergänzung und Leistungsmotivation. *Zeitschrift für Sozialpsychologie*, 18, 50-58.
- Roth, E. (1998a). Der Intelligenzbegriff. In E. Roth (Hrsg.), *Intelligenz. Grundlagen und neuere Forschung* (S. 9-20). Berlin: Kohlhammer.
- Roth, E. (1998b). Geschichte der Intelligenzforschung. In E. Roth (Hrsg.), *Intelligenz. Grundlagen und neuere Forschung* (S. 21-36). Berlin: Kohlhammer.

- Rotter, J. B. (1955). The role of the psychological situation in determining the direction of human behaviour. In R. J. Marshall (Hrsg.), *Nebraska Symposium on motivation* (S. 245-269). Lincoln: University of Nebraska Press.
- Rudolph, U. (2003). *Motivationspsychologie*. Weinheim: Beltz.
- Ruebush, B. K. (1960). Interfering and facilitating effects of test anxiety. *Journal of Abnormal and Social Psychology*, 60 (2), 205-212.
- Sarason, I. G. (1959). Intellectual and personality correlates of test anxiety. *Journal of Abnormal Psychology*, 59, 272-275.
- Sarason, I. G. (1960). Empirical findings and theoretical problems in the use of anxiety scale. *Psychological Bulletin*, 57 (5), 403-415.
- Sarason, S. B. & Mandler, G. (1952). Some correlates of test anxiety. *Journal of Abnormal Social Psychology*, 47, 810-817.
- Sarason, S. B., Mandler, G. & Craighill, P. G. (1952). The effect of differential instruction on anxiety and learning. *Journal of Abnormal Social Psychology*, 47, 561-565.
- Sauer, J. & Gattringer, H. (1986). Zur Aufklärung der Schulleistung durch spezifische und gemeinsam Varianzanteile von Intelligenz und Motivation. In K. Daumenlang & Sauer, J. (Hrsg.), *Aspekte psychologischer Forschung. Festschrift zum 60. Geburtstag von Erwin Roth* (S. 237-257). Göttingen: Hogrefe.
- Schachter, S. & Singer, J. E. (1962) Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69, 379-399.
- Schleich, M. (2007). *Leistungsmotivation und Anspruchsniveau*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Schmalt, H. D. (1976). Methoden der Leistungsmotivmessung. In H. D. Schmalt & W. U. Meyer (Hrsg.), *Leistungsmotivation und Verhalten* (S. 101-136). Stuttgart: Klett.
- Schneider, K. & Gallitz, H. (1973). Leistungsänderung nach Erfolg und Misserfolg bei leichten und schwierigen Aufgaben. In K. Schneider (Hrsg.), *Motivation unter Erfolgsrisiko* (S. 80-160). Göttingen: Hogrefe.
- Schneider, K. & Meise, C. (1973). Leistungs- und anschlussmotiviertes Risikoverhalten bei der Aufgabenwahl. In K. Schneider (Hrsg.), *Motivation unter Erfolgsrisiko* (S. 212-238). Göttingen: Hogrefe.
- Schneider, K. & Schmalt, H. D. (2000). *Motivation* (3. Auflage). Stuttgart: Kohlhammer.
- Schneider, K. (1973). *Motivation unter Erfolgsrisiko*. Göttingen: Hogrefe.
- Schulz, P. & Schönplug, W. (1982). Regulatory activity during states of stress. In H. W. Krohne & L. Laux (Hrsg.), *Achievement, stress, and anxiety* (S. 51-73). Washington: Hemisphere.

- Sears, R. R. (1937). Initiation of the repression sequence by experienced failure. *Journal of Experimental Psychology*, 20, 570-580.
- Selye, H. (1974). *Stress. Bewältigung und Lebensgewinn*. München: Piper. (Original erschienen 1974: Stress without distress)
- Sowa, D. (2009). *Leistungsmotivation und Selbstkonzept – Eine Validierungsstudie zum RLMI-K/J*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Spearman, C. E. (1904). „General intelligence,“ objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Stadler, M. (2010). *Die Beeinflussbarkeit der Intelligenztestleistung durch Testdurchführungssituation, Leistungsmotiv und implizite Intelligenztheorie*. Universität Regensburg: Unveröffentlichte Diplomarbeit.
- Stapf K. H., Fischer, P. M. & Degner, U. (1986). Über die Informationelle und motivationale Wirkung verschiedener Rückmeldungsmodalitäten beim Lernen. In K. Dauenlang & J. Sauer (Hrsg.), *Aspekte der psychologischen Forschung. Festschrift zum 60. Geburtstag von Erwin Roth* (S. 221-236). Göttingen: Hogrefe.
- Steinmayr, R. & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, 19, 80-90.
- Stern, W. (1912). *Die psychologischen Methoden der Intelligenzprüfung und deren Anwendung von Schulkindern*. Leipzig: Barth.
- Sternberg, R. J. (1984). A contextualist view of the nature of intelligence. *International Journal of Psychology*, 19, 301-334.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Sternberg, R. J. (2003). A broad view of intelligence. The theory of successful intelligence. *Consulting Psychology Journal: Practice & Research*, 55 (3), 139-154.
- Sternberg, R. J. & Berg, C. A. (1986). Quantitative integration: Definitions of intelligence: A comparison of the 1921 and 1986 Symposia. In R. J. Sternberg & D. K. Detterman (Hrsg.), *What is intelligence?* (S. 155-166). Norwood: Ablex.
- Sternberg, R. J. & Detterman, D. K. (1986). *What is intelligence?* Norwood: Ablex.
- Sullivan, E. B. (1927). Attitude in relation to learning. *Psychological Monographs*, 36 (3), 1-141.
- Taylor, J. A. (1958). Anxiety level, stress, and verbal learning. *The Journal of Abnormal and Social Psychology*, 57, 55-60.
- Taylor, J. A. & Rechsaffen, A. (1959). Manifest anxiety and reversed alphabet printing. *Journal of Abnormal Psychology*, 58, 221-224.

- Thorndike, E. L. & Woodyard, E. (1934). The influence of the relative frequency of success and frustrations upon intellectual achievement. *The Journal of Educational Psychology*, 25 (4), 241-250.
- Thorndike, E. L., Bell, J. C., Pintner, R., Terman, L. M., Henmon, V. A. C., Ruml, B., Thurstone, L. L., Pressy, S. L., Buckingham, B. R., Colvin, S. S., Dearborn, W. F., Freeman, F. N., Haggerty, M. E., Peterson, J. Woodrow, H., W & Yerkes, R. M. (1921). Intelligence and its measurement: A symposium. *Journal of Educational Psychology*, 12, 123-147, 195-216, 271-275.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Vogel, W., Raymond, S. & Lazarus, R. S. (1959). Intrinsic motivation and psychological stress. *Journal of Abnormal Psychology*, 58, 225-233.
- Wechsler, D. (1965). *Die Messung der Intelligenz Erwachsener*. Bern: Huber.
- Wegner, R. (1973). *Eine Untersuchung über die Beziehung zwischen Leistungsmotivation, Angst und Leistung*. Universität Köln: Unveröffentlichte Dissertation.
- Weiner, B. (1976a). Attributionstheoretische Analyse von Erwartung X Nutzen-Theorien. In H. D. Schmalz & W. U. Meyer (Hrsg.), *Leistungsmotivation und Verhalten* (S. 81-100). Stuttgart: Klett.
- Weiner, B. (1976b). *Theorien der Motivation*. Stuttgart: Klett. (Original erschienen 1972: Theories of motivation: from mechanism to cognition)
- Weiβ, R. H. (2006). *CFT 20-R. Grundintelligenztest Skala 2 – Revision*. Göttingen: Hogrefe.
- Williams, M. (1947). An experimental study of intellectual control under stress and associated Rorschach factors. *Journal of Consulting Psychology*, 11, 21-29.
- Winterbottom, M. R. (1953). The relation of need for achievement to learning experiences and mastery. In J. W. Atkinson (Hrsg.), *Motives in fantasy, action and society. A method of assessment and study* (S. 453-478). Princeton: Van Nostrand.
- Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology*, 18, 459-482.
- Zander, A. F. (1944). A study of experimental frustration. *Psychological Monographs*, 56 (3), 1-37.
- Zeller, A. F. (1950). An experimental analogue of repression. II. The effect of individual failure and success on memory measured by relearning. *Journal of Experimental Psychology*, 40 (4), 411-422.
- Zweibelson, I. (1956). Test anxiety and intelligence test performance, *Journal of Consulting Psychology*, 20, 479-481.

Erklärung

Ich versichere hiermit, dass ich die anliegende Arbeit mit dem Thema

Der Einfluss von Misserfolgsstress auf die Intelligenztestleistung in Abhängigkeit des Leistungsmotivs

selbstständig verfasst und keine anderen Hilfsmittel als die angegebenen benutzt habe. Die Stellen, die anderen Werke dem Wortlaut oder dem Sinn nach entnommen sind, habe ich in jedem Fall durch die Angaben der Quelle, auch der benutzten Sekundärliteratur, als Entlehnung kenntlich gemacht.

Regensburg, den 17.05.2010

Nicole Eiseler

Anhang

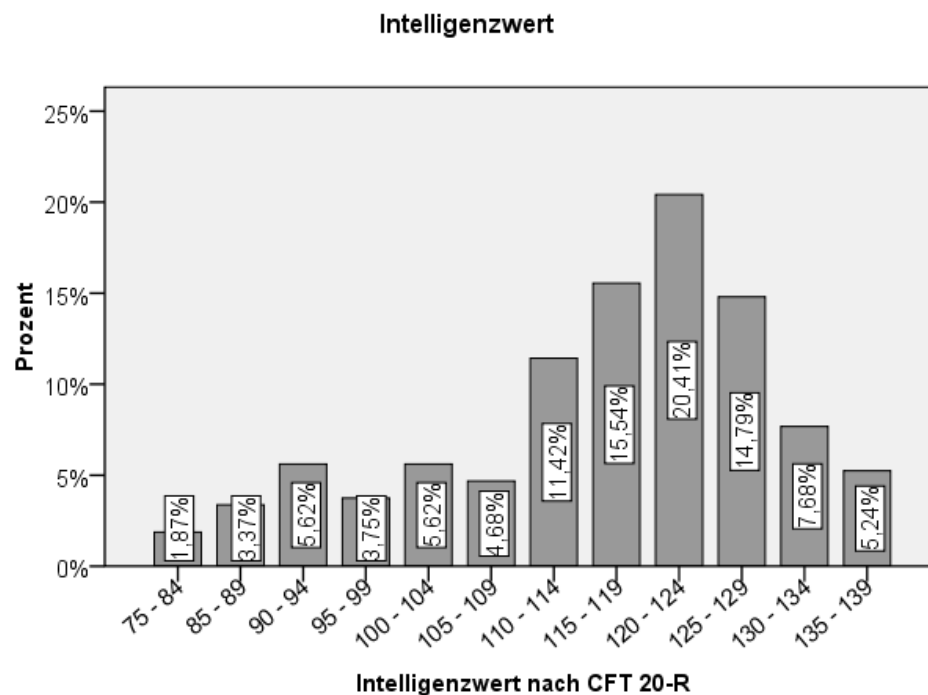
Anhang A

A-1: Instruktion zum LPS-neu für die Experimentalgruppe

Liebe / Lieber _____,

ich freue mich, dass Du heute zum dritten Teil dieses Experiments erschienen bist.

In den letzten zwei Jahren wurde an der Universität Regensburg durch den Lehrstuhl von Herrn Prof. Lukesch an Studenten verschiedener Fachrichtungen die Intelligenzleistung mit Hilfe des CFT 20-R bestimmt. Die Verteilung der Intelligenzwerte ist in folgendem Graph dargestellt. Die Ergebnisse dieser Diplomarbeit fließen in diese Untersuchung mit ein.



	N	Minimum	Maximum	Mittelwert	Standardabweichung
CFT IQ-Wert	534	79	138	117,62	13,99
Gültige Werte (Listenweise)	534				

Seit unserem ersten Treffen habe ich den ersten Intelligenztest bereits ausgewertet. Dein Wert beträgt ____ und somit fällst Du in die Gruppe ____ - ____ . Es hatten ____ % der 534 Studenten einen Intelligenzwert unter und ____ % einen Wert über Deiner Intelligenzleistung.

A-2: Debriefing der Kontrollgruppe

Liebe Versuchsperson,

vielen Dank für Deine Teilnahme am Experiment.

Das Experiment ist Teil meiner Diplomarbeit, die sich damit beschäftigt, welchen Einfluss das Leistungsmotiv und die Durchführungssituation auf das Ergebnis in einem Intelligenztest hat.

Deshalb wurde zu Beginn mit der Onlinetestung von allen Versuchspersonen das Leistungsmotiv erhoben.

Du wurdest der Kontrollgruppe zugeteilt. Das heißt, dass Du den ersten und den zweiten Intelligenztest in einer neutralen Situation bearbeitet hast. Versuchspersonen der Experimentalgruppe bearbeiteten hingegen den ersten Intelligenztest in einer neutralen Situation und den zweiten unter einer Stressbedingung.

Die Daten, die von Dir erhoben wurden, werden ausschließlich anonymisiert verarbeitet. Wenn Du daran Interesse hast, bekommst Du per Email Deine Ergebnisse der Intelligenztests mitgeteilt.

Bitte sprich nicht mit Kommilitonen oder Freunden über das Experiment bzw. die Testaufgaben, da diese eventuell noch selber daran teilnehmen werden und deswegen noch nichts darüber erfahren sollten. Da die Datenerhebung noch bis ca. November 2009 laufen wird, möchte ich Dich bitten, bis dahin Stillschweigen darüber zu bewahren, damit die Datenerhebung nicht gefährdet wird.

Hiermit erkläre ich mich damit einverstanden, dass meine Daten anonymisiert ausgewertet und für die Studie verwendet werden.

Unterschrift

Ich möchte per Email über meine Testergebnisse informiert werden.

ja

nein

A-3: Debriefing der Experimentalgruppe

Liebe Versuchsperson,

vielen Dank für Deine Teilnahme am Experiment.

Das Experiment ist Teil meiner Diplomarbeit, die sich damit beschäftigt, welchen Einfluss das Leistungsmotiv und die Durchführungssituation auf das Ergebnis in einem Intelligenztest hat.

Deshalb wurde zu Beginn mit der Onlinetestung von allen Versuchspersonen das Leistungsmotiv erhoben.

Du wurdest der Experimentalgruppe zugeteilt. Versuchspersonen der Kontrollgruppe bearbeiteten beide Tests in neutraler Situation. Du hast den ersten Intelligenztest in einer neutralen Situation und den zweiten unter einer Stressbedingung bearbeitet. Die Ergebnisse der aufgeführten Untersuchung sind somit nur fiktive Werte, sowie dein angegebener Intelligenzwert frei erfunden ist. Dein wahres Ergebnis beträgt: ____ . Deswegen beurteile bitte noch kurz deine Reaktion auf das Ergebnis:

Antwortmuster: 1 = trifft gar nicht zu 2 = trifft eher nicht zu 3 = teils teils 4 = trifft eher zu 5 = trifft sehr zu					
1. Das Ergebnis meines Intelligenzwerts hat mich frustriert.	1	2	3	4	5
2. Ich habe mich über das Ergebnis meiner Leistung geärgert	1	2	3	4	5
3. Ich habe mich durch den Wert unter Druck gesetzt gefühlt.	1	2	3	4	5

Die Daten, die von Dir erhoben wurden, werden ausschließlich anonymisiert bearbeitet. Wenn Du daran Interesse hast, bekommst Du die Ergebnisse der Intelligenztests mitgeteilt.

Bitte sprich nicht mit Kommilitonen oder Freunden über das Experiment bzw. die Testaufgaben, da diese eventuell noch selber daran teilnehmen werden und deswegen noch nichts darüber erfahren sollten. Da die Datenerhebung noch bis ca. November 2009 laufen wird, möchte ich Dich bitten, bis dahin Stillschweigen darüber zu bewahren, damit die Datenerhebung nicht gefährdet wird.

Hiermit erkläre ich mich damit einverstanden, dass meine Daten anonymisiert ausgewertet und für die Studie verwendet werden.

Unterschrift

Ich möchte per Email über meine Testergebnisse informiert werden.

ja nein

Anhang B

B-1: Fragebogen zur Person

Fragebogen zur Person

	Datum:	VP-Code:
1	Vorname:	
2	Nachname:	
3	Geschlecht:	<input type="checkbox"/> weiblich <input type="checkbox"/> männlich
4	Geburtsdatum:	
5	Alter:	
6	Studienfach / Beruf:	
7	Semester:	
8	Händigkeit:	<input type="checkbox"/> Rechtshänder <input type="checkbox"/> Linkshänder <input type="checkbox"/> Beidhänder
9	Muttersprache:	<input type="checkbox"/> Deutsch <input type="checkbox"/> andere falls andere, welche:
10	Hochschulreife:	<input type="checkbox"/> Allgemeine Hochschulreife <input type="checkbox"/> Fachabitur
11	Abiturnote:	

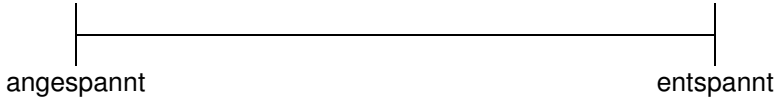
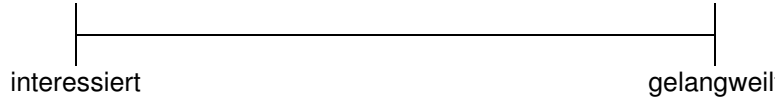
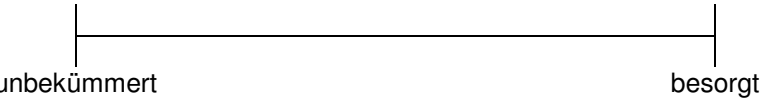
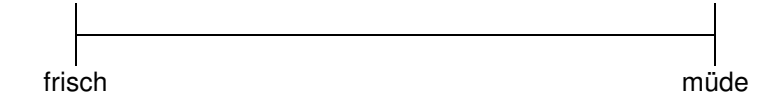
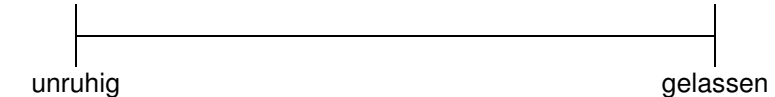
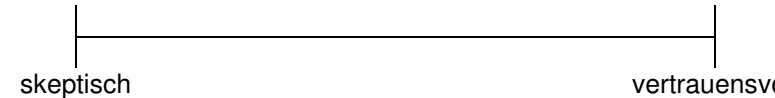
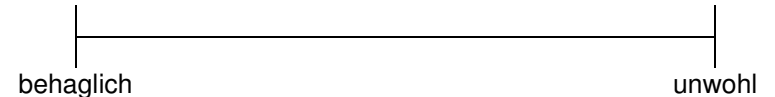
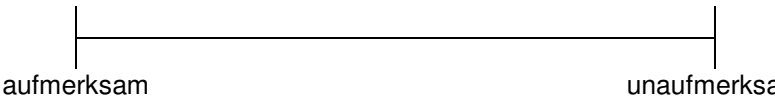
B-2: Fragebogen zum aktuellen Wohlbefinden**Fragebogen zum aktuellen Wohlbefinden - 1**

Name: _____ Testdatum: _____

VP-Code: _____

Bitte beurteile Dein aktuelles Wohlbefinden. Setze jeweils an der Stelle einen Strich auf die Linie, die Dein aktuelles Befinden zwischen den zwei gegebenen Polen am besten beschreibt. Setze den Strich z. B. in der ersten Einschätzungsaufgabe ganz links, wenn Du Dich total angespannt fühlst. Wenn Du Dich dagegen total gelassen fühlst, zeichne den Strich ganz rechts. Platziere den Strich, je nachdem, wie Du Dich gerade fühlst, zwischen den beiden jeweiligen Extremwerten.

Bitte bearbeite alle acht Aufgaben. Denk dabei nicht zu lange nach, sondern gib Deine Einschätzung möglichst spontan ab.

1. 
2. 
3. 
4. 
5. 
6. 
7. 
8. 

B-3: Evaluationsfragebogen**Fragebogen zur Testevaluation des CFT 20-R**

Name: _____ Testdatum: _____

VP-Code: _____

Danke für die Teilnahme am Intelligenztest. Bitte beurteile den gerade durchgeführten Test nach den unten beschriebenen Kriterien. Bitte antworte immer entsprechend Deiner persönlichen Meinung und scheue Dich nicht davor, auch extreme Antworten zu wählen. Es gibt dabei keine „richtigen“ oder „falschen“ Antworten.

Wenn Du einer Aussage gar nicht zustimmst, kreuze bitte die 1 an. Stimmt Du einer Aussage hingegen vollständig zu, kreuze die 5 an. Zwischen 1 und 5 kannst Du beliebig abstufen, je nachdem, in welchem Ausmaß Du der Aussage zustimmst.

Bitte beantworte den Fragebogen vollständig und lasse keine Frage aus. Kreuze bitte immer nur eine Zahl pro Aussage an.

Antwortmuster:

Antwortmuster:

- 1 = trifft gar nicht zu
- 2 = trifft eher nicht zu
- 3 = teils teils
- 4 = trifft eher zu
- 5 = trifft sehr zu

- | | | | | | | |
|--|--|---|---|---|---|---|
| 1. Die Anweisungen zu den Aufgaben sind verständlich. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |
| 2. Die Schwierigkeit der Aufgaben ist zu hoch. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |
| 3. Die Schwierigkeit der Aufgaben ist zu niedrig. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |
| 4. Die für die einzelnen Aufgaben verfügbare Zeit ist angemessen. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |
| 5. Die Anordnung der Aufgaben im Testbogen ist übersichtlich. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |
| 6. Die Beantwortung der Aufgaben auf dem Antwortbogen ist übersichtlich. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |
| 7. Die Testzeit insgesamt ist zu lang. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |
| 8. Die Aufgaben des Tests sind interessant. | <table border="1" style="display: inline-table; border-collapse: collapse;"><tr><td style="width: 20px; height: 20px; text-align: center;">1</td><td style="width: 20px; height: 20px; text-align: center;">2</td><td style="width: 20px; height: 20px; text-align: center;">3</td><td style="width: 20px; height: 20px; text-align: center;">4</td><td style="width: 20px; height: 20px; text-align: center;">5</td></tr></table> | 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 | | |

Anhang C

Tabelle C-1: Skalenwerte des RLMI-E

	Items	<i>M</i>	<i>SD</i>	Schiefe	Exzess
Hoffnung auf Erfolg					
Ausbildung	6	102.58	8.34	0.02	-0.84
Beruf	6	102.33	7.08	0.05	-0.52
Freizeit	6	102.97	9.03	0.39	0.92
Aussehen	6	100.96	8.57	-0.26	-0.16
Gesamt	24	101.85	7.98	0.36	-0.37
Furcht vor Misserfolg					
Ausbildung	6	103.35	10.59	-0.05	-0.59
Beruf	6	105.87	11.76	0.13	-0.35
Freizeit	6	102.94	10.77	0.11	-0.23
Aussehen	6	103.74	8.15	-0.15	0.89
Gesamt	24	104.00	11.45	-0.13	0.21
Furcht vor Erfolg					
Ausbildung	6	102.40	7.90	0.14	-0.68
Beruf	6	99.64	8.54	0.14	-0.59
Freizeit	6	100.28	7.80	0.09	0.23
Aussehen	6	100.82	7.50	0.20	-0.27
Gesamt	24	99.29	8.70	-0.29	0.07
Hoffnung auf Misserfolg					
Ausbildung	6	101.18	6.16	0.21	-1.03
Beruf	6	99.49	7.90	0.12	-0.39
Freizeit	6	98.21	8.34	-0.00	-0.73
Aussehen	6	101.50	8.15	0.22	-0.26
Gesamt	24	98.62	8.45	-0.42	0.05

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, α = Cronbachs Alpha

Tabelle C-2: Skalenwerte des CFT

	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Exzess
Gesamte Stichprobe						
Reihenfortsetzen	13.39	1.58	7	15	-1.60	3.15
Klassifikationen	11.76	1.71	8	15	-0.39	-0.38
Matrizen	12.54	1.47	9	15	-0.52	-0.31
Topologien	8.18	1.61	4	11	-0.43	0.08
Gesamt	45.88	4.63	31	54	-0.94	0.79
IQ-Wert	115.69	13.72	79	145	-0.28	0.07
Standardwert	11.26	9.10	86	13	-0.28	0.11
Kontrollgruppe						
Reihenfortsetzen	13.08	1.79	7	15	-1.61	2.71
Klassifikationen	11.58	1.75	8	14	-0.47	-0.45
Matrizen	12.58	1.61	9	15	-0.57	-0.66
Topologien	7.69	1.60	4	10	-0.31	-0.65
Gesamt	44.94	5.12	31	53	-0.89	0.29
IQ-Wert	110.22	12.95	79	138	-0.44	-0.14
Standardwert	106.63	8.56	86	125	-0.46	-0.11
Experimentalgruppe						
Reihenfortsetzen	13.69	1.28	10	15	-1.10	1.00
Klassifikationen	11.94	1.67	8	15	-0.30	-0.30
Matrizen	12.50	1.34	9	15	-0.49	0.42
Topologien	8.67	1.49	4	11	-0.59	1.58
Gesamt	46.81	3.94	36	54	-0.71	0.74
IQ-Wert	121.17	12.34	92	145	-0.15	-0.03
Standardwert	113.89	8.22	94	13	-0.17	0.10
Hoffnung auf Erfolg						
Reihenfortsetzen	13.19	1.87	7	15	-1.74	2.98
Klassifikationen	11.55	1.71	8	14	-0.47	-0.48
Matrizen	12.45	1.39	9	15	-0.49	0.10
Topologien	8.48	1.73	4	11	-0.94	1.29
Gesamt	45.68	5.03	31	53	-1.10	1.38
IQ-Wert	115.97	14.80	79	142	-0.40	0.24
Standardwert	110.42	9.81	86	128	-0.42	0.27
Furcht vor Misserfolg						
Reihenfortsetzen	13.54	1.32	10	15	-0.96	0.80
Klassifikationen	11.93	1.71	8	15	-0.35	-0.26
Matrizen	12.61	1.55	9	15	-0.58	-0.41
Topologien	7.95	1.50	5	11	-0.05	-0.46
Gesamt	46.02	4.37	36	54	-0.76	0.12
IQ-Wert	115.49	13.03	89	145	-0.17	0.00
Standardwert	110.13	8.65	92.50	13	-0.15	0.08

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, Min = Minimum, Max = Maximum, α = Cronbachs Alpha

Anhang D

Tabelle D-1: Ergebnisse der Faktorenanalyse des Fragebogens zum aktuellen Wohlbefinden 1 und 2

	Wohlbefinden 1				Wohlbefinden 1 (2 Faktoren)			Wohlbefinden 2		
	h^2	Faktor 1	Faktor 2	Faktor 3	h^2	Faktor 1	Faktor 2	h^2	Faktor 1	Faktor 2
Entspannung	.73	.85	.10	.07	.82	.82	.05	.53	.70	.21
Interesse	.73	-.01	.59	.62	.12	.12	.82	.64	.06	.80
Unbekümmertheit	.67	.71	.15	.38	.78	.78	.27	.75	.87	.05
Frische	.75	.24	.83	-.09	.14	.14	.65	.72	.26	.81
Gelassenheit	.77	.86	.11	.10	.85	.85	.08	.56	.73	.17
Vertrauen	.78	.16	-.05	.87	.40	.40	.39	.66	.81	.06
Behaglichkeit	.75	.42	.33	.68	.57	.57	.60	.72	.79	.30
Aufmerksamkeit	.78	.08	.84	.25	.08	.08	.84	.79	.16	.88
Eigenwert		3.52	1.38	1.06		3.52	1.38		3.79	1.58
Erklärte Varianz in %		44.01	17.27	13.26		44.01	17.27		47.38	19.80

Anmerkungen: h^2 = Kommunalität

Fortsetzung Tabelle D-1: Ergebnisse der Faktorenanalyse des Fragebogens zum aktuellen Wohlbefinden 3 bis 5

	Wohlbefinden 3			Wohlbefinden 4			Wohlbefinden 5		
	h^2	Faktor 1	Faktor 2	h^2	Faktor 1	Faktor 2	h^2	Faktor 1	Faktor 2
Entspannung	.69	.80	.25	.78	.85	.21	.61	.76	.19
Interesse	.65	.18	.78	.28	.37	.38	.68	.15	.81
Unbekümmertheit	.79	.84	.29	.89	.93	.12	.83	.89	.18
Frische	.60	.19	.75	.83	.02	.91	.68	.32	.76
Gelassenheit	.83	.91	.00	.90	.95	.05	.78	.88	.06
Vertrauen	.58	.67	.36	.67	.78	.25	.66	.74	.33
Behaglichkeit	.79	.64	.61	.80	.89	.11	.81	.76	.48
Aufmerksamkeit	.79	.18	.87	.79	.18	.87	.87	.15	.92
Eigenwert		4.40	1.33		4.42	1.52		4.51	1.42
Erklärte Varianz in %		54.97	16.56		55.21	18.93		56.33	17.70

Anmerkungen: h^2 = Kommunalität

Tabelle D-2: Ergebnisse der Faktorenanalyse des Evaluationsfragebogens

	Evaluation CFT				Evaluation CFT (2 Faktoren)			Evaluation LPS		
	h^2	Faktor 1	Faktor 2	Faktor 3	h^2	Faktor 1	Faktor 2	h^2	Faktor 1	Faktor 2
Verständlichkeit der Anweisungen	.52	.70	.16	.03	.34	.56	-.14	.41	.58	.27
Hohe Schwierigkeit	.68	.02	.12	.82	.60	.14	.76	.40	.54	-.33
Niedrige Schwierigkeit	.57	.02	.34	-.67	.59	.32	-.70	.52	-.33	.64
Angemessenheit der Zeit	.49	.49	.20	.46	.44	.08	.66	.48	.70	.03
Übersichtlichkeit des Testbogens	.47	.36	.58	-.02	.60	.72	.29	.47	.38	.56
Übersichtlichkeit des Antwortbogens	.76	.87	-.03	-.02	.66	.78	.21	.45	.64	.17
Testzeit zu lang	.47	.10	.67	.06	.51	.54	.47	.36	.24	.55
Interessante Aufgaben	.64	-.04	.80	-.10	.41	.60	-.22	.49	.06	.70
Eigenwert		2.09	1.49	1.03		2.09	1.49		2.45	1.70
Erklärte Varianz in %		26.06	18.68	12.82		26.06	18.68		30.63	21.18

Anmerkungen: h^2 = Kommunalität

Anhang E

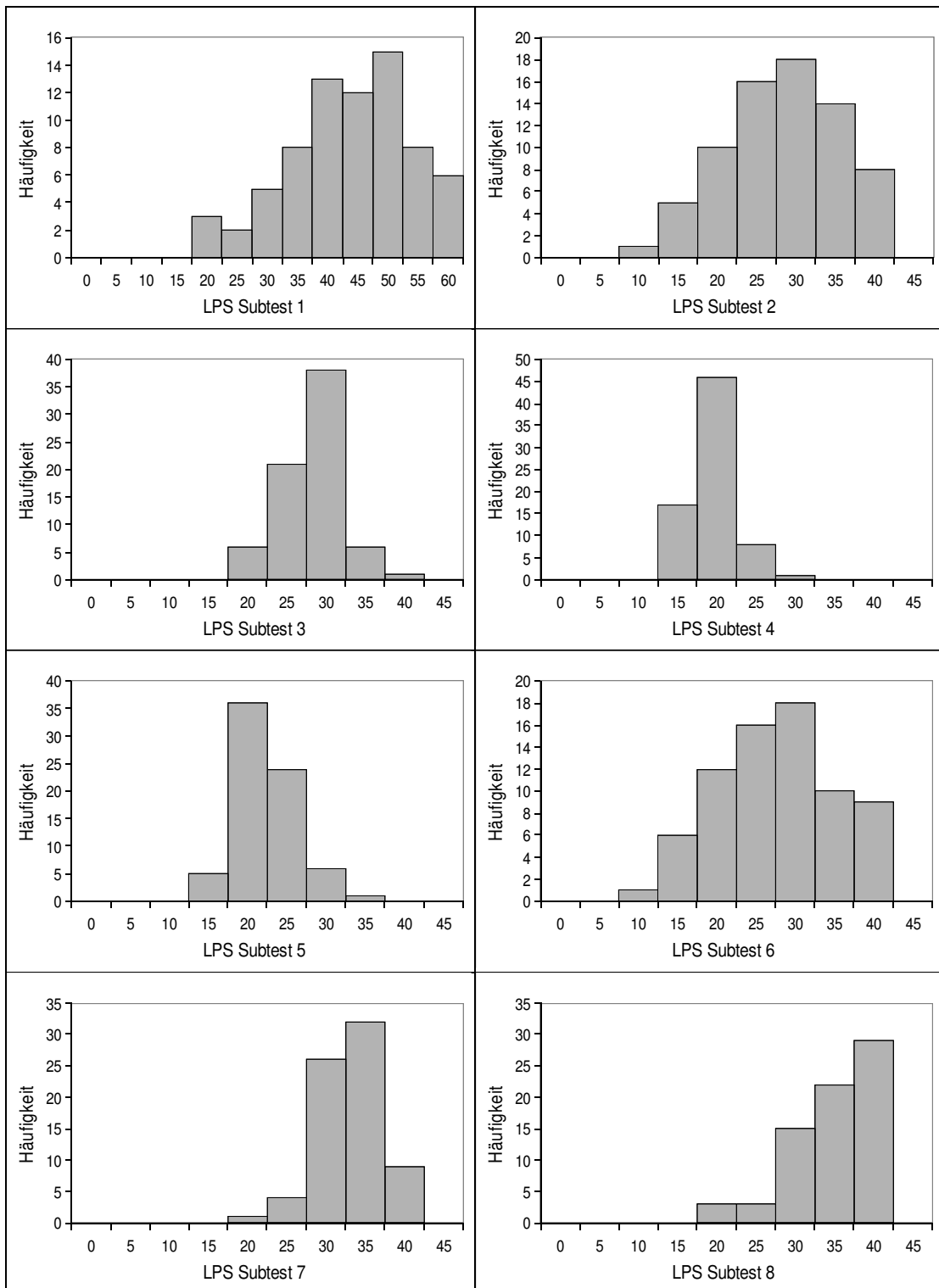
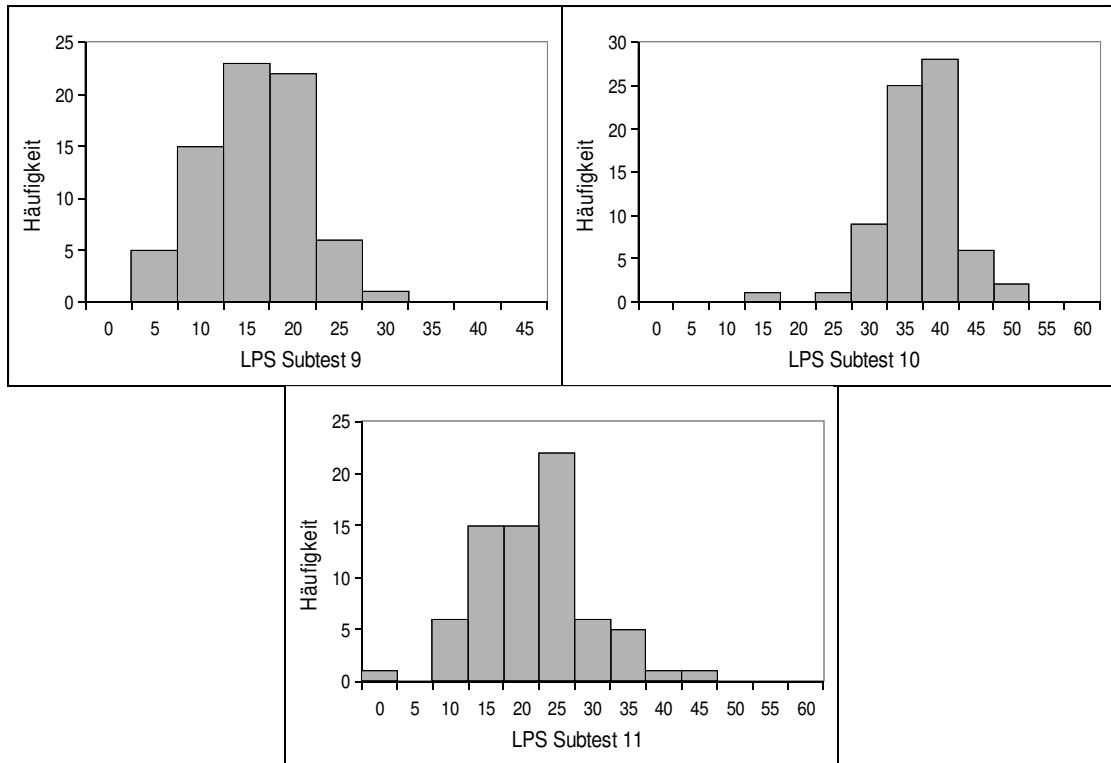


Abbildung E-1: Verteilung der LPS-neu Subtests



Fortsetzung Abbildung E-1: Verteilung der LPS-neu Subtests

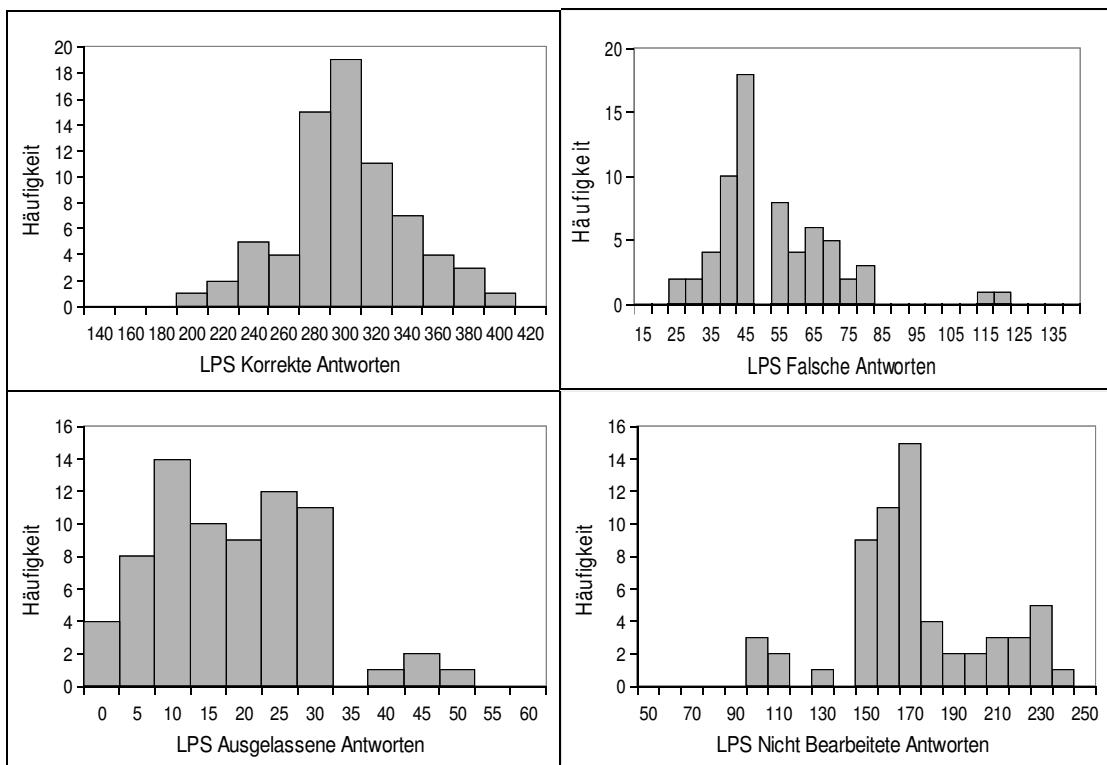
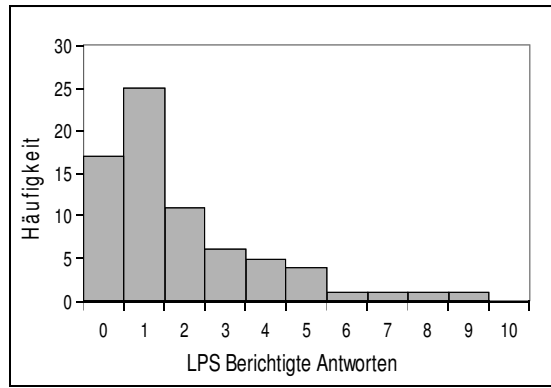


Abbildung E-2: Verteilung weiterer berechneter Skalen des LPS-neu



Fortsetzung Abbildung E-2: Verteilung weiterer berechneter Skalen des LPS-neu

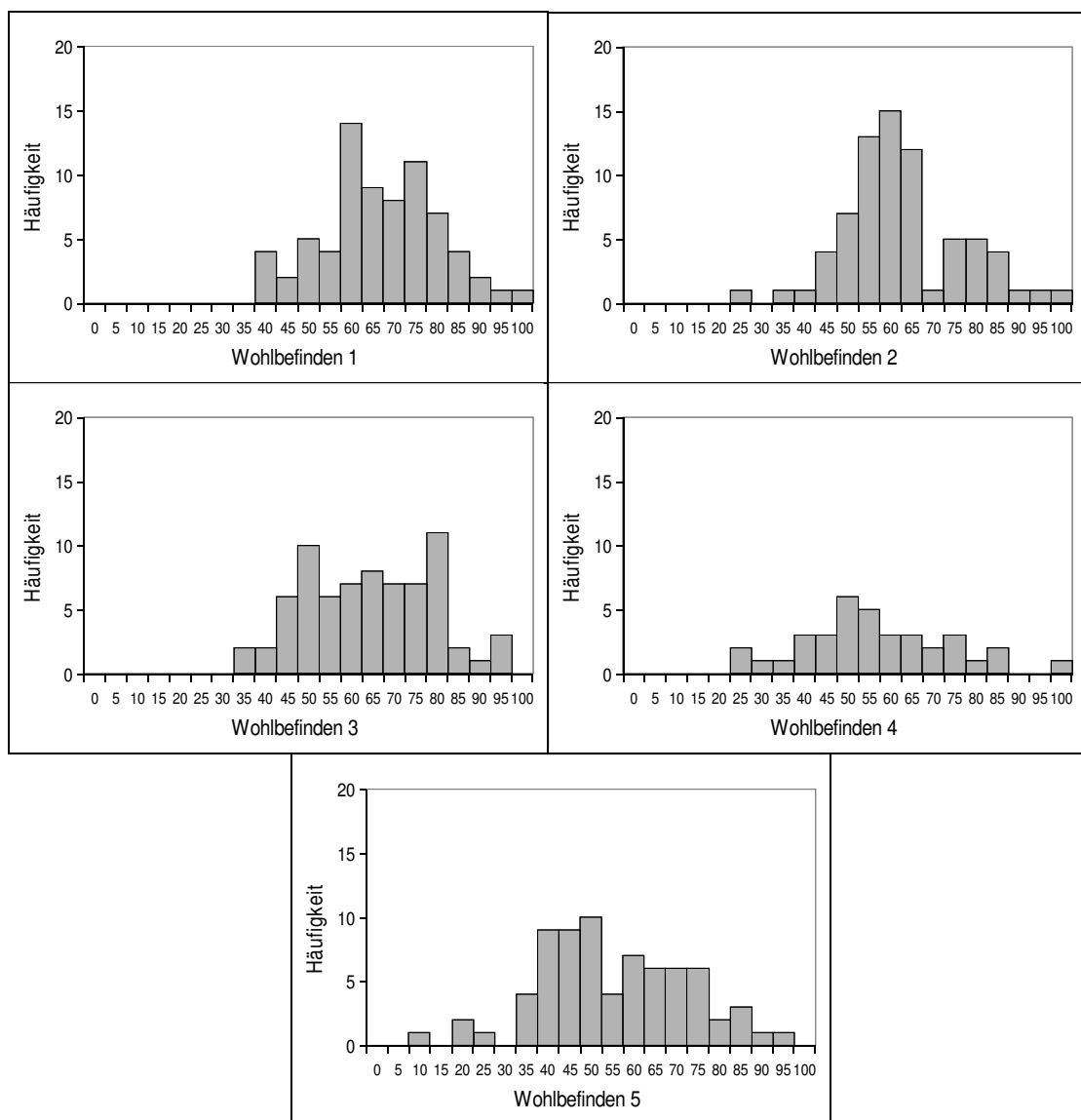


Abbildung E-3: Verteilung der Wohlbefindensfragebögen

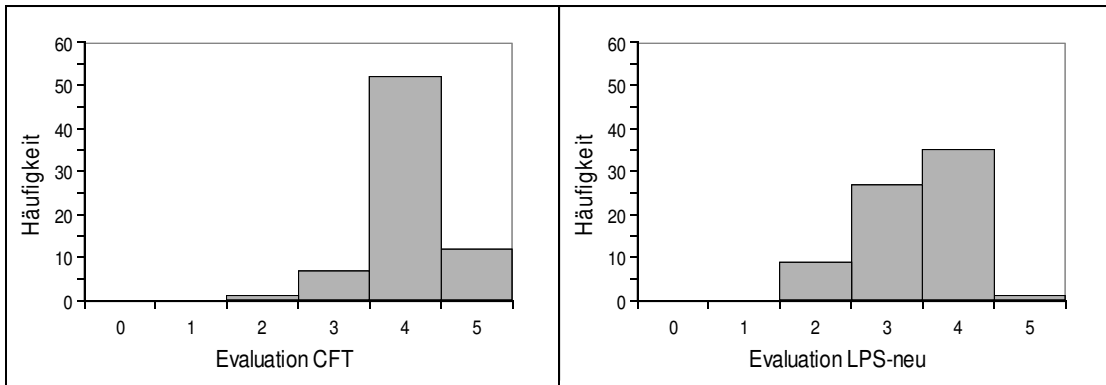


Abbildung E-4: Verteilung der Evaluationsfragebögen

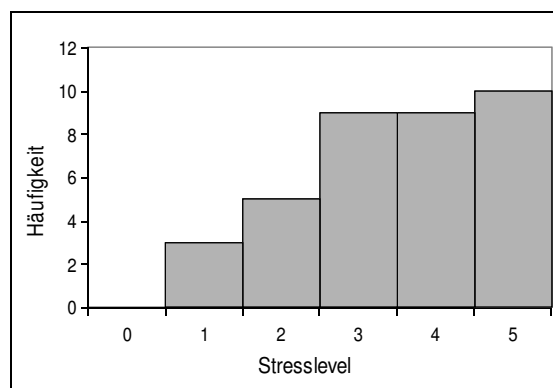


Abbildung E-5: Verteilung des gefühlten Stresslevels

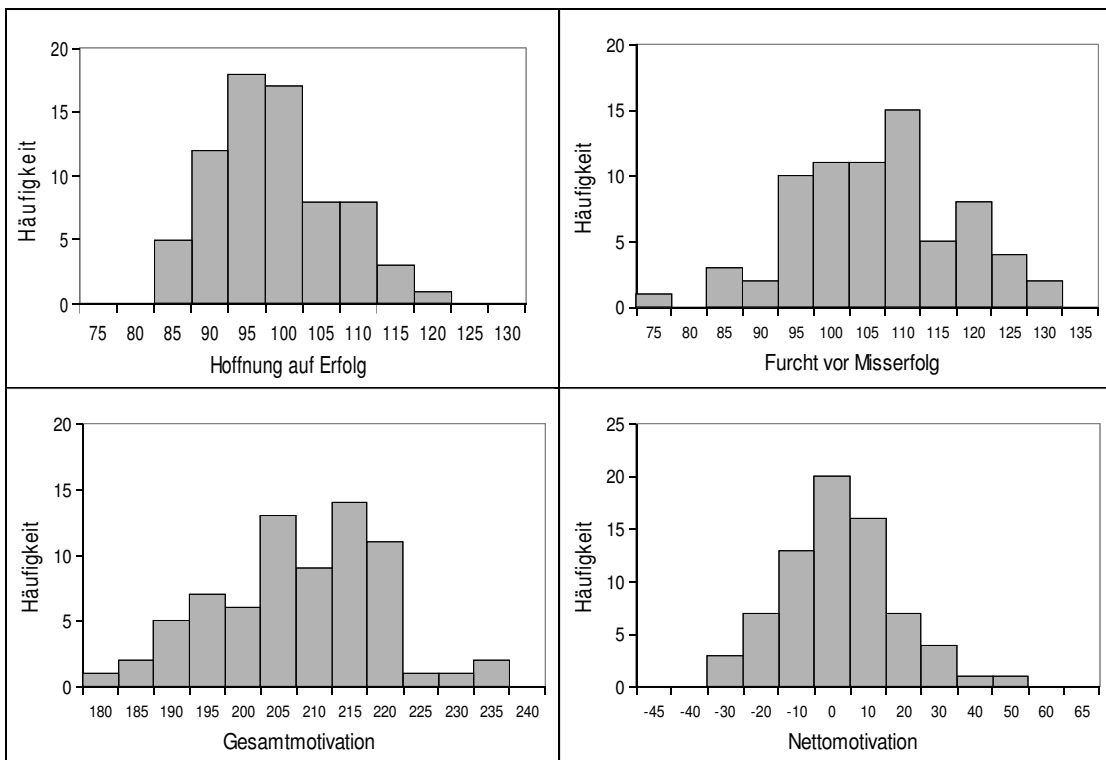


Abbildung E-6: Verteilung der RLMI-E Skalen

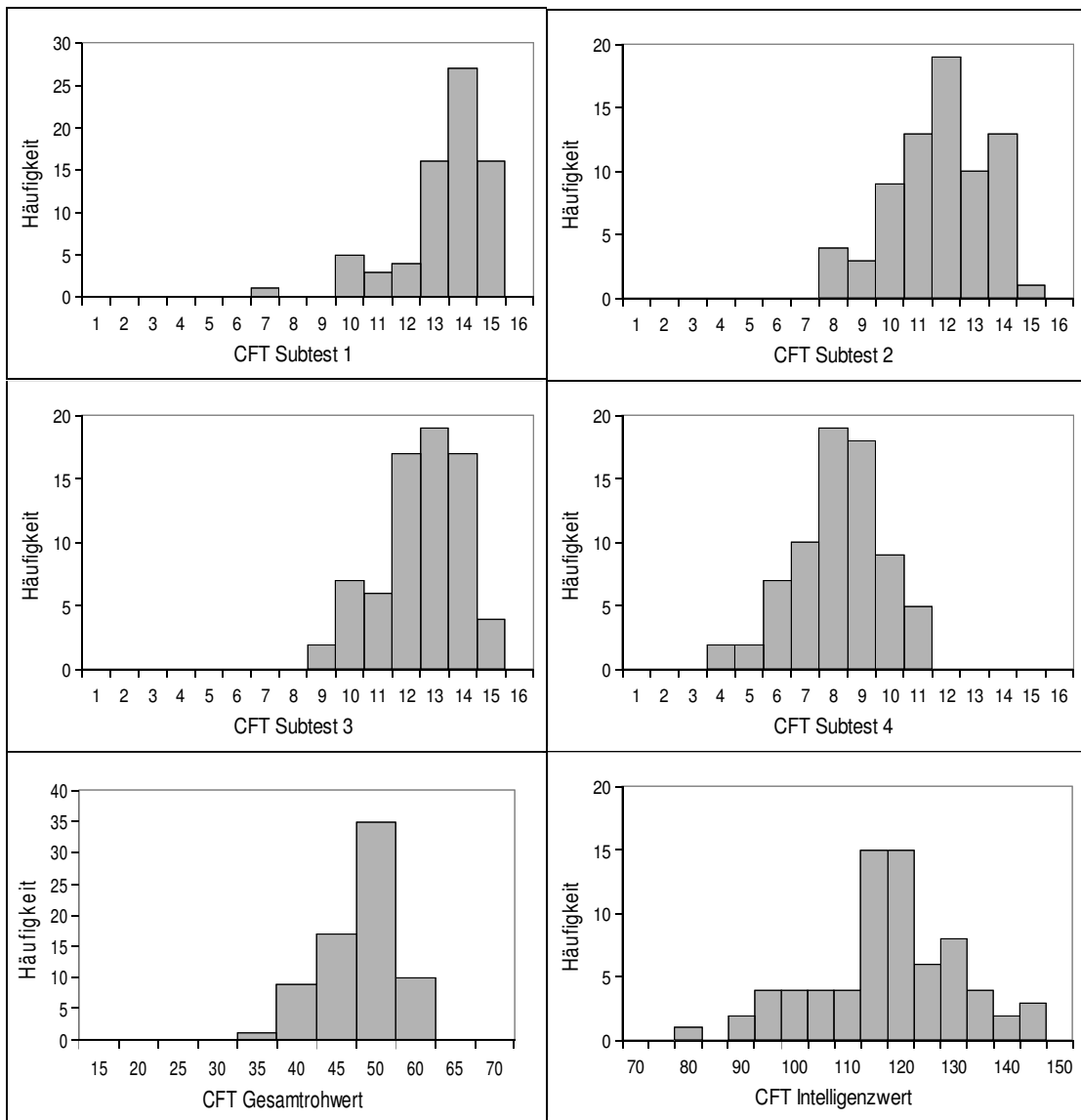


Abbildung E-7: Verteilung der CFT 20-R Subtests und des Gesamtwerts

Anhang F

Tabelle F-1: Univariate Varianzanalysen der Kontroll- und Experimentalgruppe bezüglich des Wohlbefindens

	KG, n = 36 M (SD)	EG, n = 36 M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Wohlbefinden 1	64.78 (13.95)	64.00 (13.34)	0.06	.81	.00	11.08	186.28
Wohlbefinden 2	60.04 (14.28)	60.38 (13.33)	0.01	.92	.00	2.08	190.81
Wohlbefinden 3	64.23 (14.82)	60.57 (15.27)	1.06	.31	.02	240.17	226.37
Wohlbefinden 5	56.68 (15.68)	49.19 (18.08)	3.53	.07°	.05	1009.69	286.46

Anmerkungen: KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 70, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-2: Univariate Varianzanalysen zur Prüfung der Leistungsmotivgruppe bezüglich des Wohlbefindens

	HE, n = 31 M (SD)	FM, n = 41 M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Wohlbefinden 1	67.62 (14.65)	61.95 (12.29)	0.64	.43	.02	115.02	179.90
Wohlbefinden 2	62.06 (16.92)	58.80 (10.71)	0.27	.61	.01	49.46	181.57
Wohlbefinden 3	63.08 (16.72)	61.88 (13.85)	0.06	.81	.00	14.13	239.64
Wohlbefinden 4	56.98 (19.41)	52.30 (14.21)	0.68	.41	.02	196.12	284.19
Wohlbefinden 5	55.70 (19.54)	50.85 (15.16)	1.20	.28	.03	391.32	325.14

Anmerkungen: M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 34, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers

Tabelle F-3: Univariate Varianzanalysen der Leistungsmotivgruppen bezüglich des gefühlten Stresslevels

	HE, n = 17 M (SD)	FM, n = 19 M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Frustration	3.24 (1.20)	4.26 (1.05)	7.54	.01**	.18	9.48	1.26
Ärger	2.88 (1.36)	3.89 (1.15)	5.84	.02*	.15	9.20	1.58
Druck	2.76 (1.52)	3.63 (1.26)	3.50	.07°	.09	6.74	1.93
Stress Gesamt	2.96 (1.18)	3.93 (1.03)	6.95	.01**	.17	8.425	1.212

Anmerkungen: M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 34, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers
**) sehr signifikant auf dem Niveau von .01 (zweiseitig)
*) signifikant auf dem Niveau von .05 (zweiseitig)
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-4: Zweifaktorielle Kovarianzanalyse der Kontroll- und Experimentalgruppe und der Leistungsmotivgruppen bezüglich der Bewertung

CFT 20-R		n	M	SD	F	p	$\rho\eta^2$	MQS
Gruppe	KG	36	4.08	0.49	0.10	.75	.00	0.02
	EG	36	4.06	0.46				
LM-Gruppe	HE	41	4.08	0.49	0.16	.69	.00	0.04
	FM	31	4.09	0.40				
Interaktion					0.05	.82	.00	0.07
Kovariate					0.30	.59	.00	0.01
LPS-neu		n	M	SD	F	p	$\rho\eta^2$	MQS
Gruppe	KG	36	3.50	0.54	1.86	.18	.03	0.70
	EG	36	3.32	0.66				
LM-Gruppe	HE	41	3.34	0.77	0.45	.50	.01	0.17
	FM	31	3.46	0.45				
Interaktion					0.05	.82	.00	0.02
Kovariate					0.44	.51	.01	0.16

Anmerkungen: n = Anzahl der Vpn, LM-Gruppe = Leistungsmotivgruppe, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Kovariate = CFT 20-R IQ-Wert, MQS = Mittlere Quadratsumme, Mittlere Quadratsumme des Fehlers = 0.38

Tabelle F-5: Univariate Varianzanalysen der Gruppen HE (n = 14) und FM (n = 22) bezüglich LPS-neu Leistung in neutraler Situation

	HE, M(SD)	FM, M(SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Korrekt	288.57 (42.46)	282.27 (41.59)	0.19	.66	.01	339.43	1757.88
Fehler	51.86 (13.04)	51.68 (17.85)	0.00	.97	.00	0.26	261.78
Ausgelassen	18.21 (10.96)	16.09 (9.51)	0.38	.54	.01	38.57	101.77
Nicht Bearbeitet	161.36 (33.20)	169.32 (38.95)	0.40	.53	.01	542.24	1358.59
Berichtigt	1.79 (1.76)	1.50 (1.37)	0.30	.59	.01	0.70	2.35
	M(SD)	M(SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Subtest 1	39.50 (8.27)	39.73 (9.90)	0.01	.94	.00	0.44	86.64
Subtest 2	21.86 (7.53)	26.27 (7.45)	2.98	.09°	.08	166.81	55.94
Subtest 3	27.79 (3.77)	25.36 (3.55)	3.80	.06°	.10	50.19	13.22
Subtest 4	16.86 (3.13)	17.00 (2.33)	0.02	.88	.00	0.17	7.11
Subtest 5	20.93 (3.27)	20.18 (3.28)	0.45	.51	.01	4.77	10.71
Subtest 6	26.29 (7.54)	25.36 (7.64)	0.13	.72	.00	7.27	57.76
Subtest 7	32.14 (4.50)	31.27 (3.12)	0.47	.50	.01	6.48	13.77
Subtest 8	33.71 (7.16)	32.05 (4.88)	0.69	.41	.02	23.83	34.35
Subtest 9	14.07 (3.85)	12.68 (6.37)	0.54	.47	.02	16.52	30.76
Subtest 10	36.14 (5.28)	33.82 (6.68)	1.21	.28	.03	46.24	38.21
Subtest 11	19.29 (4.94)	18.55 (8.75)	0.08	.78	.00	4.69	56.66

Anmerkungen: M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 34, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers
°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-6: Univariate Kovarianzanalysen der experimentellen Gruppen und der Motivgruppen bezüglich der Leistung im LPS-neu

	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>F</i>	<i>p</i>	$\rho\eta^2$	<i>MQS</i>	<i>MQS_E</i>
Gruppe	KG (n = 36)	EG (n =36)					
Korrekt	284.72 (41.44)	298.19 (35.34)	0.49	.49	.01	502.21	1027.45
Fehler	51.75 (15.95)	50.58 (18.14)	0.51	.48	.01	141.09	277.58
Ausgelassen	16.92 (10.00)	15.42 (11.87)	0.05	.83	.00	5.77	122.61
Nicht Bearbeitet	166.22 (36.54)	155.81 (32.98)	0.20	.65	.00	197.01	965.50
Berichtigt	1.61 (1.52)	2.14 (2.36)	0.77	.38	.01	3.13	4.07
Motiv-Gruppe	HE (n = 31)	FM (n = 41)					
Korrekt	295.90 (43.37)	288.10 (35.21)	0.94	.34	.01	963.61	1027.45
Fehler	52.32 (19.54)	50.29 (14.93)	0.24	.63	.00	65.87	277.58
Ausgelassen	17.52 (12.03)	15.15 (10.04)	0.87	.36	.01	106.54	122.61
Nicht Bearbeitet	154.26 (37.11)	166.12 (32.77)	2.40	.13	.04	2317.80	965.50
Berichtigt	2.10 (2.23)	1.71 (1.79)	0.52	.47	.01	2.11	4.07
Interaktion¹⁾							
Korrekt			0.21	.65	.00	219.96	1027.45
Fehler			0.09	.76	.00	26.17	277.58
Ausgelassen			0.01	.94	.00	0.76	122.61
Nicht Bearbeitet			0.51	.48	.01	489.62	965.50
Berichtigt			0.02	.89	.00	0.08	4.07
Kovariante IQ-Wert							
Korrekt			33.27	<.01**	.33	34178.28	1027.45
Fehler			6.03	.02**	.08	1674.39	277.58
Ausgelassen			0.84	.36	.01	103.07	122.61
Nicht Bearbeitet			18.45	<.01**	.22	17818.01	965.50
Berichtigt			0.05	.82	.00	0.22	4.07

Anmerkungen: KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Testwert, *p* = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, *MQS* = Mittlere Quadratsumme, *MQS_E* = Mittlere Quadratsumme des Fehlers, Kovariante = CFT 20-R IQ-Wert
¹⁾ für *M* und *SD* der Interaktion vgl. Tabelle 4.20

Tabelle F-7: Univariate Kovarianzanalysen der experimentellen Gruppen und der Motivgruppen bezüglich der LPS-neu Tests

	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>F</i>	<i>p</i>	$\rho\eta^2$	<i>MQS</i>	<i>MQS_E</i>	
Gruppe	KG (n = 36)	EG (n = 36)						
Subtest 1	39.64 (9.18)	44.12 (10.32)	1.29	.26	.02	123.53	95.68	
Subtest 2	24.56 (7.69)	27.72 (6.78)	0.25	.62	.00	10.31	40.48	
Subtest 3	26.31 (3.78)	26.08 (3.21)	3.36	.07°	.05	35.11	10.45	
Subtest 4	16.94 (2.63)	18.08 (2.95)	0.10	.75	.00	0.55	5.54	
Subtest 5	20.47 (3.25)	20.50 (4.02)	4.01	.05*	.06	40.02	9.99	
Subtest 6	25.72 (7.51)	25.94 (7.35)	0.29	.59	.00	15.68	53.46	
Subtest 7	31.61 (3.68)	30.83 (4.12)	7.74	.01**	.10	93.02	12.02	
Subtest 8	32.69 (5.84)	33.72 (4.61)	0.55	.46	.01	12.98	23.67	
Subtest 9	13.22 (5.51)	14.33 (5.10)	0.62	.44	.01	14.63	23.74	
Subtest 10	34.72 (6.20)	35.89 (4.40)	0.06	.80	.00	1.66	26.73	
Subtest 11	18.83 (7.43)	20.94 (7.54)	0.04	.84	.00	2.02	49.08	
Motivgruppe	HE (n =31)	FM (n = 41)						
Subtest 1	42.03 (10.11)	41.78 (9.96)	0.00	.98	.00	0.09	95.68	
Subtest 2	25.39 (8.08)	26.71 (6.83)	1.05	.31	.02	42.45	40.48	
Subtest 3	26.97 (3.66)	25.61 (3.27)	3.53	.06°	.05	36.88	10.45	
Subtest 4	17.52 (3.33)	17.51 (2.44)	0.01	.94	.00	0.03	5.54	
Subtest 5	20.97 (4.08)	20.12 (3.25)	1.45	.23	.02	14.46	9.99	
Subtest 6	27.16 (7.66)	24.83 (7.08)	1.80	.18	.03	96.29	53.46	
Subtest 7	31.35 (4.67)	31.12 (3.25)	0.21	.65	.00	2.48	12.02	
Subtest 8	33.10 (6.39)	33.29 (4.27)	0.02	.88	.00	0.53	23.67	
Subtest 9	14.10 (4.68)	13.54 (5.77)	0.24	.63	.00	5.58	23.74	
Subtest 10	36.48 (5.20)	34.41 (5.39)	2.73	.10°	.04	72.98	26.73	
Subtest 11	20.84 (6.96)	19.17 (7.91)	0.89	.35	.01	43.48	49.08	
Interaktion¹⁾				Kovariate IQ-Wert				
	<i>F</i>	<i>p</i>	$\rho\eta^2$	<i>MQS</i>	<i>F</i>	<i>p</i>	$\rho\eta^2$	<i>MQS</i>
Subtest 1	0.02	.88	.00	2.37	2.72	.10°	.04	259.91
Subtest 2	4.49	.04*	.06	181.65	19.47	<.01**	.23	788.39
Subtest 3	1.16	.29	.02	12.12	10.37	<.01**	.13	108.41
Subtest 4	0.24	.63	.00	1.31	31.71	<.01**	.32	175.54
Subtest 5	0.24	.63	.00	2.40	25.24	<.01**	.27	252.08
Subtest 6	0.82	.37	.01	43.76	2.83	.10°	.04	151.27
Subtest 7	0.12	.73	.00	1.40	21.21	<.01**	.24	255.05
Subtest 8	1.93	.17	.03	45.77	12.07	<.01**	.15	285.78
Subtest 9	0.24	.62	.00	5.74	15.33	<.01**	.19	363.81
Subtest 10	0.01	.93	.00	0.24	6.02	.02*	.08	160.91
Subtest 11	0.47	.49	.01	23.21	11.93	<.01**	.15	585.72

Anmerkungen: KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Testwert, *p* = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 67, *MQS* = Mittlere Quadratsumme, *MQS_E* = Mittlere Quadratsumme des Fehlers, Kovariate = CFT 20-R IQ-Wert
¹⁾ für *M* und *SD* der Interaktion vgl. Tabelle 4.28

***) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-8: Univariate Kovarianzanalysen der Leistungsmotivgruppen bezüglich der LPS-neu Leistung in der Stressbedingung

	HE, n = 17 M (SD)	FM, n = 19 M (SD)	F	p	$p\eta^2$	MQS	MQS _E
Korrekt	301.94 (44.45)	294.84 (25.43)	1.36	.25	.04	1087.61	797.15
Fehler	52.71 (24.02)	48.68 (10.90)	0.27	.61	.01	86.89	318.22
Ausgelassen	16.94 (13.15)	14.05 (10.78)	0.37	.55	.01	52.25	140.11
Nicht Bearbeitet	148.41 (40.08)	162.42 (24.28)	2.79	.10°	.08	2453.15	879.67
Berichtigt	2.35(2.57)	1.95 (2.20)	0.25	.62	.01	1.44	5.84
Kovariante IQ-Wert							
Korrekt			21.26	<.01**	.39	16943.55	797.15
Fehler			2.74	.11	.08	870.50	318.22
Ausgelassen			1.67	.20	.05	234.30	140.11
Nicht Bearbeitet			8.28	.01*	.20	7285.72	879.67
Berichtigt			0.00	.96	.00	0.01	5.84
	M (SD)	M (SD)	F	p	$p\eta^2$	MQS	MQS _E
Subtest 1	44.12 (11.21)	44.16 (9.75)	0.01	.94	.00	0.55	110.10
Subtest 2	28.29 (7.51)	27.21 (6.21)	0.78	.38	.02	26.94	34.32
Subtest 3	26.29 (3.53)	25.89 (2.98)	0.36	.55	.01	3.40	9.32
Subtest 4	18.06 (3.47)	18.11 (2.49)	0.11	.74	.00	0.66	5.92
Subtest 5	21.00 (4.74)	20.05 (3.31)	1.89	.18	.05	17.98	9.50
Subtest 6	27.88 (7.92)	24.21 (6.52)	2.57	.12	.07	134.08	52.16
Subtest 7	30.71 (4.84)	30.95 (3.47)	0.00	.96	.00	0.04	14.80
Subtest 8	32.59 (5.85)	34.74 (2.94)	1.65	.21	.05	30.31	18.32
Subtest 9	14.12 (5.38)	14.53 (4.97)	0.00	.99	.00	0.00	21.50
Subtest 10	36.76 (5.27)	35.11 (3.40)	1.99	.17	.06	33.83	17.02
Subtest 11	22.12 (8.19)	19.89 (6.97)	1.23	.28	.04	63.93	52.15
Kovariante IQ-Wert							
Subtest 1			21.26	<.01**	.39	90.90	110.10
Subtest 2			0.27	.61	.01	466.12	34.32
Subtest 3			21.26	<.01**	.39	51.80	9.32
Subtest 4			0.27	.61	.01	109.37	5.92
Subtest 5			0.37	.55	.01	243.47	9.50
Subtest 6			0.27	.61	.01	47.68	52.16
Subtest 7			0.37	.55	.01	104.15	14.80
Subtest 8			2.74	.11	.08	99.39	18.32
Subtest 9			1.67	.20	.05	198.98	21.50
Subtest 10			8.28	.01**	.20	91.21	17.02
Subtest 11			0.00	.96	.00	226.46	52.15

Anmerkungen: n = Anzahl der Vpn, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $p\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers, Kovariante = CFT 20-R IQ-Wert

***) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-9: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Leistung der HE-Gruppe

	KG, n = 14 M (SD)	EG, n = 17 M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Korrekt	288.57 (42.46)	301.94 (44.45)	0.14	.71	.01	153.08	1096.05
Fehler	51.86 (13.04)	52.71 (24.02)	0.46	.50	.02	168.17	364.58
Ausgelassen	18.21 (10.96)	16.94 (13.15)	0.00	.99	.00	0.01	150.09
Nicht Bearbeitet	161.36 (33.20)	148.41 (40.08)	0.00	.98	.00	0.46	1000.25
Berichtigt	1.79 (1.76)	2.35(2.57)	0.68	.42	.02	3.50	5.17
Kovariante IQ-Wert							
Korrekt			22.23	<.01**	.44	24361.04	1096.05
Fehler			3.39	.08°	.11	1234.97	364.58
Ausgelassen			0.85	.37	.03	126.88	150.09
Nicht Bearbeitet			12.03	<.01**	.30	12028.43	1000.25
Berichtigt			0.29	.59	.01	1.51	5.17
	M (SD)	M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Subtest 1	39.50 (8.27)	44.12 (11.21)	0.33	.57	.01	28.28	85.55
Subtest 2	21.86 (7.53)	28.29 (7.51)	2.91	.10°	.09	141.80	48.74
Subtest 3	27.79 (3.77)	26.29 (3.53)	4.02	.05*	.13	43.96	10.93
Subtest 4	16.86 (3.13)	18.06 (3.47)	0.00	.96	.00	0.02	7.29
Subtest 5	20.93 (3.27)	21.00 (4.74)	1.95	.17	.07	19.16	9.83
Subtest 6	26.29 (7.54)	27.88 (7.92)	0.00	.96	.00	0.13	54.21
Subtest 7	32.14 (4.50)	30.71 (4.84)	3.74	.06°	.12	61.75	16.49
Subtest 8	33.71 (7.16)	32.59 (5.85)	2.00	.17	.07	66.24	33.16
Subtest 9	14.07 (3.85)	14.12 (5.38)	0.29	.59	.01	6.06	20.90
Subtest 10	36.14 (5.27)	36.76 (5.27)	0.37	.55	.01	7.89	21.22
Subtest 11	19.29 (4.94)	22.12 (8.19)	0.24	.63	.01	10.02	42.56
Kovariante IQ-Wert							
Subtest 1			5.91	.02*	.17	505.79	85.55
Subtest 2			5.67	.02*	.17	276.49	48.74
Subtest 3			7.13	.01**	.20	77.93	10.93
Subtest 4			15.99	<.01**	.36	116.56	7.29
Subtest 5			22.75	<.01**	.45	223.66	9.83
Subtest 6			4.14	.05*	.13	224.69	54.21
Subtest 7			10.76	<.01**	.28	177.41	16.49
Subtest 8			8.64	.01**	.24	286.55	33.16
Subtest 9			3.43	.07°	.11	71.59	20.90
Subtest 10			10.01	<.01**	.26	212.48	21.22
Subtest 11			4.68	.04*	.14	199.05	42.56

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 28, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers, Kovariante = CFT 20-R IQ-Wert

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-10: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Leistung der FM-Gruppe

	KG, n = 22 M (SD)	EG, n = 19 M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Korrekt	282.27 (41.59)	294.84 (25.43)	0.24	.62	.01	233.12	979.15
Fehler	51.68 (17.85)	48.68 (10.90)	0.03	.87	.00	6.09	219.23
Ausgelassen	16.09 (9.51)	14.05 (10.78)	0.18	.68	.01	18.70	104.76
Nicht Bearbeitet	169.32 (38.95)	162.42 (24.28)	0.36	.55	.01	348.78	957.04
Berichtigt	1.50 (1.37)	1.95 (2.20)	0.02	.88	.00	0.07	3.22
Kovariante IQ-Wert							
Korrekt			10.99	<.01**	.22	10759.08	979.15
Fehler			2.27	.14	.06	498.19	219.23
Ausgelassen			0.08	.79	.00	7.91	104.76
Nicht Bearbeitet			6.38	.02*	.14	6104.04	957.04
Berichtigt			1.25	.27	.03	4.03	3.22
	M (SD)	M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Subtest 1	39.73 (9.90)	44.16 (9.75)	1.60	.21	.04	158.61	99.18
Subtest 2	26.27 (7.45)	27.21 (6.21)	2.19	.15	.05	75.87	34.61
Subtest 3	25.36 (3.55)	25.89 (2.98)	0.17	.68	.00	1.78	10.30
Subtest 4	17.00 (2.33)	18.11 (2.49)	0.12	.73	.00	0.53	4.32
Subtest 5	20.18 (3.28)	20.05 (3.31)	1.55	.22	.04	15.04	9.73
Subtest 6	25.36 (7.64)	24.21 (6.52)	0.33	.57	.01	17.42	52.27
Subtest 7	31.27 (3.12)	30.95 (3.47)	3.21	.08°	.08	28.55	8.89
Subtest 8	32.05 (4.88)	34.74 (2.94)	1.14	.29	.03	18.40	16.12
Subtest 9	12.68 (6.37)	14.53 (4.97)	0.65	.42	.02	16.16	24.76
Subtest 10	33.82 (6.68)	35.11 (3.40)	0.15	.70	.00	4.62	29.90
Subtest 11	18.55 (8.75)	19.89 (6.97)	0.69	.41	.02	37.38	54.45
Kovariante IQ-Wert							
Subtest 1			0.00	.96	.00	0.22	99.18
Subtest 2			15.73	<.01**	.29	544.51	34.61
Subtest 3			3.27	.08°	.08	33.66	10.30
Subtest 4			14.27	<.01**	.27	61.63	4.32
Subtest 5			5.40	.03*	.12	52.51	9.73
Subtest 6			0.07	.79	.00	3.87	52.27
Subtest 7			9.39	<.01**	.20	83.45	8.89
Subtest 8			2.74	.11	.07	44.09	16.12
Subtest 9			14.41	<.01**	.27	356.73	24.76
Subtest 10			0.30	.59	.01	8.85	29.90
Subtest 11			7.61	.01**	.17	414.24	54.45

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 38, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers, Kovariante = CFT 20-R IQ-Wert

***) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-11: Univariate Kovarianzanalysen der Gruppen niedriger GM (n = 36) und hoher GM (n = 36) bezüglich LPS-neu Leistung

	Niedrige GM <i>M</i> (<i>SD</i>)	Hohe GM <i>M</i> (<i>SD</i>)	<i>F</i>	<i>p</i>	$\rho\eta^2$	<i>MQS</i>	<i>MQS_E</i>
Korrekt	293.00 (41.70)	289.92 (36.27)	0.38	.54	.01	382.85	1016.26
Fehler	53.36 (20.00)	48.97 (13.20)	2.52	.12	.04	663.96	263.41
Ausgelassen	16.17 (11.54)	16.17 (10.43)	0.04	.85	.00	4.32	120.60
Nicht Bearbeitet	157.47 (33.86)	164.56 (36.15)	0.05	.83	.00	45.98	980.45
Berichtigt	2.00 (2.24)	1.75 (1.71)	0.18	.67	.00	0.73	4.02
Kovariante IQ-Wert							
Korrekt			36.19	<.01**	.34	36782.68	1016.26
Fehler			7.29	.01**	.10	1919.80	263.41
Ausgelassen			1.25	.27	.02	150.87	120.60
Nicht Bearbeitet			18.57	<.01**	.21	18208.67	980.45
Berichtigt			0.32	.58	.00	1.27	4.02
Subtest							
Subtest 1	42.67 (10.11)	41.11 (9.88)	0.09	.77	.00	8.14	94.58
Subtest 2	26.39 (8.21)	25.89 (6.54)	0.22	.64	.00	9.41	42.39
Subtest 3	26.17 (3.44)	26.22 (3.57)	0.30	.58	.00	3.38	11.18
Subtest 4	17.67 (3.00)	17.36 (2.68)	0.22	.64	.00	1.20	5.39
Subtest 5	20.92 (4.14)	20.06 (3.02)	0.14	.71	.00	1.50	10.48
Subtest 6	26.47 (6.91)	25.19 (7.86)	0.24	.63	.00	12.95	54.01
Subtest 7	31.31 (4.37)	31.14 (3.42)	0.19	.66	.00	2.50	12.99
Subtest 8	33.00 (5.85)	33.42 (4.64)	1.00	.32	.01	23.41	23.45
Subtest 9	13.25 (5.49)	14.31 (5.13)	2.77	.10°	.04	62.36	22.48
Subtest 10	35.06 (5.74)	35.56 (5.05)	0.75	.39	.01	20.03	26.73
Subtest 11	20.11 (7.89)	19.67 (7.21)	0.12	.73	.00	5.99	48.58
Kovariante IQ-Wert							
Subtest 1			4.92	.03*	.07	465.31	94.58
Subtest 2			21.92	<.01**	.24	929.12	42.39
Subtest 3			8.05	.01**	.10	90.02	11.18
Subtest 4			36.47	<.01**	.35	196.50	5.39
Subtest 5			18.82	<.01**	.21	197.30	10.48
Subtest 6			1.97	.17	.03	106.25	54.01
Subtest 7			13.95	<.01**	.17	181.32	12.99
Subtest 8			14.27	<.01**	.17	334.69	23.45
Subtest 9			18.85	<.01**	.21	423.60	22.48
Subtest 10			7.43	.01**	.10	198.54	26.73
Subtest 11			13.33	<.01**	.16	647.40	48.58

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Testwert, *p* = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 69, *MQS* = Mittlere Quadratsumme, *MQS_E* = Mittlere Quadratsumme des Fehlers, Kovariante = CFT 20-R IQ-Wert

***) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-12: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Leistung der Gruppe mit hoher GM

	KG, n = 19 M (SD)	EG, n = 17 M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Korrekt	283.74 (37.02)	296.82 (35.21)	0.20	.66	.01	207.22	1028.95
Fehler	48.89 (11.35)	49.06 (15.37)	0.24	.63	.01	43.13	179.71
Ausgelassen	17.32 (11.01)	14.88 (9.93)	0.03	.87	.00	3.01	104.99
Nicht Bearbeitet	169.32 (37.93)	159.24 (34.39)	0.05	.82	.00	64.80	1197.81
Berichtigt	1.53 (1.26)	2.00 (2.12)	0.01	.92	.00	0.03	2.78
Kovariante IQ-Wert							
Korrekt			10.26	<.01**	.24	10552.90	1028.95
Fehler			0.93	.34	.03	166.28	179.71
Ausgelassen			2.79	.10°	.08	293.22	104.99
Nicht Bearbeitet			4.42	.04*	.12	5291.54	1197.81
Berichtigt			3.23	.08*	.09	8.99	2.78
Subtest							
Subtest 1	39.47 (9.13)	42.94 (10.63)	0.57	.46	.02	56.92	99.89
Subtest 2	24.89 (5.47)	27.00 (7.58)	0.37	.55	.01	12.34	33.13
Subtest 3	26.37 (3.76)	26.06 (3.46)	1.88	.18	.05	21.62	11.53
Subtest 4	16.68 (2.69)	18.12 (2.55)	0.02	.88	.00	0.11	4.82
Subtest 5	19.79 (2.76)	20.35 (3.35)	0.04	.84	.00	0.36	9.03
Subtest 6	24.95 (8.40)	25.47 (7.47)	0.01	.90	.00	0.94	64.75
Subtest 7	31.32 (3.53)	30.94 (3.40)	6.00	.02*	.15	46.81	7.81
Subtest 8	32.47 (5.28)	34.47 (3.68)	0.07	.79	.00	1.18	16.40
Subtest 9	13.68 (5.55)	15.00 (4.68)	0.04	.83	.00	1.09	24.67
Subtest 10	35.37 (5.80)	35.76 (4.21)	0.10	.75	.00	2.69	25.99
Subtest 11	18.74 (5.65)	20.71 (8.70)	0.03	.86	.00	1.59	48.47
Kovariante IQ-Wert							
Subtest 1			0.11	.74	.00	11.19	99.89
Subtest 2			10.95	<.01**	.25	362.62	33.13
Subtest 3			5.63	.02*	.15	64.95	11.53
Subtest 4			15.48	<.01**	.32	74.68	4.82
Subtest 5			2.13	.15	.06	19.21	9.03
Subtest 6			0.38	.54	.01	24.38	64.75
Subtest 7			19.40	<.01**	.37	151.43	7.81
Subtest 8			10.85	<.01**	.25	177.87	16.40
Subtest 9			3.64	.07°	.10	89.89	24.67
Subtest 10			1.23	.28	.04	31.86	25.99
Subtest 11			3.79	.06°	.10	183.64	48.47

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers, Kovariante = CFT 20-R IQ-Wert

**) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

°) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Tabelle F-13: Univariate Kovarianzanalysen der experimentellen Gruppen bezüglich der LPS-neu Leistung der Gruppe mit niedriger GM

	KG, n = 17 M (SD)	EG, n = 19 M (SD)	F	p	$\rho\eta^2$	MQS	MQS _E
Korrekt	285.82 (47.03)	299.42 (36.37)	0.14	.71	.00	141.11	1017.71
Fehler	54.94 (19.77)	51.95 (20.63)	0.19	.66	.01	64.80	338.38
Ausgelassen	16.47 (9.06)	15.89 (13.63)	0.04	.84	.00	6.14	140.94
Nicht Bearbeitet	162.76 (35.76)	152.74 (32.29)	0.05	.83	.00	39.77	820.28
Berichtigt	1.71 (1.79)	2.26 (2.60)	1.14	.29	.03	5.77	5.05
Kovariante IQ-Wert							
Korrekt			25.17	<.01**	.43	25.17	1017.71
Fehler			8.13	.01*	.20	8.13	338.38
Ausgelassen			0.05	.82	.00	0.05	140.94
Nicht Bearbeitet			14.82	<.01**	.31	14.82	820.28
Berichtigt			1.33	.26	.04	1.33	5.05
Subtest							
Subtest 1	39.82 (9.50)	45.21 (10.20)	0.96	.33	.03	85.78	89.35
Subtest 2	24.18 (9.77)	28.37 (6.12)	0.54	.47	.02	29.14	53.78
Subtest 3	26.24 (3.91)	26.11 (3.07)	0.65	.43	.02	7.12	10.96
Subtest 4	17.24 (2.61)	18.05 (3.34)	0.11	.75	.00	0.67	6.39
Subtest 5	21.24 (3.65)	20.63 (4.62)	5.50	.03*	.14	50.70	9.22
Subtest 6	26.59 (6.51)	26.37 (7.42)	0.32	.57	.01	15.35	47.27
Subtest 7	31.94 (3.93)	30.74 (4.76)	2.71	.11	.08	44.86	16.58
Subtest 8	32.94 (6.55)	33.05 (5.33)	0.35	.56	.01	11.34	32.17
Subtest 9	12.71 (5.59)	13.74 (5.51)	0.49	.49	.01	10.00	20.39
Subtest 10	34.00 (6.72)	36.00 (4.68)	0.07	.79	.00	1.94	28.22
Subtest 11	18.94 (9.21)	21.16 (6.59)	0.00	.98	.00	0.03	52.01
Kovariante IQ-Wert							
Subtest 1			4.11	.05*	.11	367.22	89.35
Subtest 2			7.92	.01**	.19	426.03	53.78
Subtest 3			4.84	.03*	.13	53.04	10.96
Subtest 4			15.51	<.01**	.32	99.13	6.39
Subtest 5			31.77	<.01**	.49	293.08	9.22
Subtest 6			2.30	.14	.07	108.79	47.27
Subtest 7			6.49	.02*	.16	107.62	16.58
Subtest 8			4.23	.05*	.11	136.19	32.17
Subtest 9			18.27	<.01**	.36	372.46	20.39
Subtest 10			6.55	.02*	.17	184.87	28.22
Subtest 11			8.10	.01**	.20	421.08	52.01

Anmerkungen: n = Anzahl der Vpn, KG = Kontrollgruppe, EG = Experimentalgruppe, M = Mittelwert, SD = Standardabweichung, F = Testwert, p = Signifikanzwert, $\rho\eta^2$ = Effektstärke partielles Eta-Quadrat, Anzahl der Freiheitsgrade = 1, Anzahl der Fehlerfreiheitsgrade = 33, MQS = Mittlere Quadratsumme, MQS_E = Mittlere Quadratsumme des Fehlers, Kovariante = CFT 20-R IQ-Wert

***) sehr signifikant auf dem Niveau von .01 (zweiseitig)

*) signifikant auf dem Niveau von .05 (zweiseitig)

) tendenziell signifikant auf dem Niveau von .10 (zweiseitig)

Anhang G

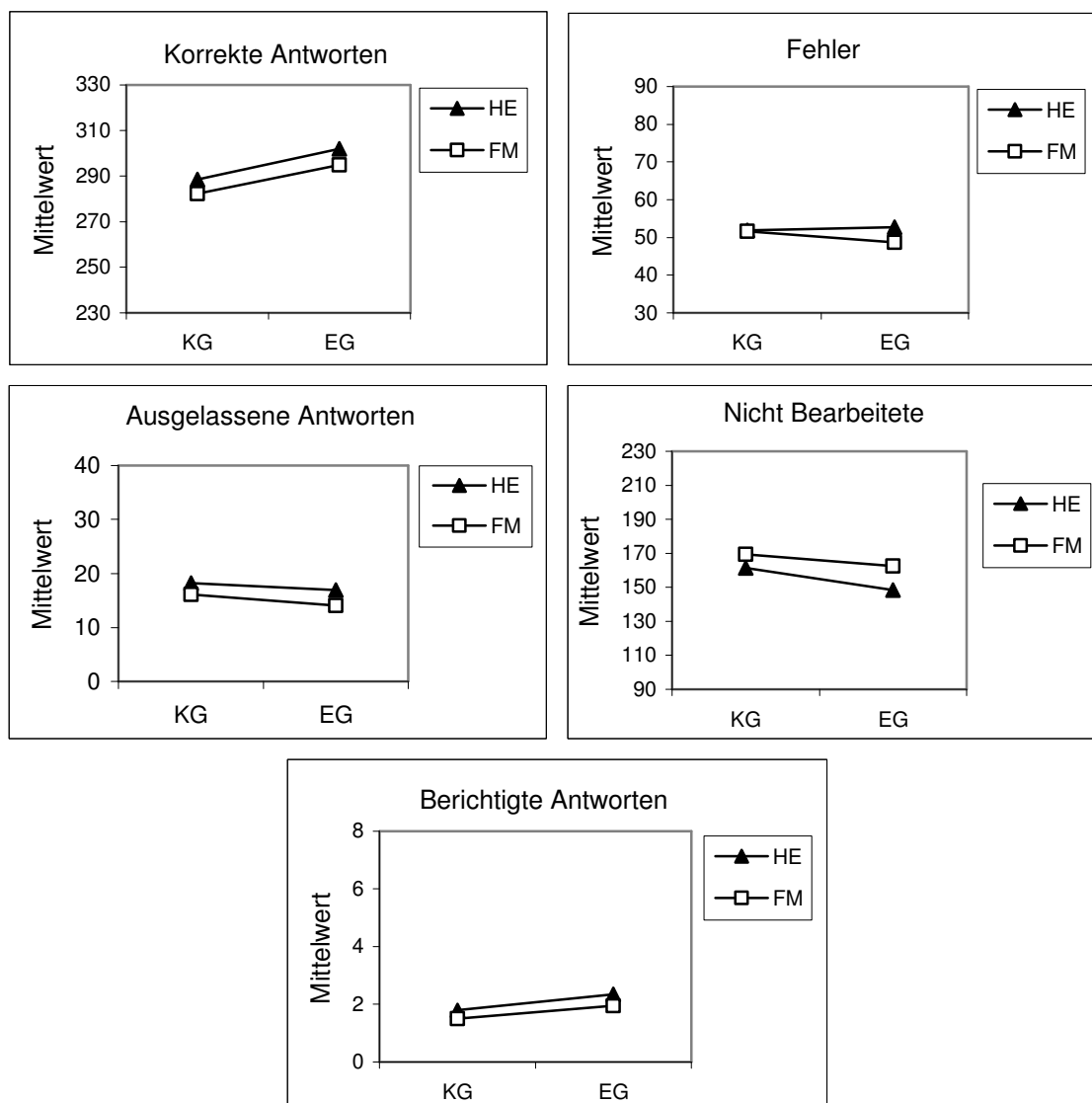


Abbildung G-1: Vergleiche der Leistungen des LPS-neu der experimentellen Gruppen und der Leistungsmotivgruppen

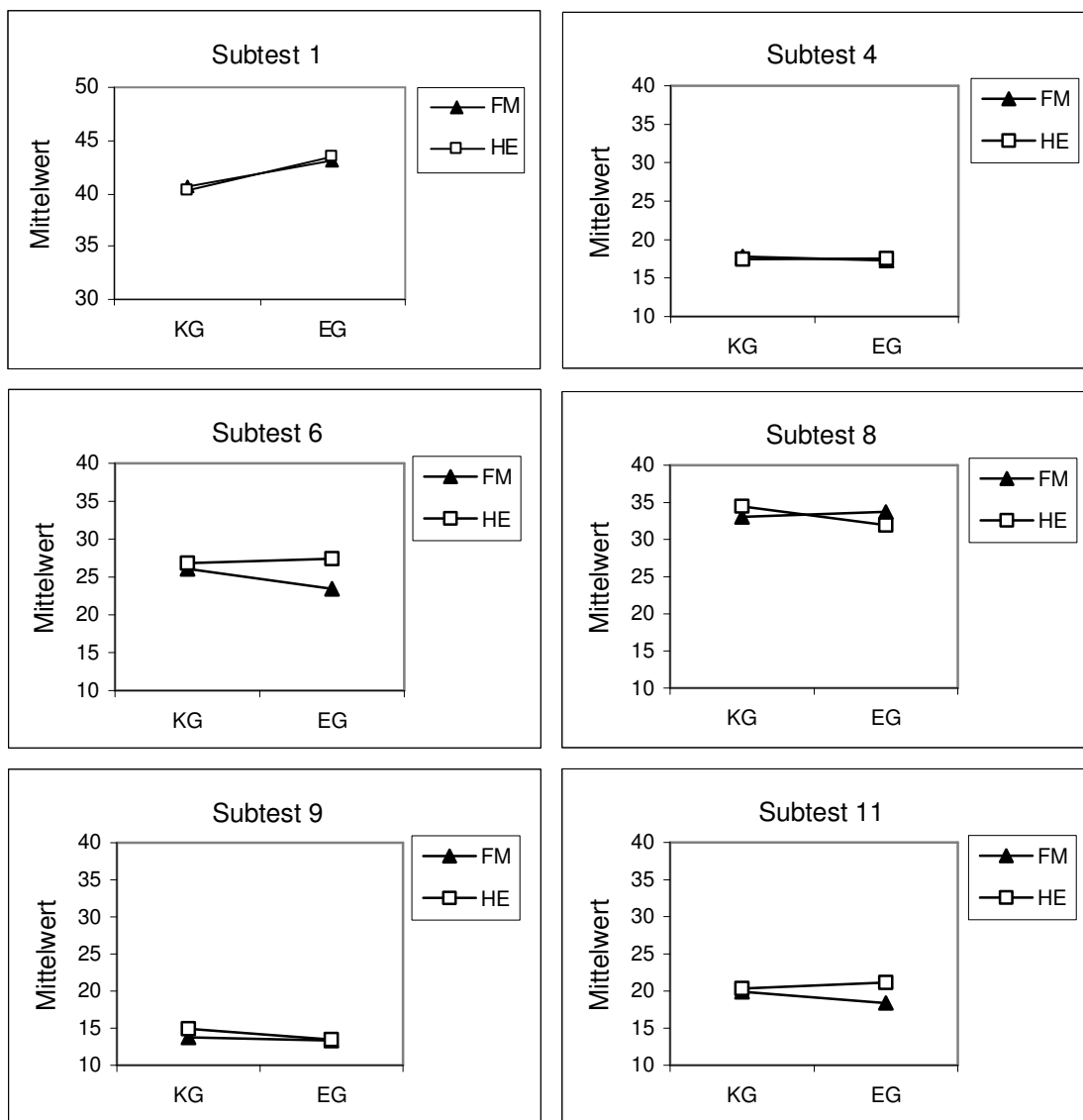


Abbildung G-2: Vergleich der LPS-neu Subtests der experimentellen Gruppen und der Leistungsmotivgruppen unter dem Einfluss der Kovariate

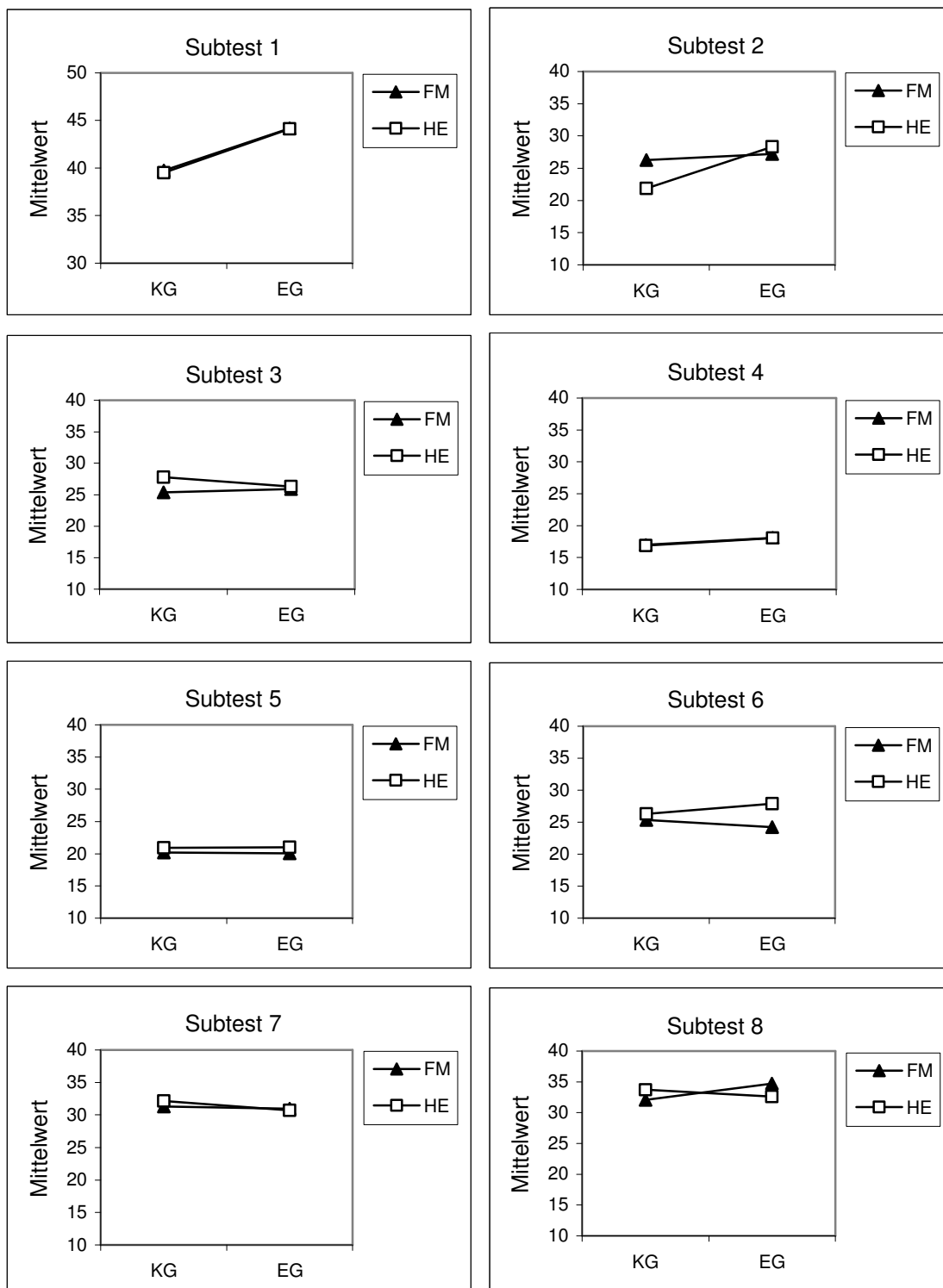
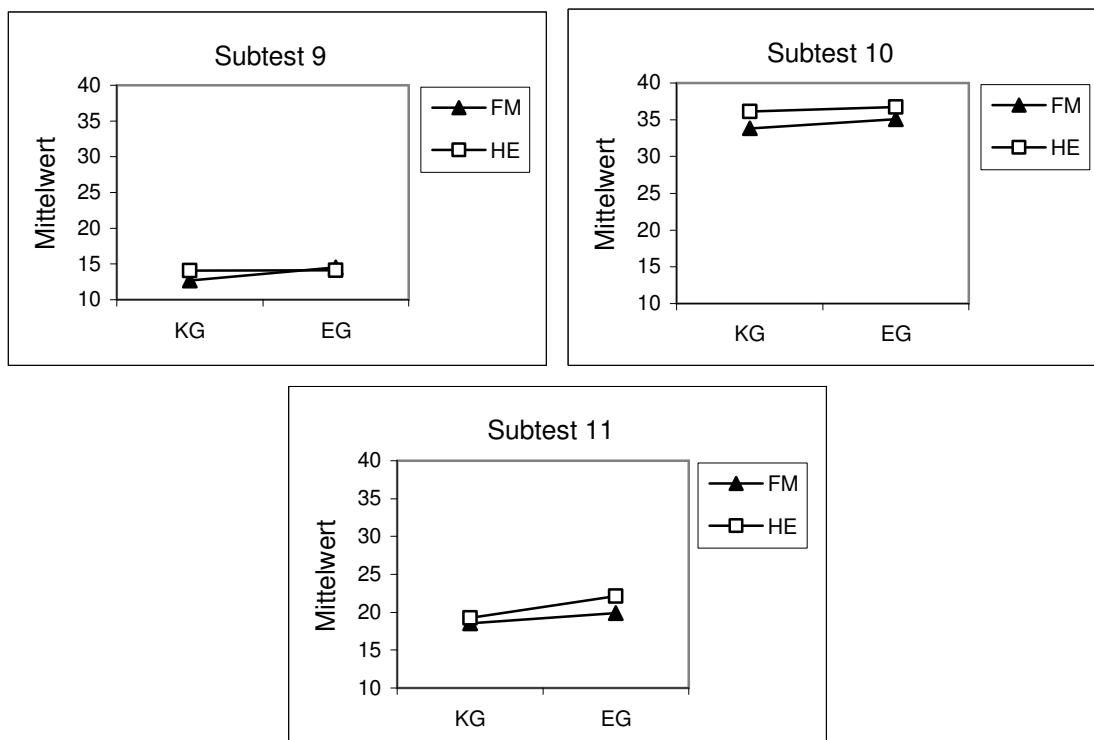


Abbildung G-3: Vergleich der LPS-neu Subtests der experimentellen Gruppen und der Leistungsmotivgruppen



Fortsetzung Abbildung G-3: Vergleich der LPS-neu Subtests der experimentellen Gruppen und der Leistungsmotivgruppen