

Der Effekt von Polarität der Textdarstellung
und Umgebungsbeleuchtung
auf die Performanz und psychophysiologische
Beanspruchung bei Bildschirmarbeit

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine Universität

Düsseldorf

vorgelegt von

Nadine Baumgartner

aus Oberhausen

Düsseldorf 2004

Aus dem Institut für Experimentelle Psychologie
der Heinrich-Heine Universität Düsseldorf
(Prof. Dr. A. Buchner)

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. A. Buchner

1. Korreferent: Prof. Dr. R. Pietrowsky

Tag der mündlichen Prüfung: 12.7.2004

Inhaltsverzeichnis

Zusammenfassung

1	Einleitung.....	8
1.1	Belastung und Beanspruchung bei Bildschirmarbeit.....	10
1.2	Psychophysiologische Analyse mentaler Beanspruchung.....	13
1.2.1	Indikatoren aus dem kardiovaskulären Bereich.....	13
1.2.1.1	Herzrate.....	14
1.2.1.2	Herzratenvariabilität.....	16
1.2.2	Elektrodermale Aktivität.....	22
1.2.2.1	Physiologie der elektrodermalen Aktivität.....	24
1.2.2.2	Parametrisierung der elektrodermalen Aktivität.....	25
1.2.2.3	Experimentelle Befunde.....	26
1.2.3	Atmung.....	27
1.2.3.1	Experimentelle Befunde.....	27
1.2.4	Psychophysiologische Beanspruchungsanalysen im Rahmen der.....	28
	Bildschirmarbeit	
1.3	Textdarstellung am Computerbildschirm.....	30
1.3.1	Vergleich zwischen Papier und Bildschirm.....	30
1.3.2	Einsatz von Farben im Rahmen der Bildschirmarbeit.....	33
1.3.2.1	Grundlagen der Farbwahrnehmung.....	33
1.3.2.1	Farbige Gestaltung von Text am Bildschirm.....	36
1.3.3	Polarität.....	39
1.3.3.1	Experimentelle Befunde zur Polarität.....	41
1.3.3.2	Polarität bei Dunkelheit.....	43
1.4	Zusammenfassung des aktuellen Forschungsstands.....	45
1.5	Hypothesen.....	46

1.5.1	Einfluss der Polarität auf die Performanz.....	46
1.5.2	Beanspruchung bei Bildschirmarbeit.....	46
2	Experiment 1	48
2.1	Methode	48
2.1.1	Stichprobenbeschreibung.....	48
2.1.2	Versuchsaufbau.....	48
2.1.2.1	Objektive Belastungsindikatoren.....	50
2.1.2.2	Subjektive Befindlichkeit und physische Beschwerden.....	51
2.1.3	Beschreibung der Stimuli.....	53
2.1.4	Versuchsablauf.....	54
2.1.5	Design.....	57
2.1.6	Auswertung.....	59
2.1.6.1	Verhaltensdaten.....	59
2.1.6.2	Physiologische Daten.....	59
2.1.6.3	Subjektive Beanspruchung.....	61
2.1.6.4	Teststärkeüberlegungen.....	62
2.1.6.5	Alpha-Adjustierung.....	62
2.2	Ergebnisse	63
2.2.1	Verhaltensdaten.....	63
2.2.1.1	Absolute Anzahl entdeckter Fehler.....	63
2.2.1.2	Anzahl gelesener Wörter.....	65
2.2.1.3	Relative Anzahl entdeckter Fehler (Genauigkeit).....	66
2.2.1.4	Fehlerarten.....	68
2.2.1.5	Genauigkeit bei der Entdeckung von Grammatikfehlern.....	69
2.2.1.6	Falsche Alarme und „Nicht übernehmen“.....	69
2.2.1.7	Subjektive Angaben.....	70
2.2.2	Objektive Beanspruchung.....	71
2.2.2.1	Herzrate.....	71
2.2.2.2	Herzratenvariabilität.....	73
2.2.2.3	Elektrodermale Aktivität.....	82
2.2.2.4	Atmung.....	83

2.2.3	Subjektive Beanspruchung.....	85
2.2.3.1	Physische Beschwerden.....	85
2.2.3.2	Befindlichkeit.....	86
2.3	Interpretation und Diskussion.....	88
2.3.1	Performanz.....	88
2.3.2	Objektive Beanspruchung.....	91
2.3.3	Subjektive Beanspruchung.....	94
3	Experiment 2.....	97
3.1	Methode.....	97
3.1.1	Stichprobenbeschreibung.....	97
3.1.2	Versuchsaufbau.....	97
3.1.2.1	Subjektive Befindlichkeit und physische Beschwerden.....	98
3.1.3	Beschreibung der Stimuli.....	98
3.1.4	Versuchsablauf.....	99
3.1.5	Design.....	99
3.1.6	Auswertung.....	99
3.1.6.1	Teststärkeüberlegungen.....	99
3.2	Ergebnisse.....	101
3.2.1	Verhaltensdaten.....	101
3.2.1.1	Absolute Anzahl entdeckter Fehler.....	101
3.2.1.2	Anzahl gelesener Wörter.....	103
3.2.1.3	Relative Anzahl entdeckter Fehler (Genauigkeit).....	104
3.2.1.4	Fehlerarten.....	105
3.2.1.5	Genauigkeit bei der Entdeckung von Grammatikfehlern.....	106
3.2.1.6	Falsche Alarme und „Nicht übernehmen“	107
3.2.1.7	Subjektive Angaben.....	107
3.2.2	Objektive Beanspruchung.....	108
3.2.2.1	Herzrate.....	108
3.2.2.2	Herzratenvariabilität.....	110

3.2.2.3	Elektrodermale Aktivität.....	117
3.2.2.4	Atmung.....	118
3.2.3	Subjektive Beanspruchung.....	120
3.2.3.1	Physische Beschwerden.....	120
3.2.3.2	Befindlichkeit.....	121
3.3	Interpretation und Diskussion.....	122
3.3.1	Performanz.....	122
3.3.2	Objektive Beanspruchung.....	124
3.3.3	Subjektive Beanspruchung.....	126
4	Experiment 3.....	128
4.1	Methode.....	128
4.1.1	Stichprobenbeschreibung.....	128
4.1.2	Versuchsaufbau.....	128
4.1.2.1	Subjektives Befinden und physische Beschwerden.....	129
4.1.3	Beschreibung der Stimuli.....	129
4.1.4	Versuchsablauf.....	129
4.1.5	Design.....	130
4.1.6	Teststärkeüberlegungen.....	131
4.2	Ergebnisse.....	132
4.2.1	Verhaltensdaten.....	132
4.2.1.1	Absolute Anzahl entdeckter Fehler.....	132
4.2.1.1	Anzahl gelesener Wörter.....	133
4.2.1.1	Relative Anzahl entdeckter Fehler (Genauigkeit).....	135
4.2.1.1	Fehlerarten.....	137
4.2.1.1	Genauigkeit bei der Entdeckung von Grammatikfehlern.....	138
4.2.1.1	Falsche Alarme und „Nicht übernehmen“	138
4.2.1.1	Subjektive Angaben.....	139
4.2.2	Subjektive Beschwerden.....	140

4.2.2.1	Physische Beschwerden.....	140
4.2.2.2	Befindlichkeit.....	141
4.3	Interpretation und Diskussion.....	143
4.3.1	Performanz.....	143
4.3.2	Subjektive Beanspruchung.....	145
5	Abschlussdiskussion.....	147
5.1	Polaritätseffekt.....	147
5.2	Objektive und subjektive Beanspruchung.....	149
5.3	Weiterführende Perspektiven.....	152
5.4	Schlussfolgerungen.....	153
6	Literatur.....	154

Zusammenfassung

Im Rahmen der Bildschirmarbeit spielt die Präsentation von Texten am Monitor eine große Rolle. Seit Beginn der 80iger Jahre hat sich im Bereich der Textverarbeitung zunehmend eine Darstellung in positiver Polarität (dunkle Schrift auf hellem Hintergrund) durchgesetzt. Es gibt allerdings Hinweise darauf, dass bei dunkler Umgebung eine Negativdarstellung günstiger ist, da in diesem Fall weniger retinale Hell-Dunkeladaptation erforderlich sind (Taptagaporn & Saito, 1990). Diese Annahme wurde bisher auf Performanzebene allerdings nicht systematisch untersucht.

Ziel der vorliegenden Untersuchung war es, unser Wissen über eine optimale Präsentation von Texten am Bildschirm in Abhängigkeit von den Lichtbedingungen zu erweitern. Außerdem sollten neue Erkenntnisse über die beanspruchende Wirkung der Bildschirmarbeit gewonnen werden. Bisher konnte zwar auch auf objektiver Ebene mittels physiologischer Messungen gezeigt werden, dass die Arbeit am Computer beanspruchend ist, allerdings fand in diesen Studien die motorische Aktivität der Probanden keine ausreichende Berücksichtigung (z.B. Schleifer & Ley, 1994).

In der vorliegenden Arbeit wurde die Leistung der Probanden bei einer Lese-Korrekturaufgabe am Bildschirm untersucht. Es wurden in zufälliger Abfolge insgesamt 15 Kurzgeschichten, die jeweils eine Bildschirmseite füllten, präsentiert. Die Gesamtlesezeit betrug 45 Minuten. In jeder Geschichte befanden sich durchschnittlich 30 Fehler, wobei zwischen fünf Fehlerarten unterschieden werden konnte. Von besonderer Bedeutung waren dabei die Grammatikfehler, die dazu dienten, die Lesestrategie der Probanden zu überprüfen.

Im ersten Experiment wurden die Texte in positiver Polarität (schwarze Schrift auf weißem Hintergrund) und in negativen Kontrastverhältnissen (weiße Schrift auf schwarzem Hintergrund) dargestellt, wobei der Raum entweder abgedunkelt war (5 Lux) oder eine künstliche Beleuchtung von 550 Lux vorlag. Als objektive Beanspruchungsindikatoren wurden die Herzrate, verschiedene Parameter der Herzratenvariabilität, die Atemfrequenz sowie die Hautleitfähigkeit erfasst. Vor und nach dem Versuch fanden Ruhemessungen statt, während derer die Versuchsteilnehmer Mausbewegungen simulieren mussten, um die motorische Aktivität zu kontrollieren. Die Ergebnisse zeigten, dass eine Textpräsentation in negativer Polarität unabhängig von den Beleuchtungsbedingungen zu einer deutlich

schlechteren Leistung bei der Fehlerentdeckung führte als positive Kontrastverhältnisse, was in erster Linie durch ein wesentlich langsames Lese- bzw. Bearbeitungstempo begründet war. Außerdem konnte festgestellt werden, dass die Aufgabenbearbeitung sowohl auf subjektiver als auch objektiver Ebene eine deutlich messbare Beanspruchung hervorrief, wobei aber kein Unterschied zwischen den Gruppen vorlag.

Experiment 2 wurde durchgeführt, um das Ausmaß der schlechten Leistung bei einer Textdarstellung in negativen Kontrastverhältnissen vor allem bei Dunkelheit besser einordnen zu können. Aus diesem Grund wurde die Performanz bei einer sehr schlecht lesbaren, äquilibriumanten Farbkombination (rote Schrift auf grünem Hintergrund) mit einer Textpräsentation in positiver Polarität bei niedriger Umgebungsbeleuchtung verglichen. Es zeigte sich erwartungsgemäß eine deutlich schlechtere Leistung bezüglich der Fehlerentdeckung in der chromatischen Bedingung, die vergleichbar war mit der schlechten Leistung bei negativem Kontrast in Experiment 1. Im Ausmaß der subjektiven und objektiven Beanspruchung unterschieden sich die Bedingungen aber nicht.

Im dritten Experiment sollte die Frage untersucht werden, ob der Polaritätseffekt ausschließlich für die typische Schwarz/Weiß-Darstellung gilt, oder ob er sich auf chromatische Bedingungen verallgemeinern lässt. Die Ergebnisse zeigen, dass auch bei einer farbigen Textpräsentation (Blau und Gelb) die Leistung der Probanden unabhängig von der Beleuchtung deutlich schlechter war unter negativen Kontrastverhältnissen.

Zusammenfassend kann die Schlussfolgerung gezogen werden, generell auf eine Textdarstellung in negativer Polarität zu verzichten. Außerdem konnte in der vorliegenden Arbeit im Rahmen einer mehrdimensionalen Beanspruchungsanalyse gezeigt werden, dass typische Bildschirmarbeit wie die Bearbeitung einer Lese-Korrektur-Aufgabe unter Kontrolle der motorischen Aktivität zu einer nachweisbaren Beanspruchung führt.

1 Einleitung

Durch den Einsatz neuer Technologien, die Einführung neuer Organisations- und Arbeitsformen sowie Softwareentwicklungen hat die Arbeitswelt in den letzten Jahren eine starke Wandlung erfahren. Dies äußert sich in einer Arbeitsverdichtung mit gesteigerter geistiger und kommunikativer Arbeit. Die Bildschirmarbeit ist dabei der Bereich, der besonders stark zugenommen hat und in vielen Berufen zur Regel geworden ist. Unter Bildschirmarbeit versteht man die Tätigkeit, die der Benutzer bei der Erledigung der Arbeitsaufgaben am Bildschirmarbeitsplatz ausführt. Zum Bildschirmarbeitsplatz zählen neben dem eigentlichen Bildschirm auch sämtliche Soft- und Hardware sowie die Arbeitsumgebung. Bildschirmarbeitsplätze sind in Büros und Verwaltungen Normalität und auch in Produktions- und Dienstleistungsbereichen stark vertreten. Bereits 1997 stellten Ertel, Junghanns, Pech und Ullsperger fest, dass über 72 % der Büroangestellten bei ihrer Arbeit einen Computer verwenden. Dieser Einsatz von Computern hat einen starken Einfluss vor allem auf die Inhalte der Arbeit, da sich viele Arbeitsabläufe dadurch verändern. Zudem kann man davon ausgehen, dass mögliche Qualifikationsdefizite viel eher augenscheinlich werden.

Im Rahmen von Bildschirmarbeit kann es sowohl zu psychischen als auch physischen Belastungen und Beanspruchungen kommen. Die Begriffe „Belastung“ und „Beanspruchung“ gehen zurück auf das Belastungs-Beanspruchungskonzept von Rohmert aus dem Jahre 1984 und stammen ursprünglich aus der Mechanik. In der arbeitspsychologischen Belastungsforschung hat das Belastungs-Beanspruchungskonzept auch durch seinen Eingang in deutsche und internationale Normen Bedeutung erlangt. Das deutsche Institut für Normung e.V. (DIN) versteht Normen als Regeln der Technik. Sie dienen unter anderem der Qualitätssicherung sowie der Sicherheit. Speziell zu Bildschirmarbeitsplätzen existieren zahlreiche DIN-Normen, in denen Maßnahmen zur ergonomischen Gestaltung der Bildschirmarbeit empfohlen werden. Laut der Norm DIN 33 405 zur psychischen Belastung und Beanspruchung aus dem Jahre 1987 wird „**Belastung**“ definiert als die Gesamtheit aller erfassbaren Einflüsse, die von außen auf den Menschen zukommen und psychisch auf ihn einwirken.

Diese Einflüsse lassen sich in die Bereiche informatorische Belastungen, das heißt Aspekte der Arbeitsaufgabe, organisatorische Belastungen wie Arbeitsmittel und Arbeitsumgebung

sowie soziale Belastungen einteilen. Zu den sozialen Belastungen zählen beispielsweise soziale Beziehungen im Betrieb oder auch betriebliche Hierarchien (vgl. Udris & Frese, 1988). Zu berücksichtigen ist die Tatsache, dass bei diesem Konzept der Begriff Belastung im Gegensatz zur Umgangssprache eine wertneutrale Bedeutung aufweist. Der Begriff „**Beanspruchung**“ bezeichnet laut DIN 33 405 (1987) die individuelle, zeitlich unmittelbare und nicht langfristige Auswirkung der Belastung auf die Einzelperson.

Gleiche Belastungen können zu individuell unterschiedlichen Beanspruchungen mit positiven sowie negativen Folgen führen. Für den Grad der Beanspruchung sind die Belastungen sowie die Persönlichkeitsmerkmale einer Person von großer Bedeutung. Je höher die Belastung, desto stärker müssen individuelle Leistungsvoraussetzungen und Ressourcen eingesetzt werden, um das Ausmaß der Beanspruchung möglichst gering zu halten und die Aufgabe adäquat zu bewältigen. Hierbei wird zwischen organisationalen Ressourcen wie beispielsweise eine große Aufgabenvielfalt, sozialen und personalen Ressourcen wie zum Beispiel Optimismus oder ein hohes Selbstwertgefühl unterschieden (vgl. Udris, Kraft & Musmann, 1992). Die resultierenden Beanspruchungen können sich in verschiedenen Bereichen manifestieren, wobei die Beanspruchungsfolgen entweder kurz- oder langfristig auftreten können. Nach Greif, Bamberg und Semmer (1991) kann zwischen einer physiologisch-somatischen, einer psychisch-subjektiven Ebene sowie einer behavioralen Ebene differenziert werden (vgl. Tabelle 1.1).

Tabelle 1.1

Klassifizierung der Beanspruchungsfolgen

	Kurzfristige Beanspruchungsfolgen	Mittel- bis langfristige Beanspruchungsfolgen
physiologisch-somatisch	Herzrate steigt, Blutdruck steigt, Adrenalinpiegel steigt, Muskelspannung nimmt zu	psychosomatische Beschwerden
psychisch-subjektiv	Anspannung, Ärger, Ermüdung, Sättigung, Monotonie	Unzufriedenheit, Resignation, Depression
behavioral:		
individuell	Fehler nehmen zu, Konzentration sinkt	Substanzkonsum, erhöhte Fehlzeiten
sozial	Konflikte, Streit, Aggression gegen andere, Rückzug (Isolation)	

Im Arbeitsalltag geht es allerdings nicht darum, sämtliche Belastungen zu reduzieren, da diese durchaus auch positive Folgen haben können. Dazu zählen neben dem Erreichen der geforderten Arbeitsleistung auch der Erhalt und Erwerb geistiger Fähigkeiten sowie eine gesteigerte Arbeitsmotivation und Zufriedenheit. Vielmehr sollte im Rahmen der Beanspruchungsoptimierung das Bestreben bestehen, nur diejenigen Belastungen zu reduzieren, aus denen dysfunktionale Beanspruchungen wie zum Beispiel körperliche Beschwerden oder Leistungseinbußen resultieren. Belastungen, aus denen funktionale Beanspruchungen entstehen, sollen dagegen intensiviert werden (vgl. Wieland-Eckelmann, Baggen, Saßmannshausen, Schmitz, Schwarz, Ademmer & Rose, 1996).

1.1 Belastung und Beanspruchung bei der Bildschirmarbeit

Im Bereich der Bildschirmarbeit treten in der Regel vielfältige körperliche und psychische Belastungsfaktoren auf, die sich gegenseitig verstärken können. Psychische Belastungen resultieren beispielsweise aus erhöhten Anforderungen an Gedächtnis- und Aufmerksamkeitsprozesse, an die Informationsverarbeitung sowie Entscheidungsfindung. Sie können aber auch durch organisatorische Faktoren wie zum Beispiel ein mangelhaftes Informationsmanagement und direkte Probleme im Dialogsystem Mensch-Maschine – dazu zählen beispielsweise Systemausfälle – bedingt sein.

Physische Belastungen bei der Bildschirmarbeit sind in vielen Fällen darin begründet, dass die Gestaltung der Arbeitsplätze hinsichtlich ergonomischer Aspekte oft sehr unzureichend ist. Dementsprechend konnte in vielen Untersuchungen gezeigt werden, dass bei Bildschirmarbeit typische Beanspruchungen vorliegen, die sich häufig auf der physiologisch-somatischen Ebene manifestieren. Ertel et al. (1997) beispielsweise befragten mittels eines standardisierten Fragebogens zur Gesundheit am Bildschirmarbeitsplatz (GESBI) von Ertel, Junghanns und Ullsperger aus dem Jahre 1994 insgesamt 280 Bildschirmarbeitnehmer bezüglich körperlicher Beschwerden. Es zeigte sich, dass etwa 65 Prozent der Untersuchungsteilnehmer Beschwerden im Stütz- und Bewegungsapparat aufwiesen. Fast die Hälfte der untersuchten Personen litt unter Kopfschmerzen und etwa 40 Prozent klagten

über Augenprobleme. Augenbeschwerden bei Bildschirmarbeit äußern sich meist in einem ganzen Symptomkomplex. Die verschiedenen Symptome wie Verschwommensehen, Tränen, Brennen, Rötung, Lidflattern sowie Doppelbilder werden üblicherweise unter dem Begriff astenopische Beschwerden zusammengefasst. Sie betreffen weniger den eigentlichen Sehvorgang als die peripheren Teile des Auges. Die Ursache für die Symptome liegt oft in einer schlechten photometrischen Bildschirmqualität (Flimmern, Schärfe), einer ungünstigen Gestaltung der Informationsdarstellung sowie Blendungen.

Neben diesen drei Hauptbeschwerdebereichen zeigten sich in der Untersuchung von Ertel et al. (1997) noch weitere Symptome. Dazu zählten Erschöpfungszustände, Konzentrationsstörungen, innere Unruhe, Schlafstörungen, Magen-Darm Beschwerden, Hautprobleme und Beschwerden an Händen, Armen und Beinen. Es hat sich gezeigt, dass sämtliche Beschwerden einen Bezug zur konkret zu erledigenden Aufgabe haben. Laut Junghanns, Ertel und Ullsperger (1998; 1999) sind sie bei qualifizierten, abwechslungsreichen und anspruchsvolleren Tätigkeiten geringer als bei weniger qualifizierten. Ursache für die genannten Gesundheitsprobleme ist – wie bereits erwähnt – in vielen Fällen eine unzureichende ergonomische Gestaltung des Arbeitsplatzes. Beispielsweise erlauben die verfügbaren Arbeitsmöbel oftmals keine entspannten und ermüdungsfreien Körperhaltungen. Neuere Forschungsergebnisse belegen aber inzwischen, dass auch psychische Belastungen wie Termin- oder Zeitdruck, inadäquater Handlungs- und Entscheidungsspielraum, widersprüchliche Anweisungen, fehlende Anerkennung oder Qualifizierungsdefizite bei der Entstehung der gesundheitlichen Beschwerden eine Rolle spielen (Ertel et al., 1997).

Der Gesetzgeber hat die entstandene Problematik im Zusammenhang mit der starken Zunahme der Bildschirmarbeit mittlerweile erkannt. Lange Zeit galt der gesetzliche Arbeitsschutz zur Bildschirmarbeit als unzureichend, da es keine gesetzliche Regelung gab, die sich ihrem Geltungsbereich nach ausschließlich mit Bildschirmarbeitsplätzen befasste. Erst im Jahre 1990 hat der Rat der Europäischen Gemeinschaft eine Richtlinie zur Arbeit an Bildschirmgeräten erlassen. Sie wurde in Deutschland am 20.12.1996 als Bildschirmarbeitsverordnung (BildscharbV) umgesetzt und dient der Sicherheit und dem Gesundheitsschutz bei der Arbeit an Bildschirmgeräten.

„Ein Bildschirmarbeitsplatz im Sinne dieser Verordnung ist ein Arbeitsplatz mit einem Bildschirmgerät, der ausgestattet sein kann mit

1. Einrichtungen zur Erfassung von Daten,

1. Software, die den Beschäftigten bei der Ausführung ihrer Arbeitsaufgaben zur Verfügung steht,
1. Zusatzgeräten und Elementen, die zum Betreiben oder Benutzen des Bildschirmgeräts gehören, oder
1. sonstigen Arbeitsmitteln,

sowie die unmittelbare Arbeitsumgebung.“ (§2 Abs.2 BildscharbV, 1996)

Gültigkeit hat die Verordnung für alle Beschäftigten, „die gewöhnlich bei einem nicht unwesentlichen Teil ihrer normalen Arbeit ein Bildschirmgerät benutzen.“ (§2 Abs.3, BildscharbV, 1996). Ausnahmen davon bilden Beschäftigte, die täglich weniger als 5-25 Prozent ihrer normalen Arbeitszeit am Bildschirm verbringen. Keine Gültigkeit hat die Verordnung auch für Arbeitnehmer, für deren Tätigkeit der Bildschirm nicht zwingend erforderlich ist, die keine besonderen Kenntnisse benötigen, denen dabei auch keine Aufmerksamkeit abverlangt wird oder die weniger als mindestens eine Stunde täglich am Bildschirm arbeiten. Weiterhin ausgenommen von dieser Regelung sind Bedienerplätze von Maschinen; Bildschirmgeräte an Bord von Verkehrsmitteln; Datenverarbeitungsanlagen, die hauptsächlich zur Benutzung durch die Öffentlichkeit bestimmt sind; Bildschirmgeräten für den ortsveränderlichen Gebrauch, sofern sie nicht regelmäßig an einem Arbeitsplatz eingesetzt werden; Rechenmaschinen, Registrierkassen oder andere Arbeitsmittel mit einer kleinen Daten- oder Messwertanzeigevorrichtung, die zur unmittelbaren Benutzung des Arbeitsmittels erforderlich ist; sowie Schreibmaschinen klassischer Bauart mit einem Display (§1, Abs. 2 BildscharbV, 1996).

Die BildscharbV schreibt ein Mindestmaß an Sicherheit und Gesundheitsschutz am Bildschirmarbeitsplatz vor. Der Bildschirmarbeitsplatz soll ergonomischen Mindestanforderungen, die im Anhang der Richtlinie aufgeführt sind, genügen. Arbeitet ein Arbeitnehmer einen „nicht unwesentlichen Teil seiner normalen Arbeit“ (§2, Abs.3 BildscharbV, 1996) an einem Bildschirmarbeitsplatz, so sind Arbeitgeber verpflichtet, Arbeitsplatzanalysen durchzuführen. Im Rahmen dieser Analysen werden neben der ergonomischen Gestaltung des Bildschirmarbeitsplatzes auch die psychischen Belastungen sowie der Gesundheitsstatus der Bildschirmarbeitnehmer ermittelt und bewertet. Zur Beurteilung der Arbeitsbedingungen kommen dabei verschiedene Instrumente wie beispielsweise Checklisten, Arbeitsplatzbegehungen, Interviews, Messungen, medizinische Untersuchungen usw. zum Einsatz.

1.2 Psychophysiologische Analyse mentaler Beanspruchung

In vielen verbreiteten Theorien wird psychische oder mentale Beanspruchung definiert als die Intensität der menschlichen Informationsverarbeitung (Kahnemann, 1973). Das Ausmaß dieser Beanspruchung wird allerdings häufig nur auf subjektiver Ebene beschrieben, da in den meisten Fällen verschiedene Fragebögen zur Ermittlung der psychischen und physischen Beanspruchung verwendet werden. Dabei stellt sich die Frage, welche Validität solche subjektiven Aussagen aufweisen. Es bleibt offen, inwieweit sich die Beanspruchung bei Bildschirmarbeit auch objektiv – beispielsweise durch physiologische Messungen, die Indikatoren für psychische Beanspruchung darstellen – nachweisen lässt.

Die psychophysiologische Analyse psychischer Beanspruchung hat das Ziel, körperliche Indikatoren dieser Informationsverarbeitungsprozesse und die Anstrengung der Person zu identifizieren. Zur physiologischen Beanspruchungsanalyse können verschiedene Indikatoren herangezogen werden. Häufig stammen diese aus dem Bereich der kardiovaskulären und elektrodermalen Aktivität. Daneben findet auch die Atmung Anwendung im Rahmen der Beanspruchungsanalyse. Im Folgenden werden die verschiedenen Indikatoren näher beschrieben.

1.2.1 Indikatoren aus dem kardiovaskulären Bereich

Es gibt verschiedene Beanspruchungsindikatoren aus dem kardiovaskulären Bereich. In der vorliegenden Arbeit soll der Fokus auf der Herzrate sowie der Herzratenvariabilität liegen.

1.2.1.1 Herzrate

Die Herzrate gehört zu den am häufigsten verwendeten psychophysiologischen Parametern. Üblicherweise wird sie als Herzschläge pro Minute (engl.: **beats per minute**, bpm) angegeben.

Es gibt eine Reihe von Untersuchungen, in denen ein Zusammenhang zwischen mentaler Belastung und Veränderungen der Herzrate aufgezeigt werden konnte. Allerdings ist die Befundlage zur Herzrate als Beanspruchungsindikator nicht vollkommen homogen. Es liegen einige Feldstudien vor, die den erwarteten positiven Zusammenhang zwischen Beanspruchung und Herzrate nicht zeigen konnten. Beispielsweise war dies der Fall in einer Flugsimulator-Untersuchung von Casali und Wierwille (1983), in der 30 Piloten insgesamt drei verschiedene Flüge absolvieren mussten. Die Flüge unterschieden sich im Schwierigkeitsgrad der Kommunikationsanforderungen, was sich aber nicht in der Herzrate widerspiegelte, die während des Flugs erhoben wurde.

Im Folgenden sollen aus der Vielzahl der Untersuchungen verschiedene Laborexperimente mit unterschiedlichen Aufgaben unter anderem am Computerbildschirm näher erläutert werden. Ein Anstieg der Herzrate wird üblicherweise in experimentellen Situationen beobachtet, die hohe Anspannung oder geistige Anstrengung erfordern. Ohsuga, Shimono und Genno (2001) konnten bei einer psychomotorischen Tracking-Aufgabe einen signifikanten Anstieg der Herzrate nachweisen. Wurden die insgesamt 100 Versuchspersonen zusätzlich unter Stress gesetzt, indem man ihnen bei Fehlern eine verlängerte Versuchszeit androhte, war das Ausmaß der Herzratenzunahme noch größer. Auch bei simulierten Überwachungstätigkeiten am PC mit unterschiedlichem Schwierigkeitsgrad diente die Herzrate als zuverlässiger Beanspruchungsindikator (Fournier, Wilson & Swain, 1999). Diese Überwachungstätigkeit bestand aus insgesamt vier verschiedenen Aufgaben. Beispielsweise mussten die Versuchspersonen Anzeigeelemente überwachen und auf bestimmte Lichtsignale reagieren. Es zeigte sich, dass die Herzrate umso höher war, je mehr Aufgaben gleichzeitig bearbeitet werden mussten.

In einer Reihe von Experimenten wurde die Auswirkung verschiedener Rechenaufgaben wie beispielsweise serielle Subtraktions- und Additionsaufgaben am Computerbildschirm auf die Herzrate untersucht. Es zeigte sich, dass auch bei diesem experimentellen Setting die Herzrate die mentale Beanspruchung reliabel abbildete (Al 'Absi, Everson & Lovallo,

1995; Oshima, Miyano, Yamashita, Owashi, Suzuki, Sakano & Higuchi, 2001; Vögele, 1998). Carroll, Turner und Hellawell (1986) stellten fest, dass der Anstieg der Herzrate bei schwierigen arithmetischen Aufgaben stärker ausgeprägt war als bei einfachen Aufgaben. Es zeigte sich ebenfalls, dass bei Bearbeitung eines kognitiv herausfordernden Computerspiels („Space Invaders“) die Herzrate signifikant zunahm (Turner, Carroll & Courtney, 1983). Auch bei Bearbeitung des Stroop-Tests konnte ein signifikanter Anstieg der Herzrate im Vergleich zur Ruhebedingung festgestellt werden (Aslan, Nelson, Carruthers & Lader, 1981; Moya Albiol, Salvador, Costa, Martinez Sanchis, Gonzalez Bono, Ricarte & Arnedo, 2001; Renaud & Blondin, 1997; Tulen, Moleman, Van Steenis & Boomsma, 1989). Darüber hinaus konnten Scher, Furedy und Heslegrave (1984) beim Abruf von Zahlenspannen aus dem Arbeitsgedächtnis einen Zusammenhang aufzeigen zwischen dem Anstieg der Herzrate und der Aufgabenschwierigkeit, die durch die Anzahl der zu memorierenden Items variiert wurde.

Die hier aufgeführten Studien unterstützen die Annahme, dass die Herzrate einen geeigneten Indikator für mentale Beanspruchung darstellt und der Einsatz dieses Parameters auch in der vorliegenden Untersuchung gerechtfertigt erscheint.

Während der Bearbeitung kognitiver Aufgaben kann es aber auch durchaus zu einer Abnahme der Herzfrequenz kommen. Ein Absinken der Herzrate wurde oft beobachtet unter Müdigkeit, Monotonie, nachlassender Aufmerksamkeit und/oder geringer Aktivierung. Beispielsweise zeigte sich bei Zugführern und Busfahrern mit zunehmender Fahrtzeit eine kontinuierliche Abnahme der Herzrate (Milosevic, 1997). In einer Studie von Hörhold und Walschburger (1997) bearbeiteten die Versuchspersonen Konzentrationsaufgaben. Es wurde eine Matrix präsentiert und per Tastendruck musste entschieden werden, ob diese Matrix ein bestimmtes Zielitem enthielt oder nicht. Es zeigte sich, dass nach einem initialen Anstieg die Herzrate bei wiederholten Belastungsbedingungen als Zeichen einer Anpassung bzw. Übung wieder abnahm. Auch Heslegrave und Furedy (1979) stellten fest, dass während der Bearbeitung mathematischer Aufgaben die Herzrate fast zum Ruhenniveau zurückkehrte und interpretieren dieses ebenfalls im Sinne einer Habituation.

1.2.1.2 Herzratenvariabilität

Eine große Bedeutung als kardiovaskulärer Indikator mentaler Beanspruchung wird seit den Arbeiten von Kalsbeek und Ettema (1963) auch der Herzratenvariabilität (HRV) zugeschrieben. Darunter versteht man die Variabilität der R-Zacken Abstände im Elektrokardiogramm (EKG), welche auch als Sinusarrhythmie bezeichnet wird (siehe Abbildung 1.1).

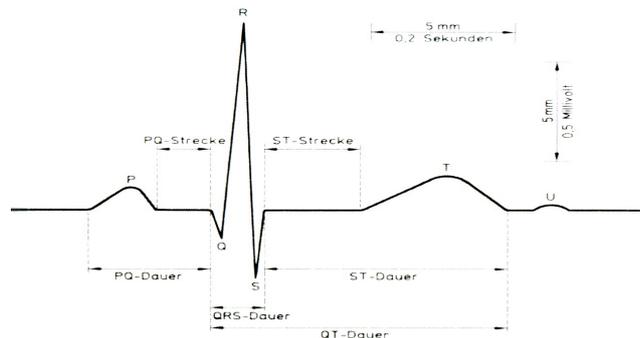


Abbildung 1.1

Darstellung eines QRS-Komplexes im Elektrokardiogramm. Als Herzratenvariabilität wird die Variabilität der R-Zacken Abschnitte bezeichnet. (modifiziert nach: Schandry, 1988)

Kalsbeek und Ettema konnten zeigen, dass die HRV durch zeitliche Schwankungen der Aktivierung des autonomen Nervensystems belastungsabhängig variiert. Steigerungen der mentalen Beanspruchung durch Erhöhungen der Signalrate bei binären Wahlreaktionsaufgaben führten zu einer signifikanten Abnahme der Variabilität der R-Zacken Abstände im EKG, das heißt der Herzschlag wurde regelmäßiger. Ein Anstieg der tonischen Herzrate konnte allerdings nicht nachgewiesen werden. Die Korrelation zwischen HRV und Herzrate liegt bei -0.4 und ist somit nur mäßig.

Parametrisierung der Herzratenvariabilität: Die Parametrisierung der HRV erfolgt sowohl über zeit- als auch frequenzbezogene Größen. Im Rahmen der zeitbezogenen Analyse kann die Variabilität der R-Zacken Intervalle in einem bestimmten Zeitbereich über verschiedene globale Streuungsmaße beschrieben werden. Dazu werden zunächst die so genannten „Normal-to-Normal“-Intervalle (NN-Intervalle) bestimmt. Darunter versteht man alle In-

tervale zwischen im EKG erkennbaren benachbarten QRS-Komplexen, die aus einer Depolarisation des Sinusknoten resultieren.

Als Streuungsmaße werden standardmäßig berechnet die Standardabweichung aller NN-Intervalle (SDNN), der Prozentsatz der Intervalle mit mindestens 50ms Abweichung vom vorausgehenden Intervall (PNN50) und die Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD), welche gegenüber der PNN50 aufgrund besserer statistischer Eigenschaften oft bevorzugt wird. Unter Belastungsbedingungen sinken diese Streuungsmaße ab, während höhere Kennwerte auf vermehrte parasymphatische Aktivität hinweisen.

Seit Ende der 60iger Jahre werden neben der zeitlichen Betrachtung auch verschiedene spektralanalytische Methoden zur Analyse der frequenzbezogenen Größen der Herzratenvariabilität herangezogen. Spektralanalytische Verfahren gehen von der Grundüberlegung aus, dass eine gegebene Zeitreihe aus trigonometrischen Teilschwingungen zusammengesetzt ist. Das Ziel der Spektralanalyse besteht darin, die charakteristischen Parameter (Amplitude, Frequenz und Phase) der Schwingungen zu ermitteln. Als Ergebnis der Analyse entsteht ein Amplitudenspektrum, das direkt in ein Leistungsdichtespektrum umgerechnet werden kann (siehe Abbildung 1.2). In die Berechnung des Spektrums gehen alle Amplituden größer Null ein. Spektralanalysen von RR-Intervall Zeitreihen belegen, dass es sich bei der HRV um ein komplexes Biosignal handelt, das den Einfluss mehrerer Faktoren auf die Herztätigkeit abbildet und eine entsprechend differenzierte Parametrisierung erfordert. Es können drei Frequenzbänder unterschieden werden, die mit unterschiedlichen biologischen Kontrollmechanismen assoziiert sind (vgl. Mulder, 1985).

0,02 – 0,06 Hz: Fluktuationen in diesem unteren Frequenzband hängen mit der Regulierung der Körpertemperatur (Kitney, 1975) und mit der Aktivität des Renin-Angiotensin-Systems (Akselrod, Gordon, Madwed, Snidman, Shannon & Cohen, 1985) zusammen.

0,07 – 0,14 Hz: Das mittlere Frequenzband wird auch als 0,1 Hz Komponente bezeichnet und Fluktuationen reflektieren im Wesentlichen die Kurzzeit-Blutdruckregulation (Wesseling & Settels, 1985).

0,15 – 0,4 Hz: Das obere Frequenzband bildet im Wesentlichen die respiratorische Aktivität, das heißt die jeweils dominante Atemfrequenz ab (siehe Abbildung 1.2). Oszillationen in diesem Band spiegeln fast ausschließlich parasymphatische Einflüsse auf die Herzrate wider (Grossman, 1992) und werden auch als respiratorische Sinusrhythmie bezeichnet.

Die Leistung im jeweiligen Frequenzband gibt an, wie groß der durch das Band erklärte Anteil an der Gesamtvarianz der Zeitreihe ist. Die Gesamtleistung in allen Frequenzbändern entspricht mathematisch betrachtet der Gesamtvarianz.

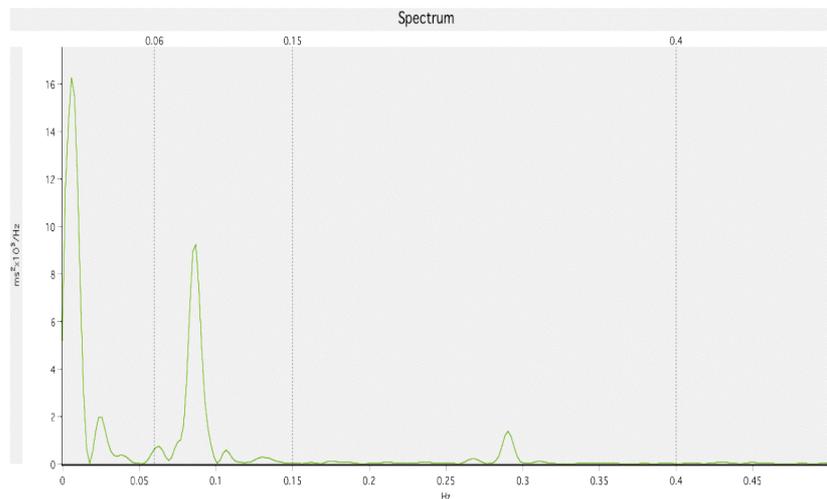


Abbildung 1.2: Darstellung eines Leistungsdichtespektrums, das mit Hilfe einer Spektralanalyse aus einer Herzfrequenz-Zeitreihe berechnet wurde. Das Spektrum wird standardmäßig in drei Frequenzbänder unterteilt. Deutlich erkennbar ist der Einfluss der Atmung bei ca. 0,3 Hz.

Mulder und Mulder (1981) untersuchten in einer Reihe von Experimenten sehr systematisch die Zusammenhänge zwischen dem Ausmaß mentaler Beanspruchung und spektralanalytisch abgeleiteten Kennwerten der HRV. In einem dieser Experimente wurden beispielsweise Veränderungen im Leistungsspektrum bei der Bearbeitung unterschiedlich schwieriger Satzverifikationsaufgaben untersucht. Es zeigte sich, dass bei mentaler Beanspruchung die Leistung in allen drei Frequenzbändern, aber vor allem im mittleren Band, in Abhängigkeit von der Aufgabenschwierigkeit reduziert wurde. Andere Arbeiten sind zu ähnlichen Ergebnissen gekommen (Aasman, Mulder & Mulder, 1987; Backs & Seljos,

1994; Redondo, 1992), obwohl das klare Befundmuster von Mulder und Mulder aus dem Jahre 1981 nur selten repliziert werden konnte. In diesen Untersuchungen spiegelte das Ausmaß der Reduktion der HRV im mittleren Frequenzband die Aufgabenschwierigkeit nur bei sehr großen Unterschieden in der Beanspruchung wider. Jorna (1992) zog daraus den Schluss, dass mit Hilfe der 0,1 Hz Komponente relativ zuverlässig zwischen Ruhe und Beanspruchung differenziert werden kann. Unterschiede in der Aufgabenschwierigkeit werden aber weniger deutlich abgebildet. Es wird vermutet, dass das mittlere Frequenzband nicht die Aufgabenkomplexität an sich widerspiegelt, sondern einen Indikator für den Anstrengungsaufwand (effort) darstellt. Dafür spricht die Beobachtung, dass die Leistung in diesem Band nach anfänglicher Reduktion wieder zunimmt, wenn die Anstrengungsbereitschaft der Versuchspersonen aufgrund von Überforderung sinkt (Aasman et al., 1987). Das mittlere Frequenzband scheint allerdings kein spezifischer Indikator für mentale Beanspruchung zu sein. Es konnte gezeigt werden, dass auch emotionale Belastungsfaktoren wie Angst oder Zeitdruck zu Reduktionen der 0,1 Hz Komponente führten (Jorna, 1992). Ebenso hatte körperliche Aktivität, also physische Beanspruchung wie beispielsweise Radfahren, einen Einfluss auf die Herzratenvariabilität (Perini, Milesi, Fisher, Pendergast & Veicsteinas, 2000; Yamamoto, Hughson & Peterson, 1991).

Neben den oben erwähnten Forschungsergebnissen sprechen aber auch Gründe rein praktischer und theoretischer Natur dafür, die 0,1 Hz Komponente als Hauptindikator für die Variationen der Herzratenvariabilität auszuwählen. Die Leistung im unteren Frequenzband ist relativ unreliabel, da lineare und nicht lineare Trends diese unteren Frequenzen stark beeinflussen können. Außerdem sind die üblichen Kurzzeit-Aufnahmen von zu kurzer Dauer, um zuverlässig die Leistung in diesem Frequenzbereich zu bestimmen. Damit reliable Ergebnisse erzielt werden, sollte die Aufnahmedauer mindestens die zehnfache Wellenlänge der untersten Frequenz des untersuchten Bandes betragen. Um dieser Anforderung zu genügen, müsste für eine Analyse des unteren Frequenzbandes die Aufnahmedauer mindestens 50 Minuten betragen. Basierend auf diesem Kriterium werden analog für die Analyse des oberen Frequenzbandes Zeitfenster von mindestens einer Minute empfohlen. Die Mindestdauer für das mittlere Band beträgt zwei Minuten (Task Force, 1996). Ebenso wie die sehr niedrigen Frequenzen ist auch das obere Frequenzband, also die respiratorische Komponente, kein zuverlässiger Indikator für mentale Beanspruchung, da es

nicht stabil ist und durch Veränderungen der Atemfrequenz seine Position im Spektrum variiert.

Das mittlere Band ist dagegen unter normalen Bedingungen und regulärer Atmung (üblicherweise 0,2-0,5 Hz) weit gehend unabhängig von der Atemfrequenz. Bei sehr geringer Atemfrequenz verschiebt sich die respiratorische Komponente allerdings in das mittlere Band und erhöht hier die Leistung. Es konnte gezeigt werden, dass der mittlere Frequenzbereich nur dann einen geeigneten Index für mentale Beanspruchung darstellt, wenn er nicht durch respiratorische Aktivität beeinflusst wird (Althaus, Mulder, Mulder, van Roon & Minderaa, 1998). Aus diesen Gründen ist eine unabhängige Überwachung der Atmung unabdingbar, um solche Effekte zu kontrollieren.

Als Erklärung für das Absinken der Herzratenvariabilität im mittleren Frequenzband unter Belastung werden in der Literatur verschiedene Mechanismen diskutiert. Unter anderem wird in diesem Zusammenhang eine Verringerung der Sensitivität des Barorezeptorenreflexes genannt. Bei diesem Reflex handelt es sich um einen Kontrollprozess, mit dem das kardiovaskuläre System auf Blutdruckveränderungen reagiert. Barorezeptoren sind Presso- und Volumenrezeptoren, die in den Wänden des Karotissinus und der Aorta liegen. Sie sind beteiligt an der Herz-Kreislauf Regulation über ein negatives Feedback-System. Bei einem Blutdruckanstieg werden die Rezeptoren stimuliert, was zu einer Aktivierung des parasympathischen und einer Hemmung des sympathischen Nervensystems führt. Als Folge dieses Regulationsmechanismus sinkt der Blutdruck wieder auf das normale Niveau ab. Analog dazu wird bei einem Blutdruckabfall der Sympathikus aktiviert und der Druck steigt wieder an. Die Tatsache, dass der Barorezeptorenreflex für die Regulation der HRV eine wichtige Rolle spielt, konnte durch experimentelle Befunde gestützt werden. Lucini, Guzzetti, Casiraghi und Pagani (2002) wiesen signifikante Korrelationen zwischen der Sensitivität des Barorezeptorenreflexes und dem Prozentsatz der Intervalle mit mindestens 50ms Abweichung vom vorausgehenden Intervall (PNN50) sowie der Gesamtleistung im Spektrum nach. Pagani et al. (1991) stellten eine Reduktion der Barorezeptorenreflex-Sensitivität während der Bearbeitung von Rechenaufgaben fest, die mit der Veränderung der Leistung im mittleren Frequenzband korreliert war.

Die Angabe der Leistungswerte der einzelnen Frequenzbänder in absoluten Einheiten, wie es früher – und teilweise heute auch noch – praktiziert wurde, bringt allerdings Probleme

mit sich. Bei der Berechnung der absoluten Werte findet die Gesamtleistung des Spektrums keine Berücksichtigung. Im Zuge einer Beanspruchung kommt es allerdings nicht nur zu einer Veränderung in den einzelnen Frequenzbändern, sondern auch zu einer Reduktion der Gesamtleistung, die zudem interindividuell sehr variabel ist. Aus diesen Gründen bietet es sich an, die Leistung in den einzelnen Bändern unter Berücksichtigung der Gesamtleistung in normalisierten Einheiten anzugeben. Dazu wird die absolute Leistung im oberen bzw. mittleren Band dividiert durch die Gesamtleistung minus die Leistung im unteren Frequenzband und das Ergebnis mit 100 multipliziert. Die Umrechnung in normalisierte Einheiten ermöglicht den Vergleich zwischen verschiedenen Probanden und experimentellen Bedingungen auch dann, wenn eine große Variabilität in der Gesamtleistung aller Frequenzbänder vorliegt.

Die Leistung im oberen Frequenzband wird maßgeblich durch die Vagusaktivität determiniert. Dies konnte durch zahlreiche Experimente, in denen Stimulationen, Rezeptorblockaden und Vagotomie zur Anwendung kamen, belegt werden (Akselrod, Gordon, Ubel, Shannon, Berger & Cohen, 1981; Malliani, Pagani, Lombardi & Cerutti, 1991; Pomeranz, Macaulay, Caudill, Kutz & Adam, 1985). Somit stellt die Leistung im oberen Band einen Indikator für die Aktivität des parasympathischen Teils des autonomen Nervensystems dar. Die Interpretation der mittleren Frequenzkomponente ist dagegen kontrovers. Einige Forscher betrachten die Leistung in diesem Frequenzbereich – vor allem wenn sie in normalisierten Einheiten angegeben wird – als Marker für die Aktivität des sympathischen Nervensystems (Kamath & Fallen, 1993; Malliani et al. 1991; Montano, Ruscone, Porta, Lombardi, Pagani & Malliani, 1994). Als Folge repräsentiert das Verhältnis zwischen der Leistung im mittleren und oberen Frequenzband die Balance zwischen den sympathischen und parasympathischen Anteilen des autonomen Nervensystems. Andere wiederum sind der Ansicht, dass sich in dieser Komponente sympathische *und* vagale Einflüsse widerspiegeln (Akselrod et al., 1981; Appel, Berger, Saul, Smith & Cohen, 1989).

Die Leistung im mittleren und oberen Frequenzbereich, ausgedrückt in normalisierten Einheiten, kann unter verschiedenen Bedingungen ansteigen. Kontrollierte Atmung verursachte einen Anstieg im oberen Band (Malliani et al., 1991), während körperliche Aktivität und mentale Belastung zu höheren Werten im mittleren Band bei einem gleichzeitigen Abfall der Leistung im oberen Frequenzbereich führten (zum Beispiel Yamamoto et al.,

1991). Wood, Maraj, Lee und Reyes (2002) stellten fest, dass bei einfachen Reaktionszeitaufgaben die normalisierte Leistung im oberen Band abfiel, während sie im mittleren Frequenzbereich anstieg. Man konnte auch zeigen, dass mentaler Stress, experimentell induziert durch ein computergestütztes Reaktionszeitparadigma und ein stressreiches Interview, zu einem signifikanten Anstieg der Leistung in normalisierten Einheiten im mittleren Frequenzband führte (Pagani, Rimoldi, Pizzinelli, Furlan & Crivellaro, 1991). Auch bei der Bearbeitung von Rechenaufgaben stieg die Leistung im mittleren Frequenzband an, während sie im oberen Frequenzbereich abfiel (Kamada, Sato, Miyake & Kumashiro, 1992; Pagani, Furlan, Pizzinelli, Crivellaro & Malliani, 1989). Delaney und Brodie (2000) verwendeten den Stroop-Test als mentalen Stressor. Sie fanden ebenfalls einen signifikanten Anstieg der Leistung im mittleren Frequenzband, während die Leistung im Bereich der oberen Frequenzen abfiel. Zusammenfassend kann man also feststellen, dass bei körperlicher Aktivität und mentaler Belastung die Leistung im mittleren Frequenzbereich ansteigt und im oberen Frequenzband analog dazu abnimmt.

Neben diesen kardiovaskulären Parametern werden häufig im Rahmen einer mehrdimensionalen Analyse auch andere psychophysiologische Kennwerte eingesetzt. Im folgenden Abschnitt soll näher auf die Messung der elektrodermalen Aktivität als weiterer Beanspruchungsindikator eingegangen werden.

1.2.2 Elektrodermale Aktivität

Der Begriff „Elektrodermale Aktivität (EDA)“ wurde von Johnson und Lubin (1966) als Sammelbegriff für die vorher unter den unterschiedlichsten Bezeichnungen beschriebenen elektrischen Erscheinungen der Haut eingeführt und bezeichnet systematische Veränderungen des elektrischen Leitwerts. Wie in Abbildung 1.3 erkennbar bildet die Haut keine strukturelle Einheit, sondern besteht aus verschiedenen, lichtmikroskopisch gut unterscheidbaren vertikalen Schichten. Die fünfschichtige Epidermis (Oberhaut) besteht aus Epithelgewebe, das nach außen verhornt ist und als Stratum corneum bezeichnet wird. Die Epidermis besitzt an den mechanisch stark beanspruchten palmaren und plantaren Hautflächen eine Stärke von einem Millimeter, während sie sonst nur 50-200 μm dick ist. Die dar-

unter liegende Dermis setzt sich aus straffem, faserreichen Gewebe zusammen und ist erheblich dicker als die Epidermis. Sie bewirkt die Festigkeit der Haut, daher der deutsche Name Lederhaut. Unterhalb der Dermis findet man die Subcutis (Unterhaut), die aus einer lockeren Bindegewebeschicht besteht und die oberen Hautschichten mit dem Muskel- und Zellgewebe des Körpers verbindet.

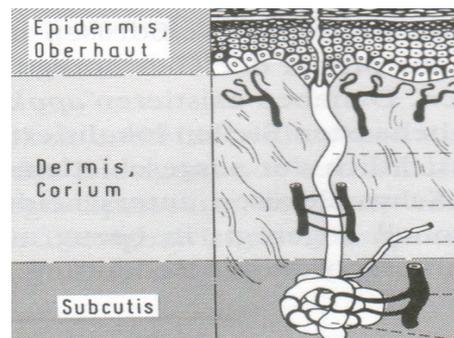


Abbildung 1.3:

Schematischer Schnitt durch die Haut (modifiziert nach: Schandry, 1988).

In der menschlichen Haut befinden sich etwa drei Millionen Schweißdrüsen, die für die Entstehung der elektrodermalen Aktivität sehr wichtig sind. Es handelt sich dabei um exokrine Drüsen, die ihr Sekret unmittelbar an der Hautoberfläche abgeben. Die Verteilung der Schweißdrüsen in der Haut ist stark unterschiedlich. Eine besonders große Anzahl findet sich auf Handflächen, Fußsohlen und Stirn. Der knäuelartige sekretorische Teil der Schweißdrüse liegt in der Subcutis und hat einen Durchmesser von etwa 0,4 mm. Daran schließt sich der Ausführungsgang (Ductus) an, der in der Dermis relativ gerade verläuft und danach in den korkenzieherartig gewundenen epidermalen Ductus übergeht, der an der Hautoberfläche in einer kleinen Pore mündet. Der sekretorische Teil der Schweißdrüse wird von stark verzweigten sympathischen Nervenendigungen innerviert, die als einzige postganglionäre sympathische Fasern nicht adrenerg, sondern cholinerg versorgt werden. Die Schweißdrüsen spielen eine wichtige Rolle bei der Thermoregulation. Dabei wird zwischen der Perspiratio Sensibilis (sichtbare Wasserabgabe durch Schwitzen) und der Perspiratio Insensibilis (unmerkliche Verdunstung von Wasser) unterschieden. Zu vermehrter

Schweißdrüsenaktivität, dem so genannten emotionalen Schwitzen, kommt es aber auch als Begleiterscheinung psychischer Zustände. Verantwortlich dafür sind die Schweißdrüsen an den palmaren und plantaren Hautflächen.

1.2.2.1 Physiologie der elektrodermalen Aktivität

Als wesentlicher Aspekt im Geschehen der elektrodermalen Aktivität wird die Aktivität der Schweißdrüsen betrachtet. Für das Entstehen der elektrodermalen Eigenschaften der Haut ist allerdings kein Schweißaustritt aus der Pore erforderlich. Die Leitfähigkeitsreaktion kann bereits eine Sekunde vor Erhöhung der Oberflächenfeuchtigkeit gemessen werden. Wichtig für die Entstehung des Biosignals sind die unterschiedlichen elektrischen Leitfähigkeiten der Hautschichten aufgrund verschiedener Feuchtigkeiten. Während Dermis und Subcutis stark durchblutet und reichlich mit Interstitialflüssigkeit versorgt sind, handelt es sich zumindest bei den oberen Schichten der Epidermis um trockene, verhornte Zellstrukturen, die schlecht leiten. Erhält der sekretorische Teil der Schweißdrüsen suderisekretorische Impulse und wird dadurch zur Schweißabsonderung angeregt, füllen sich zunächst die dermalen und dann die epidermalen Anteile der Ducti mit dem Sekret. Durch den hohen Druck im Ductus und/oder Diffusionsvorgänge wird das Sekret an die Epidermis abgegeben, wodurch es zu einer Befeuchtung des Stratum corneum kommt. Da es sich bei der Schweißabsonderung um eine verdünnte Salzlösung mit überwiegend Natriumchlorid-Ionen handelt, die gut elektrisch leitet, wird die elektrische Leitfähigkeit der Epidermis durch die Schweißdrüsenaktivität bestimmt. Je aktiver diese sind, desto bessere Leiter stellen die äußeren Hautschichten dar. Daneben spielen für die EDA auch die Ducti der Schweißdrüsen selbst mit ihrem variablen Füllungsstand eine bedeutsame Rolle als elektrischer Leiter. Bei der Erforschung der Ursachen elektrodermalen Aktivität müssen neben den unterschiedlichen Leitfähigkeiten der Hautschichten auch die an verschiedenen Stellen der Schweißdrüse stattfindenden Membranprozesse berücksichtigt werden. Durch den aktiven Transport von Ionen durch semipermeable Membranen werden zusätzliche Polarisationskapazitäten aufgebaut.

Außerdem haben auch humorale Veränderungen einen Einfluss auf die Polarisation von Membranen. Dazu zählt beispielsweise die Freisetzung des Gewebshormons Bradykinin,

das stark vasodilatatorisch wirkt, aber auch die Ausschüttung von Katecholaminen und Acetylcholin.

1.2.2.2 Parametrisierung der EDA

Die Kennwerte der elektrodermalen Aktivität lassen sich in tonische und phasische Maße einteilen. Zu den tonischen Parametern wird der Hautleitwert (engl.: skin conductance level, **SCL**) gezählt, der üblicherweise im Bereich zwischen 2-100 $\mu\text{S}/\text{cm}^2$ liegt. Durch kognitive und emotionale Belastung kommt es zu einer Erhöhung des Werts. Außerdem werden die so genannten Spontanfluktuationen zu den tonischen Kennwerten gerechnet. Darunter versteht man spontane phasische Leitwerterhöhungen ohne äußeren Anlass oder solche, die nicht Folge unregelmäßiger Atmung sind. Unter Ruhebedingungen liegt die absolute Anzahl bei 3-7/min und im aktivierten Zustand steigt die Zahl auf 10-15/min.

Ein wichtiger phasischer Kennwert der EDA stellt die Hautleitwertsreaktion (engl.: skin conductance reaction, **SCR**) dar. Dabei handelt es sich um eine Erhöhung des Leitwerts als Reaktion auf einen einzelnen Reiz, wie zum Beispiel eine Orientierungsreaktion auf einen Schreckreiz. Das Amplitudenkriterium wird abhängig von der Verstärkung üblicherweise auf 0,01-0,1 $\mu\text{S}/\text{cm}^2$ festgelegt. Nur wenn das registrierte Signal dieses Kriterium überschreitet, wird es als Leitwertsreaktion gewertet. Als Indikatoren zur Beschreibung der Hautleitwertsreaktion dienen die Amplitude, die Latenz, die Anstiegs- sowie Erholungszeit. Die Latenz einer elektrodermalen Reaktion bezeichnet die Zeit zwischen Reizeinsatz und Beginn der Reaktion und liegt zwischen ein bis drei Sekunden. 50 Prozent dieser Zeit gehen auf eine Verzögerung bei der zentralen Verarbeitung zurück, 10 bis 20 Prozent auf die neuronale Weiterleitung und weitere 25-50 Prozent auf die Freisetzung und Diffusion des Botenstoffs Acetylcholin. Unter der Anstiegszeit versteht man die Dauer vom ersten Abheben von der Grundlinie bis zum Erreichen des Gipfelpunkts. Sie nimmt Werte zwischen 0,5 und 3,5 Sekunden an. Als Erholungszeit wird die Zeit zwischen Gipfelpunkt und Erreichen des 50%-Werts der Amplitude bezeichnet.

1.2.2.3 Experimentelle Befunde

Die elektrodermale Aktivität wurde bisher in zahlreichen experimentell kontrollierten Belastungsuntersuchungen als Beanspruchungsindikator eingesetzt, wobei in den meisten Fällen die tonischen Parameter der EDA als Kennwerte verwendet wurden. Viele Forscher trennten dabei zwischen emotionalen und kognitiven Aspekten der Informationsverarbeitung, die sich in der elektrodermalen Aktivität widerspiegeln (Boucsein, 1992). Die Ergebnisse von Kirby (1999) deuteten allerdings darauf hin, dass sowohl kognitive als auch emotionale Komponenten der EDA letztendlich auf kognitive bzw. Aufmerksamkeitsprozesse zurückzuführen sind und keine eindeutige Trennung möglich ist.

Die Mehrzahl der Untersuchungen zur EDA als Beanspruchungsindikator stammt aus Laborexperimenten, da im Labor eine bessere Artefaktkontrolle möglich ist. In einer Reihe von Studien konnte gezeigt werden, dass es bei der Bearbeitung von Arithmetikaufgaben zu einem Anstieg des Hautleitwerts kam, was eine erhöhte Beanspruchung im Vergleich zur Ruhebedingung widerspiegelte (Albus, Müller-Spahn, Ackenheil & Engel, 1987; Oshima et al., 2001; Vögele, 1998).

In anderen Untersuchungen wurde der Stroop-Test als mentaler Stressor verwendet. Es zeigte sich, dass während der Aufgabebearbeitung der Hautleitwert signifikant anstieg (Aslan et al., 1981; Moya Albiol et al., 2001; Tulen et al., 1989). Auch in Reaktionszeitexperimenten wurde ein Anstieg der elektrodermalen Aktivität festgestellt (Fichera & Andreassi, 1998; Morell, 1989).

Zusammenfassend kann man aus den verschiedenen Untersuchungen folgern, dass tonische EDA-Maße während des Ablaufens von Arbeitsprozessen als sensible und valide Indikatoren kognitiver Beanspruchungen geeignet sind und neben kardiovaskulären Parametern im Sinne einer mehrdimensionalen Beanspruchungsmessung eingesetzt werden können.

1.2.3 Atmung

In der psychophysiologischen Beanspruchungsforschung erfreut sich die Messung respiratorischer Kennwerte keiner großen Beliebtheit, da viele traditionelle Techniken wie beispielsweise die Messung des Atemzugvolumens, des Atemwiderstands oder auch des Sauerstoffverbrauchs sehr aufwendig und artefaktanfällig sind. Der bekannteste und bis heute am häufigsten eingesetzte respiratorische Kennwert ist die Atemfrequenz, die üblicherweise bei Aktivierung im Vergleich zur Ruhebedingung ansteigt.

1.2.3.1 Experimentelle Befunde

Es gibt eine Reihe von neueren Untersuchungen, die eine Erhöhung der Atemfrequenz bei mentaler Belastung aufzeigen. Carroll, Turner und Hellawell (1986) ließen ihre Versuchspersonen arithmetische Aufgaben mit unterschiedlichem Schwierigkeitsgrad bearbeiten. Dazu wurden mithilfe einer Tonbandaufnahme mathematische Aufgaben und dazugehörige Antworten präsentiert. Die Versuchsperson musste entscheiden, ob die jeweilige Antwort richtig war oder nicht. Es zeigte sich in allen Bedingungen ein signifikanter Anstieg der Atemfrequenz, der mit zunehmender Aufgabenschwierigkeit stärker ausgeprägt war. Ein ebenfalls signifikanter Anstieg der Atemfrequenz fand sich bei Bearbeitung eines mental herausfordernden Computerspiels (Sims, Carroll, Turner & Hewitt, 1988; Turner et al., 1983, Turner & Carroll, 1985). Svebak (1982) untersuchte, welchen Effekt die Bearbeitung eines perzeptuell-motorischen Paradigmas auf verschiedene psychophysiologische Parameter hat. Die Aufgabe der Probanden bestand darin, am Bildschirm mittels Joystickbewegungen Bälle aufzufangen, die mit verschiedenen Geschwindigkeiten erschienen. Der Hälfte der Versuchspersonen wurden Elektroschocks bei mangelnder Aufgabenbewältigung angedroht. Es zeigt sich, dass es sowohl bei Androhung von Elektroschocks als auch ohne zu einem Anstieg der Atemfrequenz bei Aufgabenbearbeitung kam. Auch bei psychomotorischen Tracking-Aufgaben am Bildschirm stieg die Atemfrequenz signifikant an. Wurden die Probanden zusätzlich unter Stress gesetzt, indem man ihnen eine verlängerte Versuchsdauer bei Fehlern androhte, stieg die Atemfrequenz allerdings nicht noch weiter an (Ohsuga et al., 2001). In einer anderen Studie konnte gezeigt werden, dass bei einer

Aufgabe mit hohen Anforderungen an das Arbeitgedächtnis die Atemfrequenz ebenfalls signifikant anstieg. Die Höhe des Anstiegs war dabei abhängig von der Aufgabenschwierigkeit, die durch eine Erhöhung der Anzahl der Zielitems als auch durch eine Variation des Interstimulusintervalls manipuliert wurde (Backs & Seljos, 1994).

1.2.4 Psychophysiologische Beanspruchungsanalysen im Rahmen der Bildschirmarbeit

In den bisher erwähnten Untersuchungen wurde in erster Linie der Einfluss mentaler Aufgaben am Bildschirm, zum Beispiel Rechenaufgaben, auf die physiologischen Beanspruchungsindikatoren, nicht aber die Effekte typischer Bildschirmarbeit wie Textverarbeitung oder Dateneingabe untersucht. Die Zahl der Studien auf diesem Gebiet ist relativ umfangreich, allerdings konzentrieren sich viele dieser Untersuchungen auf die durch Bildschirmarbeit hervorgerufene Muskelaktivität, die mit Hilfe des Elektromyogramms erfasst wird. Birch, Juul Kristensen, Jensen, Finsen und Christensen (2000) untersuchten die Auswirkungen von Zeitdruck während simulierter Computerarbeit wie beispielsweise das Erstellen von Zeichnungen. Sie stellten fest, dass hoher Zeitdruck kombiniert mit geringen kognitiven Anforderungen die Muskelaktivität im Schulter- und Oberarmbereich signifikant erhöhte. In einer anderen Studie bestand die Aufgabe der Probanden darin, Texte – sowohl mit als auch ohne Zeitdruck – am Bildschirm zu editieren (Wahlström, Hagberg, Johnson, Svensson & Rempel, 2002). Es zeigte sich, dass die Aufgabenbearbeitung unter Zeitdruck zu stärkeren physiologischen Reaktionen führte im Vergleich zur Kontrollbedingung, in der ohne zeitliche Beschränkung Text editiert werden musste. Die Muskelaktivität war größer und auch Herzrate und Blutdruck stiegen stärker an. Zu kritisieren ist allerdings, dass bei dieser Untersuchung keine Ruhemessung stattgefunden hat und somit unklar bleibt, ob bereits die Kontrollbedingung einen Anstieg der Beanspruchungsindikatoren aufweist. Man kann aus den Ergebnissen nicht schließen, ob die Bildschirmarbeit ohne Zeitdruck auch zu physiologischen Beanspruchungen führte.

Kuhmann, Schäfer & Boucsein (1990) untersuchten die Beanspruchungswirkungen von Intra-Task-Wartezeiten bei einfachen Suchaufgaben (Sterzinger-Aufgaben), die mit einer

durch die experimentelle Wartezeitdauer bestimmten Verzögerung getrennt hintereinander dargeboten wurden. Es zeigte sich, dass kürzere Wartezeiten dabei zu höheren Herzfrequenzen und zu häufigeren Hautleitwertsreaktionen führten als lange Warteintervalle. Schleifer und Okogbaa (1990) analysierten die psychophysiologischen Effekte von Intra-Task-Wartezeiten während einer Dateneingabe-Aufgabe. Unabhängig von der Länge der Warteintervalle fanden sie eine reduzierte Herzrate sowie eine erhöhte Herzratenvariabilität im Vergleich zur Ruhebedingung, was im Sinne einer herabgesetzten Anstrengungsbereitschaft bzw. ansteigender Müdigkeit interpretiert wurde. Eine andere Untersuchung konnte zeigen, dass während der Eingabe von Zahldaten mithilfe der Tastatur über einen Arbeitszeitraum von insgesamt 120 Minuten sowohl die Herz- als auch die Atemfrequenz im Vergleich zu einer Ruhemessung und einer Entspannungsübung anstieg (Schleifer & Ley, 1994).

Zu kritisieren bei diesen Untersuchungen ist allerdings die Tatsache, dass die motorische Aktivität der Versuchspersonen keine Berücksichtigung fand. Während der Dateneingabe kommt es sowohl zu physischen als auch psychischen Anforderungen, die beide einen Einfluss auf die physiologischen Beanspruchungsindikatoren ausüben können. Garde, Laurson, Jorgensen und Jensen (2002) stellten beispielsweise bei der Analyse der Herzratenvariabilität fest, dass bei Bildschirmarbeit allein durch physische Aktivität – wie zum Beispiel Mausbewegungen oder Bedienung der Tastatur – ohne mentale Ansprüche die Leistung im mittleren Frequenzband signifikant anstieg. Wurden zusätzlich noch durch Bearbeitung des Stroop-Tests mentale Anforderungen an die Versuchsperson gestellt, stieg die Leistung im mittleren Frequenzbereich nicht weiter an. Allerdings erreichte die Herzrate während Aufgabenbearbeitung im Vergleich zur rein motorischen Kontrollbedingung höhere Werte. Kohlisch und Schäfer (1996) stellten dagegen fest, dass motorische Aktivität nicht per se einen Einfluss auf die physiologischen Kennwerte hat. Die Versuchspersonen mussten mithilfe kompensatorischer Tastaturdruckbewegungen in verschiedenen Geschwindigkeiten am Bildschirm eine sich bewegende Markierung innerhalb eines vorgegebenen Gebiets halten. Dabei wurden zusätzlich beispielsweise die Anforderungen an die Gedächtnisleistung variiert. Es zeigte sich, dass nur bei sehr schnellen Tastendruckintervallen ($<300\text{ms}$) die motorische Aktivität einen Einfluss auf die physiologischen Parameter hatte.

In den Studien von Schleifer und Okogbaa (1990) sowie Schleifer und Ley (1994) entspannten die Probanden während der Ruhemessung vollkommen und infolgedessen kann nicht ausgeschlossen werden, dass die Veränderungen der physiologischen Beanspruchungsindikatoren während der Aufgabenbearbeitung auch durch die motorische Aktivität bedingt waren.

1.3 Textdarstellung am Computerbildschirm

Bei der Arbeit mit Bürocomputern spielt die Präsentation von Texten am Bildschirm eine ganz besondere Rolle. Nur sehr wenige Abläufe können vollkommen ohne Textdarstellung auskommen, einige basieren sogar ausschließlich auf Text. Neben der Textdarstellung am Computer wird auch der normale Tagesablauf zusätzlich durch eine kontinuierlich zunehmende Anzahl an Mensch-Bildschirm-Schnittstellen bestimmt, wie beispielsweise bei der Benutzung von Geldautomaten oder diverser elektronischer Geräte.

1.3.1 Vergleich zwischen Bildschirm und Papier

In zahlreichen Studien hat sich gezeigt, dass das Lesen am Bildschirm weniger effizient ist im Vergleich zu gedrucktem Text auf Papier. Muter, Latremouille, Treurniet und Beam (1982) verglichen Buch und Bildschirm in Hinblick auf Lesegeschwindigkeit und Textverständnis. Es zeigte sich, dass das Lesen von Texten am Bildschirm um 28,5% langsamer war im Vergleich zum Buch, das heißt es wurden weniger Seiten gelesen. Bezüglich Textverständnis und des Auftretens physischer Beschwerden beim Lesen lag allerdings kein Unterschied zwischen den beiden Medien vor.

Aufgrund der weit verbreiteten Anwendung von Textverarbeitungsprogrammen stellt sich die Frage, auf welchem Wege die Texte am erfolgreichsten bezüglich Tipp- und Schreibfehlern kontrolliert werden können. Sollen die Texte am Bildschirm korrigiert werden?

Oder soll ein weiterer Schritt unternommen und der Text auf Papier gedruckt werden? Um diese Fragen zu beantworten, wählten Wright und Lickorish (1983) das Paradigma einer Lese-Korrektur-Aufgabe und untersuchten die Leistungsunterschiede zwischen Bildschirm und Papier. Insgesamt wurden vier Texte mit jeweils 1500 Wörtern präsentiert. Die Darstellung erfolgte in beiden Bedingungen in negativen Kontrastverhältnissen, das heißt weiße Schrift auf schwarzem Hintergrund. In jedem Text befanden sich neben Tipp- auch Grammatikfehler, um zu überprüfen, ob beim Lesen die Intention bestand, die Bedeutung des Textes zu erfassen. Auch hier stellte sich heraus, dass die Probanden in der Bildschirmbedingung signifikant langsamer und auch ungenauer waren als in der Papierbedingung, das heißt es wurden weniger Fehler gefunden. Wright und Lickorish schlugen vor, dass das Ergebnis auf Probleme bei der Lesbarkeit von Text am Bildschirm zurückzuführen sei, zumal die Lesekorrekturen-Leistung stark auf der Erkennung der Wortgestalt beruht (Monk & Hulme, 1983).

In einer Untersuchung von Gould und Grischkowsky (1984) mussten die Versuchsteilnehmer je einen gesamten Arbeitstag Fehlertexte am Bildschirm beziehungsweise auf Papier lesen und korrigieren. Auch in dieser Studie waren die Probanden signifikant langsamer in der Bildschirmbedingung. Die Genauigkeit war ebenfalls geringer, allerdings war dieser Effekt statistisch nicht signifikant. Bezüglich der physischen Beanspruchung – in erster Linie Augenbeschwerden – und psychischen Beanspruchung wie beispielsweise Müdigkeit unterschieden sich die beiden Bedingungen nicht.

Creed, Dennis und Newstead (1987) verglichen ebenfalls Papier und Bildschirm bei einer Lese-Korrektur-Aufgabe. Die Aufgabe der Probanden bestand darin, drei Texte zu lesen, die insgesamt neun Seiten füllten. Jeder Text enthielt 18 Fehler, die in drei Fehlerkategorien eingeordnet werden konnten. Dabei handelte es sich zum einen um grammatikalische Fehler und zum anderen um Tippfehler, wobei ein Buchstabe entweder durch einen visuell ähnlichen oder unähnlichen anderen Buchstaben ersetzt wurde. Die Texte wurden in negativer Polarität, das heißt weiße Schrift auf schwarzem Hintergrund präsentiert. Es zeigte sich auch in dieser Studie, dass in der Papierbedingung signifikant mehr Fehler entdeckt wurden. Visuell ähnliche Fehler waren dabei in beiden Bedingungen schwerer zu entdecken als die anderen Fehlerarten. Man konnte allerdings keinen Unterschied in der Lesegeschwindigkeit feststellen.

In einer weiteren Studie mussten die Probanden an vier aufeinander folgenden Tagen für jeweils 50 Minuten eine Lese-Korrektur-Aufgabe bearbeiten. An zwei Tagen lag die Pa-

pierbedingung vor, an den anderen beiden Tagen wurde der Text am Bildschirm in hellgrüner Schrift auf dunkelgrünem Hintergrund, das heißt in negativer Polarität, präsentiert. Es gab insgesamt sechs verschiedene Darbietungsreihenfolgen, die gleichmäßig auf die 24 Probanden verteilt wurden. Die Studie wurde in einem fensterlosen Raum mit künstlicher Beleuchtung durchgeführt, wobei allerdings keine Angaben über die Beleuchtungsstärke gemacht wurden. Auch in dieser Untersuchung konnte man feststellen, dass in der Bildschirmbedingung sowohl weniger Fehler entdeckt als auch weniger Seiten gelesen wurden. Die Anzahl der nicht entdeckten Fehler stieg in beiden Bedingungen während der 50-minütigen Aufgabendauer. Dieser Effekt war größer in der Bildschirmbedingung, was als ein stärkerer Anstieg des Müdigkeitsempfindens im Vergleich zur Papierbedingung interpretiert wurde (Wilkinson & Robinshaw, 1987).

Fasst man diese Studien zusammen, lässt sich sagen, dass die Lesegeschwindigkeiten am Bildschirm im Vergleich zur Papierbedingung reduziert sind. Die Bearbeitungsgenauigkeit scheint nur bei kognitiv anspruchsvolleren Aufgaben – beispielsweise das Entdecken von Grammatikfehlern im Rahmen von Lesekorrektur-Aufgaben – herabgesetzt zu sein. Bezüglich des Textverständnisses, das beispielsweise mithilfe von Fragen zum Text überprüft wurde, unterschieden sich die beiden Bedingungen aber nicht (vgl. Dillon, 1992, Mills & Weldon, 1987).

Es stellt sich die Frage, wodurch die Unterschiede in der Lesegeschwindigkeit zwischen Papier und Bildschirm hervorgerufen wurden. Es wurden einige Faktoren vorgeschlagen, von denen man annimmt, dass sie die Ergebnisse der verschiedenen Studien erklären können. Die Unterschiede konnten zum einen durch die Aufgabengestaltung oder zum anderen durch technische Aspekte des Bildschirms und der Textdarstellung wie beispielsweise Kontrastverhältnisse oder Schriftgröße begründet sein. Auch persönliche Faktoren wie Dauer der Bildschirmerfahrung wurden als mögliche Ursachen erwähnt.

Gould (1987) konnte aber zeigen, dass keiner der vorgeschlagenen Faktoren allein, sondern vielmehr eine Kombination der verschiedenen Variablen die Leistungsunterschiede zwischen der Papier- und Bildschirmdarstellung erklären kann. Dabei wurde die Bildqualität als wahrscheinlichste Ursache für die differente Lesegeschwindigkeit angegeben. In einer weiteren Untersuchung wurde festgestellt, dass durch eine Textdarstellung in positiver Polarität (dunkle Schrift auf hellem Hintergrund), die Verwendung eines Bildschirms

mit hoher Auflösung, sowie die Anwendung des Anti-Aliasing Verfahrens zur Bildkanten-glättung der Buchstaben die Unterschiede zwischen Papier und Bildschirm bezüglich der Leseleistung eliminiert werden konnten (Gould, Alfaro, Finn & Haupt, 1987). Auch Muter und Maurutto (1991) stellten im Rahmen einer Leseaufgabe fest, dass bei Verwendung eines Röhrenmonitors mit hoher Auflösung und positiver Polarität die Leseleistung – Anzahl der gelesenen Wörter sowie Textverständnis – sich nicht zwischen der Papier- und Bildschirmbedingung unterschied.

1.3.2 Einsatz von Farben im Rahmen der Bildschirmarbeit

Im Zeitalter des Internets und chromatischer Displays kommt es neben einer achromatischen Darstellung immer häufiger zum Einsatz von Farben im Bereich der Textpräsentation. Aus diesem Grund ist es wichtig zu untersuchen, wie die Verwendung von Farben die Textdarstellung am Bildschirm beeinflusst.

1.3.2.1 Grundlagen der Farbwahrnehmung

Für die Beschreibung des Farbempfindens ist es in erster Linie notwendig, zwischen achromatischen und chromatischen Farben zu unterscheiden. Während Weiß, Schwarz und alle Grauschattierungen zu den achromatischen Farben zählen, sind Blau, Rot, Grün und Gelb chromatische Farben. Gouras (1991) stellte fest, dass der Beobachter im sichtbaren Spektrum etwa 200 Farbabstufungen differenzieren kann. Diese Zahl lässt sich auf zwei Arten vervielfältigen. Zum einen kann durch Variationen der Intensität die Helligkeit der Farbe verändert werden, wobei der Mensch zwischen etwa 500 Helligkeitsabstufungen unterscheiden kann. Zum anderen besteht die Möglichkeit, die Sättigung zu verändern. Die Sättigung beschreibt, wie rein oder „ausgewaschen“ eine Farbe erscheint. Pastellfarben sind beispielsweise niedrig gesättigt. Physikalisch bezieht sich die Sättigung auf die Breite des Wellenlängenspektrums. Ist nur eine Wellenlänge vorhanden, ist die Farbe gesättigt.

Da jede Farbe bis zu 20 vom Menschen unterscheidbare Sättigungswerte aufweist, kann als Folge zwischen zwei Millionen Farben diskriminiert werden.

Die Photorezeptoren im menschlichen Auge sprechen auf elektromagnetische Wellen im Bereich von 350-750 nm an. Von zentraler Bedeutung für das Farbsehen sind die ca. fünf Millionen Zapfen, die sowohl in der Fovea als auch der Netzhautperipherie zu finden sind. Die Zapfen lassen sich aufgrund unterschiedlicher spektraler Absorptionsspektren, die durch drei verschiedene Pigmentarten bedingt sind, in drei Klassen einteilen: Es gibt Rezeptoren für den kurz-, mittel- und langwelligen Bereich des Spektrums.

Bei der Repräsentation von Farben im Sehsystem des Menschen spielen verschiedene Modelle eine Rolle, die auf unterschiedlichen Ebenen wirksam sind. Thomas Young (1773-1829) stellte bei Farbgleichexperimenten fest, dass aus drei Farben unterschiedlicher Wellenlänge eine Mischfarbe hergestellt werden kann, die einer vorgegebenen Vergleichsfarbe mit nur einer Wellenlänge entspricht. Aufgrund dieser Beobachtungen schlug er 1802 die trichromatische oder Dreifarbentheorie der Farbwahrnehmung vor, die 1852 von Hermann von Helmholtz (1821-1894) weiter ausgearbeitet wurde. Dieser Theorie zufolge beruht die Farbwahrnehmung auf drei Rezeptorsystemen mit jeweils unterschiedlicher spektraler Empfindlichkeit. Licht mit einer bestimmten Wellenlänge erregt die drei Rezeptortypen verschieden stark und das Aktivitätsmuster in den drei Systemen führt zur Farbwahrnehmung, wobei jede Farbe durch ein individuelles Aktivitätsmuster codiert wird.

Ewald Hering (1834-1918) entwickelte etwa zur gleichen Zeit die Gegenfarbentheorie. Er ging von vier so genannten „Urfarben“ (Rot, Gelb, Grün und Blau) aus und kam aufgrund phänomenologischer Beobachtungen zu dem Schluss, dass Rot und Grün sowie Blau und Gelb Farbpaare bilden. Im Zuge der Gegenfarbentheorie postulierte Hering drei Mechanismen, die jeweils antagonistisch auf Licht unterschiedlicher Wellenlänge oder Intensität reagieren. Der Schwarz-Weiß-Mechanismus reagiert positiv auf weißes Licht und negativ auf das Fehlen von Licht. Rot-Grün reagiert positiv auf Rot und negativ auf Grün und Blau-Gelb negativ auf Blau und positiv auf Gelb. Solche Gegenfarbencellen (opponent cells) konnten beispielsweise von Russell DeValois (1960) im Corpus geniculatum laterale des Rhesusaffen gefunden werden. Auf corticaler Ebene konnten doppelte Gegenfarbencellen (double-opponent-cell) nachgewiesen werden, deren Zentrum und Peripherie antagonistisch organisiert sind (Livingstone & Hubel, 1984). Beispielsweise wirken im Zent-

rum des rezeptiven Feldes die Zapfen für langwelliges Licht exzitatorisch und die Rezeptoren für mittlere Wellenlängen inhibitorisch. In der Peripherie verhält es sich umgekehrt. Solche doppelten Gegenfarbenzellen sind für Farbkontraste hoch empfindlich. Livingstone und Hubel (1984) konnten zeigen, dass die doppelten Rot/Grün-Gegenfarbenzellen beim Affen den häufigsten Farbzelltypus darstellen.

Die physiologische Forschung hat belegt, dass beide Theorien richtig sind und sich auf physiologischer Ebene miteinander vereinen lassen. Die trichromatische Theorie gilt demzufolge auf der Ebene der Photorezeptoren, die Gegenfarbentheorie dagegen für die weitere neuronale Verschaltung (siehe Abbildung 1.4).

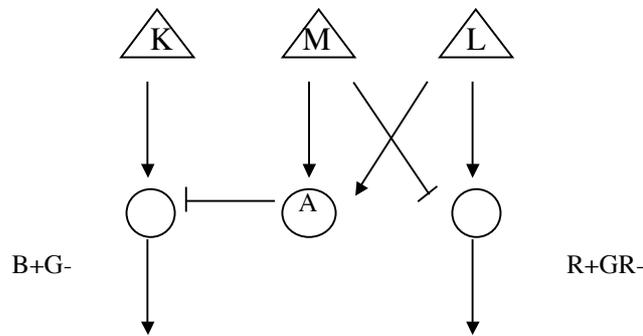


Abbildung 1.4: Schema eines neuronalen Schaltkreises, der aus den exzitatorischen und inhibitorischen Signalen der drei Zapfentypen die Blau-Gelb-, Rot-Grün- und Hell-Dunkel-Mechanismen erzeugt.

Wie bereits erwähnt wird eine wahrgenommene Farbe determiniert durch den Farbton, der sich im physikalischen Sinne auf die dominante Wellenlänge bezieht, die Sättigung und die Helligkeit (Luminanz). Bei gleich bleibendem Farbton kann eine Farbe dunkler oder heller erscheinen. Physikalisch versteht man unter der Helligkeit die Menge des Lichts, die im Auge eintrifft. Gerade bei der Darstellung von Text am Bildschirm spielen die Kontrastverhältnisse, das heißt sowohl Luminanz- als auch Chromatizitätskontrast, zwischen Text- und Hintergrundfarbe eine wichtige Rolle. Nach der Michelson-Definition wird der Luminanzkontrast nach folgender Formel berechnet:

$$\frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$$

Dabei stellen L_{\max} und L_{\min} den maximalen bzw. minimalen Luminanzwert von Text- und Hintergrundfarbe dar. Der Kontrast nimmt Werte zwischen 0 und 1 an.

Für das Lesen von Texten ist ein ausreichender Luminanzkontrast essentiell, da sich bei reduziertem Kontrast zwischen Text und Hintergrund die Worterkennung rapide verschlechtert. Snyder (1988) geht davon aus, dass der Kontrast mindestens bei .667, also einem Verhältnis von 3:1, und für eine bessere Leseleistung sogar bei .857 (7:1) liegen sollte. Zahlreiche Studien konnten zeigen, dass für die Leseleistung am Bildschirm der Luminanzkontrast von größerer Bedeutung zu sein scheint als der Farbkontrast und in erster Linie die Leserate determiniert. Solange der Luminanzkontrast ausreichend hoch war, schien jede gewählte Farbwahl von Text und Hintergrund geeignet zu sein, das heißt, die verschiedenen Farbkombinationen unterschieden sich nicht bezüglich der Leseleistung (Legge, Rubin & Luebker, 1987; Matthews, 1987; Travis, Bowles, Seton & Peppe, 1990). Des Weiteren konnte festgestellt werden, dass das Vorhandensein von sowohl Luminanz- als auch Farbkontrast keinen additiven Effekt auf die Leseleistung hat. Knoblauch, Arditi und Szlyk (1988) zeigten, dass bei ausreichend hohem Luminanzkontrast die zusätzliche Verwendung eines Farbkontrasts die Leserate nicht erhöhte. Auch Legge, Parish, Luebker und Wurm (1990) fanden keine Interaktion und folgerten, dass Farb- und Luminanzkontrast im visuellen System zwar ähnlich, aber in getrennten neuronalen Wegen verarbeitet werden.

1.3.2.2 Farbige Gestaltung von Text am Bildschirm

Eine farbige Darstellung im Bereich der Bildschirmarbeit ist allerdings als nicht unkritisch zu betrachten, da die Verwendung von Farben aus verschiedenen Gründen zu erhöhten Ansprüchen führt. Zum einen besitzt das menschliche Auge eine Farbfehlsichtigkeit, die als chromatische Aberration bezeichnet wird. Darunter versteht man die Tatsache, dass die Brechkraft des Auges von der Wellenlänge des Lichts abhängt. Kurzwelliges Licht wird stärker gebrochen als langwelliges. Diese Fehlsichtigkeit wird mit Hilfe des Ziliarmuskels ausgeglichen, was jedoch zusätzliche Einstellarbeit erfordert. Zum anderen wird durch Farbe der Kontrast zwischen Schriftzeichen und Hintergrund entweder zu groß oder zu klein. Dadurch kann es zu Blendungen oder mangelnder Lesbarkeit kommen.

Die Verwendung von Farben weist allerdings nicht nur Nachteile auf. Beispielsweise können Informationen durch Farben besser gegliedert werden als dies bei achromatischer Dar-

stellung möglich wäre. Auch das schnelle Identifizieren wichtiger Informationen wird erleichtert. Allerdings sollten dabei einige Faktoren berücksichtigt werden. Es sollten nur maximal vier verschiedene Farbtöne verwendet werden, die ausreichend voneinander unterscheidbar sind. Dabei ist es wichtig, Alltagserfahrungen im Umgang mit Farben zu berücksichtigen, zum Beispiel Rot für Gefahr (vgl. Richenhagen, Prümper & Wagner, 1998).

Vor allem im Bereich der Entwicklung von Internetpräsenzen zeigen sich die Designer sehr experimentierfreudig und man findet zahlreiche Buchstaben-Hintergrund-Farbkombinationen. Es ist allerdings davon auszugehen, dass – wenn auf Farbe nicht verzichtet werden kann – bei der Darstellung von Text am Bildschirm bezüglich ergonomischer Kriterien nicht alle Farbkombinationen gleichermaßen geeignet sind.

So untersuchte Matthews (1987) die Effekte acht verschiedener Farbkombinationen (Weiß, Gelb, Grün, Rot und Blau auf schwarzem Hintergrund, sowie Rot auf Blau, Blau auf Rot und Gelb auf Grün) auf die Leistung bei einer Lese-Korrektur-Aufgabe, wobei die Probanden die Helligkeit des Monitors selbst einstellen durften. Über die genauen Luminanzkontrastwerte wurden keine Angaben gemacht. Die Bearbeitungsdauer betrug insgesamt 150 Minuten und die Texte sollten so schnell und so genau wie möglich gelesen werden. Es zeigte sich, dass bei roter und blauer Textfarbe – sowohl in Kombination mit einem schwarzen als auch einem farbigen Hintergrund – die Leistung schlechter war als in den anderen getesteten Farbkombinationen. Das heißt, es wurde langsamer gelesen, und es wurden auch weniger Fehler gefunden. In einer Folgeuntersuchung kamen die gleichen Stimuli zur Anwendung, allerdings wurde der Luminanzkontrast von Text und Hintergrund im Vergleich zur ersten Studie dieses Mal variiert; er lag nun zwischen 0,18 und 0,97. Es stellte sich heraus, dass nur bei niedrigem Kontrast Leistungsunterschiede bezüglich Genauigkeit und Schnelligkeit zwischen den verschiedenen Farbkombinationen augenscheinlich wurden. Bei hohem Kontrast hatte die Farbwahl nur sehr wenig Einfluss auf die Leistung (Matthews, 1989). Matthews, Lovasik und Mertins (1989) analysierten den Einfluss sieben verschiedener Farbkombinationen (Grün, Rot und Blau auf Schwarz; Weiß auf Schwarz sowie Rot auf Blau, Rot auf Grün und Blau auf Grün) auf die Performanz bei einer visuellen Suchaufgabe. Die Versuchspersonen mussten über eine Dauer von vier Stunden bestimmte Zielitems innerhalb einer Menge von Distraktoren finden. Es zeigten sich nur geringe Unterschiede in Hinblick auf die Diskriminationsgenauigkeit und die Re-

aktionszeiten zwischen den verschiedenen Farbbedingungen. Die Rot/Grün-Bedingung wies allerdings signifikant längere Reaktionszeiten auf als die anderen Farbkombinationen und zeigte auch eine geringere Genauigkeit. In eine Studie von Pastoor (1990) musste eine Textdarbietung in insgesamt 800 unterschiedlichen Farbkombinationen mit festgesetztem Luminanzkontrast von 10:1 bzw. 6,5:1 bezüglich der subjektiven Präferenz eingestuft werden. Es zeigte sich, dass die Farbsättigung den größten Einfluss auf die Bewertungen hatte. Alle ungesättigten Farbkombinationen wurden für die Textpräsentation als geeignet eingeschätzt. Im zweiten Teil der Untersuchung wurden insgesamt neun der 800 Farbkombinationen ausgewählt und sowohl in positiver (dunkle Schrift auf hellem Hintergrund) als auch negativer Polarität (helle Schrift auf dunklem Hintergrund) präsentiert. Aufgabe der Probanden war es, so schnell wie möglich bestimmte Zielworte innerhalb von Distraktoren zu finden. Es stellte sich heraus, dass weder die Farbwahl noch die Polarität der verschiedenen Kombinationen einen Effekt auf die Lesegeschwindigkeit und die Auffindegenauigkeit hatte. Die subjektive Bewertung der Farbkombinationen ergab allerdings ein differentes Bild. Die Schwarz/Weiß-Darstellung wurde bezüglich ästhetischer Faktoren und Lesbarkeit deutlich positiver eingestuft als die Textdarstellung in Gegenfarben wie beispielsweise Rot/Grün, die wiederum zu mehr physischen Beschwerden führte. Der nicht vorhandene Leistungsunterschied zwischen den verschiedenen Farbkombinationen wurde von Pastoor durch eine erhöhte Anstrengungsbereitschaft bei schlechterer Lesbarkeit erklärt.

Shieh, Chen und Chuang (1997) präsentierten mit einer Darbietungsdauer von 250 msec chinesische Schriftzeichen in insgesamt fünf verschiedenen Farbkombinationen (Weiß auf Blau, Gelb auf Blau, Weiß auf Schwarz, Schwarz auf Weiß sowie Rot auf Grün). Es zeigte sich, dass der Prozentsatz richtig erkannter Zeichen sich zwischen den Farbkombinationen nicht unterschied. Das Ergebnis der subjektiven Bewertung wies allerdings auch in dieser Studie ein anderes Bild auf. Weiße Schrift auf blauem Hintergrund wurde am besten bewertet, während die Rot/Grün-Kombination als am wenigsten angenehm eingestuft wurde. In einer weiteren Studie untersuchten Shieh und Chen (1997) den Sehabstand zum Bildschirm bei einer visuellen Suchaufgabe in Abhängigkeit von der gewählten Farbkombination und stellten fest, dass der Abstand bei roter Schrift auf grünem Hintergrund signifikant kürzer war als in den anderen Bedingungen, die sich nicht voneinander unterschieden. In früheren Studien konnte gezeigt werden, dass der Sehabstand mit dem Auftreten visueller Beschwerden korrelierte. Bei kurzen Abständen waren die Beschwerden stärker ausgeprägt

(Jaschinski-Kruza, 1991). Die Sehdistanz kann somit als ein Indikator für die Einschätzung der Effektivität bestimmter Farbkombinationen herangezogen werden. Wie bereits in einigen Untersuchungen gezeigt, scheint auch in der Studie von Shieh und Chen die Rot/Grün-Farbkombination den anderen Darstellungsmöglichkeiten unterlegen zu sein, da in dieser Bedingung der Sehabstand am geringsten war.

1.3.3 Polarität

Texte können auf verschiedene Arten präsentiert werden. Neben der farblichen Darstellung spielt der Kontrast zwischen Schrift und Hintergrund eine wichtige Rolle. Im Bereich der Textverarbeitung hat sich seit Beginn der 80iger Jahre zunehmend eine Textdarstellung in positiver Polarität bzw. positivem Kontrast (dunkle Schrift auf hellem Hintergrund) durchgesetzt und wird mittlerweile standardmäßig eingesetzt und auch nach den Sicherheitsregeln für Bildschirmarbeitsplätze empfohlen. Begründet wird diese Tatsache beispielsweise dadurch, dass bei positiven Kontrastverhältnissen die Bildschirmanzeige weniger anfällig für Blendungen ist als bei negativer Polarität (vgl. Krüger, 1984). Hier erhöhen Reflektionen die mittlere Luminanz des Monitors, wodurch allerdings der Kontrast zwischen Text und Hintergrund reduziert wird, was einen negativen Effekt auf die Lesbarkeit hat. Außerdem weist die Positivdarstellung eine höhere Leuchtdichte auf als die Negativdarstellung. Die höhere Leuchtdichte der Positivdarstellung führt im Zusammenhang mit den für andere Tätigkeiten erforderlichen Beleuchtungsstärken zu einer Verringerung beanspruchender Hell-Dunkel Adaptationen. Auch kommt es zu einer Angleichung der Leuchtdichten insbesondere von Papier-Vorlage und Bildschirmanzeige (vgl. Richenhagen, Prümper & Wagner, 1998).

Weiterhin wird die Lesbarkeit der Zeichen verbessert, weil die höhere Leuchtdichte der Positivdarstellung eine Pupillenverengung bewirkt, wodurch sich die Qualität des retinalen Bildes erhöht (Krüger, 1984). Dies geschieht zum einen durch eine Reduzierung der sphärischen Aberration, worunter man eine mangelnde Bildschärfe durch kleine Brechungsfehler in Linse und Cornea versteht. Außerdem wird durch den geringeren Lichteinfall die visuelle Feldtiefe vermindert, und Objekte am Bildschirm können leichter fokussiert werden (Taptagaporn & Saito, 1990). Da darüber hinaus in der Regel sämtliche Printmedien

wie Bücher, Zeitungen etc. in positiver Polarität dargestellt werden, kann man davon ausgehen, dass diese Art der Textdarstellung wesentlich vertrauter ist, was einen zusätzlichen Vorteil darstellt.

Die Positivdarstellung weist aber auch einen Nachteil auf. Aufgrund der höheren Leuchtdichte kann ein Flimmern wahrgenommen werden, das bei negativem Kontrast weniger stark ausgeprägt ist (vgl. Krüger, 1984). Wichtig ist deshalb die Wahl eines Röhrenmonitors mit einer ausreichend hohen Bildwiederholrate von mindestens 90 Hz oder eines TFT-Monitors. TFT-Bildschirme weisen gegenüber herkömmlichen CRT-Monitoren einige technische Vorteile auf. Aufgrund der konstanten Hintergrundbeleuchtung und der Tatsache, dass die Transistoren stabil sind, arbeiten sie flimmerfrei. Darüber hinaus ist bei TFT-Monitoren im Vergleich zum Röhrenmonitor der Unterschied bezüglich der vertikalen und horizontalen Kontrastübertragung – das heißt die senkrechten Linien von Schriftzeichen sind dunkler als die horizontalen Linien – deutlich geringer. Diese Tatsache ist von großer Bedeutung für die Messung der Luminanzwerte von Text und Hintergrund, da eine größere Fläche gemessen und dann auf die Schriftzeichen extrapoliert wird. Dieses Vorgehen würde bei Röhrenmonitoren aufgrund der unterschiedlichen horizontalen und vertikalen Kontrastübertragung zu sehr fehlerhaften Werten führen. Neben diesen technischen Argumenten zeigten sich in einigen Studien auch auf Performanzebene deutliche Vorteile für TFT-Monitore. Menozzi, Lang, Naepflin, Zeller und Krüger (2001) verglichen Röhren- und TFT-Monitore im Zuge einer visuellen Suchaufgabe. Aufgabe der Versuchspersonen war es, einen Zielbuchstaben unter Distraktoren zu entdecken. Es konnte gezeigt werden, dass bei Nutzung des TFT-Monitors die Fehlerhäufigkeit um 22 Prozent geringer war als bei Verwendung eines CRT-Bildschirms mit einer Bildwiederholrate von 75 Hz. Eine weitere Vergleichsstudie wurde von Naesaenen, Karlsson und Ojanpää (2001) durchgeführt. Im Rahmen einer visuellen Diskriminationsaufgabe zeigten sich bei hohem Kontrast keine Unterschiede zwischen CRT- und TFT-Bildschirmen. Bei geringerem Kontrast und kleinerer Schriftgröße waren die Suchzeiten bei Verwendung des Röhrenmonitors deutlich länger.

1.3.3.1 Experimentelle Befunde zur Polarität

In zahlreichen Untersuchungen wurde bei achromatischer Textdarstellung mittels verschiedener Paradigmen der Einfluss der Polarität auf die Leseleistung untersucht, das heißt, es wurde entweder schwarze Schrift auf weißem Hintergrund präsentiert oder umgekehrt. Die Resultate dieser Studien weisen allerdings kein homogenes Bild auf.

In einer Untersuchung von Bauer und Cavonius (1980) wurden einzelne Wörter, die jeweils aus vier Buchstaben bestanden und keinen Sinn ergaben, kurz am Bildschirm präsentiert. Die Aufgabe der Probanden bestand darin, die Wörter zu lesen und anschließend aufzuschreiben. Es wurde ein Monitor mit einer Bildwiederholfrequenz von 100 Hz verwendet, um das Entstehen von Flimmern zu vermeiden. Es zeigte sich, dass bei positiven Kontrastverhältnissen 26 Prozent weniger Fehler gemacht wurden als in der Negativdarstellung. Im zweiten Teil der Untersuchung wurde erneut negativer und positiver Kontrast verglichen, wobei die Negativdarstellung auf einem 50 Hz Monitor und die Positivdarstellung auf einem 100 Hz Bildschirm erfolgte. Die Versuchspersonen mussten Unterschiede entdecken zwischen Wörtern, die kurz am Bildschirm dargeboten wurden und auf Papier gedruckt waren. Auch hier war die Leistung der Probanden in der Positivdarstellung besser. Es wurden 23 Prozent weniger Fehler gemacht und die Bearbeitungszeit war um acht Prozent schneller. Darüber hinaus wurde von 95 Prozent der Versuchspersonen die positive Polarität subjektiv bevorzugt.

Cushman (1986) untersuchte den Einfluss der Polarität auf das Lesen von Texten am Bildschirm. Es zeigte sich, dass nach 80minütiger Lesezeit die Versuchspersonen weniger visuelle Beschwerden berichteten, die den Text in Negativdarstellung lasen im Vergleich zur positiven Polarität. Dieses Ergebnis sah Cushman darin begründet, dass bei der Untersuchung Bildschirme mit einer Bildwiederholrate von 60 Hz verwendet wurden und diese bei positiven Kontrastverhältnissen zu einer Wahrnehmung von Flimmern führen können. Es zeigten sich allerdings keine signifikanten Unterschiede bezüglich Lesegeschwindigkeit und Textverständnis zwischen den beiden Bedingungen. Legge, Rubin und Luebker (1987) untersuchten bei insgesamt nur zwei Versuchspersonen den Einfluss des Kontrasts auf das Lesen eines Fließtextes am Bildschirm. Die Fließgeschwindigkeit wurde solange erhöht, bis die Versuchsperson begann, Fehler zu machen. Auch in dieser Studie hatte die Wahl

der Polarität keinen Einfluss auf die Leistung, das heißt die Leserate unterschied sich nicht bei positiver und negativer Darstellung.

Creed, Dennis und Newstead (1988) verwendeten das Paradigma einer Lese-Korrektur-Aufgabe. In vier Texten, die jeweils 800 Wörter umfassten, mussten sowohl Tipp- als auch Grammatikfehler mittels eines Joysticks markiert werden. Die Versuchspersonen wurden instruiert, so schnell und so genau wie möglich zu arbeiten. Eine zeitliche Begrenzung bestand nicht. Die Ergebnisse zeigten keine Unterschiede zwischen positivem und negativem Kontrast, weder in der durchschnittlichen Anzahl gefundener Fehler noch in der benötigten Zeit. Allerdings wurden in dieser Untersuchung Monitore mit geringer Bildwiederholrate verwendet, was eine Erklärung für den nicht vorhandenen Unterschied darstellen kann. Gould et al. (1987) untersuchten ebenfalls den Einfluss der Polarität bei einer Lese-Korrektur-Aufgabe. Insgesamt mussten fünf Artikel mit jeweils 1000 Wörtern gelesen werden und es zeigte sich, dass die Lesegeschwindigkeit bei positiver Darstellung etwas höher war.

Fasst man die Ergebnisse der verschiedenen Untersuchungen zusammen, zeigte sich auf Performanzebene entweder ein Vorteil für eine Textdarstellung in positiver Polarität (Bauer & Cavonius, 1980; Gould et al., 1987) oder kein Unterschied zwischen negativem und positivem Kontrast (Creed et al., 1988; Cushman, 1986; Legge et al., 1987). In keiner Studie konnte allerdings eine bessere Leistung bei einer Textdarstellung in negativen Kontrastverhältnissen festgestellt werden. Dass die Studien zur Polarität ein heterogenes Bild aufweisen und sich nicht immer ein Vorteil für eine Textdarstellung in positiver Polarität – wie sie heute standardmäßig eingesetzt wird – zeigte, kann verschiedene Ursachen haben. Zum einen kann dieses Ergebnis auf die verwendeten Monitore zurückgeführt werden. Darüber hinaus war eventuell in den Experimenten, in denen sich kein Unterschied zwischen positiver und negativer Polarität zeigte, die Anzahl der Versuchspersonen, die zwischen 2 und 76 lag, hinsichtlich Teststärkeüberlegungen zu gering, um mögliche Effekte zu entdecken.

1.3.3.2 Polarität bei Dunkelheit

Es wird allerdings argumentiert, dass der Vorteil einer Textdarstellung in positiver Polarität nicht uneingeschränkt gültig ist und bei dunkler Umgebung eine Negativdarstellung günstiger zu sein scheint. Diese Empfehlungen stammen allerdings in erster Linie von Internetseiten zum Thema Ergonomie und Bildschirmarbeit. „Ein Kriterium für die Auswahl einer Polarität ist die Reduzierung des Helligkeitskontrasts zwischen Bildschirm und Umgebung. Bei hellerer Umgebung ist die positive Polarität günstiger. Soll eine Dunkeladaptation der Benutzer erhalten bleiben, dann ist die negative Polarität angemessener.“ (DBAG Forschungsteam MMI, 1998).

Während der Dunkeladaptation nimmt die absolute Empfindlichkeit des Sehsystems langsam zu. Dieser Prozess setzt sich aus einem schnellen Anfangsstadium, das durch eine Zapfenadaptation gekennzeichnet ist, und eine spätere langsamere Phase, während derer die Stäbchen adaptieren, zusammen. Nach etwa 20 bis 30 Minuten ist die maximale Empfindlichkeit erreicht, wobei das Stäbchensystem eine deutlich höhere Empfindlichkeit aufweist als das Zapfensystem. Die absolute Sehschärfe im dunkeladaptierten Zustand bleibt jedoch erheblich unter den Werten des Tageslichtsehens.

Bei Dunkelheit oder nur schemenhaft beleuchteten Räumen könnte ein heller Hintergrund zu Blendungen und daraus resultierenden Augenbeschwerden führen. „If the room is dimly illuminated, then a bright white screen can be a glare source and create eye discomfort.“ (Ergo/Gero, 1998-2001). Diese beiden Angaben sollen exemplarisch für die zahlreichen Hinweise dargestellt werden, die sich zu diesem Thema im Internet finden.

Trotz einer umfangreichen Literaturrecherche konnte nur eine Arbeit gefunden werden, die nicht bloß pauschale Empfehlungen ohne offensichtliche Grundlage gibt, sondern sich auch wissenschaftlich mit den Auswirkungen der Kontrastwahl bei Bildschirmarbeit in Räumen mit niedriger Umgebungsbeleuchtung beschäftigt hat. Taptagaporn und Saito (1990) untersuchten den Einfluss der Bildschirm-Polarität unter verschiedenen Lichtbedingungen auf die Pupillengröße bei Computer-Nutzern. Der erste Teil der Studie bestand aus einer visuellen Suchaufgabe am Bildschirm, die unter drei Beleuchtungsbedingungen (dunkel, 500 lx, 1200 lx) durchgeführt wurde. Es zeigte sich, dass in Bedingungen mit negativem Kontrast der Pupillendurchmesser größer war im Vergleich zur Positivdarstellung. Bei einer Textpräsentation in negativer Polarität variierte der Pupillendurchmesser in Ab-

hängigkeit von der Raumbeleuchtung. Mit zunehmender Beleuchtungsstärke nahm die Pupillenweite ab. Bei einer Darstellung in positiven Kontrastverhältnissen zeigte sich kein Unterschied im Pupillendurchmesser zwischen den verschiedenen Lichtbedingungen. Im zweiten Teil der Untersuchung mussten die Probanden mit einem Sehabstand von 50 cm abwechselnd auf Monitor, Vorlage und Tastatur blicken. Man konnte feststellen, dass bei Dunkelheit die Variationen im Pupillendurchmesser beim Blickwechsel deutlich größer waren bei positivem als bei negativem Kontrast. Bei Licht verhielt es sich umgekehrt, das heißt die Pupillenweite variierte stärker bei negativer Polarität. Aus diesen Ergebnissen wurde die Schlussfolgerung gezogen, dass bei Beleuchtung die Arbeit unter negativen Kontrastverhältnissen eine deutlich stärkere retinale Hell-Dunkeladaptation erfordert, was ein Grund für Augenbeschwerden sein kann. In dunklen Räumen wird dagegen bezüglich der retinalen Adaptation der negative Kontrast als vorteilhafter angesehen. Die Dunkeladaptation kann besser beibehalten werden und die Variationen im Pupillendurchmesser sind geringer, da der Leuchtdichteunterschied zwischen Umgebung und Bildschirm kleinere Werte aufweist.

Aus diesen Befunden leiteten Taptagaporn und Saito (1990) die allgemeine Empfehlung ab, bei heller Umgebung den positiven Kontrast zu bevorzugen und bei Dunkelheit auf negative Kontrastverhältnisse zurückzugreifen. Dieser Bezug ist oberflächlich durchaus nachvollziehbar. Allerdings ist er indirekt und im Grunde nicht gerechtfertigt, weil die entscheidende Variable – die Performanz – selbst nicht gemessen wurde. Es wird lediglich aus Indikatoren – Variationen im Pupillendurchmesser – ein Rückschluss auf das eigentliche Problem gezogen. Dabei bleibt unklar, welchen Einfluss die Kombination von Polarität und Umgebungsbeleuchtung auf die Leistung der Probanden hat. Es stellt sich die Frage, ob bei geringer Beleuchtung die Leistung bei einer Textdarstellung in negativer Polarität tatsächlich besser ist als unter positiven Kontrastverhältnissen. Nach bisherigem Forschungsstand liegt keine Studie vor, in der dieser Fragestellung nachgegangen und der Effekt auf die Performanz bei typischer Bildschirmarbeit systematisch untersucht wurde.

1.4 Zusammenfassung des aktuellen Forschungsstands

Im Bereich der Bildschirmarbeit spielt die Präsentation von Texten am Bildschirm eine ganz besondere Rolle, und vor allem der Kontrast zwischen Schrift und Hintergrund ist dabei von Bedeutung. Bezüglich der Textverarbeitung hat sich seit Beginn der 80iger Jahre zunehmend eine Textdarstellung in positiver Polarität durchgesetzt und wird mittlerweile standardmäßig eingesetzt. Es wurde allerdings die Empfehlung ausgesprochen, bei Dunkelheit negativen Kontrast zu bevorzugen, was bisher auf Performanzebene aber nicht experimentell untersucht worden ist.

In verschiedenen Studien konnte außerdem gezeigt werden, dass es im Rahmen von Bildschirmarbeit zum Auftreten sowohl psychischer als auch physischer Belastungen und Beanspruchungen kommt. Das Ausmaß der Beanspruchung wird allerdings häufig nur auf subjektiver Ebene beschrieben, und es bleibt offen, inwieweit sich die Beanspruchung bei Bildschirmarbeit auch objektiv mit Hilfe psychophysiologischer Beanspruchungsindikatoren nachweisen lässt. Die Studien, in denen physiologische Indikatoren zur Beanspruchungsanalyse eingesetzt wurden, weisen zwei entscheidende Nachteile auf. Sie konzentrieren sich lediglich auf die durch Bildschirmarbeit hervorgerufene Muskelaktivität und berücksichtigen nicht den Einfluss körperlicher Aktivität auf physiologische Kennwerte.

Ausgehend vom momentanen Forschungsstand sollen die folgenden Experimente einen Beitrag dazu leisten, unser Wissen über eine optimale Präsentation von Texten am Bildschirm in Abhängigkeit von der Umgebungsbeleuchtung zu erweitern. Außerdem sollen neue Erkenntnisse über die beanspruchende Wirkung der Bildschirmarbeit auf physiologischer Ebene gewonnen werden. Für die vorliegende Arbeit ergeben sich folgende Hypothesen.

1.5 Fragestellungen und Hypothesen

1.5.1 Einfluss der Polarität auf die Performanz

Bei einer hohen Umgebungsbeleuchtung von 550 Lux zeigt sich im Rahmen einer Lesekorrekture-Aufgabe auf Performanzebene ein Vorteil für eine Textdarstellung in positiver Polarität. Bei einer niedrigen Umgebungsbeleuchtung von 5 Lux dreht sich der Polaritätseffekt um, das heißt, es zeigt sich eine bessere Leistung für eine Textpräsentation in negativen Kontrastverhältnissen. Darüber hinaus wird erwartet, dass der Polaritätseffekt nicht nur unter achromatischen Bedingungen zu finden ist, sondern auch auf farbige Textpräsentationen verallgemeinert werden kann.

1.5.2 Beanspruchung bei Bildschirmarbeit

Die beanspruchende Wirkung der Bildschirmarbeit kann unter Kontrolle der motorischen Aktivität auch objektiv mit Hilfe einer mehrdimensionalen psychophysiologischen Beanspruchungsanalyse, in deren Rahmen kardiovaskuläre und elektrodermale Parameter sowie die Atmung erfasst werden, nachgewiesen werden. Bei der Bildschirmarbeit zeigen sich darüber hinaus psychische und physische Beschwerden. Es wird die Vermutung aufgestellt, dass sich die Probanden im Ausmaß ihrer objektiven und subjektiven Beanspruchung in Abhängigkeit von den verschiedenen Arten der Textpräsentation unterscheiden. Für den Fall, dass sich alle Probanden gleich stark anstrengen, werden keine Display-abhängigen Unterschiede in den Beanspruchungsindikatoren, sondern nur auf Performanzebene erwartet.

Insgesamt wurden drei Experimente durchgeführt, um die genannten Fragestellungen zu klären. Die ersten beiden Experimente beschäftigen sich mit dem Polaritätseffekt bei hoher und niedriger Umgebungsbeleuchtung, sowie der subjektiven und objektiven Beanspruchung.

chung bei Bildschirmarbeit. Experiment 3 untersucht die Generalisierung des Polaritätseffekts auf chromatische Bedingungen.

2 Experiment 1

2.1 Methode

In den folgenden Abschnitten erfolgt eine Darstellung des ersten Experiments. Neben der untersuchten Stichprobe werden der Versuchsaufbau, die verwendeten Stimuli, das experimentelle Design und der Versuchsablauf beschrieben.

2.1.1 Stichprobenbeschreibung

Insgesamt nahmen am ersten Experiment 80 Probanden (48 weiblich, 32 männlich) im Alter von 19 - 55 Jahren ($M = 28,48$; $SD = 8,28$) teil. Die Zusammensetzung der verschiedenen Gruppen wurde bezüglich Alter, Geschlecht und Tageszeit angeglichen, da diese Faktoren einen Einfluss auf die physiologischen Beanspruchungsindikatoren aufweisen. 77 der Teilnehmer waren nach eigener Einschätzung Rechtshänder und drei Linkshänder. Die durchschnittliche Erfahrung mit Bildschirmarbeit betrug 8,4 Jahre ($SD = 5,11$). Alle Probanden hatten Deutsch als Muttersprache und litten nach eigenen Angaben nicht unter kardiovaskulären Erkrankungen. Die Teilnehmer wurden gebeten, mindestens sechs Stunden vor dem Versuch kein Koffein zu konsumieren und auf sehr anstrengende körperliche Aktivitäten zu verzichten. Alle Probanden hatten eine normale oder korrigierte Sehfähigkeit. Für die Teilnahme am Experiment wurde eine Entlohnung von 5 Euro gezahlt.

2.1.2 Versuchsaufbau

Zur Präsentation der Stimuli wurde ein Programm mithilfe der Metacard Version 2.4.3 erstellt. Die Durchführung des Experiments erfolgte an einem Apple Power Mac G4 Computer, und die Stimuli wurden an einem 17 Zoll TFT (=Thin Film Transistor) Farbmonitor

dargeboten. Die Ausmessung der Luminanzwerte sowie Chromatizitäts-Koordinaten der Vorder- und Hintergrundfarben erfolgte mithilfe eines Minolta Colormeter CS-100.

Der Versuchsraum wurde durch Jalousien vollkommen abgedunkelt, und der Einfall von Tageslicht wurde somit verhindert. Die Raumtemperatur lag gleich bleibend bei 23 Grad Celsius. Diese konstante Temperatur war von besonderer Bedeutung für die Messung der elektrodermalen Aktivität, da es bei einer hohen Körperkerntemperatur zu einer Erhöhung des Leitwerts kommt.

Wie in Abbildung 2.1 erkennbar saß der Proband während des Versuchs auf einem höhenverstellbaren Stuhl mit einem Sehabstand von 56 cm vor dem Bildschirm. Um eine ruhige Sitzhaltung sicherzustellen und zugleich den Sehabstand für alle Versuchspersonen zu standardisieren, wurde eine Kinnstütze verwendet. Sowohl die Höhe der Kinnstütze als auch des Stuhls konnten vor Versuchsbeginn vom Probanden individuell adjustiert werden. Die Bedienung der Computermaus erfolgte für alle Versuchsteilnehmer – auch für Linkshänder – mit der rechten Hand.

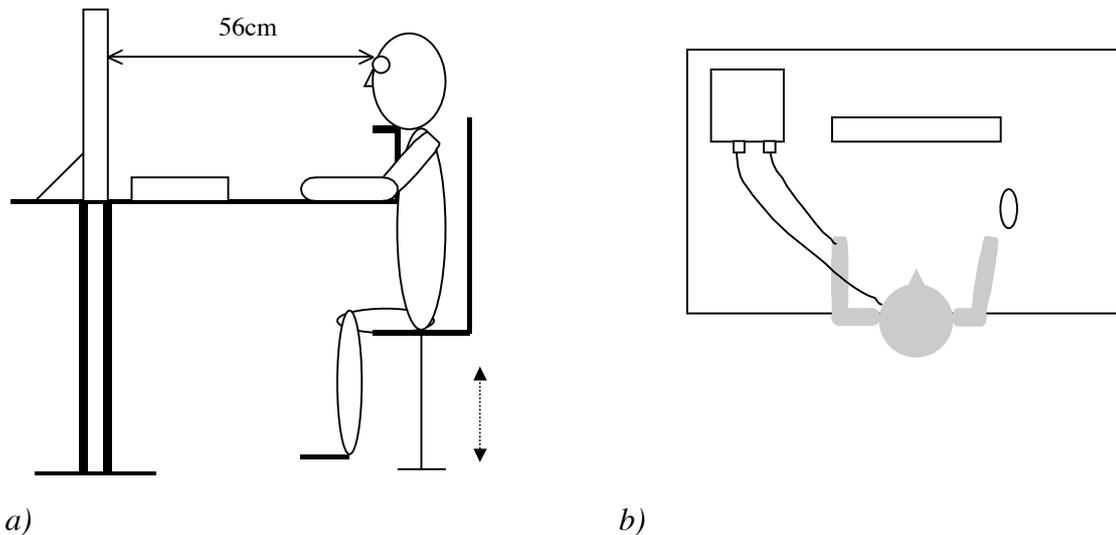


Abbildung 2.1: Schematische Abbildung des Versuchsaufbaus aus der Seitenperspektive (a) und Aufsicht (b). Die Versuchsperson saß mit einem Sehabstand von 56 cm vor dem Monitor. Das physiologische Messsystem „Powerlab“ befand sich ebenfalls auf dem Arbeitstisch.

2.1.2.1 Objektive Belastungsindikatoren

Die physiologischen Messdaten wurden erhoben mit dem Messsystem Powerlab/8SP, einem 16 Bit Analog-Digital-Wandler. Das Messgerät hat acht analoge Eingänge und kann Signale im Bereich von einigen Mikrovolt bis zu einem Maximum von etwa 10 Volt erfassen.

Zur Berechnung der kardiovaskulären Parameter wurde der Puls durch einen Infrarot-Puls-Plethysmographen (Modell 1020 EC) erfasst. Mit diesem Plethysmographen können durchblutungsbedingte Volumenschwankungen gemessen und in elektrische Spannung umgewandelt werden. Das analoge Eingangssignal wurde mit einem 200 Hz Tiefpass-Filter gefiltert und einer Abtastrate von 1000 Hz digitalisiert. Für die vorliegende Arbeit wurde ein Plethysmograph gewählt, der mittels eines Clips am Ohrläppchen der Versuchsperson befestigt werden konnte.

Die Messung der Hautleitfähigkeit erfolgte durch einen Vorverstärker (ML 116), der mit dem Powerlab verbunden war. Die bipolaren, nicht polarisierbaren Messelektroden wurden an den mittleren Segmenten von Zeige- und Mittelfinger der linken Hand befestigt. Der elektrisch isolierte Verstärker gab eine konstante Erregungsspannung ($22\text{mV}_{\text{rms}}@75\text{Hz}$) zu einer der beiden Elektroden aus. Hatte die Haut eine messbare elektrische Leitfähigkeit, floss Strom über die andere Elektrode in den Eingangskanal des Verstärkers. Anschließend wurde das Signal korrigiert und mit einem 1-Hz Tiefpass-Filter zweiter Ordnung gefiltert, wodurch hochfrequente Fluktuationen eliminiert wurden. In dem so geschaffenen Stromkreis variierte bei konstant angelegter Spannung die Leitfähigkeit, die in der Einheit Microsiemens angegeben wird, direkt mit der gemessenen Stromstärke („Konstantspannungsprinzip“) nach der Formel

$$I = U \times C.$$

wobei I = Stromstärke, U = Spannung und C = Leitwert.

Die Abtastrate wurde auf 10 Hz festgelegt. Die Hand der Versuchsperson lag bequem und leicht gekrümmt in natürlicher Ruhestellung auf einer weichen Unterlage auf dem Tisch.

Die Messung der Atemfrequenz erfolgte über einen Atemgurt (MLT 1132), der direkt mit dem Powerlab verbunden war, mit einer Abtastrate von 40 Hz. Der Gurt maß atmungsbedingte Veränderungen im Brust- und Bauchumfang. Ermöglicht wurde dies durch einen Sensor, der zwischen zwei elastischen Bändern befestigt war. Eine Zunahme des Brustumfangs bewirkte eine Dehnung des Sensors, wodurch linear zur Längenzunahme eine positive Spannung erzeugt wurde.

Um die Registrierung der physiologischen Parameter mit der Versuchsdurchführung zu synchronisieren, wurde bei Präsentation einer neuen Textseite von einer im Versuchsrechner eingebauten I/O-Karte von National Instruments ein externes elektrisches Auslösesignal mit einer Spannung von 5 Volt an das physiologische Messgerät gesendet, dass daraufhin mit der Datenaufnahme begann.

2.1.2.2 Subjektive Befindlichkeit und physische Beschwerden

Zur Messung der psychischen Beanspruchung wurde die in zwei Versionen vorliegende Kurzform des Mehrdimensionalen Befindlichkeitsfragebogens (MDBF) (Steyer, Schwenkmezger, Notz & Eid, 1997) verwendet. Der MDBF ist ein Instrument dreier bipolar konzipierter Dimensionen zur Erfassung der aktuellen psychischen Befindlichkeit: „Gute/Schlechte Stimmung (GS)“, „Wachheit/Müdigkeit (WM)“ und „Ruhe-/Unruhe (RU)“. Jede der drei Skalen setzt sich aus insgesamt vier Items zusammen. Die Items bestehen aus einfachen Adjektiven wie „munter“, „entspannt“ etc., die auf einer fünfstufigen Antwortskala mit den Endpunkten 1 („überhaupt nicht“) und 5 („sehr“) beurteilt werden. Die Items aller drei Skalen sind in Tabelle 2.1 dargestellt.

Tabelle 2.1

Zuordnung der Items zu den Skalen und den Kurzformen		
Skala	Kurzform A	Kurzform B
GS	zufrieden	wohl
	gut	glücklich
	schlecht	unglücklich
	unwohl	unzufrieden
WM	ausgeruht	wach
	munter	frisch
	schlapp	schläfrig
	müde	ermattet
RU	gelassen	ruhig
	entspannt	ausgeglichen
	ruhelos	angespannt
	unruhig	nervös

Physische Beschwerden wurden mit einem Fragebogen von Heuer, Hollendieck, Kröger und Römer (1989) erfasst. Dieser Fragebogen besteht aus den drei Skalen Sehbeschwerden (7 Items), Kopfschmerz-Symptome (3 Items) und Muskel- und Haltungsbeschwerden (4 Items). Die Items werden auf einer siebenstufigen Antwortskala mit den Endpunkten 1 („überhaupt nicht“) bis 7 („sehr stark“) beurteilt. Die Items sind in Tabelle 2.2 dargestellt.

Tabelle 2.2

Zuordnung der Items zu den Skalen	
Skala	
Augenbeschwerden	<ol style="list-style-type: none"> 1. Das Sehen macht mir Schwierigkeiten. 2. Meine Augenlider sind schwer. 3. Meine Augen schmerzen. 4. Meine Augen tränen. 5. Meine Augen brennen. 6. Ich habe ein komisches Gefühl um die Augen herum. 7. Meine Augen jucken.
Kopfschmerzsymptome	<ol style="list-style-type: none"> 1. Ich fühle mich benommen. 2. Ich fühle mich schwindelig. 3. Ich habe Kopfschmerzen.
Muskel- und Haltungsbeschwerden	<ol style="list-style-type: none"> 1. Ich habe Schmerzen in den Armen. 2. Ich habe Schmerzen im Nacken. 3. Ich habe Schmerzen im Rücken. 4. Ich habe Schmerzen in den Schultern.

2.1.3 Beschreibung der Stimuli

In allen berichteten Experimenten wurde das Paradigma einer Lese-Korrektur-Aufgabe eingesetzt. Insgesamt wurden 15 Kurzgeschichten verschiedener Autoren verwendet. Da die Geschichten unterschiedliche Schwierigkeitsgrade aufwiesen, wurde die Abfolge der Geschichten für jede Versuchsperson individuell randomisiert. Jede Geschichte wurde auf 875 Wörter gekürzt und füllte eine Bildschirmseite. Die Zeilenanzahl variierte aufgrund unterschiedlicher Wortlängen zwischen 35 und 40. Der Text wurde in der Schriftart Helvetica mit einer Schriftgröße von 14 Punkt präsentiert. In jeder Geschichte befanden sich durchschnittlich 30 Fehler, wobei zwischen fünf Fehlerarten unterschieden werden konnte (siehe Tabelle 2.3). Die Grammatikfehler wurden eingefügt, um die Lesestrategie der Probanden zu überprüfen. Sie konnten nur entdeckt werden, wenn beim Lesen die Intention bestand, die Bedeutung des Textes zu erfassen. Eine Analyse dieser Fehlerart ließ Rückschlüsse darüber zu, ob der Text nur flüchtig überflogen wurde.

Tabelle 2.3

Beispiele für die verschiedenen Fehlerarten	
Fehlerart	Beispiel
Buchstabe zuviel	„nicht“ → „nichtt“
Buchstabe zuwenig	„geöffnet“ → „geöffet“
Verdrehte Buchstabenreihenfolge	„Angst“ → „Ansgt“
Falscher Buchstabe	„Käfig“ → „Köfig“
Grammatikfehler	Die Sonne scheinen .

Es gab zwei Arten der Textdarstellung. Der Text wurde entweder in positiver Polarität (schwarze Schrift auf weißem Hintergrund) oder in negativer Polarität (weiße Schrift auf schwarzem Hintergrund) präsentiert, wobei der Raum abgedunkelt war (5 Lux) oder eine künstliche Beleuchtung von 550 Lux vorlag. Die Chromatizitäts-Koordinaten sowie Luminanzwerte der verwendeten Farben sind in Tabelle 2.4 aufgeführt.

Tabelle 2.4

CIE Chromatizitäts-Koordinaten (x,y) und Luminanzwerte (L) der verwendeten Farben. Die Berechnung des Luminanzkontrastes erfolgte nach der Michelson Definition $L_{\max}-L_{\min}/L_{\max}+L_{\min}$.

Farbe	x	y	Luminanz (cd/m ²)	Luminanz- kontrast
Schwarz	.294	.296	0.53	0.98
Weiß	.316	.335	192	

2.1.4 Versuchsablauf

Das Experiment wurde im Einzelversuch durchgeführt und dauerte insgesamt circa 75 Minuten. Die schematische Abbildung des Ablaufs ist in Abbildung 2.2 dargestellt.

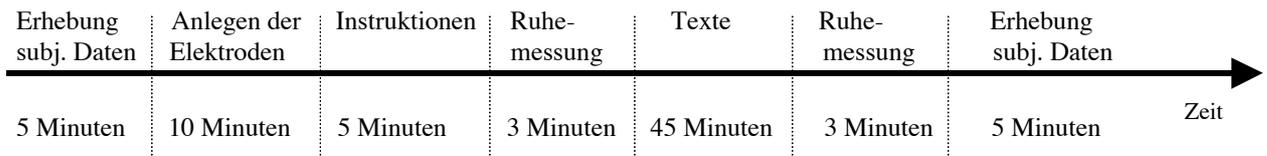


Abbildung 2.2: Schematische Darstellung des Versuchsablaufs

Vor dem Versuch wurden zunächst die subjektiven Daten erhoben. Der Proband bekam die Kurzform A bzw. B des MDBF vorgelegt, wobei die Reihenfolge ausbalanciert wurde. Das heißt, die Hälfte der Probanden bekam in der Vorerhebung Version A und in der Nachmessung Version B ausgehändigt. Für die andere Hälfte der Teilnehmer war die Reihenfolge umgekehrt. Darüber hinaus mussten ein Fragebogen zu physischen Beschwerden sowie allgemeine Fragen zu sportlichen Aktivitäten, Erkrankungen etc. beantwortet werden. Ein erwünschter Nebeneffekt dieser Datenerhebung war es, dass alle Teilnehmer vor dem eigentlichen Versuch eine kurze Pause hatten und die physiologischen Indikatoren ein Ruheniveau einnehmen konnten. Anschließend bestand die Möglichkeit, Kinnstütze und Stuhl individuell zu adjustieren.

Danach erfolgten die Vorbereitungen für die physiologischen Messungen. Zunächst wurde der Puls-Plethysmograph am linken Ohrläppchen befestigt. Eventuell vorhandener Ohr-

schmuck musste entfernt werden. Im Anschluss daran wurde der Atemgurt befestigt und zwar so, dass der Gurt bei maximaler Inspiration bereits stark gedehnt war. Dadurch war eine Messung der Atmung mit maximaler Sensitivität und Linearität möglich. Zum Schluss wurden die Elektroden zur Messung der Hautleitfähigkeit befestigt. Um die Haut optimal vorzubereiten und die Bedingungen für alle Probanden zu standardisieren, wurden die Teilnehmer vorher gebeten, sich mit warmem Wasser und ohne Seife die Hände zu waschen. Die bipolaren Elektroden wurden am mittleren Segment von Zeige- und Mittelfinger der linken Hand befestigt, wobei die Polarität keine Rolle spielte. Bei den Linkshändern wurde vorher sichergestellt, dass sie die Maus auch mit der rechten Hand bedienen können. Anschließend wurden die Probanden gebeten, für wenige Minuten still zu sitzen, nicht zu sprechen und die linke Hand nicht zu bewegen, um die absolute Leitfähigkeit der Haut zu bestimmen. Dieser Wert diente als Nullwert und Abweichungen wurden in Relation dazu gesetzt.

Im Anschluss daran folgten die Instruktionen, die standardisiert für alle Versuchsteilnehmer am Bildschirm dargeboten wurden. Die Probanden wurden zunächst darauf hingewiesen, während des gesamten Versuchs ruhig zu sitzen, nicht zu sprechen und nach Möglichkeit Husten und Räuspern zu unterdrücken, um die physiologischen Ableitungen nicht zu stören und keine Artefakte zu erzeugen. Weiterhin wurden sie instruiert, die Texte so schnell und so genau wie möglich zu lesen und dabei sowohl auf Tipp- als auch Grammatikfehler zu achten. Darüber hinaus wurden die Probanden gebeten, zeilenweise zu lesen und nicht im Text zu springen. Außerdem wurden die Teilnehmer darüber in Kenntnis gesetzt, dass Fehler in der Zeichensetzung, Groß- und Kleinschreibung sowie Silbentrennung nicht relevant sind. Sie erhielten keine Information über die Anzahl der Texte. Anschließend folgte ein Übungsbeispiel, um mit der Aufgabe und der Bedienung des Programms vertraut zu werden. Falls keine Unklarheiten mehr bestanden, konnte mit dem eigentlichen Versuch begonnen werden.

Zunächst wurde eine Ruhemessung von drei Minuten durchgeführt. In dieser Zeit wurde ein schwarzer Bildschirm präsentiert. Die Versuchsperson wurde gebeten, auf den Bildschirm zu blicken und Mausbewegungen zu simulieren, das heißt, mit dem Mauszeiger über den Bildschirm zu fahren. Veränderungen in den physiologischen Parametern während der Aufgabenbearbeitung konnten somit nicht auf motorische Aktivität zurückgeführt

werden. Es wurde vorher nochmals darauf hingewiesen, ruhig zu sitzen, nicht zu sprechen und normal zu atmen.

Im Anschluss daran erfolgte die Präsentation der Texte. Entdeckte Fehler konnten mit dem Mauszeiger markiert werden und wurden dann unterstrichen dargestellt. Falls die Versuchsperson sicher war, einen Fehler gefunden zu haben, konnte sie dieses durch Anklicken des Buttons „Übernehmen“, der sich rechts neben dem Text befand, bestätigen. Das markierte Wort war nun gespeichert. Für den Fall, dass ein Wort irrtümlich markiert wurde, bestand die Möglichkeit, dies durch Anklicken des Buttons „Nicht übernehmen“ rückgängig zu machen. Solange ein Wort markiert war, ohne dieses durch Betätigung einer der Buttons zu bestätigen oder rückgängig zu machen, konnte kein anderes Wort markiert werden. Damit wurde verhindert, dass entdeckte Fehler irrtümlich nicht gespeichert wurden. Jede Geschichte wurde für drei Minuten dargeboten. Zur besseren zeitlichen Orientierung ertönte nach der Hälfte der Zeit ein akustisches Signal, das zehn Sekunden vor Ablauf der Darbietungszeit wiederholt wurde. Die Textlänge war bewusst so gewählt, dass es kaum möglich war, die komplette Geschichte in drei Minuten zu bearbeiten. Nach Ende der Darbietung konnten keine Fehler mehr markiert werden. Neben dem Text erschien ein Zeilenzähler, mit dem der Proband die Zeile kenntlich machen sollte, die er zuletzt komplett gelesen hatte. Danach wurde sofort ein neuer Text dargeboten.

Nach Bearbeitung der Aufgabe wurde erneut eine Ruhemessung durchgeführt, die der ersten Baseline entsprach. Zu Beginn dieser Messungen und bei Präsentation einer neuen Textseite wurde jedes Mal automatisch ein elektrisches Auslösesignal an das Powerlab geschickt, das dann jeweils für eine vorher festgelegte Dauer von drei Minuten die physiologischen Daten aufnahm. In den Pausen zwischen den Texten erfolgte keine Messung. Somit ergaben sich insgesamt 17 isolierte Messabschnitte von jeweils drei Minuten.

Während der gesamten Versuchsdauer saß der Versuchsleiter im gleichen Raum und protokollierte sämtliche Ereignisse, die einen Einfluss auf die physiologischen Messdaten haben und Artefakte erzeugen konnten. Dazu zählten beispielsweise Husten, Sprechen, starke Bewegungen etc. Nach der zweiten Ruhemessung wurden erneut die subjektiven Daten erhoben. Der Proband bekam die Kurzform A bzw. B des MDBF, den Fragebogen zu physischen Beschwerden sowie einige allgemeine Fragen bezüglich Textdarstellung, Textschwierigkeit etc. vorgelegt.

2.1.5 Design

Tabelle 2.5

Versuchsplan		
Polarität	Positiver Kontrast	Negativer Kontrast
Beleuchtung		
Dunkel	n=20	n=20
Hell	n=20	n=20

In Tabelle 2.5 ist der Versuchsplan abgebildet. Es handelte sich um ein 2×2 Design mit den Faktoren „Polarität“ und „Beleuchtung“. Der Faktor „Polarität“ lag in den Ausprägungen „Positiver Kontrast“ (=schwarze Schrift auf weißem Hintergrund) und „Negativer Kontrast“ (=weiße Schrift auf schwarzem Hintergrund) vor. Der Faktor „Beleuchtung“ hatte die Abstufungen „hell“ (künstliche Raumbeleuchtung von 550 Lux in Höhe des Arbeitstisches) und „dunkel“ (5 Lux). Die Versuchspersonen wurden den Bedingungen zufällig zugeordnet.

Die abhängigen Variablen gliederten sich in die drei Bereiche Leistung, objektive Beanspruchung (physiologische Messungen) und subjektive Beanspruchung (Befindlichkeit und physische Beschwerden). Die einzelnen Variablen sind in Tabelle 2.6 aufgeführt.

Tabelle 2.6

 Abhängige Variablen getrennt für die Bereiche Performanz, objektive und subjektive Beanspruchung

 Performanz

Absolute Anzahl entdeckter Fehler pro Text

Anzahl gelesener Wörter pro Text

Genauigkeit (Anzahl entdeckter Fehler relativ zur Anzahl tatsächlich vorhandener Fehler im gelesenen Abschnitt)

Anzahl falscher Alarme und falsch markierter Wörter

Objektive Beanspruchung

Herzrate (bpm)

Herzratenvariabilität:

Gesamtleistung (ms^2)Leistung im mittleren Frequenzband (0,07-0,14 Hz) in absoluten (ms^2) und normalisierten Einheiten (NU)Leistung im oberen Frequenzband (0,15-0,4 Hz) in absoluten (ms^2) und normalisierten Einheiten (NU)

Prozentsatz der Intervalle mit mindestens 50ms Differenz zum vorausgehenden Intervall (PNN50)

Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD)

Tonische Hautleitfähigkeit (Microsiemens)

Subjektive Beanspruchung

Physische Beschwerden

Befindlichkeit

2.1.6 Auswertung

Im Folgenden wird das Verfahren zur Auswertung der Verhaltensdaten, der physiologischen Indikatoren sowie der subjektiv berichteten Beanspruchung dargestellt.

2.1.6.1 Verhaltensdaten

Die Anzahl gelesener Wörter, falscher Alarme, nicht übernommener Wörter sowie die absolute Anzahl entdeckter Fehler getrennt nach Fehlerarten wurden automatisch für jede Versuchsperson in einer Ergebnisdatei gespeichert. Anschließend erfolgte eine manuelle Berechnung der Fehlerentdeckungs-Genauigkeit. Dazu wurden die entdeckten Fehler in Relation gesetzt zur tatsächlichen Fehleranzahl im gelesenen Textabschnitt.

2.1.6.2 Physiologische Daten

Die Analyse der physiologischen Daten wurde für alle 17 Messabschnitte (2 Ruhemessungen und 15 Texte) getrennt durchgeführt. Die Auswertung der kardiovaskulären Parameter erfolgte mithilfe der Software „Chart v4.0“ (ADInstruments Pty Ltd., 2000) sowie einer Extension zur Berechnung der Herzratenvariabilität. Als Rohsignal für die Berechnung der HRV diente das Spannungssignal, das mittels des Puls-Plethysmographen aufgenommen worden war. Für die Analyse wurden die gesamten drei Minuten Aufnahmezeit eines jeden Abschnitts verwendet. Vor der Berechnung musste zuerst mithilfe der Software ein Algorithmus konfiguriert werden, um den Zeitpunkt jedes einzelnen Herzschlags festzustellen. In diesem Zusammenhang konnte ein Schwellenwert manuell eingestellt werden. Jedes Mal, wenn das Rohsignal diese Schwelle überschritt, wurde ein Herzschlag registriert. Im Anschluss daran wurde der Bereich festgelegt, in dem die gemessenen Intervalle zwischen zwei Herzschlägen als normal galten. Außerhalb dieses Bereichs liegende Intervalle wurden als Artefakte bewertet. Diese Artefakte waren beispielsweise bedingt durch Bewegungen oder auch eine falsche Konfiguration des Algorithmus.

Anschließend wurden von der Software automatisch die verschiedenen Parameter der Herzratenvariabilität berechnet. Die spektralanalytische Auswertung wurde mit der Kurzzeit-FOURIER Transformation mit einem Überlappungsgrad von 50 % und einer Länge von 512 Punkten durchgeführt, wobei als spektrales Fenster das so genannte Hann-Fenster Anwendung fand. Dieser Art der Auswertung lag die Annahme zugrunde, dass die Herzfrequenz-Zeitreihe aus Teilschwingungen zusammengesetzt ist und die Varianz der Zeitreihe vollständig durch diese Teilschwingungen erklärt werden kann. Das Ziel der Spektralanalyse war die Ermittlung von Amplitude, Frequenz und Phase der beteiligten Schwingungen sowie eine Umrechnung in ein Leistungsspektrum.

Nach Berechnung der Parameter war es notwendig, eine visuelle Artefaktkontrolle durchzuführen. Nachdem die Artefakte gelöscht bzw. fehlende Herzschläge durch Interpolation zwischen dem vorhergehenden und nachfolgenden als normal geltenden Intervall ergänzt waren, wurden die Parameter der HRV erneut berechnet.

Die Auswertung der elektrodermalen Aktivität erfolgte ebenfalls mithilfe der Software „Chart v4.0“ (ADInstruments Pty Ltd., 2000). Das Hauptinteresse galt der tonischen Hautleitfähigkeit als Beanspruchungsindikator. Bei der Analyse eines längeren Zeitabschnitts □ in diesem Fall drei Minuten □ trat allerdings folgendes Problem auf. Die tonische Leitfähigkeit war immer wieder von Hautleitwertsreaktionen, die nach einem Reizergebnis oder auch als Spontanfluktuation auftraten, überlagert. Es konnte darüber hinaus davon ausgegangen werden, dass in den Bedingungen, in denen mehr Fehler gefunden wurden, auch mehr Hautleitwertsreaktionen vorlagen. Hätte man einfach über den gesamten Messabschnitt von drei Minuten das Signal gemittelt, wären phasische und tonische Aktivität konfundiert worden, wodurch es zu einer Verfälschung der Ergebnisse gekommen wäre. Für jede Versuchsperson wurde deshalb automatisch protokolliert, zu welchem Zeitpunkt Fehler im Text markiert wurden. Für die Berechnung der tonischen Hautleitfähigkeit wurde letztendlich ein variables, relativ kurzes Intervall von etwa 10 Sekunden zu Beginn und am Ende des Messabschnitts gewählt, in dem keine phasischen Hautleitwertsreaktionen vorlagen. Dieses Intervall wurde für jede Versuchsperson und jeden Messabschnitt individuell bestimmt. Es hat sich allerdings herausgestellt, dass bei fast allen Probanden aufgrund tiefer Einatmung („Seufzen“) zu Beginn einer neuen Textseite, das vorde-re Intervall für die weitere Analyse nicht verwendet werden konnte. In Abbildung 2.3 ist beispielhaft wiedergegeben, wie das tiefe Einatmen mit einer kurzen Latenz zu einer star-

ken Hautleitwertsreaktion mit relativ langer Erholungszeit führte. Aus diesem Grund wurde bei der weiteren Auswertung nur das hintere Intervall gegen Ende des Messabschnitts berücksichtigt. Dieses Intervall lag bei dem meisten Probanden im Bereich zwischen 20-40 Sekunden vor Ende des dreiminütigen Messabschnitts.

Die Auswertung der Atemfrequenz erfolgte ebenfalls mithilfe der Software „Chart v4.0“ (ADInstruments Pty Ltd., 2000). Dazu wurde im gesamten Messintervall die Anzahl der Atemzüge ausgezählt und in die mittlere Atemfrequenz umgerechnet.

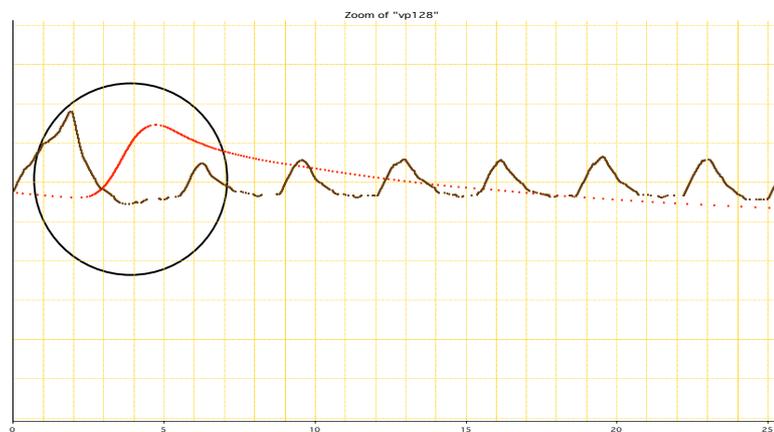


Abbildung 2.3: Parallele Erfassung der Atemfrequenz (dunkle Kurve) sowie der elektrodermalen Aktivität (helle Kurve). Man sieht, dass infolge eines tiefen Atemzugs die Hautleitfähigkeit mit einer Latenz von einigen Sekunden stark ansteigt.

2.1.6.3 Subjektive Beanspruchung

Für jede Teilskala des Mehrdimensionalen Befindlichkeitsfragebogens wurden die Werte der entsprechenden Items zu Werten der dazugehörigen Skala summiert. Die Werte der negativ formulierten Items mussten für die Skalenwertberechnung vorher umgepolt werden. Um den neuen Itemwert zu berechnen, wurde der alte Wert von der Zahl sechs subtrahiert. Für jede Skala ergab sich somit ein maximaler Wert von 20.

Da der MDBF die momentane Befindlichkeit der Probanden erfassen sollte und zudem nur der Vergleich der experimentellen Gruppen inhaltlich interessant war, wurde ein Vergleich mit Normwerten als nicht sinnvoll betrachtet. Ein hoher Wert der Skala „Gute/Schlechte

Stimmung (GS)“ wurde als positive Stimmungslage interpretiert. Hohe Werte der Skala „Wachheit/Müdigkeit (WM)“ resultierten bei wachen und ausgeruhten Personen. Bezüglich der Skala „Ruhe/Unruhe (RU)“ deutete ein hoher Wert darauf hin, dass sich der Proband innerlich ruhig und gelassen fühlte. Anschließend wurde durch Differenzbildung Messzeitpunkt A (vor der Aufgabenbearbeitung) mit Messzeitpunkt B (nach der Aufgabenbearbeitung) verglichen.

Für die Auswertung der physischen Beschwerden wurden für jede der drei Teilskalen die Werte der entsprechenden Items zu einem Summenwert der entsprechenden Skala summiert. Anschließend wurde ebenfalls durch Differenzbildung der Messzeitpunkt vor der Aufgabenbearbeitung mit dem Messzeitpunkt nach der Aufgabenbearbeitung verglichen.

2.1.6.4 Teststärkeüberlegungen

Unter praktischen Gesichtspunkten sollte der Unterschied zwischen den experimentellen Bedingungen sowie deren Interaktion mindestens „groß“ ($d = 0.8$) im Sinne der Konvention von Cohen (1977) sein. Eine A-priori Teststärkeanalyse mithilfe des Programms GPower (Erdfelder, Faul & Buchner, 1996) ergab, dass damit bei $\alpha = \beta = .05$ insgesamt mindestens 84 Untersuchungsteilnehmer benötigt wurden. Es konnten allerdings nur 80 Probanden erhoben werden, so dass die Teststärke mit $(1 - \beta) = .942$ geringfügig unter der angestrebten Teststärke von .95 lag.

2.1.6.5 Alpha-Adjustierung

An den statistischen Schlussfolgerungen würde sich durch eine Bonferoni-Holm-Korrektur (Bortz, 1999) innerhalb der drei Gruppen von abhängigen Variablen (Performanz, objektive und subjektive Beanspruchung) nichts ändern. Aus diesem Grund wurde wegen einer größeren Übersichtlichkeit auf eine α -Adjustierung in den jeweiligen Bereichen verzichtet. Die einzige Ausnahme in diesem Zusammenhang stellt der Polaritätseffekt bei der absoluten Anzahl richtig erkannter Wörter in Experiment 3 dar, der bei einer α -Adjustierung eventuell knapp scheitern würde.

2.2 Ergebnisse

Die Ergebnisse werden getrennt für die drei Bereiche Verhaltensdaten, objektive Beanspruchung (physiologische Messungen) und subjektive Beanspruchung dargestellt. Das α -Fehler-Niveau wird auf .05 festgesetzt. Als Maß für die Größe eines Effekts wird das partielle R^2 berichtet, welches den Anteil aufgeklärter Varianz an der Varianz präsentiert, die nicht durch andere Variablen des Designs aufgeklärt wird.

2.2.1 Verhaltensdaten

Im Folgenden wird die Leistung der Probanden in den vier Experimentalbedingungen dargestellt. Dazu zählt die absolute Anzahl entdeckter Fehler, die Genauigkeit bei der Fehlerentdeckung sowie die Anzahl gelesener Wörter.

2.2.1.1 Absolute Anzahl entdeckter Fehler

Abbildung 2.4 zeigt die Mittelwerte sowie die Standardfehler der absoluten Anzahl entdeckter Fehler, die in der vorliegenden Untersuchung das wichtigste Performanzmaß darstellen, pro Textseite in den vier Gruppen. Es ist eindeutig zu erkennen, dass die mittlere Anzahl absolut entdeckter Fehler in der Negativ Hell-Gruppe ($M = 8,79$) und der Negativ Dunkel-Gruppe ($M = 8,55$) geringer war als in der Positiv Hell- ($M = 11,84$) und der Positiv Dunkel-Bedingung ($M = 12,64$). Eine $2 \times 2 \times 15$ MANOVA mit den Gruppenvariablen Beleuchtung (hell/dunkel) und Polarität (positiv/negativ) sowie der messwiederholten Variable Zeitverlauf (15 aufeinanderfolgende Texte) zeigte einen signifikanten Effekt der Polarität auf die absolute Anzahl entdeckter Fehler ($F(1,76) = 22,23, p < .001, R^2 = .23$). Der Faktor Beleuchtung hatte keinen Effekt ($F < 1$) und es zeigte sich auch keine Interaktion zwischen den beiden Variablen ($F < 1$).

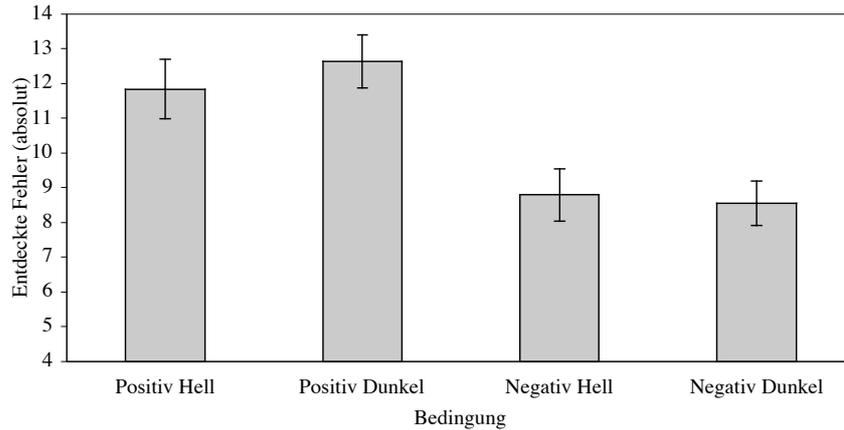


Abbildung 2.4: Durchschnittliche Anzahl richtig entdeckter Fehler pro Textseite. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Abbildung 2.5 zeigt den Verlauf der Anzahl richtig entdeckter Fehler über die 15 Texte. Es ist zu erkennen, dass die Anzahl der entdeckten Fehler pro Text relativ konstant war und die Leistung gegen Ende der Aufgabenbearbeitung nicht abnahm. Eine multivariate Varianzanalyse zeigte keinen Effekt des Zeitverlaufs ($F(14,63) = 1.53$). Auch konnte weder eine Interaktion der Polarität ($F(14,63) = 1,43$), noch der Beleuchtung ($F(14,63) = 1,22$) mit dem Faktor Zeitverlauf festgestellt werden. Es lag auch keine Interaktion zwischen dem Zeitverlauf und den anderen beiden Variablen vor ($F < 1$).

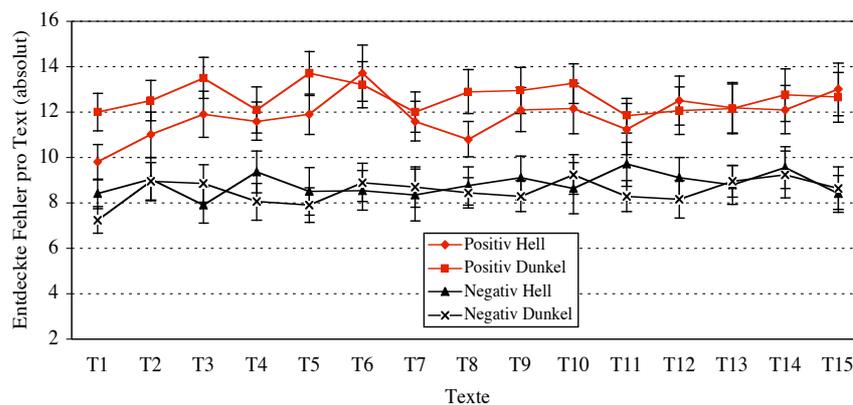


Abbildung 2.5: Die Abbildung zeigt den Verlauf der Anzahl richtig entdeckter Fehler über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.1.2 Anzahl gelesener Wörter

Die Tatsache, dass unter negativen Kontrastverhältnissen deutlich weniger Fehler entdeckt wurden als unter positivem Kontrast, kann zwei Gründe haben. Zum einen war es möglich, dass bei negativer Darstellung weniger Wörter gelesen wurden. Zum anderen konnte eine geringere Genauigkeit die Ursache für dieses Ergebnis sein.

Abbildung 2.6 zeigt die Mittelwerte sowie Standardfehler der Anzahl gelesener Wörter pro Textseite in den vier Gruppen. Es ist eindeutig zu erkennen, dass die mittlere Anzahl gelesener Wörter am geringsten war in der Negativ Dunkel-Gruppe ($M = 454$), gefolgt von der Negativ Hell-Gruppe ($M = 504$), der Positiv Dunkel- ($M = 586$) sowie der Positiv Hell-Bedingung ($M = 598$). Eine $2 \times 2 \times 15$ MANOVA zeigte einen signifikanten Effekt der Polarität ($F(1,76) = 12,2, p = .001, R^2 = .14$). Der Faktor Beleuchtung hatte keinen Effekt ($F < 1$) und es gab auch keine Interaktion zwischen den beiden Variablen ($F < 1$).

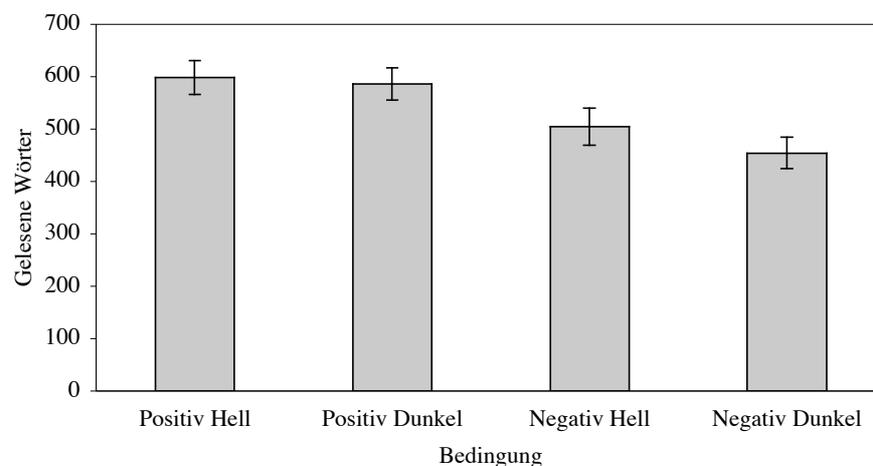


Abbildung 2.6: Illustration der durchschnittlichen Anzahl gelesener Wörter pro Textseite. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Abbildung 2.7 zeigt den Verlauf der Anzahl gelesener Wörter über die 15 Texte. Es ist zu erkennen, dass die Anzahl der gelesenen Wörter pro Text relativ konstant war und die Leistung gegen Ende der Aufgabenbearbeitung nicht abnahm. Eine multivariate Varianz-

analyse zeigte einen Effekt des Zeitverlaufs ($F(14,63) = 4,57, p < .001, R^2 = .5$). Dieser Effekt war allerdings durch den Anstieg der Leserate von Text 1 zu Text 3 bedingt, da mithilfe orthogonaler Kontraste gezeigt werden konnte, dass sich bezüglich der Anzahl gelesener Wörter nur der erste ($F(1,76) = 67,89, p < .001, R^2 = .47$) und der zweite Text ($F(14,63) = 12,64, p = .001, R^2 = .14$) von den anderen Texten unterschieden. Es konnte keine Interaktion der Polarität ($F(14,63) = 1,13$) und der Beleuchtung ($F < 1$) mit dem Faktor Zeitverlauf festgestellt werden. Darüber hinaus lag auch keine Interaktion zwischen dem Zeitverlauf und den anderen beiden Variablen vor ($F(14,63) = 2,03$).

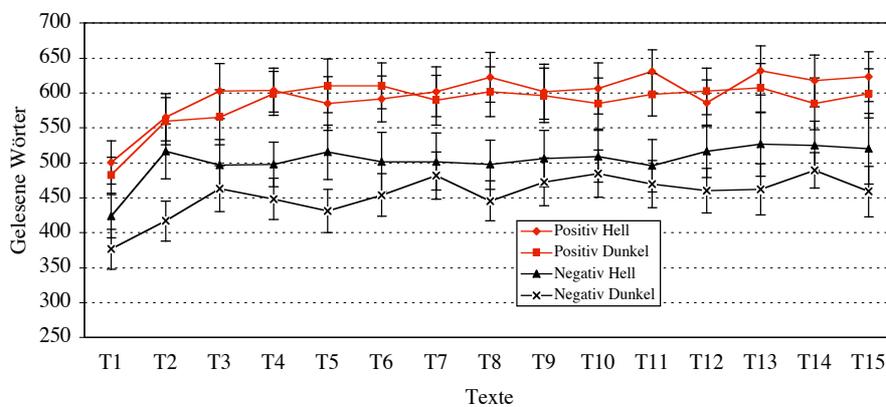


Abbildung 2.7: Die Abbildung zeigt den Verlauf der Anzahl gelesener Wörter über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.1.3 Relative Anzahl gelesener Wörter (Genauigkeit)

Abbildung 2.8 zeigt die Mittelwerte und Standardfehler der Fehlerentdeckungs-Genauigkeit pro Textseite in den vier Gruppen. Man sieht, dass die Genauigkeit am geringsten war in der Negativ Hell-Gruppe ($M = 53\%$), gefolgt von der Negativ Dunkel-Gruppe ($M = 57\%$), der Positiv Hell- ($M = 58\%$) und Positiv Dunkel-Bedingung ($M = 64\%$). Das Ergebnis einer $2 \times 2 \times 15$ MANOVA zeigte allerdings keinen signifikanten Effekt der Polarität ($F(1,76) = 3,54, p = .064$). Der Faktor Beleuchtung hatte ebenfalls keinen Effekt ($F(1,76) = 1,92$) und es gab auch keine Interaktion zwischen den beiden Variablen ($F < 1$).

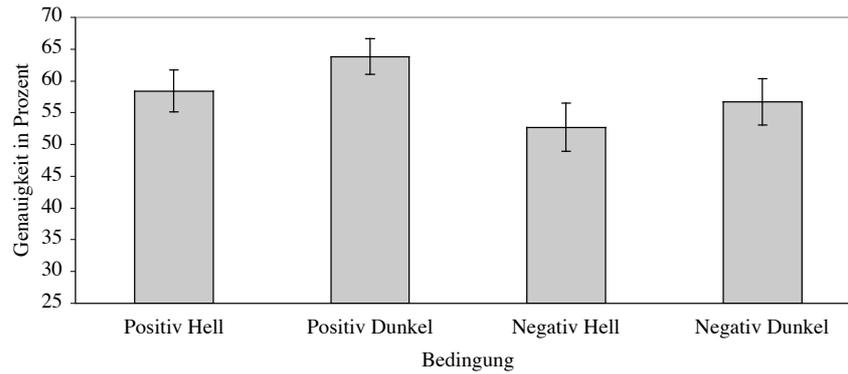


Abbildung 2.8: Durchschnittliche Genauigkeit bei der Fehlerentdeckung. Die Anzahl entdeckter Fehler wird hierbei in Relation gesetzt zur Anzahl tatsächlich vorhandener Fehler im gelesenen Abschnitt. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Abbildung 2.9 zeigt den Verlauf der Fehlerentdeckungs-Genauigkeit über die 15 Texte. Es ist zu erkennen, dass der Verlauf relativ konstant war und die Genauigkeit gegen Ende der Aufgabenbearbeitung nicht abnahm. Ein Effekt des Zeitverlaufs konnte nicht festgestellt werden ($F(14,63) = 1,87$). Es zeigte sich auch weder eine Interaktion der Polarität ($F(14,63) = 1,68$), noch der Beleuchtung ($F < 1$) mit Zeitverlauf. Darüber hinaus lag auch keine Interaktion zwischen dem Zeitverlauf und den anderen beiden Variablen vor ($F < 1$). Rein deskriptiv zeigte sich eine geringere Genauigkeit in der Negativ Hell-Gruppe im Vergleich zur Negativ Dunkel-Bedingung, die aber durch eine höhere Anzahl gelesener Wörter in der Negativ Hell-Gruppe wieder ausgeglichen wurde.

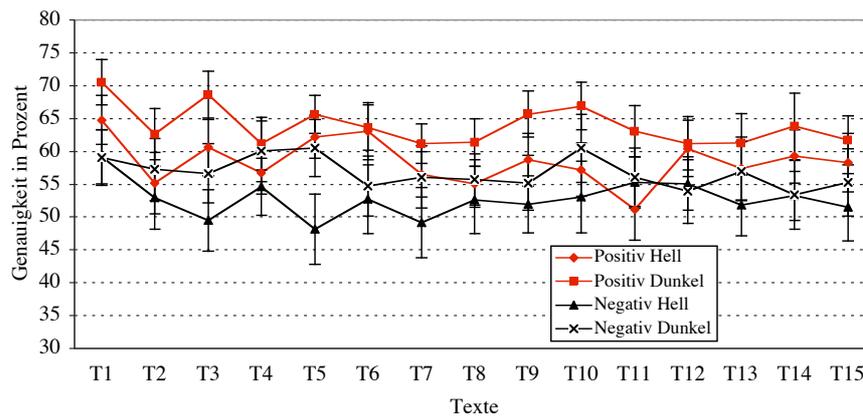


Abbildung 2.9: Verlauf der Fehlerentdeckungs-Genauigkeit über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.1.4 Fehlerarten

In Abbildung 2.10 ist die Genauigkeit bei der Fehlerentdeckung getrennt nach den fünf Fehlerarten dargestellt. Man kann erkennen, dass in allen vier Bedingungen Grammatikfehler weniger genau erkannt wurden als die vier anderen Fehlerarten. Mithilfe orthogonaler Kontraste, die die Genauigkeit für Grammatikfehler mit dem Mittelwert der vier anderen Fehlerarten verglichen, konnte dies sowohl für negativen Kontrast ($F(1,38) = 5,08, p = .03, R^2 = .12$) als auch für positiven Kontrast ($F(1,38) = 34,98, p < .001, R^2 = .47$) gezeigt werden.

Klammert man die Grammatikfehler von der weiteren Betrachtung aus, zeigte sich, dass auch die anderen vier Fehlerarten mit unterschiedlicher Genauigkeit erkannt wurden. Die Fehlerart „falscher Buchstabe“ wurde mit höherer Genauigkeit erkannt als die Fehlerarten „Buchstabe zuviel“, „Buchstabe zuwenig“ und „Buchstabe verdreht“. Dies konnte sowohl für negativen Kontrast ($F(1,38) = 123,94, p < .001, R^2 = .76$) als auch positiven Kontrast ($F(1,38) = 88,78, p < .001, R^2 = .69$) gezeigt werden.

Die Fehlerart „Buchstabe zuwenig“ wurde mit geringster Genauigkeit entdeckt. Mittels orthogonaler Kontraste konnte dies sowohl für negative Polarität ($F(1,38) = 77,33, p < .001, R^2 = .66$) als auch positive Polarität ($F(1,38) = 180,63, p < .001, R^2 = .82$) nachgewiesen werden.

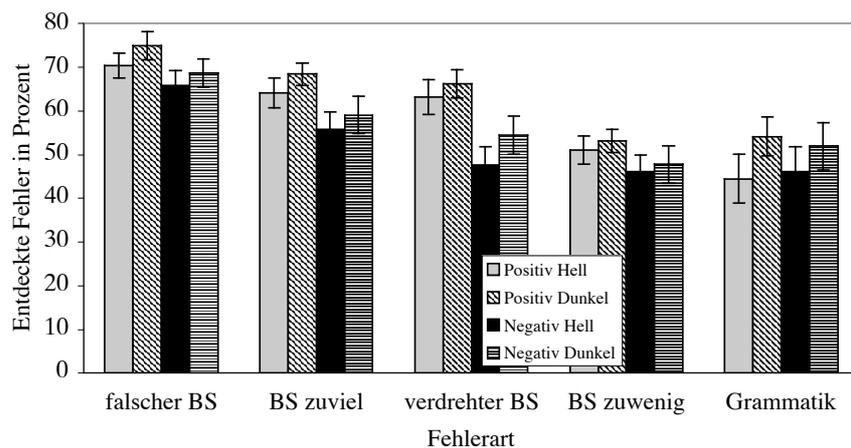


Abbildung 2.10: Die Abbildung zeigt die Fehlerentdeckungs-Genauigkeit getrennt nach Fehlerarten. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.1.5 Genauigkeit bei der Entdeckung von Grammatikfehlern

Abbildung 2.11 zeigt den Verlauf der Entdeckungs-Genauigkeit für Grammatikfehler über die 15 Texte. Es zeigte sich, dass für alle Gruppen die Kurve leicht schwankend verlief, allerdings die Genauigkeit gegen Ende der Aufgabenbearbeitung nicht abnahm. Ein Effekt des Zeitverlaufs konnte nicht festgestellt werden ($F < 1$). Außerdem zeigte sich keine Interaktion zwischen der Polarität ($F(14,63) = 1,55$) und dem Zeitverlauf und der Beleuchtung ($F(14,63) = 1,49$) und dem Zeitverlauf. Darüber hinaus lag auch keine Interaktion zwischen dem Zeitverlauf und den anderen beiden Faktoren vor ($F < 1$).

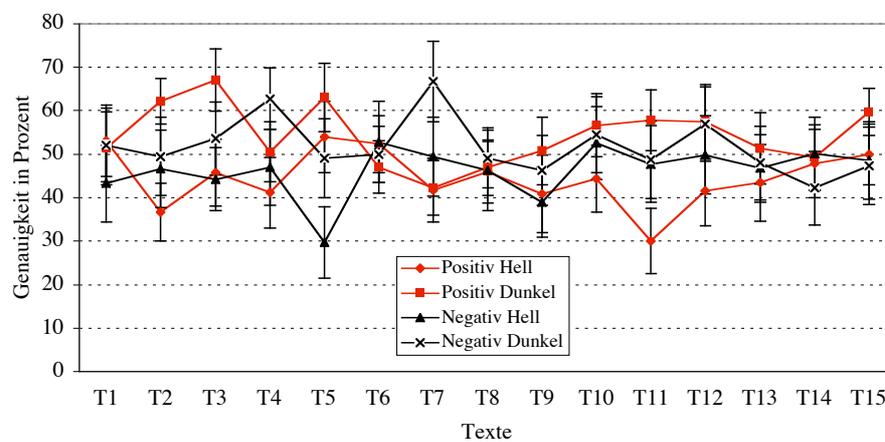


Abbildung 2.11: Dargestellt ist der Verlauf der Entdeckungs-Genauigkeit für Grammatikfehler. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.1.6 Falsche Alarmer und „Nicht übernehmen“

Tabelle 2.7

Durchschnittliche Zahl falscher Alarmer und nicht übernommener Wörter pro Seite. In Klammern ist die Standardabweichung angegeben.

	Falsche Alarmer	„Nicht übernehmen“
Positiv Hell	0,49 (0,08)	0,37 (0,05)
Positiv Dunkel	0,46 (0,06)	0,38 (0,06)
Negativ Hell	0,48 (0,07)	0,38 (0,06)
Negativ Dunkel	0,68 (0,15)	0,25 (0,06)

In Tabelle 2.7 sind die Mittelwerte und Standardabweichungen der durchschnittlichen Zahl falscher Alarme und nicht übernommener Wörter pro Textseite für die verschiedenen Bedingungen dargestellt. Für die Anzahl der falschen Alarme zeigte eine Varianzanalyse weder einen Effekt der Polarität ($F(1,76) = 1,26$), noch der Beleuchtung ($F < 1$). Auch lag keine Interaktion zwischen den beiden Faktoren vor ($F(1,76) = 1,72$). Gleiches galt für die Anzahl nicht übernommener Wörter. Auch hier konnte kein Effekt der Polarität ($F(1,76) = 1,18$) und der Beleuchtung ($F(1,76) = 1,39$) festgestellt werden. Es zeigte sich auch keine Interaktion zwischen den beiden Faktoren ($F(1,76) = 1,64$).

2.2.1.7 Subjektive Angaben

Zum Abschluss des Versuchs wurden die Probanden gebeten, die Textschwierigkeit, die Textdarstellung, den empfundenen Zeitdruck sowie die eigene Sicherheit in der Rechtschreibung einzuschätzen. Die Mittelwerte und Standardabweichungen der Ergebnisse sind in Tabelle 2.8 dargestellt. Bei der Einschätzung des Textniveaus zeigte sich kein Effekt der Polarität ($F < 1$), der Helligkeit ($F < 1$) und auch keine Interaktion zwischen den beiden Variablen ($F < 1$). Auch für die Beurteilung der Textdarstellung ergab sich kein Effekt der Polarität ($F < 1$) und der Helligkeit ($F < 1$). Auch zeigte sich keine Interaktion zwischen den Variablen. Für den empfundenen Zeitdruck konnte ebenfalls kein Effekt der Polarität ($F < 1$) und der Helligkeit ($F < 1$), sowie auch keine Interaktion festgestellt werden ($F < 1$). Bezüglich der eigenen Rechtschreibsicherheit zeigte sich weder ein Effekt der Polarität ($F < 1$), noch der Helligkeit ($F < 1$). Ebenso lag auch keine Interaktion zwischen den beiden Variablen vor ($F < 1$).

Tabelle 2.8

Subjektive Einschätzung des Textniveaus, der Textdarstellung, des empfundenen Zeitdrucks sowie der eigenen „Rechtschreibsicherheit“. In Klammern ist die Standardabweichung angegeben.

		Positiv Hell	Positiv Dunkel	Negativ Hell	Negativ Dunkel
Textschwierigkeit	1=sehr einfach bis 7=sehr schwierig	2,9 (0,14)	2,6 (0,22)	2,7 (0,31)	2,7 (0,29)
Textsdarstellung	1=sehr angenehm bis 7=sehr unangenehm	5,2 (0,32)	4,5 (0,33)	4,9 (0,44)	5,2 (0,37)
Zeitdruck	1=gar nicht bis 7=sehr stark	4,4 (0,33)	4,2 (0,33)	3,9 (0,35)	3,9 (0,38)
Rechtschreibung	1=sehr sicher bis 7=sehr unsicher	3,4 (0,37)	3,1 (0,37)	2,8 (0,39)	3,2 (0,36)

2.2.2 Objektive Beanspruchung

Die Darstellung der Ergebnisse für die objektive Beanspruchung erfolgt getrennt für die Bereiche Herzrate, Herzratenvariabilität, elektrodermale Aktivität sowie Atmung.

2.2.2.1 Herzrate

Um zu überprüfen, ob es während der Aufgabenbearbeitung zu einer psychophysiologischen Beanspruchung kam, wurde der Mittelwert aus den 15 Versuchsabschnitten mit den Werten der Ruhemessungen verglichen. In Abbildung 2.12 sind die Mittelwerte und Standardfehler dargestellt und man sieht deutlich, dass die mittlere Herzrate in allen vier Gruppen bei der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung anstieg und dann in der zweiten Ruhemessung wieder abfiel. Es wurde eine 2×2×3 MANOVA mit den Gruppenvariablen Beleuchtung (hell/dunkel) und Polarität (positiv/negativ) sowie der

messwiederholten Variable Zeitpunkt (2 Ruhemessungen und Mittelwert der Aufgabenbearbeitung) berechnet. Mithilfe orthogonaler Kontraste konnte gezeigt werden, dass der Anstieg der Herzrate während der Aufgabenbearbeitung im Vergleich zum Mittelwert der beiden Ruhemessungen signifikant war ($F(1,76) = 86,91, p < .001, R^2 = .53$). Es konnte allerdings weder eine Interaktion der Polarität ($F(1,76) = 2,59$), noch der Helligkeit ($F(1,76) = 1,05$) mit dem Zeitpunkt, noch eine Interaktion zwischen den beiden Variablen und dem Zeitpunkt festgestellt werden ($F(1,76) = 2,04$). Vergleicht man mittels orthogonaler Kontraste den Wert der Herzrate in Ruhemessung 1 und 2, zeigte sich kein Unterschied ($F < 1$). Auch konnte keine Interaktion zwischen der Polarität und dem Zeitpunkt ($F < 1$) und der Helligkeit ($F < 1$) und dem Zeitpunkt, sowie auch keine Interaktion zwischen den beiden Faktoren und dem Zeitpunkt ($F(1,76) = 2,15$) nachgewiesen werden.

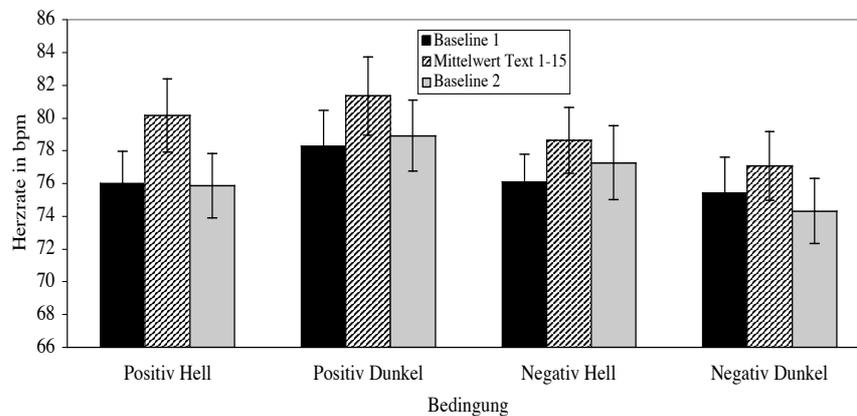


Abbildung 2.12: Mittlere Herzrate während Ruhe (Baseline 1 und 2) sowie Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Neben einem Vergleich zwischen Ruhemessung und Aufgabenbearbeitung ist es unter explorativen Gesichtspunkten interessant, den Verlauf der Herzrate über die 15 Texte hinweg zu betrachten. Schaut man sich die Herzrate (Differenzwerte zur ersten Ruhemessung) in Abbildung 2.13 im Verlauf an, kann man erkennen, dass die Herzrate in allen Gruppen mit fortschreitender Versuchsdauer immer weiter abfiel. Das Ergebnis einer $2 \times 2 \times 15$ MANOVA zeigte, dass diese Abnahme der Herzrate signifikant war ($F(14,63) = 2,57, p < .006, R^2 = .36$). Rein deskriptiv war dieses Absinken am stärksten in der Negativ-Dunkel Bedingung ausgeprägt. Es zeigte sich allerdings weder eine Interaktion zwischen der Polarität und dem Zeitverlauf ($F(14,63) = 1,28$), noch zwischen der Beleuchtung und dem Zeit-

verlauf ($F(14,63) = 1,39$). Ebenso lag keine Interaktion zwischen dem Zeitverlauf und den anderen beiden Faktoren vor ($F < 1$).

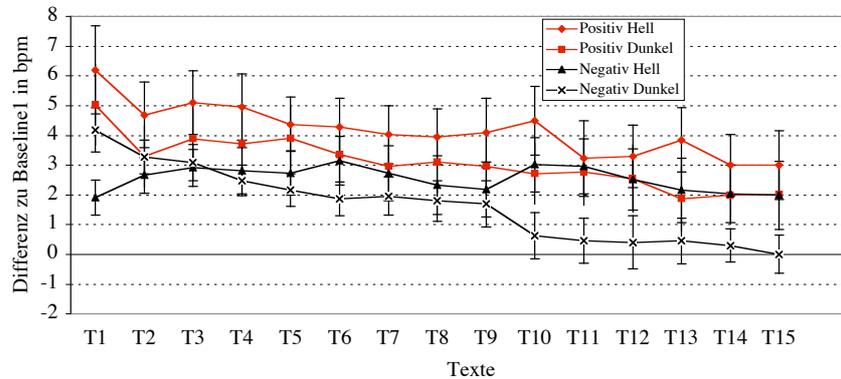


Abbildung 2.13: Verlauf der mittleren Herzrate über die 15 Texte. Dargestellt sind die Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.2.2 Herzratenvariabilität

Dargestellt werden die Ergebnisse der spektralanalytischen und zeitbezogenen Auswertung der Herzratenvariabilität. Zu den spektralanalytischen Parametern zählen die Gesamtleistung sowie die Leistung im mittleren und oberen Frequenzband. Als zeitbezogene Streuungsmaße werden der Prozentsatz der Intervalle mit mindestens 50ms Differenz zum vorausgehenden Intervall (PNN50) sowie die Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD) erläutert.

Gesamtleistung: Es zeigte sich, dass die Gesamtleistung des Leistungsspektrums bei der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung abfiel und dann in der zweiten Ruhemessung wieder anstieg (siehe Abbildung 2.14). Mithilfe orthogonaler Kontraste konnte gezeigt werden, dass dieser Abfall der Gesamtleistung im Vergleich zum Mittelwert der beiden Ruhemessungen signifikant war ($F(1,76) = 12,33$, $p = .001$, $R^2 = .14$). Es zeigte sich aber keine Interaktion der Polarität ($F < 1$) und der Beleuchtung ($F < 1$) mit dem Zeitpunkt. Ebenso lag keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt vor ($F < 1$). Vergleicht man mittels orthogonaler Kontraste die Gesamtleistung in Ruhemessung 1 und 2, zeigte sich ein signifikanter Unterschied ($F(1,76) = 5,28$, $p = .024$, $R^2 = .06$). Es konnte allerdings keine Interaktion der Polarität ($F < 1$) und der Hellig-

keit ($F < 1$) mit dem Zeitpunkt, sowie auch keine Interaktion zwischen den beiden Faktoren und dem Zeitpunkt ($F(1,76) = 2,15$) nachgewiesen werden.

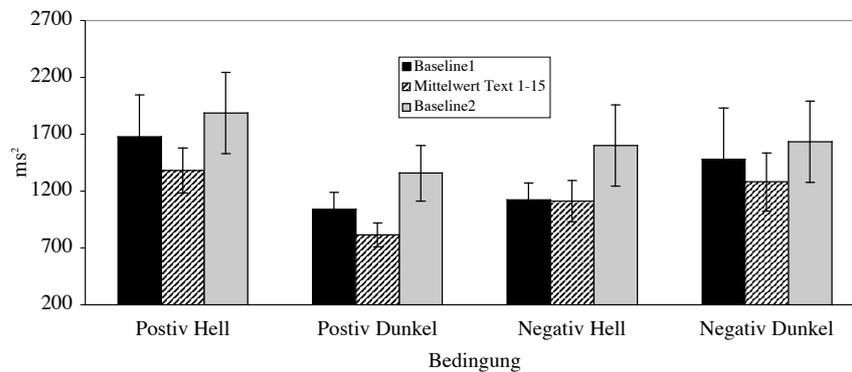


Abbildung 2.14: Mittlere Gesamtleistung aller Frequenzbänder während Ruhe (Baseline 1 und 2) sowie Aufgabebearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Mittleres Frequenzband (0,07-0,14 Hz): Die Leistung in den einzelnen Frequenzbändern des Power-Spektrums kann sowohl in absoluten (ms^2) als auch in normalisierten Einheiten angegeben werden. Zur Umrechnung in normalisierte Einheiten wird die absolute Leistung dividiert durch die Gesamtleistung minus die Leistung im unteren Frequenzband und das Ergebnis mit 100 multipliziert.

Es zeigte sich wie in Tabelle 2.9 erkennbar mithilfe orthogonaler Kontraste kein erwartungsgemäßer Abfall der Leistung in absoluten Einheiten während der Aufgabebearbeitung im Vergleich zum Mittelwert der beiden Ruhemessungen ($F(1,76) = 2,31$). Es konnte auch keine Interaktion zwischen der Polarität ($F(1,76) = 2,66$) und dem Zeitpunkt und der Beleuchtung ($F(1,76) = 1,19$) und dem Zeitpunkt festgestellt werden. Ebenso lag keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt vor ($F < 1$). Vergleicht man die Leistung in Ruhemessung 1 und 2, zeigte sich kein signifikanter Unterschied ($F(1,76) = 3,56$). Es konnte auch keine Interaktion zwischen der Polarität und dem Zeitpunkt ($F < 1$) sowie der Helligkeit und dem Zeitpunkt ($F < 1$) festgestellt werden. Auch konnte keine Interaktion zwischen den beiden Faktoren und dem Zeitpunkt ($F < 1$) nachgewiesen werden.

Bei der Darstellung der Mittelwerte und Standardfehler der Leistung in normalisierten Einheiten in Abbildung 2.15 erkennt man deutlich, dass die Leistung erwartungsgemäß in

allen vier Gruppen bei der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung anstieg. Da die Leistung in der Positiv Dunkel- und Negativ Hell-Bedingung in der zweiten Ruhemessung nicht wieder abfiel, wurde im Rahmen der statistischen Analyse nur die erste Ruhemessung mit dem Mittelwert der Aufgabenbearbeitung verglichen. Das Ergebnis einer $2 \times 2 \times 2$ MANOVA zeigte, dass dieser Anstieg signifikant war ($F(1,76) = 47,41, p < .001, R^2 = .38$). Wie in Abbildung 2.16 zu sehen ist, war der Anstieg der Leistung in den beiden Bedingungen mit negativem Kontrast deskriptiv größer, allerdings zeigte sich keine signifikante Interaktion zwischen der Polarität und dem Zeitpunkt ($F(1,76) = 2,81$). Ebenso konnte keine Interaktion mit der Beleuchtung festgestellt werden ($F < 1$). Auch lag keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt vor ($F < 1$).

Betrachtet man den Verlauf der Leistung in normalisierten Einheiten als Differenzwerte zur ersten Ruhemessung über die 15 Texte in Abbildung 2.17, zeigte das Ergebnis einer $2 \times 2 \times 15$ MANOVA keinen Effekt des Zeitverlaufs ($F(14,63) = 1,36$). Auch konnte kein Einfluss der Polarität ($F(14,63) = 1,21$) und der Beleuchtung ($F(14,63) = 1,46$) auf den Zeitverlauf festgestellt werden. Außerdem lag keine Interaktion zwischen dem Zeitverlauf und den anderen beiden Faktoren vor ($F < 1$).

Tabelle 2.9

Mittelwerte der Leistung im mittleren Frequenzband in absoluten (ms^2) und normalisierten Einheiten während Ruhe (Baseline 1 und 2) sowie Aufgabenbearbeitung. In Klammern sind die Standardfehler der Mittelwerte angegeben.

		Baseline 1	Aufgabe (Mittelwert Text 1-15)	Baseline 2
absolute Leistung (ms^2)	Positiv Hell	650,62 (192,28)	525,01 (83,78)	699,01 (166,21)
	Positiv Dunkel	355,35 (58,42)	293,88 (41,69)	405,42 (63,73)
	Negativ Hell	362,78 (70,08)	358,63 (61,29)	446,97 (86,51)
	Negativ Dunkel	315,31 (76,28)	467,51 (125,86)	510,51 (137,78)
normalisierte Leistung (NU)	Positiv Hell	48,46 (4,04)	56,79 (2,69)	52,56 (4,41)
	Positiv Dunkel	51,57 (4,29)	57,41 (3,48)	58,62 (3,87)
	Negativ Hell	44,75 (5,26)	55,19 (4,46)	56,03 (4,84)
	Negativ Dunkel	39,04 (5,21)	51,86 (4,71)	49,95 (5,17)

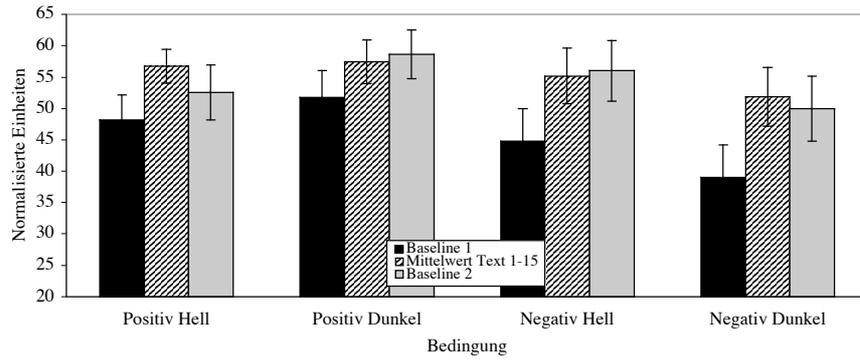


Abbildung 2.15: Durchschnittliche Leistung im mittleren Frequenzband in normalisierten Einheiten während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

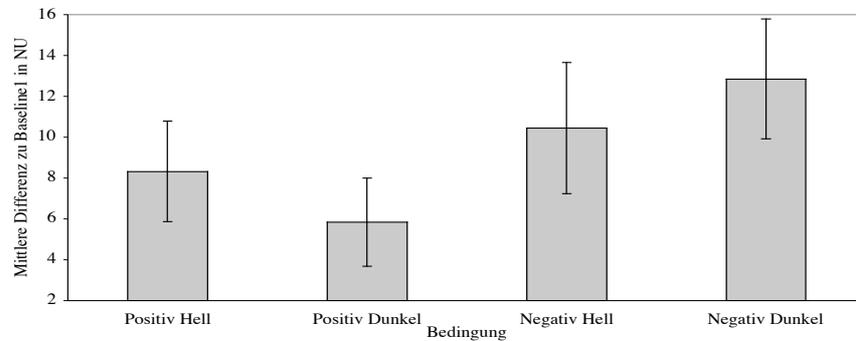


Abbildung 2.16: Anstieg der Leistung im mittleren Frequenzband im Vergleich zu Baseline 1 gemittelt über alle Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

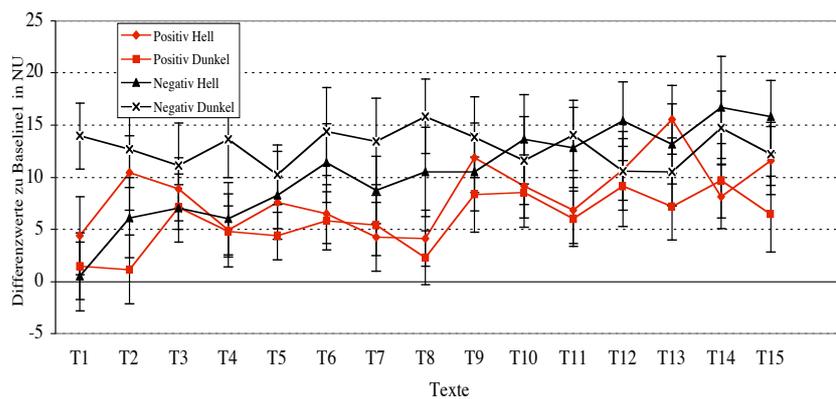


Abbildung 2.17: Dargestellt ist der Verlauf der Leistung im mittleren Frequenzband. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Oberes Frequenzband (0,15-0,4 Hz): Die Leistung im oberen Frequenzband des Leistungsspektrums wird sowohl in absoluten (ms^2) als auch in normalisierten Einheiten angegeben. Die Umrechnung in normalisierte Einheiten erfolgt analog zum unteren Frequenzband.

Mithilfe orthogonaler Kontraste konnte ein erwartungsgemäßer Abfall der Leistung in absoluten Einheiten während Aufgabenbearbeitung im Vergleich zum Mittelwert der beiden Ruhemessungen (siehe Tabelle 2.10) festgestellt werden ($F(1,76) = 7,79, p = .007, R^2 = .09$). Es zeigte sich allerdings keine Interaktion der Polarität ($F < 1$) sowie der Beleuchtung ($F < 1$) mit dem Zeitpunkt. Auch konnte keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt nachgewiesen werden ($F(1,76) = 2,2$). Vergleicht man die Leistung in Ruhemessung 1 und 2, zeigte sich kein signifikanter Unterschied ($F(1,76) = 3,66$). Es konnte auch keine Interaktion zwischen der Polarität ($F < 1$) sowie der Helligkeit ($F(1,76) = 1,26$) und dem Zeitpunkt festgestellt werden. Auch zeigte sich keine Interaktion zwischen den beiden Faktoren und dem Zeitpunkt ($F(1,76) = 1,68$).

Bei der Darstellung der Leistung in normalisierten Einheiten in Abbildung 2.18 erkennt man deutlich, dass die Leistung wie erwartet in allen vier Gruppen bei der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung abfiel und dann in der zweiten Ruhemessung in allen Bedingungen außer der Positiv Dunkel-Gruppe wieder anstieg. Über orthogonale Kontraste konnte gezeigt werden, dass der Abfall der Leistung im Vergleich zum Mittelwert der beiden Ruhemessungen signifikant war ($F(1,76) = 20,53, p < .001, R^2 = .21$). Wie in Abbildung 2.19 zu sehen, ist der Abfall der Leistung in der Negativ Dunkel-Bedingung deskriptiv am größten. Es zeigte sich allerdings weder eine Interaktion der Polarität ($F < 1$), noch der Beleuchtung ($F < 1$) mit dem Zeitpunkt. Auch lag keine Interaktion zwischen diesen beiden Variablen und dem Faktor Zeitpunkt vor ($F(1,76) = 3,55$). Vergleicht man mittels orthogonaler Kontraste die Leistung in normalisierten Einheiten in Ruhemessung 1 und 2, zeigte sich ein signifikanter Unterschied ($F(1,76) = 16,59, p < .001, R^2 = .18$). Es konnte allerdings keine Interaktion zwischen der Polarität ($F < 1$), der Helligkeit ($F(1,76) = 1,26$) sowie auch keine Interaktion zwischen den beiden Faktoren und dem Messzeitpunkt ($F(1,76) = 1,68$) festgestellt werden.

Betrachtet man den Verlauf der Leistung in normalisierten Einheiten als Differenzwerte zur ersten Ruhemessung über die 15 Texte in Abbildung 2.20, zeigte sich kein Effekt des Zeitverlaufs ($F(14,63) = 2,01$). Die Faktoren Polarität ($F(14,63) = 1,25$) und Beleuchtung

($F(14,63) = 1,03$) hatten auch keinen Effekt auf den Zeitverlauf. Es konnte darüber hinaus keine Interaktion des Zeitverlaufs mit diesen beiden Variablen festgestellt werden ($F < 1$).

Tabelle 2.10

Mittelwerte der Leistung im oberen Frequenzband in ms^2 und normalisierten Einheiten während Ruhe (Baseline 1 und 2) sowie Aufgabenbearbeitung. In Klammern sind die Standardfehler der Mittelwerte angegeben.

		Baseline 1	Aufgabe (Mittelwert Text 1-15)	Baseline 2
absolute Leistung (ms^2)	Positiv Hell	578,67 (154,83)	335,09 (55,52)	509,59 (130,28)
	Positiv Dunkel	323,09 (58,85)	191,56 (38,24)	269,32 (56,59)
	Negativ Hell	337,34 (62,09)	288,67 (72,88)	336,29 (70,48)
	Negativ Dunkel	797,83 (422,94)	399,12 (156,71)	581,62 (284,58)
normalisierte Leistung (NU)	Positiv Hell	47,06 (3,89)	38,53 (2,49)	42,86 (4,09)
	Positiv Dunkel	45,11 (4,32)	39,56 (3,41)	38,01 (3,76)
	Negativ Hell	43,66 (4,37)	36,96 (3,77)	37,86 (4,11)
	Negativ Dunkel	55,81 (4,94)	43,17 (4,36)	45,78 (4,99)

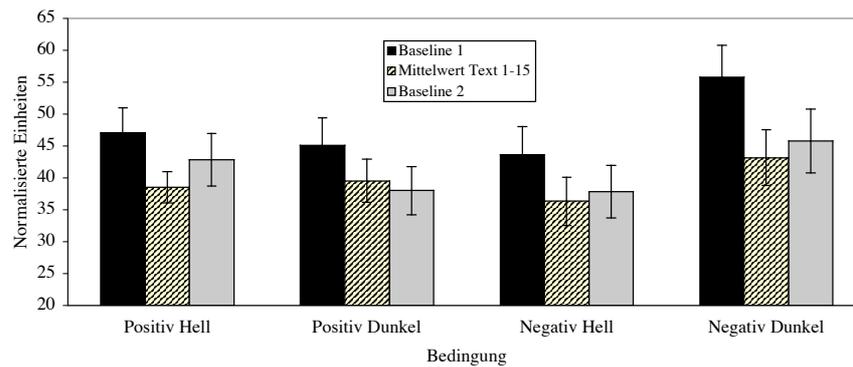


Abbildung 2.18: Durchschnittliche Leistung im oberen Frequenzband in normalisierten Einheiten (NU) während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

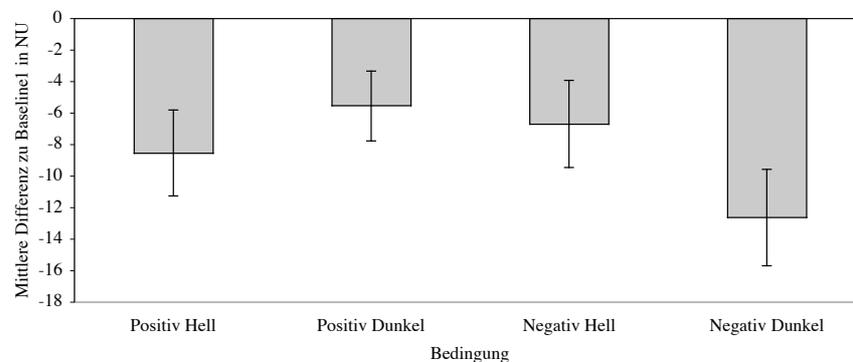


Abbildung 2.19: Abfall der Leistung im oberen Frequenzband im Vergleich zu Baseline 1 gemittelt über alle Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

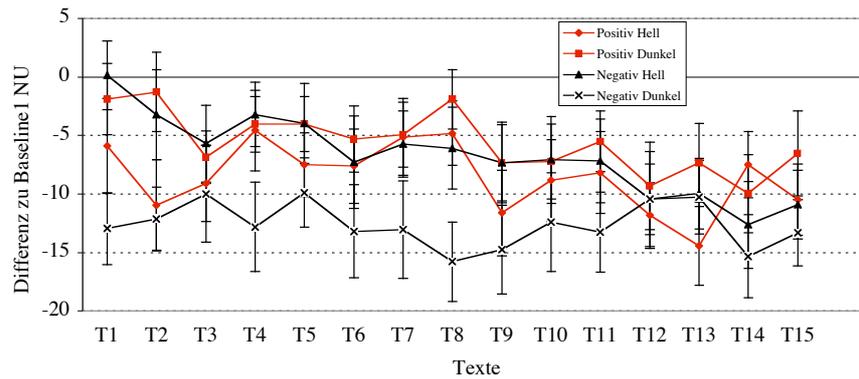


Abbildung 2.20: Dargestellt ist der Verlauf der Leistung im oberen Frequenzband. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Zeitbezogene Streuungsmaße: Abbildung 2.21 illustriert deutlich, dass der Wert der PNN50 in allen vier Gruppen wie erwartet während der Aufgabenbearbeitung abfiel und dann in der zweiten Ruhemessung wieder anstieg. Mithilfe orthogonaler Kontraste wurde der PNN50-Wert während Aufgabenbearbeitung mit dem Mittelwert der beiden Ruhemessungen verglichen. Es zeigte sich, dass der Abfall signifikant war ($F(1,76) = 54,48, p < .001, R^2 = .42$). Die Variablen Polarität ($F(1,76) = 1,28$) und Beleuchtung ($F < 1$) zeigten allerdings keine Interaktion mit dem Zeitpunkt. Auch konnte keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt festgestellt werden ($F(1,76) = 1,28$). Vergleicht man mittels orthogonaler Kontraste den PNN50-Wert in Ruhemessung 1 und 2, zeigte sich kein signifikanter Unterschied ($F < 1$). Es konnte auch keine Interaktion zwischen der Polarität ($F < 1$) sowie der Helligkeit ($F(1,76) = 2,91$) und dem Zeitpunkt festgestellt werden. Auch lag keine Interaktion zwischen den beiden Faktoren und dem Zeitpunkt vor ($F < 1$).

Der Verlauf der PNN50 über die 15 Texte als Differenzwerte zur ersten Ruhemessung ist in Abbildung 2.22 dargestellt. Es zeigte sich in einer $2 \times 2 \times 15$ MANOVA kein Effekt des Zeitverlaufs ($F(14,63) = 1,09$). Die Faktoren Polarität ($F(14,63) = 1,46$) und Beleuchtung ($F(14,63) = 1,83$) hatten keinen Effekt auf den Zeitverlauf. Auch lag keine Interaktion zwischen dem Zeitverlauf und diesen beiden Faktoren vor ($F < 1$).

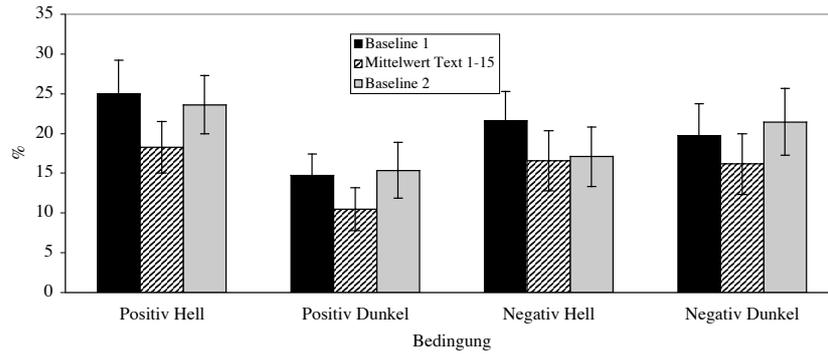


Abbildung 2.21: Mittlerer Prozentsatz der Intervalle mit mindestens 50ms Differenz zum vorausgehenden Intervall (PNN50) während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

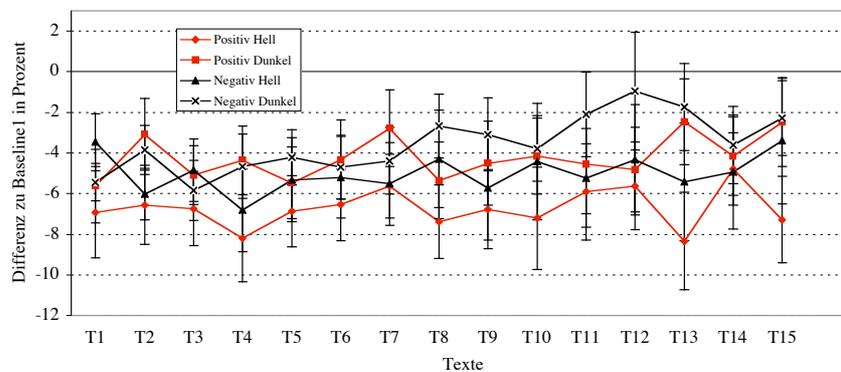


Abbildung 2.22: Illustration des PNN50-Verlaufs über die 15 Texte. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

In Abbildung 2.23 sieht man deutlich, dass der Wert der Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD) in allen vier Gruppen wie erwartet während der Aufgabenbearbeitung abfiel und dann ausgenommen der Negativ Hell-Gruppe in der zweiten Ruhemessung wieder anstieg. Es konnte mit orthogonalen Kontrasten gezeigt werden, dass dieser Abfall im Vergleich zum Mittelwert der beiden Ruhemessungen signifikant war ($F(1,76) = 22,59, p < .001, R^2 = .23$). Es lag allerdings keine Interaktion zwischen der Polarität ($F < 1$) sowie der Beleuchtung ($F < 1$) und dem Zeitpunkt vor. Ebenso fand sich auch keine Interaktion zwischen den beiden Faktoren und dem Zeitpunkt ($F(1,76) = 2,27$). Vergleicht man den RMSSD-Wert in Ruhemessung 1 und 2, zeigte sich kein signifikanter Unterschied ($F < 1$). Es konnte auch keine Interaktion zwischen der Polarität ($F(1,76) =$

1,63), der Helligkeit ($F(1,76) = 2,81$) sowie auch keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt ($F < 1$) festgestellt werden.

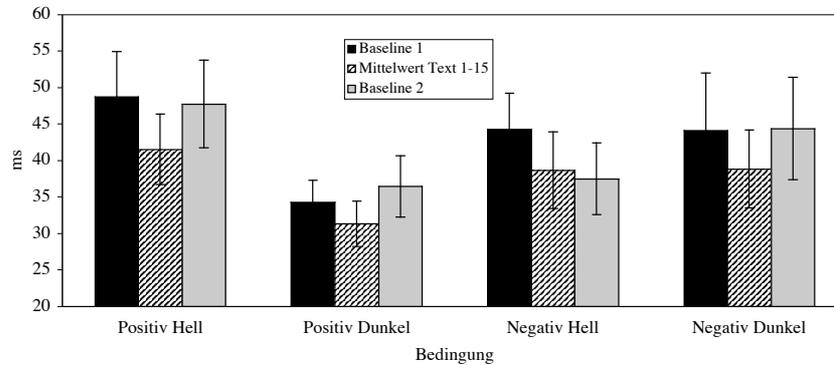


Abbildung 2.23: Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD) während Ruhe (Baseline 1 und 2) und Aufgabebearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Der Verlauf der Wurzel aus dem mittleren Quadrat sukzessiver Differenzen über die 15 Texte im Vergleich zur ersten Ruhemessung ist in Abbildung 2.24 dargestellt. Es konnte kein Effekt des Zeitverlaufs festgestellt werden ($F < 1$). Die Faktoren Polarität ($F(14,63) = 1,64$) und Beleuchtung ($F(14,63) = 1,01$) hatten keinen Effekt auf den Zeitverlauf. Ebenso fand sich keine Interaktion zwischen diesen beiden Faktoren und dem Zeitverlauf ($F < 1$).

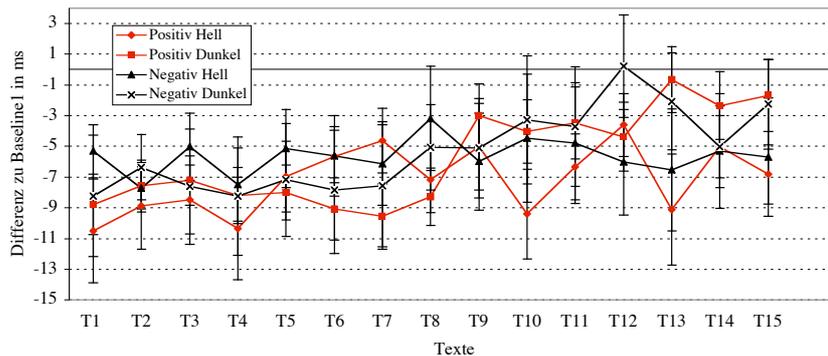


Abbildung 2.24: Illustration des RMSSD-Verlaufs über die 15 Texte. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.2.3 Elektrodermale Aktivität

In Abbildung 2.25 ist zu sehen, dass die mittlere tonische Hautleitfähigkeit in allen Gruppen während der Aufgabenbearbeitung anstieg, aber nur in der Negativ Dunkel-Gruppe in der zweiten Ruhemessung wieder abfiel. Aus diesem Grund soll für die statistische Analyse nur die erste Ruhemessung mit dem Mittelwert der Aufgabenbearbeitung verglichen werden. Es zeigte sich, dass der Anstieg der elektrodermalen Aktivität signifikant war ($F(1,76) = 85,88, p < .001, R^2 = .53$). Es konnte aber keine Interaktion zwischen der Polarität ($F(1,76) = 1,73$) sowie der Beleuchtung ($F < 1$) und dem Zeitpunkt festgestellt werden. Ebenso lag keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt vor ($F < 1$).

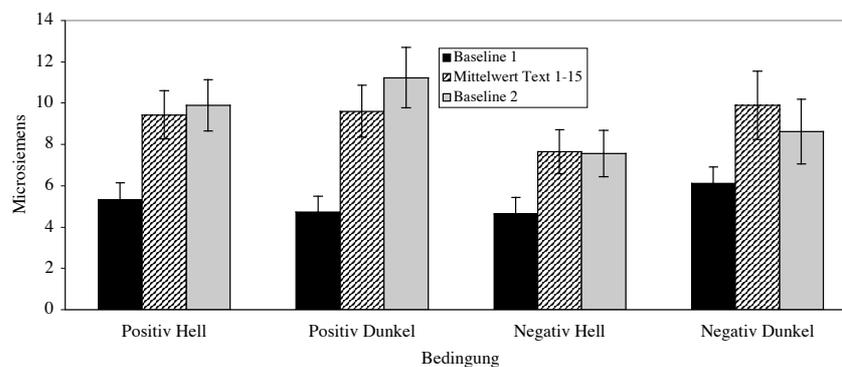


Abbildung 2.25: Tonische Hautleitfähigkeit während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die dargestellten Werte stammen aus einem variablen Zeitfenster von etwa zehn Sekunden gegen Ende der dreiminütigen Textdarbietungsdauer. Bei den Werten handelt es sich nicht um Absolutwerte, sondern um Differenzwerte zum individuellen Nullwert jeder Person, der vor dem eigentlichen Versuch bestimmt wurde. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

In Abbildung 2.26 ist der Verlauf der tonischen Leitfähigkeit über die 15 Texte dargestellt. Es handelt sich dabei um Differenzwerte zur ersten Ruhemessung. Es zeigte sich, dass die Leitfähigkeit bis etwa zur Hälfte der Texte kontinuierlich anstieg, um sich dann auf einem relativ konstanten Wert einzupendeln. Eine $2 \times 2 \times 15$ MANOVA zeigte einen signifikanten Effekt des Zeitverlaufs ($F(14,63) = 2,43; p = .008, R^2 = .35$). Es konnte aber kein Effekt der Polarität ($F < 1$) und der Beleuchtung ($F < 1$) auf den Zeitverlauf gefunden werden.

Auch lag keine Interaktion zwischen diesen beiden Faktoren und dem Zeitverlauf vor ($F < 1$).

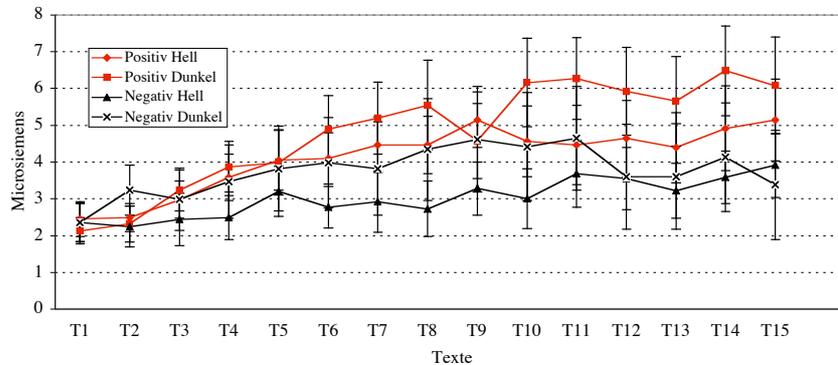


Abbildung 2.26: Verlauf der mittleren tonischen Hautleitfähigkeit über die 15 Texte. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.2.4 Atmung

Abbildung 2.27 illustriert die mittlere Atemfrequenz sowie die Standardfehler während Ruhe und Aufgabenbearbeitung. Es ist zu erkennen, dass in allen Gruppen bei Aufgabenbearbeitung die Atemfrequenz im Vergleich zur ersten Ruhemessung deutlich anstieg und dann in der zweiten Ruhemessung wieder abfiel. Mittels orthogonaler Kontraste wurde die Atemfrequenz während Aufgabenbearbeitung mit dem Mittelwert der beiden Ruhemessungen verglichen. Es konnte gezeigt werden, dass der Anstieg der Atemfrequenz signifikant war ($F(1,76) = 100,38, p < .001, R^2 = .57$). Es konnte allerdings weder eine Interaktion zwischen der Polarität ($F < 1$), noch zwischen der Beleuchtung ($F < 1$) und dem Zeitpunkt nachgewiesen werden. Auch lag keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt vor ($F(1,76) = 1,47$). Vergleicht man die Atemfrequenz in den beiden Ruhemessungen, zeigte sich kein signifikanter Unterschied ($F(1,76) = 1,32$). Es konnte auch keine Interaktion zwischen der Polarität ($F < 1$) sowie zwischen der Helligkeit ($F(1,76) = 1,43$) und dem Zeitpunkt festgestellt werden. Auch fand sich keine Interaktion zwischen diesen beiden Faktoren und dem Zeitpunkt ($F < 1$).

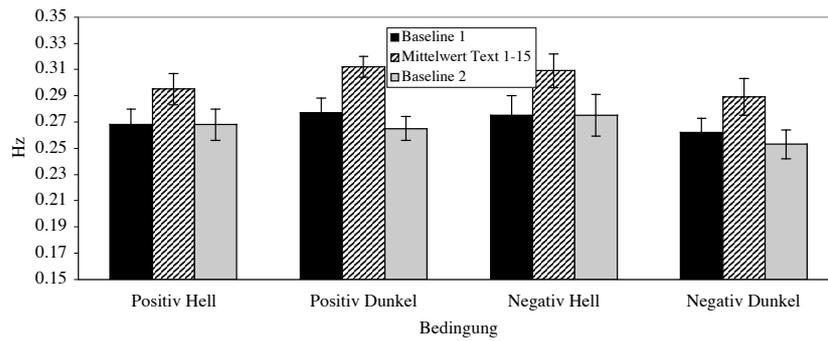


Abbildung 2.27: Mittlere Atmungsfrequenz während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Betrachtet man den Verlauf der mittleren Atemfrequenz in Relation zur ersten Ruhemes-
 sung in Abbildung 2.28, sieht man, dass mit zunehmender Versuchsdauer die Frequenz in
 allen Gruppen immer weiter absank. Eine $2 \times 2 \times 15$ MANOVA zeigte einen signifikanten
 Effekt des Zeitverlaufs ($F(14,63) = 5,35, p < .001, R^2 = .54$). Polarität ($F < 1$) und Be-
 leuchtung ($F(14,63) = 1,29$) hatten keinen Effekt auf den Zeitverlauf. Auch konnte keine
 Interaktion des Zeitverlaufs mit diesen beiden Faktoren festgestellt werden ($F(14,63) =$
 1,37).

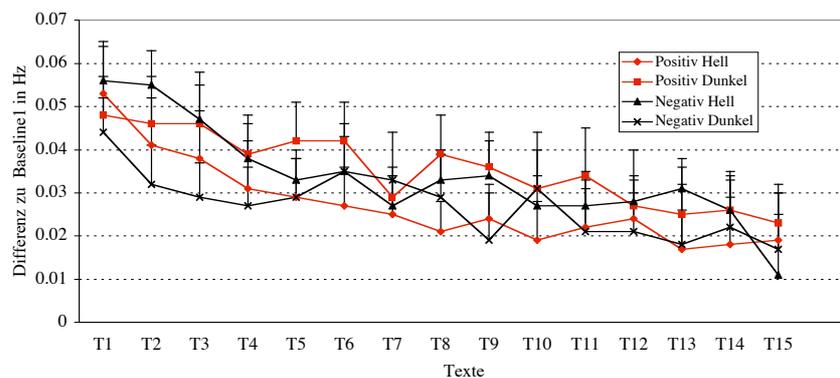


Abbildung 2.28: Verlauf der mittleren Atemfrequenz über die 15 Texte. Bei den Werten
 handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken rep-
 räsentieren die Standardfehler der Mittelwerte.

2.2.3 Subjektive Beanspruchung

Die subjektive Beanspruchung gliedert sich in die Bereiche physische Beschwerden sowie subjektive Befindlichkeit.

2.2.3.1 Physische Beschwerden

Der durchschnittliche Anstieg der Symptomkomplexe Augenbeschwerden, Kopfschmerzsymptome und Muskel- und Haltungsbeschwerden sowie die Standardfehler sind in Abbildung 2.29 dargestellt. Die Zunahme der Augenbeschwerden war in allen Gruppen signifikant ($F(1,76) = 73,01, p < .001, R^2 = .49$). Man kann erkennen, dass der Anstieg der visuellen Beschwerden in der Negativ Hell- ($M = 6,65$) und der Negativ Dunkel-Bedingung ($M = 5,95$) größer war als in der Positiv Hell- ($M = 5,35$) und der Positiv Dunkel-Gruppe ($M = 4,65$). Das Ergebnis einer Varianzanalyse mit Beleuchtung und Polarität als unabhängigen Variablen zeigte allerdings keinen signifikanten Effekt der Polarität ($F < 1$). Ebenso konnte kein Effekt der Beleuchtung ($F < 1$) und keine Interaktion zwischen den beiden Faktoren festgestellt werden ($F < 1$). Der durchschnittliche Anstieg der Kopfschmerzsymptome in allen vier Bedingungen war ebenfalls signifikant ($F(1,76) = 53,54, p < .001, R^2 = .41$). Man sieht, dass die Zunahme der Kopfbeschwerden in der Positiv Hell- ($M = 3,2$) und der Negativ Hell-Bedingung ($M = 3,6$) größer war als in der Positiv Dunkel- ($M = 1,65$) und der Negativ Dunkel-Gruppe ($M = 1,9$). Das Ergebnis einer ANOVA mit Beleuchtung und Polarität als unabhängigen Variablen zeigte einen signifikanten Effekt der Beleuchtung ($F(1,76) = 5,28, p = .024, R^2 = .06$). Es konnte allerdings kein Effekt der Polarität festgestellt werden ($F < 1$). Auch zeigte sich keine Interaktion zwischen den beiden Faktoren ($F < 1$).

Die Haltungs- und Muskelbeschwerden nahmen auch in allen vier Bedingungen signifikant zu ($F(1,76) = 76,67, p < .001, R^2 = .5$). Es lag aber kein Effekt der Polarität ($F < 1$) sowie der Beleuchtung vor ($F(1,76) = 1,25$). Ebenso konnte keine Interaktion zwischen den beiden Faktoren festgestellt werden ($F(1,76) = 1,49$).

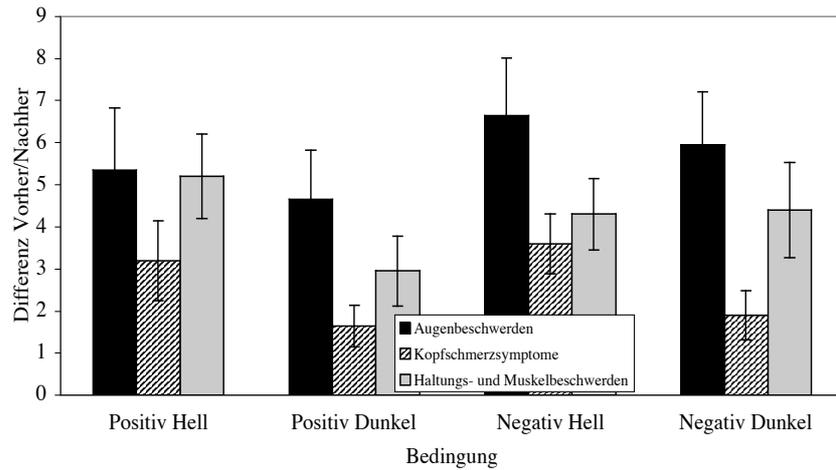


Abbildung 2.29: Durchschnittlicher Anstieg der Augenbeschwerden, Kopfschmerzsymptome sowie Haltungs- und Muskelbeschwerden nach Aufgabebearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.2.3.2 Befindlichkeit

Abbildung 2.30 zeigt die durchschnittliche Veränderung nach Aufgabebearbeitung sowie die Standardfehler auf den Skalen „Gute/Schlechte Stimmung“, „Wachheit/Müdigkeit“ sowie „Ruhe/Unruhe“ des Mehrdimensionalen Befindlichkeitsfragebogens in allen Gruppen. Man kann erkennen, dass bis auf die Negativ Dunkel-Gruppe der Skalenwert „Gute/Schlechte Stimmung“ in allen Bedingungen abnahm, das heißt die Stimmung wurde schlechter. Fasst man für die statistische Analyse alle Probanden zusammen, war dieser Effekt signifikant ($F(1,76) = 7,66, p = .007, R^2 = .09$). Das Ergebnis einer Varianzanalyse mit Beleuchtung und Polarität als unabhängigen Variablen zeigte allerdings weder einen signifikanten Effekt der Polarität ($F(1,76) = 3,07$), noch der Beleuchtung ($F(1,76) = 1,67$). Auch konnte keine Interaktion zwischen den beiden Faktoren festgestellt werden ($F < 1$). Der Wert der Skala „Wachheit/Müdigkeit“ nahm in allen Gruppen nach Aufgabebearbeitung deutlich ab ($F(1,76) = 46,86, p < .001, R^2 = .38$). Das Müdigkeitsempfinden wurde also größer. Es lag aber weder ein Haupteffekt der Polarität ($F < 1$), noch der Beleuchtung vor ($F(1,76) = 1,08$). Auch zeigte sich keine Interaktion zwischen den beiden Faktoren ($F < 1$).

Auch die Werte der Skala „Ruhe/Unruhe“ sanken in allen Bedingungen in der Nachmessung signifikant ab ($F(1,76) = 9,94, p = .002, R^2 = .12$). Die Probanden fühlten sich nach

Aufgabenbearbeitung unruhiger und unausgeglichener als vorher. Es gab weder einen Haupteffekt der Polarität ($F < 1$), noch der Beleuchtung ($F < 1$). Ebenso konnte keine Interaktion zwischen den beiden Faktoren festgestellt werden ($F(1,76) = 1,53$).

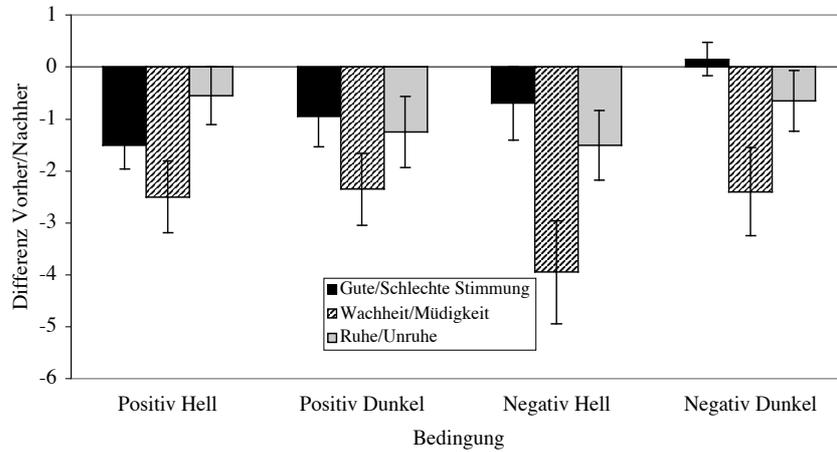


Abbildung 2.30: Mittlere Veränderung auf den Skalen „Gute/Schlechte Stimmung“, „Wachheit/Müdigkeit“ sowie „Ruhe/Unruhe“ nach Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

2.3 Interpretation und Diskussion

Die Ergebnisse werden getrennt nach den Bereichen Performanz, objektive und subjektive Beanspruchung zusammengefasst und interpretiert.

2.3.1 Performanz

Es wurde die Hypothese aufgestellt, dass sich bei einer Standardbeleuchtung im Rahmen einer Lese-Korrektur-Aufgabe eine bessere Leistung für eine Textdarstellung in positiver Polarität zeigt. Weiterhin wurde vermutet, dass sich bei Dunkelheit der Polaritätseffekt dahin gehend umdreht, dass auf Performanzebene ein Vorteil für eine Textpräsentation in negativen Kontrastverhältnissen vorliegt.

Die Ergebnisse zeigen, dass unter negativen Kontrastverhältnissen sowohl bei einer Raumbeleuchtung von 550 Lux als auch bei einer niedrigen Umgebungsbeleuchtung von 5 Lux die Leistung der Probanden deutlich schlechter war als im Fall einer Textpräsentation in positiver Polarität. Die Tatsache, dass bei negativem Kontrast gemittelt über alle 15 Texte signifikant weniger Fehler pro Textseite entdeckt wurden, konnte zwei Gründe haben. Zum einen war es möglich, dass bei negativer Darstellung das Lesetempo geringer war und infolgedessen weniger Wörter gelesen und somit auch weniger Fehler entdeckt werden konnten. Zum anderen bestand aber auch die Möglichkeit, dass eine geringere Genauigkeit bei der Fehlerentdeckung die Ursache für dieses Ergebnis war. Die Daten zeigen, dass die Anzahl der gelesenen Wörter bei negativer Polarität signifikant geringer war unabhängig von den Beleuchtungsbedingungen. Die Fehlerentdeckungs-Genauigkeit wies ebenfalls unter negativen Kontrastverhältnissen geringere Werte auf, allerdings war dieses Ergebnis statistisch nicht signifikant. Man kann also die Schlussfolgerung ziehen, dass die bessere Performanz, das heißt die höhere Anzahl entdeckter Fehler unter positiven Kontrastverhältnissen in erster Linie auf ein schnelleres Lesetempo und erst zweitrangig auf eine bessere Genauigkeit zurückzuführen ist. Darüber hinaus war von Interesse, ob sich in den vier Experimentalgruppen die Performanz im Verlauf der Aufgabenbearbeitung aufgrund von

Ermüdung oder Motivationsdefiziten möglicherweise unterschied. Die Ergebnisse zeigen allerdings, dass in allen Bedingungen die Leistung der Probanden über die gesamte Lesedauer von 45 Minuten konstant verlief. In Hinblick auf Genauigkeit, Lesetempo und die absolute Anzahl gefundener Fehler zeigten sich zwischen den Gruppen keine Unterschiede im Verlauf. Die Leistung ließ auch gegen Ende der Aufgabenbearbeitung nicht nach. Aus diesem Resultat kann die Schlussfolgerung gezogen werden, dass die Performanz unter negativen Kontrastverhältnissen konsistent während des gesamten Versuchs schlechter war. Allerdings fällt auf, dass in allen Gruppen der erste Text deutlich langsamer gelesen wurde im Vergleich zu den nachfolgenden Texten. Dieser Effekt lässt sich dadurch erklären, dass die Teilnehmer sich zunächst an die zeitliche Begrenzung gewöhnen mussten. Nach dem ersten Text wurde den Probanden bewusst, dass sie schneller lesen müssen, um die gestellten Anforderungen besser bewältigen zu können.

An dieser Stelle könnte angemerkt werden, dass sich die vier Gruppen eventuell bezüglich der Leistungsfähigkeit – wie beispielsweise sprachliche Fähigkeiten oder Rechtschreibkenntnisse – unterschieden und das den Grund für das vorliegende Ergebnis darstellt. Dagegen spricht zum einen, dass es sich bei fast allen Probanden um Studenten handelte und die vier Experimentalgruppen bezüglich ihrer Zusammensetzung homogen waren. Zum anderen wurden die Versuchspersonen befragt, wie sicher sie im Umgang mit der Rechtschreibung sind. Es zeigte sich, dass sich die Gruppen diesbezüglich nicht unterschieden. Darüber hinaus fiel auch die subjektive Bewertung der Textschwierigkeit in allen Experimentalgruppen gleich aus. Das Textniveau wurde von allen als relativ leicht eingestuft.

Im Rahmen der Fehleranalyse konnte insgesamt zwischen fünf verschiedenen Fehlerarten unterschieden werden, und es stellt sich die Frage, ob sich die Entdeckung der fünf Fehlertypen möglicherweise zwischen den Gruppen unterschied. Von besonderer Bedeutung waren dabei die eingefügten Grammatikfehler, die dazu dienten, die Lesestrategie der Probanden zu überprüfen. Sie konnten nur entdeckt werden, wenn beim Lesen die Intention bestand, die Bedeutung des Textes zu erfassen. Es gab allerdings keine Hinweise darauf, dass die Experimentalgruppen hinsichtlich ihrer Lesestrategie differierten. Es zeigte sich zwar, dass signifikant weniger Grammatikfehler erkannt wurden in Relation zu den vier anderen Fehlertypen, was aber für positive und negative Kontrastverhältnisse gleichermaßen galt. Darüber hinaus ergab sich die Fragestellung, ob mit zunehmender Versuchsdauer

eventuell die Strategie geändert und der Text nur noch flüchtig gelesen wurde. Die Daten deuten allerdings darauf hin, dass dies nicht der Fall war. In allen vier Gruppen nahm der Prozentsatz richtig erkannter Grammatikfehler mit zunehmender Versuchsdauer nicht ab, was auf eine unveränderte Lesestrategie während des gesamten Versuchs hindeutet.

Klammert man die Grammatikfehler von der weiteren Betrachtung aus, zeigte sich, dass auch die anderen vier Fehlerarten mit unterschiedlicher Genauigkeit erkannt wurden. Die Fehlerart „falscher Buchstabe“ wurde mit höherer Genauigkeit erkannt als die anderen drei Fehlerarten. Wörter, in denen ein Buchstabe fehlte, wurden dagegen am wenigsten häufig entdeckt. Creed (1987) verwendete ebenfalls das Paradigma einer Lese-Korrektur-Aufgabe und unterschied dabei zwischen visuell ähnlichen (z.B. „e“ durch „c“ ersetzt) und visuell unähnlichen Fehlern (z.B. „e“ durch „w“ ersetzt). Es stellte sich heraus, dass die visuell unähnlichen Fehler einfacher entdeckt wurden. In der vorliegenden Arbeit kann man davon ausgehen, dass die Wörter, die einen Buchstaben zuviel enthielten, ebenfalls visuell unähnliche Fehler darstellten und dadurch mit höherer Genauigkeit erkannt wurden. Wörter, die dagegen einen Buchstaben zuwenig enthielten, wurden möglicherweise schlechter erkannt, weil fehlende Buchstaben eventuell durch frühere Erfahrungen in der Worterkennung automatisch ergänzt wurden.

Zusammenfassend lässt sich sagen, dass bei heller Umgebungsbeleuchtung auf Performanzebene ein erwartungsgemäßes Ergebnis vorliegt. Wie bereits in früheren Studien gezeigt werden konnte, ist die Leistung bei einer Textpräsentation in positiver Polarität signifikant besser.

In der Vergangenheit gab es allerdings Hinweise darauf, dass bei Bildschirmarbeit in einer dunklen Umgebung die Negativdarstellung die bessere Wahl darstellt, da dadurch die Dunkeladaptation des Nutzers erhalten bleibt und beanspruchende Variationen im Pupillendurchmesser aufgrund verschiedener Leuchtdichten von Monitor und Arbeitsumgebung geringer ausfallen (Taptagaporn & Saito, 1990). Diese Empfehlung beruht allerdings in erster Linie auf theoretischen Überlegungen, die sich auf die Physiologie des visuellen Systems beziehen sowie auf Experimenten zur Variation der Pupillenweite. Es wurde bisher nicht analysiert, ob sich bei Dunkelheit die Vorteile einer Negativdarstellung auch auf die Leistung auswirken. In der vorliegenden Untersuchung wurde erstmals systematisch der Einfluss der Polarität unter verschiedenen Lichtbedingungen auf der Performanzebene

untersucht. Es zeigte sich, dass auch bei Dunkelheit die Negativdarstellung der positiven Polarität deutlich unterlegen ist, da im Rahmen der Lese-Korrektur-Aufgabe signifikant weniger Fehler gefunden wurden. Dieser Effekt war im dunklen Raum sogar noch größer als bei Beleuchtung.

An dieser Stelle kommt die Frage auf, wie die verschiedenen Arten der Textpräsentation von den Versuchsteilnehmern subjektiv beurteilt wurden und ob sich die Performanz auch in der subjektiven Bewertung widerspiegelte. Aus diesem Grund wurden die Probanden am Ende des Versuchs gebeten zu beurteilen, wie angenehm bzw. unangenehm sie die Textdarstellung empfanden. Es zeigte sich allerdings, dass sich die Gruppen bezüglich dieser Einschätzung nicht unterschieden. Von allen Probanden wurde die Textpräsentation als unangenehm wahrgenommen. Dieses Ergebnis war zum einen darin begründet, dass die Textdarstellung – unabhängig von der Polarität – aufgrund der langen Zeilen und geringen Zeilenabstände unattraktiv wirkte. Zum anderen wurden die Bedingungen zwischen den Probanden variiert und somit bestand für die einzelne Versuchsperson keine Vergleichsmöglichkeit mit den anderen Präsentationsarten.

Das Resultat der vorliegenden Untersuchung lässt die Schlussfolgerung zu, dass entgegen der Hypothese und den Annahmen von unter anderem Taptagaporn und Saito (1990) auch bei Dunkelheit auf eine Textpräsentation in negativer Polarität verzichtet werden sollte.

2.3.2 Objektive Beanspruchung

Ein weiteres Ziel der vorliegenden Untersuchung war es zu zeigen, dass sich die Beanspruchung bei Bearbeitung der Lese-Korrektur-Aufgabe auch objektiv, das heißt mithilfe psychophysiologischer Messungen, unter Kontrolle der motorischen Aktivität der Probanden nachweisen lässt. Als Beanspruchungsindikatoren wurden elektrodermale und kardiovaskuläre Parameter sowie die Atmung erfasst. Da sich die körperliche Fitness der Versuchsteilnehmer auf die psychophysiologischen Messwerte auswirken kann (vgl. Moya-Albiol et al., 2001), wurde diese Variable kontrolliert. Die Probanden wurden bezüglich

des Ausmaßes sportlicher Aktivitäten befragt und die Gruppen unterschieden sich nicht hinsichtlich dieses Faktors.

Die Ergebnisse zeigen, dass die Herzrate in allen Gruppen während der Aufgabenbearbeitung signifikant anstieg im Vergleich zur ersten Ruhemessung und dann in der zweiten Ruhemessung wieder abfiel. Betrachtet man den Verlauf der Herzrate über alle 15 Texte, kann man erkennen, dass die Herzfrequenz mit fortschreitender Versuchsdauer in allen Gruppen immer weiter absank. Dieses Phänomen war deskriptiv am stärksten ausgeprägt in der Negativ Dunkel-Bedingung. Ein Absinken der Herzrate während der Bearbeitung kognitiver Aufgaben wird in der Literatur diskutiert als Zeichen von Müdigkeit, Monotonie, nachlassender Aufmerksamkeit und /oder geringer Aktivierung (Milosevic, 1997).

Auch die Parameter der Herzratenvariabilität bildeten eine deutliche Beanspruchung während der Aufgabenbearbeitung ab. Als zeitbezogene Streuungsmaße wurden standardmäßig der Prozentsatz der Intervalle mit mindestens 50ms Abweichung vom vorausgehenden Intervall (PNN50) sowie die Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD) berechnet. Erwartungsgemäß sanken die Kennwerte beider Parameter in allen Gruppen während der Aufgabenbearbeitung signifikant ab und stiegen dann in der zweiten Ruhemessung wieder an. Die spektralanalytische Auswertung der Herzratenvariabilität spiegelte ebenfalls die psychophysiologische Beanspruchung während der Bearbeitung der Lese-Korrektur-Aufgabe wider. Es zeigte sich ein signifikanter Anstieg der normalisierten Leistung im mittleren Frequenzband (0,07-0,14 Hz) in allen Gruppen. Analog dazu fiel die Leistung in normalisierten Einheiten im oberen Band (0,15-0,4 Hz) signifikant ab. Dieses Ergebnis steht vollkommen im Einklang mit früheren Studien, die eine Leistungszunahme im mittleren Frequenzband sowie eine Abnahme im oberen Frequenzband unter anderem während der Bearbeitung von Reaktions- und Rechenaufgaben sowie des Stroop-Tests zeigen konnten (Delaney & Brodie, 2000; Pagani et al., 1989; Wood et al., 2002). Die Ergebnisse deuten darauf hin, dass diese Zu- bzw. Abnahme die stärkste Ausprägung aufwies in der Negativ Dunkel-Bedingung. Allerdings war dieser Unterschied zu den anderen Bedingungen statistisch nicht signifikant.

Als weiterer Beanspruchungsindikator stieg die mittlere tonische Hautleitfähigkeit in allen Gruppen während der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung signifikant an, wobei sich die Gruppen aber nicht unterschieden. Auch dieses Ergebnis ist er-

wartungsgemäß und entspricht den Resultaten früherer Studien, die einen Anstieg der tonischen Hautleitfähigkeit bei der Bearbeitung kognitiver Anforderungen wie beispielsweise Reaktions- und Rechenaufgaben nachweisen konnten (Albus et al., 1987; Fichera & Andreassi, 1998; Morell, 1989; Oshima et al., 2001). Betrachtet man die tonische Hautleitfähigkeit im Verlauf über die gesamten Versuchsabschnitte, zeigte sich, dass die Werte bis etwa zur Hälfte der 15 Texte anstiegen, um sich dann auf einem relativ konstanten Niveau einzupendeln. Möglicherweise spiegelt dieses Ergebnis eine Art Gewöhnung an die Versuchsdurchführung wider. Die Versuchspersonen sind zu diesem Zeitpunkt mit der Aufgabe sehr vertraut und es kommt zu keinem weiteren Anstieg der Hautleitfähigkeit.

Neben kardiovaskulären und elektrodermalen Parametern wurde noch zusätzlich die Atmung aufgezeichnet. Die Atemfrequenz diente zum einen als Kontrollparameter für die Erfassung der Herzratenvariabilität sowie der tonischen Hautleitfähigkeit, wurde aber auch als eigenständiger Indikator im Sinne einer mehrdimensionalen Beanspruchungsanalyse erfasst. Es zeigte sich in allen Gruppen, dass die Atemfrequenz während der Lese-Korrektur-Aufgabe signifikant anstieg und in der zweiten Ruhemessung wieder auf das Ursprungsniveau zurückkehrte. Betrachtet man den Verlauf der mittleren Atemfrequenz über die insgesamt 15 Texte, zeigte sich mit zunehmender Versuchsdauer ein kontinuierliches, signifikantes Absinken der Atemfrequenz in allen Gruppen. Ähnlich wie bei der Herzfrequenz kann dieses Phänomen als Müdigkeit oder Monotonie interpretiert werden. Zusammenfassend lässt sich sagen, dass unter Kontrolle der motorischen Aktivität alle psychophysiologischen Indikatoren die Beanspruchung während der Bearbeitung der Lese-Korrektur-Aufgabe deutlich abbildeten.

Ziel der vorliegenden Studie war außerdem, herauszufinden, ob sich Display-abhängige Unterschiede eventuell auch in der objektiven Beanspruchung widerspiegeln. Es hat sich allerdings gezeigt, dass sich die vier Gruppen nicht signifikant im Ausmaß der psychophysiologischen Beanspruchung unterschieden. Außerdem konnte auch auf physiologischer Ebene kein Hinweis darauf gefunden werden, dass bei Dunkelheit die negative Polarität die bessere Wahl darstellt. Wenn der negative Kontrast bei Dunkelheit tatsächlich zu bevorzugen wäre, hätte man vermuten können, dass die Probanden unabhängig von ihrer Leistung in dieser Bedingung eine geringere objektive Beanspruchung aufweisen. Dies war allerdings nicht der Fall. Neben der Performanz sprechen also auch die physiologischen

Daten nicht dafür, bei Dunkelheit auf eine Textdarstellung in negativen Kontrastverhältnissen zurückzugreifen.

Zusammenfassend kann man sagen, dass sich eine schlechte Leistung bzw. eine ungünstige Wahl der Polarität nicht in einer signifikant höheren objektiven Beanspruchung widerspiegelte. Dieses Ergebnis ist vereinbar mit der Annahme, dass sich im gegebenen experimentellen Kontext alle Probanden gleichermaßen anstregten. Display-abhängige Unterschiede spiegeln sich nur in den Performanzindikatoren wider, das heißt, ob bezüglich der Textdarstellung eine bestimmte Art der Polarität unter den verschiedenen Beleuchtungsbedingungen günstiger ist als eine andere Kontrastwahl, zeigte sich lediglich auf der Leistungsebene. Dadurch wird eine eindeutige Interpretation der Performanzindikatoren ermöglicht.

2.3.3 Subjektive Beanspruchung

Es wurde die Hypothese aufgestellt, dass im Rahmen der Bildschirmarbeit sowohl psychische als auch physische Beschwerden deutlich werden.

Die Ergebnisse zeigen, dass auf physischer Ebene bei fast allen Probanden am Ende der Aufgabenbearbeitung deutliche Beschwerden vorlagen. Vor allem konnte ein signifikanter Anstieg visueller Probleme wie Schmerzen, Tränen, Brennen und Jucken festgestellt werden. Die Zunahme der Augenbeschwerden war unabhängig von der Beleuchtung unter negativen Kontrastverhältnissen deskriptiv größer, allerdings war dieser Effekt der Polarität nicht signifikant. Wäre die negative Polarität bei Dunkelheit wirklich die bessere Alternative, hätte sich diese Tatsache vor allem in einem geringeren Ausmaß visueller Beschwerden widerspiegeln müssen, was aber nicht der Fall war. Die stärkere Zunahme visueller Probleme bei negativem Kontrast und Helligkeit stellte dagegen ein erwartungsgemäßes Ergebnis dar und bestätigte die Tatsache, dass bei normaler Beleuchtung die positive Polarität empfohlen wird. Darüber hinaus kam es in allen Gruppen zu einer signifikanten Zunahme von Kopfschmerzen, Benommenheit und Schwindel. Der Anstieg der Symptome war unabhängig von der Polarität am stärksten bei Licht und die statistische Analyse zeigte einen signifikanten Haupteffekt der Helligkeit. Dieses Ergebnis kann dadurch erklärt werden, dass in den Lichtbedingungen kein Tageslicht, sondern künstliche Beleuchtung ver-

wendet wurde. Leuchtstoffröhren flimmern allerdings im 50- bzw. 100 Hz-Takt des Wechselstroms, was zu Kopfschmerzen führen kann.

Auch die Beschwerden im Nacken-, Schulter- und Rückenbereich waren in allen Gruppen am Ende der Aufgabenbearbeitung deutlich stärker ausgeprägt. An dieser Stelle muss allerdings berücksichtigt werden, dass die Probanden aufgrund der physiologischen Ableitungen während der Versuchsdurchführung ihre Sitzhaltung nicht stark variieren konnten und ruhig sitzen mussten. Neben der eigentlichen Bildschirmarbeit hat diese Tatsache mit Sicherheit zu der deutlichen Zunahme der Haltungs- und Muskelbeschwerden beigetragen.

Neben der physischen Beanspruchung zeigten sich bei allen Probanden auch Beschwerden im Bereich der subjektiven Befindlichkeit nach Beendigung der Lese-Korrektur-Aufgabe. Auf allen drei Skalen des mehrdimensionalen Befindlichkeitsfragebogens spiegelte sich die psychische Beanspruchung signifikant wider. Im Vergleich zur Vormessung war die Stimmung schlechter, das Müdigkeitsempfinden nahm deutlich zu und die Probanden fühlten sich unruhiger und unausgeglichener als vorher. Diese Verschlechterung der subjektiven Befindlichkeit konnte bereits schon in früheren Studien nachgewiesen werden (Ertel et al., 1997). Allerdings unterschieden sich in der vorliegenden Arbeit die Gruppen nicht im Ausmaß der Veränderungen, was wiederum vereinbar ist mit der Annahme, dass sich alle Probanden im Rahmen der Aufgabenbearbeitung gleich stark anstrebten.

An diesem Punkt stellt sich nun die Frage, wie die schlechte Leistung bei einer Textdarstellung in negativen Kontrastverhältnissen vor allem bei Dunkelheit eingeordnet werden kann. Da bisher die Performanz in der besagten Bedingung nur in Relation zur ergonomisch günstigen positiven Polarität gesetzt wurde, kann nicht beurteilt werden, wie schlecht die Leistung der Probanden wirklich war. Ist bereits die untere Grenze der Leistungsfähigkeit erreicht oder besteht die Möglichkeit durch Wahl einer sehr schlecht lesbaren Kombination von Text- und Hintergrundfarbe die Leistung der Probanden nochmals zu verschlechtern?

Dazu muss zunächst geklärt werden, welche Art der Präsentation in diesem Zusammenhang für das Lesen von Texten besonders ungeeignet ist. Es konnte bisher in verschiedenen Untersuchungen gezeigt werden, dass unter anderem beim Lesen rote Schrift auf grünem Hintergrund eine sehr ungünstige Farbkombination darstellt und von den Versuchsperso-

nen als besonders unangenehm empfunden wurde (Matthews et al., 1989; Pastoor, 1990; Shieh & Chen, 1997; Shieh et al., 1997). Diese subjektive Bewertung spiegelte sich in den meisten dieser Studien allerdings nicht in der objektiv messbaren Leistung wider, das heißt die verschiedenen Farbkombinationen unterschieden sich auf Performanzebene nicht. Als Begründung für dieses Ergebnis wurde die Tatsache angeführt, dass für die Leseleistung am Bildschirm der Luminanzkontrast von größerer Bedeutung zu sein scheint als der Farbkontrast (Legge & Rubin, 1986; Travis et al., 1990). Laut Snyder (1988) sollte der Luminanzkontrast mindestens bei .667 (3:1) und für eine bessere Leseleistung sogar bei .857 (7:1) liegen. Aus diesen Studien lässt sich die Schlussfolgerung ziehen, dass eine Textdarstellung mit roter Schrift auf grünem Hintergrund, die zusätzlich noch mit einem sehr niedrigen bzw. nicht vorhandenen Luminanzkontrast verbunden ist, eine sehr schlecht lesbare Bedingung darstellt. Folglich wurde im zweiten Experiment erwartet, dass für diese Farbkombination die Leistung im Rahmen einer Lese-Korrektur-Aufgabe schlechter ist als bei einer Schwarz/Weiß-Darstellung.

3. Experiment 2

3.1 Methode

In den folgenden Abschnitten erfolgt eine genaue Beschreibung des zweiten Experiments. Neben der untersuchten Stichprobe werden der Versuchsaufbau, die verwendeten Stimuli, das experimentelle Design und der Versuchsablauf dargestellt.

3.1.1 Stichprobenbeschreibung

Insgesamt nahmen am zweiten Experiment 50 Probanden (26 weiblich, 24 männlich) im Alter von 18-50 Jahren ($M = 27,92$; $SD = 6,39$) teil. Die Zusammensetzung der verschiedenen Gruppen wurde bezüglich Alter, Geschlecht und Tageszeit angeglichen, da diese Faktoren einen Einfluss auf die physiologischen Beanspruchungsindikatoren aufweisen. 43 der Probanden waren nach eigener Einschätzung Rechtshänder und sieben Linkshänder. Die durchschnittliche Erfahrung mit Bildschirmarbeit betrug 8,65 Jahre ($SD = 4,77$). Alle Probanden hatten Deutsch als Muttersprache und litten nach eigenen Angaben nicht unter kardiovaskulären Erkrankungen. Die Teilnehmer wurden gebeten, mindestens sechs Stunden vor dem Versuch kein Koffein zu konsumieren und auf sehr anstrengende körperliche Aktivitäten zu verzichten. Mithilfe der Ishihara Farbtafeln (Ishihara, 1972) wurde sichergestellt, dass bei keinem der Versuchsteilnehmer Farbsinnstörungen vorlagen. Alle Probanden hatten eine normale oder korrigierte Sehfähigkeit. Die Teilnahme am Experiment wurde mit 5 Euro entlohnt.

3.1.2 Versuchsaufbau

Der Versuchsaufbau war identisch mit Experiment 1. Die Durchführung des Experiments erfolgte erneut an einem Apple Power Mac G4 Computer und die Stimuli wurden an einem 17 Zoll TFT (= Thin Film Transistor) Farbmonitor dargeboten.

3.1.2.1 Subjektive Befindlichkeit und physische Beschwerden

Zur Messung der psychischen Beanspruchung (Befindlichkeit) wurde erneut die Kurzform des Mehrdimensionalen Befindlichkeitsfragebogens (Steyer et al., 1997) verwendet. Die körperlichen Beschwerden wurden mit dem Fragebogen zur physischen Beanspruchung von Heuer et al. (1989) erfasst.

3.1.3 Beschreibung der Stimuli

Es wurde das gleiche Stimulusmaterial wie in Experiment 1 verwendet und die Abfolge der 15 Geschichten fand erneut für jede Versuchsperson in randomisierter Form statt. Der Text wurde in der Schriftart Helvetica mit einer Schriftgröße von 14 Punkt präsentiert. Der Versuch wurde unter einer niedrigen Umgebungsbeleuchtung von 5 Lux durchgeführt.

Es gab zwei Arten der Textdarstellung. Die Hälfte der Versuchspersonen bekam den Text in achromatischer Form und positiven Kontrastverhältnissen, das heißt schwarze Schrift auf weißem Hintergrund, präsentiert. In der anderen Bedingung wurde rote Schrift auf grünem Hintergrund eingesetzt, wobei sich die beiden Farben in ihren Luminanzwerten nicht unterschieden. Die Chromatizitäts-Koordinaten und Luminanzwerte der verwendeten Farben sind in Tabelle 3.1 aufgeführt.

Tabelle 3.1

CIE Chromatizitäts-Koordinaten (x,y) und Luminanzwerte (L) der verwendeten Farben. Die Berechnung des Luminanzkontrastes erfolgte nach der Michelson Definition ($(L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$).

Farbe	x	y	Luminanz (cd/m ²)	Luminanz- kontrast
Schwarz	.294	.296	0.53	0.98
Weiß	.316	.335	192	
Rot	.297	.548	31.5	0
Grün	.627	.353	31.5	

3.1.4 Versuchsablauf

Der Ablauf des Versuchs gestaltete sich wie in Experiment 1. Die Experimente wurden einzeln durchgeführt und dauerten insgesamt jeweils circa 75 Minuten.

3.1.5 Design

Es handelte sich um ein einfaktorielles Design mit dem Faktor „Farbkombination“, der in den Ausprägungen „achromatisch“ (=schwarze Schrift auf weißem Hintergrund) und „chromatisch“ (=rote Schrift auf grünem Hintergrund) vorlag. Jeder Versuchsbedingung wurden 25 Versuchspersonen per Zufall zugeordnet.

Die abhängigen Variablen gliederten sich wie in Experiment 1 in die drei Bereiche Performanz sowie objektive und subjektive Beanspruchung. Die einzelnen Variablen sind in Tabelle 3.2 dargestellt.

3.1.6 Auswertung

Die Auswertung der Verhaltensdaten, der psychophysiologischen Indikatoren sowie der subjektiven Angaben erfolgte analog zu Experiment 1.

3.1.6.1 Teststärkeüberlegungen

Eine Schätzung der Effektstärke aus den Ergebnissen von Experiment 1 zeigte, dass der Unterschied zwischen den experimentellen Bedingungen sowie deren Interaktion sehr groß war ($d = 1,06$). Unter Annahme solch großer Effekte ergab eine A-priori Teststärkeanalyse, dass damit bei $\alpha = \beta = .05$ für das zweite Experiment insgesamt mindestens 50 Untersuchungsteilnehmer benötigt wurden.

Tabelle 3.2

 Abhängige Variablen getrennt für die Bereiche Performanz, objektive und subjektive Beanspruchung

 Performanz

Absolute Anzahl entdeckter Fehler pro Text

Anzahl gelesener Wörter pro Text

Genauigkeit (Anzahl entdeckter Fehler relativ zur Anzahl tatsächlich vorhandener Fehler im gelesenen Abschnitt)

Anzahl falscher Alarme und falsch markierter Wörter

Objektive Beanspruchung

Herzrate (bpm)

Herzratenvariabilität:

Gesamtleistung (ms^2)Leistung im mittleren Frequenzband (0,07-0,14 Hz) in absoluten (ms^2) und normalisierten Einheiten (NU)Leistung im oberen Frequenzband (0,15-0,4 Hz) in absoluten (ms^2) und normalisierten Einheiten (NU)

Prozentsatz der Intervalle mit mindestens 50ms Differenz zum vorausgehenden Intervall (PNN50)

Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD)

Tonische Hautleitfähigkeit (Microsiemens)

Subjektive Beanspruchung

Physische Beschwerden

Befindlichkeit

3.2 Ergebnisse

Der Ergebnisteil gliedert sich in die Bereiche Verhaltensdaten, objektive Beanspruchung (physiologische Messungen) und subjektive Beanspruchung.

3.2.1 Verhaltensdaten

Im Folgenden wird die Leistung der Probanden in den beiden Experimentalbedingungen dargestellt. Dazu zählt die absolute Anzahl entdeckter Fehler, die Genauigkeit bei der Fehlerentdeckung sowie die Anzahl gelesener Wörter.

3.2.1.1 Absolute Anzahl entdeckter Fehler

Abbildung 3.1 zeigt als Hauptperformanzmaß die Mittelwerte und Standardfehler der absoluten Anzahl entdeckter Fehler pro Textseite. Es ist eindeutig zu erkennen, dass die mittlere Anzahl absolut entdeckter Fehler in der Rot/Grün-Gruppe ($M = 7,86$) geringer war als in der Schwarz/Weiß-Bedingung ($M = 13,25$). Das Ergebnis einer 2x15 MANOVA mit der Gruppenvariable Farbkombination (Rot/Grün und Schwarz/Weiß) sowie der messwiederholten Variable Zeitverlauf (15 aufeinanderfolgende Texte) zeigte einen signifikanten Effekt der Farbkombination ($F(1,48) = 27,61, p < .001, R^2 = .37$).

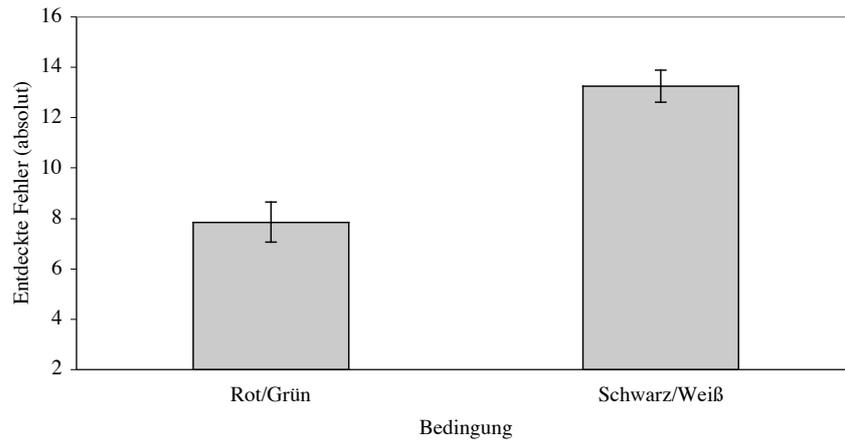


Abbildung 3.1: Durchschnittliche Anzahl richtig entdeckter Fehler pro Textseite. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

In Abbildung 3.2 ist der Verlauf der absoluten Anzahl richtig entdeckter Fehler über die 15 Texte dargestellt. Es ist zu erkennen, dass die Anzahl der entdeckten Fehler pro Text relativ konstant war und die Leistung gegen Ende der Aufgabenbearbeitung nicht abnahm. Es zeigte sich kein Effekt des Zeitverlaufs ($F(14,35) = 1,89$) und es konnte auch keine Interaktion zwischen der Farbkombination und dem Zeitverlauf festgestellt werden ($F(14,35) = 1,94$).

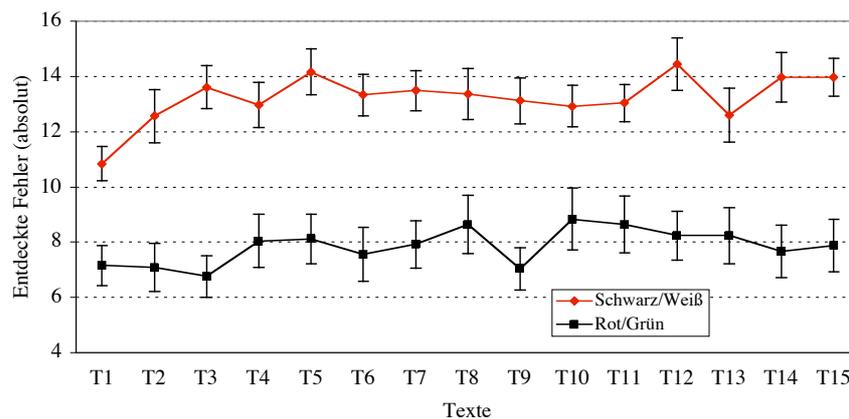


Abbildung 3.2: Die Abbildung zeigt den Verlauf der Anzahl richtig entdeckter Fehler über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.1.2 Anzahl gelesener Wörter

Die Tatsache, dass in der Rot/Grün-Bedingung deutlich weniger Fehler entdeckt wurden als in der Schwarz/Weiß-Gruppe kann mehrere Gründe haben. Zum einen war es möglich, dass bei chromatischer Darstellung weniger Wörter gelesen wurden. Zum anderen konnte auch eine geringere Genauigkeit die Ursache für dieses Ergebnis sein.

Abbildung 3.3 zeigt die Mittelwerte und Standardfehler der Anzahl gelesener Wörter pro Textseite in den beiden Gruppen. Die mittlere Anzahl gelesener Wörter in der Rot/Grün-Gruppe ($M = 484$) war eindeutig geringer als in der achromatischen Bedingung ($M = 620$). Das Ergebnis einer 2[15 MANOVA zeigte einen signifikanten Effekt der Farbkombination ($F(1,48) = 10,9, p = .002, R^2 = .14$).

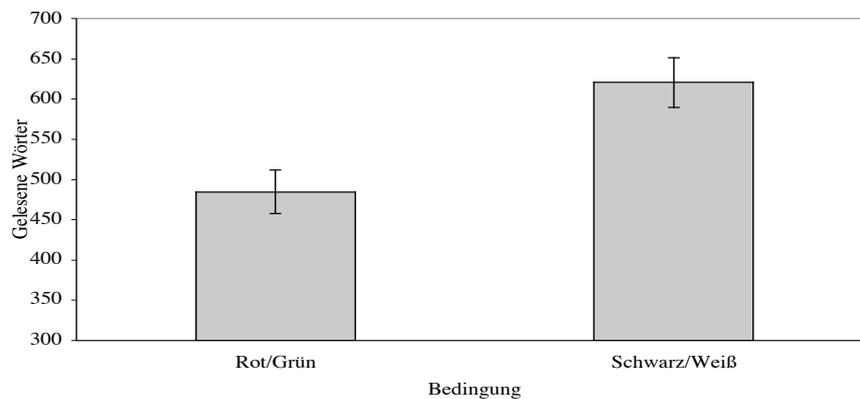


Abbildung 3.3: Illustration der durchschnittlichen Anzahl gelesener Wörter pro Textseite. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Abbildung 3.4 zeigt den Verlauf der Anzahl gelesener Wörter über die 15 Texte. Es ist zu erkennen, dass die Anzahl der gelesenen Wörter pro Text relativ konstant war und die Leistung gegen Ende der Aufgabenbearbeitung nicht abnahm. Es zeigte sich allerdings ein signifikanter Effekt des Verlaufs. Mittels orthogonaler Kontraste konnte gezeigt werden, dass dieser Effekt durch den Anstieg der Leserate von Text 1 zu den übrigen Texten bedingt war, da sich bezüglich der Anzahl gelesener Wörter nur der erste Text von den restlichen Texten unterschied ($F(1,48) = 39,39, p < .001, R^2 = .45$). Zwischen der Farbkombination und dem Zeitverlauf lag keine Interaktion vor ($F < 1$).

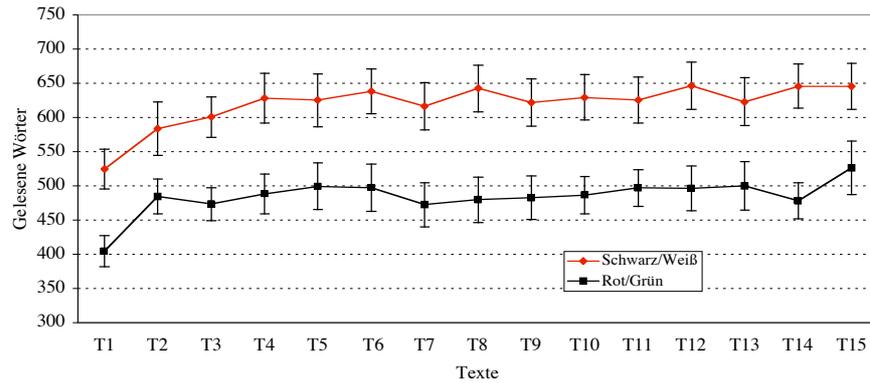


Abbildung 3.4: Die Abbildung zeigt den Verlauf der Anzahl gelesener Wörter über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.1.3 Relative Anzahl entdeckter Fehler (Genauigkeit)

Abbildung 3.5 zeigt die Fehlerentdeckungs-Genauigkeit gemittelt über alle 15 Texte sowie die Standardfehler der Mittelwerte. Es ist deutlich erkennbar, dass die Genauigkeit in der Rot/Grün-Gruppe ($M = 46\%$) geringer war als in der Schwarz/Weiß-Bedingung ($M = 63\%$). Das Ergebnis einer 2×15 MANOVA zeigte einen signifikanten Effekt der Farbkombination ($F(1,48) = 16,83, p < .001, R^2 = .26$).

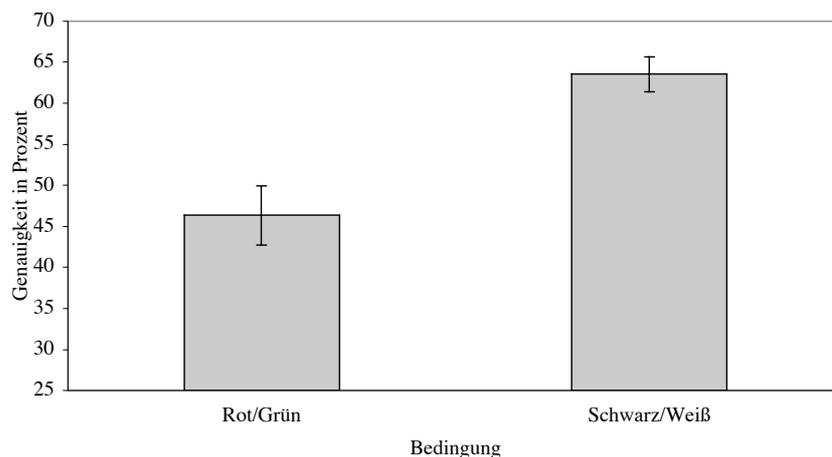


Abbildung 3.5: Durchschnittliche Genauigkeit bei der Fehlerentdeckung. Die Anzahl entdeckter Fehler wird hierbei in Relation gesetzt zur Anzahl tatsächlich vorhandener Fehler im gelesenen Abschnitt. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Abbildung 3.6 zeigt den Verlauf der Fehlerentdeckungs-Genauigkeit über die 15 Texte. Es ist zu erkennen, dass der Verlauf relativ konstant war und die Genauigkeit gegen Ende der Aufgabenbearbeitung nicht abnahm. Es zeigte sich kein Effekt des Zeitverlaufs ($F(14,35) = 1,85$) und auch keine Interaktion zwischen den Faktoren Farbkombination und Zeitverlauf ($F(14,35) = 1,49$).

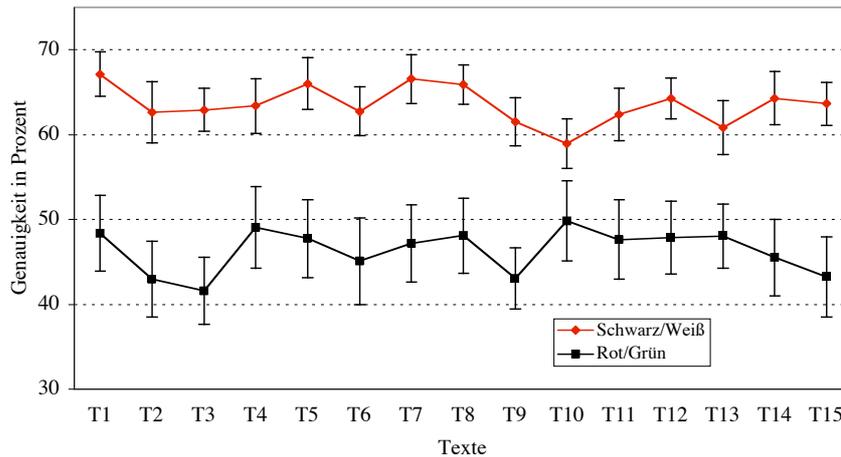


Abbildung 3.6.: Verlauf der Fehlerentdeckungs-Genauigkeit über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.1.4 Fehlerarten

In Abbildung 3.7 wird die Genauigkeit bei der Fehlerentdeckung getrennt nach den fünf Fehlerarten dargestellt. Man kann erkennen, dass in der Schwarz/Weiß-Gruppe Grammatikfehler weniger genau erkannt wurden als die vier anderen Fehlerarten. Mithilfe orthogonaler Kontraste, die die Genauigkeit für Grammatikfehler mit dem Mittelwert der vier anderen Fehlerarten verglichen, konnte gezeigt werden, dass dieser Unterschied signifikant war ($F(1,24) = 29,25$, $p < .001$, $R^2 = .55$). Für die chromatische Bedingung ergab sich kein signifikanter Unterschied zwischen Grammatikfehlern und den restlichen Fehlerarten ($F(1,24) = 2,36$). Klammert man die Grammatikfehler von der weiteren Betrachtung aus, zeigte sich, dass auch die anderen vier Fehlerarten mit unterschiedlicher Genauigkeit erkannt wurden. Die Fehlerart „falscher Buchstabe“ wurde mit höherer Genauigkeit erkannt als die Fehlerarten „Buchstabe zuviel“, „Buchstabe zuwenig“ und „Buchstabe verdreht“.

Mithilfe orthogonaler Kontraste konnte dies sowohl für die Rot/Grün-Bedingung ($F(1,24) = 62,79, p < .001, R^2 = .72$) als auch die Schwarz/Weiß-Gruppe ($F(1,24) = 79,75, p < .001, R^2 = .77$) gezeigt werden.

Die Fehlerart „Buchstabe zuwenig“ wurde mit geringster Genauigkeit entdeckt. Dies konnte sowohl für die chromatische ($F(1,24) = 31,4, p < .001, R^2 = .57$) als auch die achromatische Bedingung ($F(1,24) = 62,47, p < .001, R^2 = .72$) nachgewiesen werden.

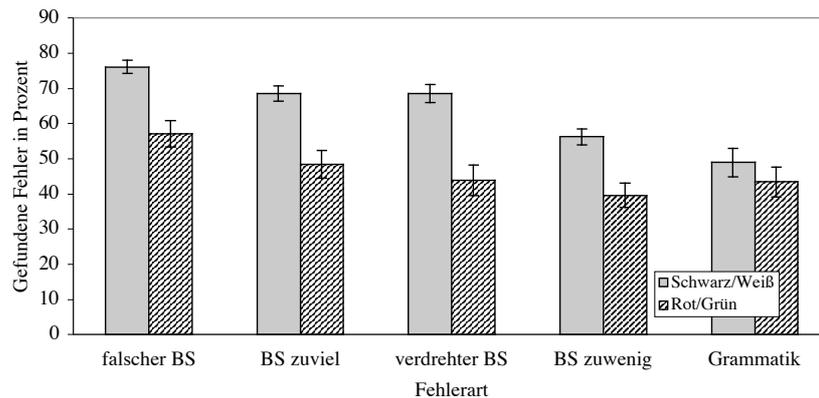


Abbildung 3.7: Die Abbildung zeigt die Fehlerentdeckungs-Genauigkeit getrennt nach Fehlerarten. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.1.5 Genauigkeit bei der Entdeckung von Grammatikfehlern

Abbildung 3.8 zeigt den Verlauf der Entdeckungs-Genauigkeit für Grammatikfehler über die 15 Texte. Man kann erkennen, dass für beide Gruppen die Kurve leicht schwankend verlief, allerdings die Genauigkeit gegen Ende der Aufgabenbearbeitung nicht abnahm. Es zeigte sich kein Effekt des Zeitverlaufs ($F(14,35) = 1,79$) und auch keine Interaktion zwischen den Faktoren Zeitverlauf und Farbkombination ($F(14,35) = 1,34$).

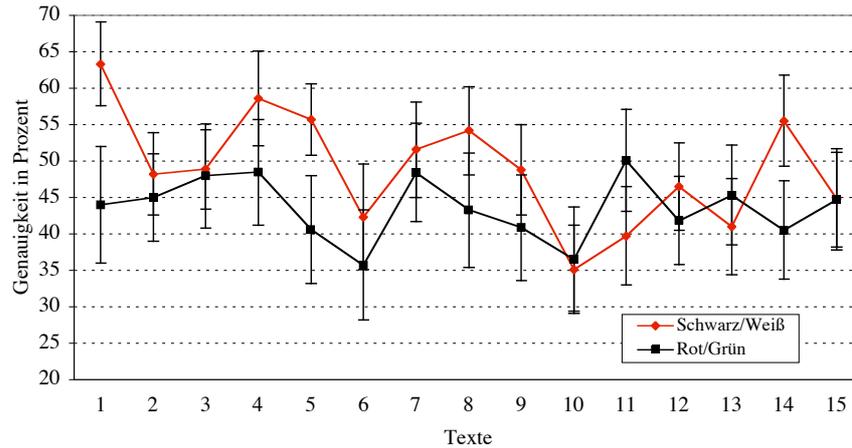


Abbildung 3.8: Dargestellt ist der Verlauf der Entdeckungs-Genauigkeit für Grammatikfehler. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.1.6 Falsche Alarmer und „Nicht übernehmen“

Tabelle 3.3

Durchschnittliche Zahl falscher Alarmer und nicht übernommener Wörter pro Seite.		
	Schwarz/Weiß	Rot/Grün
„Nicht übernehmen“	0,35 (0,07)	0,22 (0,03)
Falsche Alarmer	0,48 (0,06)	0,59 (0,31)

In Tabelle 3.3 ist die durchschnittliche Zahl falscher Alarmer und nicht übernommener Wörter pro Textseite für die beiden Bedingungen dargestellt. Die Wahl der Farbkombination hatte weder einen Effekt auf die falsche Alarmrate ($F < 1$), noch auf die Anzahl nicht übernommener Wörter ($F(1,48) = 3,41$).

3.2.1.7 Subjektive Angaben

Zum Abschluss des Versuchs wurden die Probanden gebeten, die Textschwierigkeit, die Textdarstellung, den empfundenen Zeitdruck sowie die eigene Sicherheit in der Rechtschreibung einzuschätzen. Die Ergebnisse sind in Tabelle 3.4 dargestellt. Eine Varianzanalyse zeigte einen signifikanten Effekt der Farbkombination auf die subjektive empfunden-

dene Textschwierigkeit ($F(1,48) = 6,12, p = .017, R^2 = .11$). Die Texte wurden von der Rot/Grün-Gruppe als sprachlich schwieriger gewertet. Auch die Textdarstellung wurde in der chromatischen Bedingung als signifikant unangenehmer eingestuft ($F(1,48) = 12,27, p = .001, R^2 = .2$). Die Wahl der Farbkombination hatte dagegen keinen Effekt auf die Einschätzung des empfundenen Zeitdrucks ($F < 1$) und der eigenen Rechtschreibsicherheit ($F(1,48) = 3,64$).

Tabelle 3.4

Subjektive Einschätzung des Textniveaus, der Textdarstellung, des empfundenen Zeitdrucks sowie Sicherheit in der Rechtschreibung. In Klammern sind die Standardabweichungen angegeben.

		Schwarz/Weiß	Rot/Grün
Textschwierigkeit	1=sehr einfach bis 7=sehr schwierig	2,28 (0,25)	3,28 (0,32)
Textdarstellung	1=sehr angenehm bis 7=sehr unangenehm	4,04 (0,36)	5,72 (0,31)
Zeitdruck	1=gar nicht bis 7=sehr stark	3,56 (0,71)	3,88 (0,31)
Rechtschreibung	1=sehr sicher 7=sehr unsicher	2,24 (0,27)	3,04 (0,32)

3.2.2 Objektive Beanspruchung

Die Darstellung der Ergebnisse für die objektive Beanspruchung erfolgt getrennt für die Bereiche Herzrate, Herzratenvariabilität, elektrodermale Aktivität sowie Atmung.

3.2.2.1 Herzrate

In Abbildung 3.9 sieht man deutlich, dass während der Aufgabenbearbeitung die mittlere Herzrate in beiden Gruppen im Vergleich zur ersten Ruhemessung anstieg und dann in der zweiten Ruhemessung wieder abfiel. Berechnet wurde eine 2×3 MANOVA mit der Gruppenvariable Farbkombination (Rot/Grün und Schwarz/Weiß) sowie der messwiederholten Variable Zeitpunkt (2 Ruhemessungen und Aufgabenbearbeitung). Mithilfe orthogonaler Kontraste wurde die Herzrate während Aufgabenbearbeitung mit dem Mittelwert aus den

beiden Ruhemessungen verglichen. Es zeigte sich ein signifikanter Unterschied zwischen Aufgabenbearbeitung und Ruhemessung ($F(1,48) = 43, p < .001, R^2 = .47$), allerdings konnte keine Interaktion zwischen der Farbkombination und dem Zeitpunkt festgestellt werden ($F < 1$). Darüber hinaus zeigte sich kein Unterschied zwischen der mittleren Herzrate in Ruhemessung 1 und 2 ($F < 1$), sowie auch keine Interaktion zwischen Farbkombination und Zeitpunkt ($F < 1$).

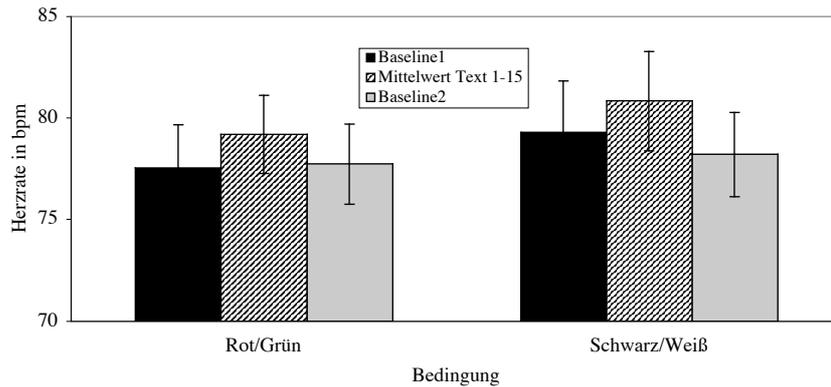


Abbildung 3.9: Mittlere Herzrate während Ruhe (Baseline1 und 2) sowie Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Schaut man sich in Abbildung 3.10 die Herzrate im Verlauf über die 15 Texte als Differenzwerte zur ersten Ruhemessung an, kann man erkennen, dass die Herzrate in beiden Gruppen mit fortschreitender Versuchsdauer immer weiter abfiel. Es zeigte sich allerdings kein Effekt des Verlaufs ($F(14,35) = 1,12$). Auch konnte keine Interaktion der beiden Faktoren Zeitpunkt und Farbkombination festgestellt werden ($F(14,35) = 1,12$).

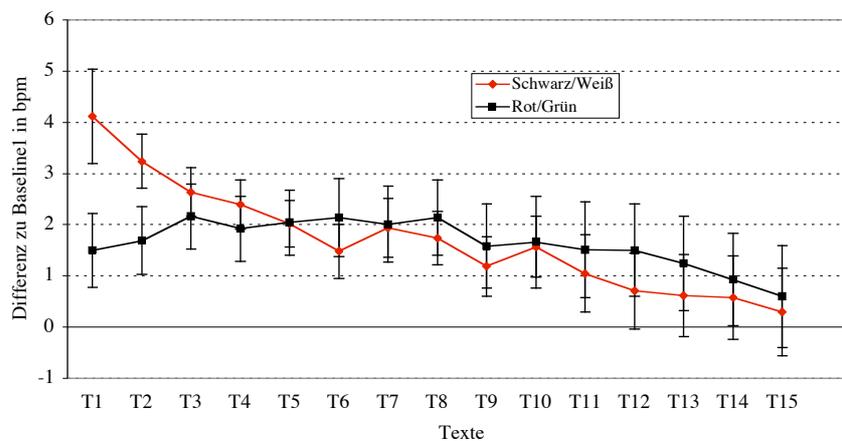


Abbildung 3.10: Verlauf der mittleren Herzrate über die 15 Texte. Dargestellt sind die Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.2.2 Herzratenvariabilität

Dargestellt werden die Ergebnisse der spektralanalytischen und zeitbezogenen Auswertung der Herzratenvariabilität. Zu den spektralanalytischen Parametern zählen die Gesamtleistung sowie die Leistung im mittleren und oberen Frequenzband. Als zeitbezogene Streuungsmaße werden der Prozentsatz der Intervalle mit mindestens 50ms Differenz zum vorausgehenden Intervall (PNN50) sowie die Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD) erläutert.

Gesamtleistung: Abbildung 3.11 zeigt, dass die Gesamtleistung des Leistungsspektrums erwartungsgemäß in beiden Gruppen während der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung abfiel und dann in der zweiten Ruhemessung wieder anstieg. Mittels orthogonaler Kontraste konnte nachgewiesen werden, dass sich die Gesamtleistung während der Aufgabenbearbeitung signifikant vom Mittelwert der beiden Ruhemessungen unterschied ($F(1,48) = 14,11, p < .001, R^2 = .23$). Es zeigte sich keine Interaktion zwischen Farbkombination und Zeitpunkt ($F < 1$). Auch konnte kein Unterschied zwischen der Leistung in Ruhemessung 1 und 2 ($F(1,48) = 2,97$) sowie keine Interaktion zwischen den beiden Faktoren festgestellt werden.

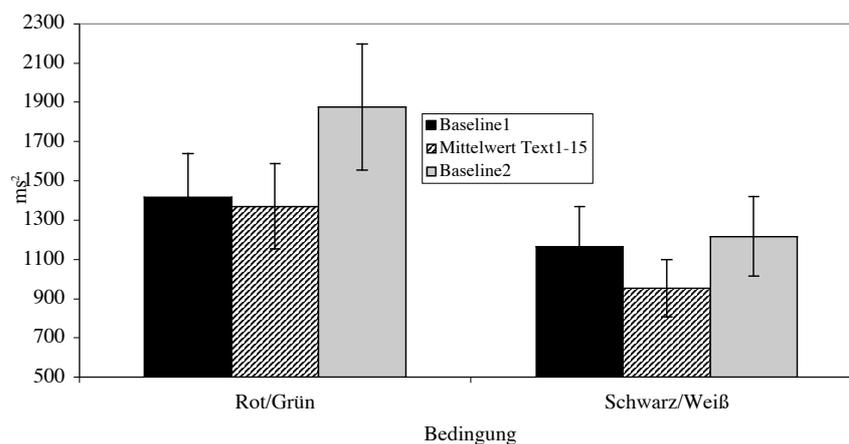


Abbildung 3.11: Mittlere Gesamtleistung aller Frequenzbänder während Ruhe (Baseline 1 und 2) sowie Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Mittleres Frequenzband: Die Leistung in den einzelnen Frequenzbändern des Power-Spektrums kann sowohl in absoluten (ms^2) als auch in normalisierten Einheiten angegeben werden.

Tabelle 3.5 macht deutlich, dass sich bei Aufgabenbearbeitung entgegen der Erwartungen kein signifikanter Abfall der Leistung in absoluten Einheiten im Vergleich zum Mittelwert der Ruhemessungen zeigte ($F(1,48) = 4,03$). Es konnte auch keine Interaktion zwischen der Farbkombination und dem Zeitpunkt festgestellt werden ($F < 1$).

Bei der Darstellung der Leistung in normalisierten Einheiten (siehe Abbildung 3.12) erkennt man, dass die Leistung im mittleren Band erwartungsgemäß in beiden Gruppen bei der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung anstieg. Da die Leistung in der zweiten Ruhemessung nicht wieder abfiel, wurde bei der Analyse nur der Anstieg von der ersten Ruhemessung zur Aufgabenbearbeitung berücksichtigt, der signifikant war ($F(1,48) = 14,96$, $p < .001$, $R^2 = .24$). Es zeigte sich allerdings keine Interaktion zwischen Farbkombination und Zeitpunkt ($F < 1$).

Betrachtet man den Verlauf der Leistung in normalisierten Einheiten als Differenzwerte zur ersten Ruhemessung über die 15 Texte in Abbildung 3.13, zeigte sich ein Effekt des Zeitverlaufs ($F(14,35) = 2,8$, $p < .007$, $R^2 = .53$), allerdings keine Interaktion zwischen Farbkombination und Zeitverlauf ($F < 1$).

Tabelle 3.5

Mittelwerte der Leistung im mittleren Band in absoluten und normalisierten Einheiten während Ruhe (Baseline 1 und 2) sowie Aufgabenbearbeitung. In Klammern sind die Standardfehler der Mittelwerte angegeben.

		Baseline 1	Aufgabe (Mittelwert Texte 1-15)	Baseline 2
Leistung (ms^2)	Schwarz/Weiß	417,51 (102,64)	348,65 (61,24)	495,37 (96,74)
	Rot/Grün	504,25 (110,6)	539,12 (111,92)	728,92 (150,78)
Leistung (normalisierte Ein- heiten)	Schwarz/Weiß	49,94 (3,67)	58,28 (3,46)	58,05 (4,28)
	Rot/Grün	52,52 (3,51)	58,45 (3,49)	61,15 (3,94)

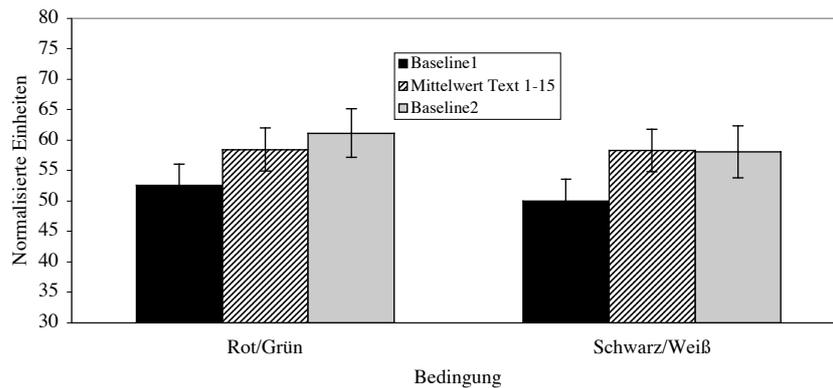


Abbildung 3.12: Durchschnittliche Leistung im mittleren Frequenzband in normalisierten Einheiten während Ruhe (Baseline 1 und 2) und Aufgabebearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

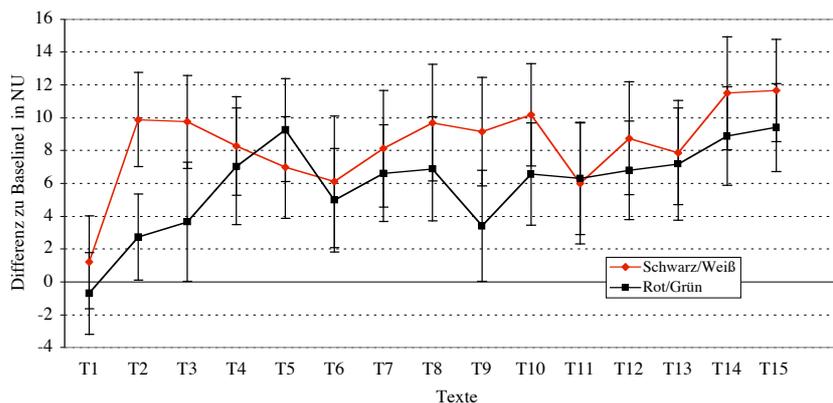


Abbildung 3.13: Dargestellt ist der Verlauf der Leistung im mittleren Frequenzband. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Oberes Frequenzband (0,15-0,4 Hz): Die Leistung im oberen Frequenzband des Power-Spektrums wird sowohl in absoluten (ms^2) als auch in normalisierten Einheiten angegeben. Ein erwartungsgemäßer Abfall der Leistung in absoluten Einheiten bei Aufgabebearbeitung sowie ein erneuter Anstieg in der zweiten Ruhemessung zeigte sich in beiden Bedingungen (siehe Tabelle 3.6). Mithilfe orthogonaler Kontraste konnte gezeigt werden, dass der Unterschied zwischen Aufgabebearbeitung und Mittelwert der Ruhemessungen signifikant war ($F(1,48) = 22,47, p < .001, R^2 = .32$), wobei aber keine Interaktion zwischen Farbkombination und Zeitpunkt vorlag ($F(1,48) = 2,52$). Es konnte mittels orthogonaler Kontraste kein Unterschied zwischen der Leistung in Ruhemessung 1 und 2 festgestellt

werden ($F < 1$). Auch zeigte sich keine Interaktion zwischen Farbkombination und Zeitpunkt ($F < 1$).

Bei der Darstellung der Leistung in normalisierten Einheiten in Abbildung 3.14 erkennt man deutlich, dass die Leistung erwartungsgemäß in beiden Gruppen bei der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung abfiel. Da die Leistung in der zweiten Ruhemessung nicht wieder anstieg, wurde bei der Analyse nur die Abnahme von der ersten Ruhemessung zur Aufgabenbearbeitung berücksichtigt, die signifikant war ($F(1,48) = 19,08$, $p < .001$, $R^2 = .28$). Es zeigte sich keine Interaktion zwischen Farbkombination und Zeitpunkt ($F < 1$).

Tabelle 3.6

Mittelwerte der Leistung im oberen Band in absoluten und normalisierten Einheiten während Ruhe (Baseline 1 und 2) sowie Aufgabenbearbeitung. In Klammern sind die Standardfehler der Mittelwerte angegeben.

		Baseline 1	Aufgabe (Mittelwert Texte 1-15)	Baseline 2
Leistung (ms^2)	Schwarz/Weiß	397,55 (77,91)	266,29 (51,79)	409,21 (93,36)
	Rot/Grün	362,54 (53,53)	284,85 (56,61)	343,78 (54,19)
Leistung (normalisierte Einheiten)	Schwarz/Weiß	47,19 (3,64)	37,69 (3,14)	38,94 (4,12)
	Rot/Grün	42,79 (2,99)	36,67 (2,63)	35,19 (3,24)

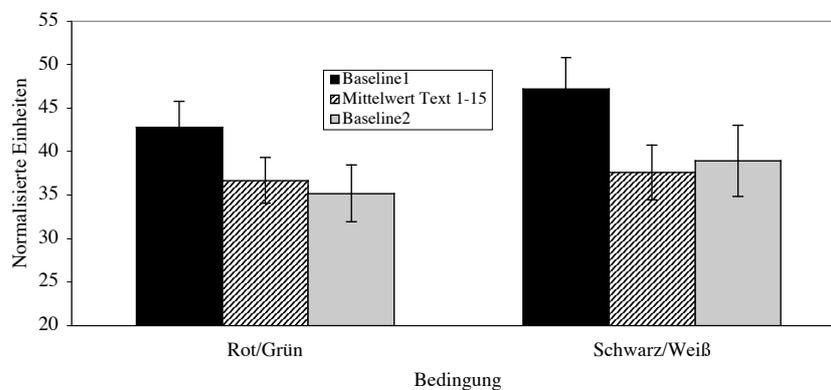


Abbildung 3.14: Durchschnittliche Leistung im oberen Frequenzband in normalisierten Einheiten während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Betrachtet man den Verlauf der Leistung in normalisierten Einheiten als Differenzwerte zur ersten Ruhemessung über die 15 Texte in Abbildung 3.15, zeigte sich kein Effekt des Zeitverlaufs ($F(14,35) = 1,46$) und auch keine Interaktion zwischen Zeitverlauf und Farbkombination ($F < 1$).

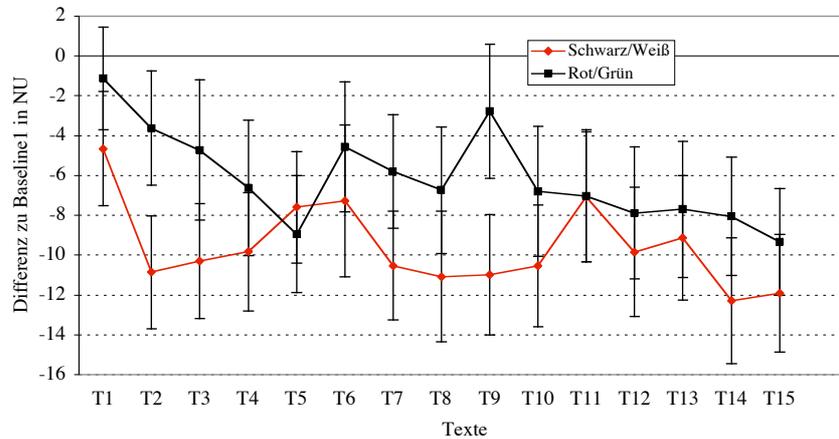


Abbildung 3.15 Dargestellt ist der Verlauf der Leistung im oberen Frequenzband. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Zeitbezogene Streuungsmaße: Abbildung 3.16 illustriert, dass der Wert der PNN50 in beiden Bedingungen wie erwartet während der Aufgabenbearbeitung abfiel und dann in der zweiten Ruhemessung wieder anstieg. Mittels orthogonaler Kontraste konnte gezeigt werden, dass der Unterschied zwischen dem Mittelwert der beiden Ruhemessungen und dem PNN50-Wert während Aufgabenbearbeitung signifikant war ($F(1,48) = 21,37$, $p < .001$, $R^2 = .31$). Es lag allerdings keine Interaktion zwischen Farbkombination und Zeitpunkt vor ($F < 1$). Darüber hinaus zeigte sich kein Unterschied zwischen den Werten der beiden Ruhemessungen ($F < 1$) und auch keine Interaktion zwischen Farbkombination und Zeitpunkt ($F < 1$).

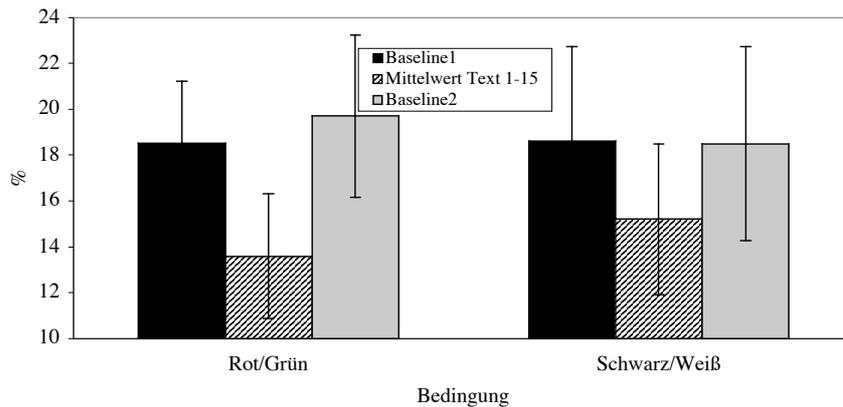


Abbildung 3.16: Mittlerer Prozentsatz der Intervalle mit mindestens 50ms Differenz zum vorausgehenden Intervall (PNN50) während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Der Verlauf der PNN50 als Differenzwert zur ersten Ruhemessung über die 15 Texte ist in Abbildung 3.17 dargestellt. Es konnte kein Effekt des Zeitverlaufs ($F(14,35) = 1,37$) und keine Interaktion zwischen Zeitverlauf und Farbkombination ($F(14,35) = 1,82$) festgestellt werden.

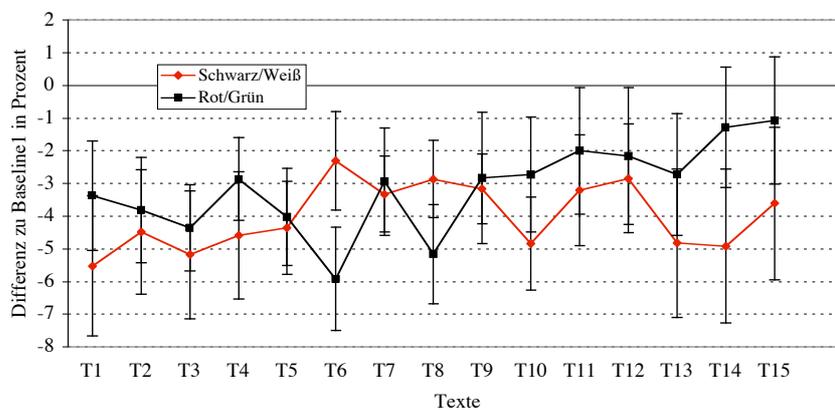


Abbildung 3.17: Illustration des PNN50-Verlaufs über die 15 Texte. Bei den Werten handelt es sich um Differenzwerte zur ersten Ruhemessung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

In Abbildung 3.18 sieht man deutlich, dass der RMSSD-Wert in beiden Gruppen wie erwartet während der Aufgabenbearbeitung abfiel und in der zweiten Ruhemessung wieder anstieg. Mittels orthogonaler Kontraste konnte nachgewiesen werden, dass der Unterschied

zwischen dem RMSSD-Wert während Aufgabenbearbeitung und dem Mittelwert der beiden Ruhemessungen signifikant war ($F(1,48) = 31,76, p < .001, R^2 = .4$). Es zeigte sich aber keine Interaktion zwischen der Farbkombination und dem Zeitpunkt ($F < 1$). Darüber hinaus konnte kein Unterschied zwischen den Werten der beiden Ruhemessungen festgestellt werden ($F < 1$).

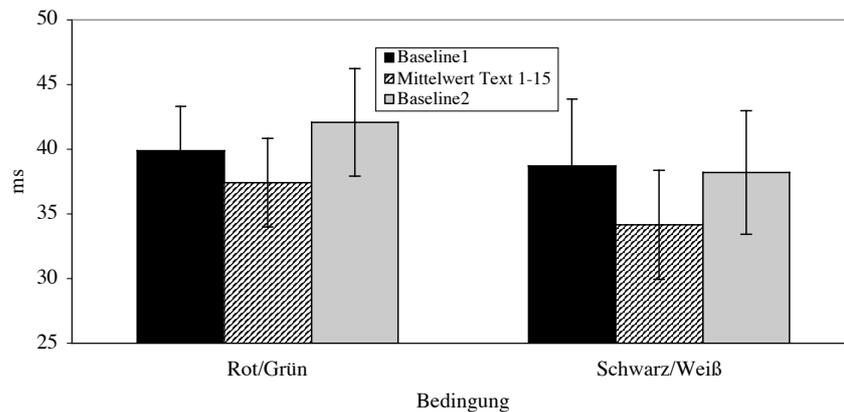


Abbildung 3.18: Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD) während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Der Verlauf der Wurzel aus dem mittleren Quadrat sukzessiver Differenzen über die 15 Texte ist in Abbildung 3.19 dargestellt. Es zeigte sich kein Effekt des Zeitverlaufs ($F(14,35) = 1,72$) und keine Interaktion zwischen Farbkombination und Zeitverlauf ($F(14,35) = 2,45$).

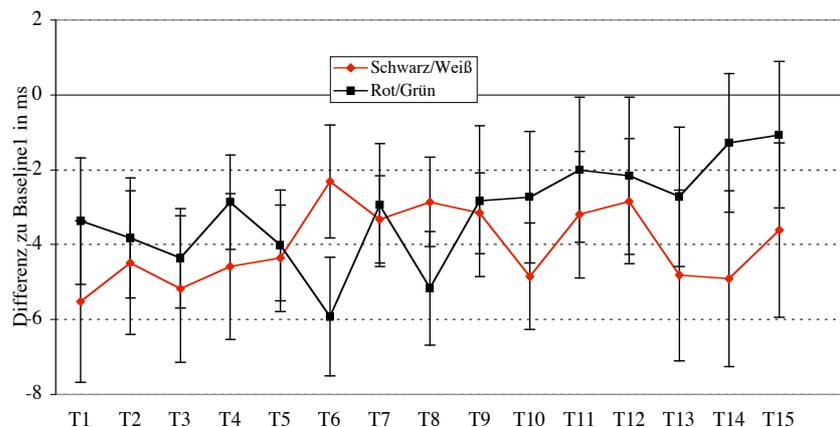


Abbildung 3.19: Illustration des RMSSD-Verlaufs über die 15 Texte. Bei den Werten handelt es sich um Differenzwerte zur ersten Ruhemessung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.2.3 Elektrodermale Aktivität

In Abbildung 3.20 ist zu erkennen, dass die mittlere tonische Hautleitfähigkeit in beiden Gruppen während der Aufgabenbearbeitung anstieg. Da die Leitfähigkeit in der zweiten Ruhemessung nicht wieder abnahm, wurde für die statistische Analyse nur der Unterschied zwischen der ersten Ruhemessung und dem Wert während Aufgabenbearbeitung berücksichtigt. Es zeigte sich, dass der Anstieg signifikant war ($F(1,48) = 53,58, p < .001, R^2 = .53$). Diese Zunahme war in der Rot/Grün- Gruppe deskriptiv größer, allerdings zeigte sich keine Interaktion zwischen Zeitpunkt und Farbkombination ($F < 1$).

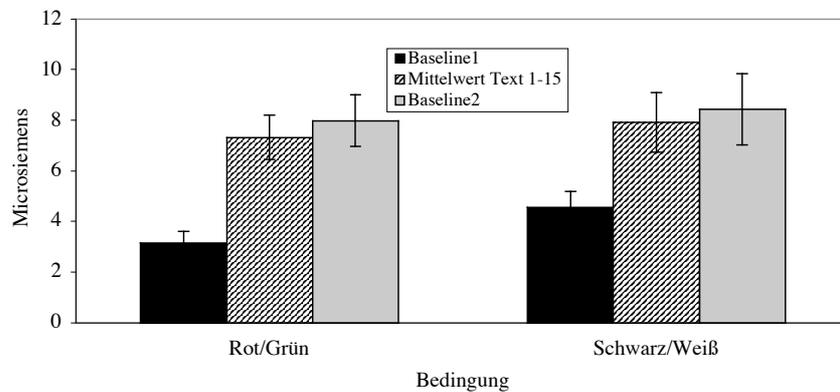


Abbildung 3.20: Tonische Hautleitfähigkeit während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die dargestellten Werte stammen aus einem variablen Zeitfenster von etwa zehn Sekunden gegen Ende der dreiminütigen Textdarbietungsdauer. Bei den Werten handelt es sich nicht um Absolutwerte, sondern um Differenzwerte zum individuellen Nullwert jeder Person, der vor dem eigentlichen Versuch bestimmt wurde. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

In Abbildung 3.21 ist der Verlauf der tonischen Leitfähigkeit über die 15 Texte dargestellt. Dabei handelt es sich um Differenzwerte zur ersten Ruhemessung. Es zeigte sich, dass die Leitfähigkeit bis etwa zur Hälfte der Texte kontinuierlich anstieg, um sich dann auf einem relativ konstanten Wert einzupendeln. Die statistische Analyse zeigte keinen Effekt des Zeitverlaufs ($F(14,35) = 1,39$) und auch keine Interaktion zwischen Zeitverlauf und Farbkombination ($F(14,35) = 1,04$).

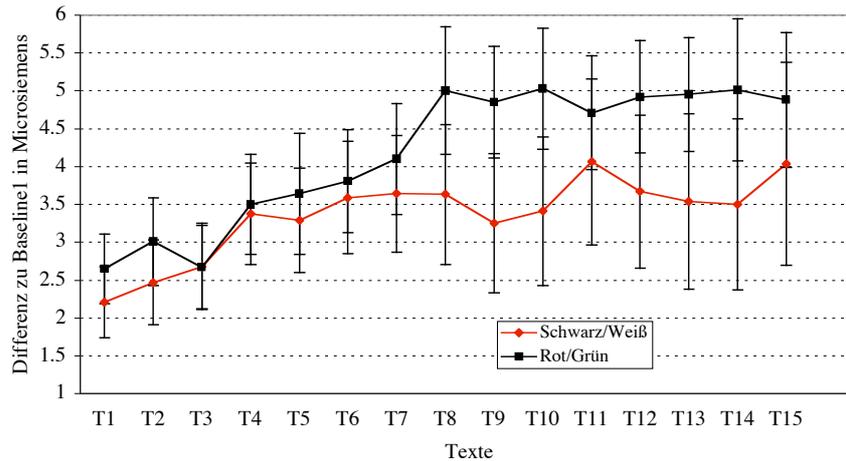


Abbildung 3.21: Verlauf der mittleren tonischen Hautleitfähigkeit über die 15 Texte. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.2.4 Atmung

Abbildung 3.22 illustriert die mittlere Atemfrequenz sowie die Standardfehler während Ruhe und Aufgabenbearbeitung. Es ist zu erkennen, dass in beiden Gruppen bei Aufgabenbearbeitung die Atemfrequenz im Vergleich zur ersten Ruhemessung deutlich zunahm und dann in der zweiten Ruhemessung wieder abfiel. Mittels orthogonaler Kontraste konnte gezeigt werden, dass der Unterschied zwischen Aufgabendurchführung und dem Mittelwert der beiden Ruhemessungen signifikant war ($F(1,48) = 41,87, p < .001, R^2 = .47$). Es fand sich aber keine Interaktion zwischen Farbkombination und Zeitpunkt ($F < 1$). Die Atemfrequenz in den beiden Ruhemessungen unterschied sich nicht ($F(1,48) = 2,09$), wobei auch keine Interaktion zwischen Farbkombination und Zeitpunkt festgestellt werden konnte ($F < 1$).

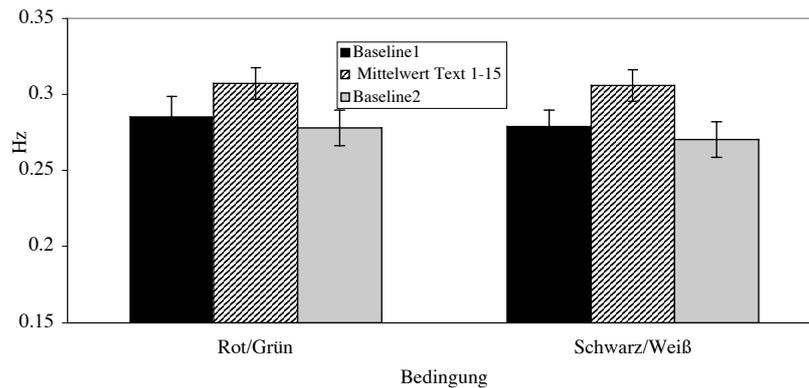


Abbildung 3.22: Mittlere Atemfrequenz während Ruhe (Baseline 1 und 2) und Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Betrachtet man den Verlauf der mittleren Atemfrequenz als Differenzwerte zur ersten Ruhemessung in Abbildung 3.23, sieht man, dass mit zunehmender Versuchsdauer die Frequenz in allen Gruppen immer weiter absank. Es zeigte sich ein signifikanter Effekt des Zeitverlaufs ($F(14,35) = 4,38, p < .001, R^2 = .64$), aber keine Interaktion zwischen Zeitverlauf und Farbkombination ($F < 1$).

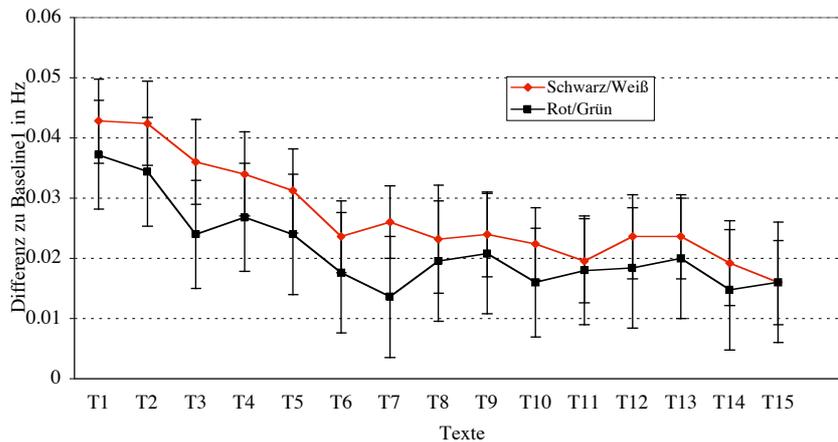


Abbildung 3.23: Verlauf der mittleren Atemfrequenz über die 15 Texte. Bei den Werten handelt es sich um Differenzwerte zu Baseline 1. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.3 Subjektive Beanspruchung

Die subjektive Beanspruchung gliedert sich in die Bereiche physische Beschwerden sowie subjektive Befindlichkeit.

3.2.3.1 Physische Beschwerden

Der durchschnittliche Anstieg der Symptomkomplexe Augenbeschwerden, Kopfschmerzsymptome und Muskel- und Haltungsbeschwerden sowie die Standardfehler sind in Abbildung 3.24 dargestellt.

Die Zunahme der Augenbeschwerden war in beiden Gruppen signifikant ($F(1,48) = 28,81$, $p < .001$, $R^2 = .38$). Man kann erkennen, dass der Anstieg der visuellen Beschwerden in der Rot/Grün-Bedingung ($M = 7,0$) größer war als in der Schwarz/Weiß-Gruppe ($M = 3,64$). Das Ergebnis einer Varianzanalyse mit Farbkombination als unabhängiger Variable zeigte allerdings, dass dieser Effekt nicht signifikant war ($F(1,48) = 2,87$). Der durchschnittliche Anstieg der Kopfschmerzsymptome in beiden Bedingungen war ebenfalls signifikant ($F(1,48) = 29,78$, $p < .001$, $R^2 = .38$), wobei sich die Gruppen nicht unterschieden ($F < 1$). Die Haltungs- und Muskelbeschwerden nahmen auch beiden Bedingungen nach Aufgabenbearbeitung signifikant zu ($F(1,48) = 33,59$, $p < .001$, $R^2 = .41$). Es lag aber kein Effekt der Farbkombination vor ($F < 1$).

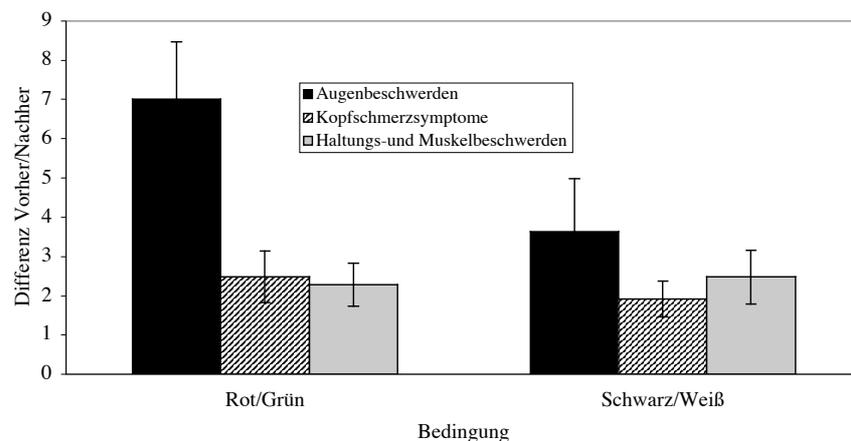


Abbildung 3.24: Durchschnittlicher Anstieg der Augenbeschwerden, Kopfschmerzsymptome sowie Haltungs- und Muskelbeschwerden nach Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.2.3.2 Befindlichkeit

Abbildung 3.25 zeigt die durchschnittliche Veränderung nach Aufgabenbearbeitung auf den Skalen „Gute/Schlechte Stimmung“, „Wachheit/Müdigkeit“ sowie „Ruhe/Unruhe“ des mehrdimensionalen Befindlichkeitsfragebogens. Man kann erkennen, dass in beiden Gruppen der Skalenwert „Gute/Schlechte Stimmung“ signifikant abnahm ($F(1,48) = 9,07, p = .004, R^2 = .16$), das heißt die Stimmung wurde schlechter. Die Abnahme war für die Rot/Grün-Gruppe deskriptiv größer, allerdings unterschieden sich die Gruppen statistisch betrachtet nicht ($F(1,48) = 1,25$).

Der Wert der Skala „Wachheit/Müdigkeit“ nahm in beiden Gruppen nach Aufgabenbearbeitung ebenfalls signifikant ab ($F(1,48) = 42,02, p < .001, R^2 = .47$). Das Müdigkeitsempfinden wurde also größer. Auch hier war der Abfall in der Rot/Grün-Bedingung deskriptiv höher. Die Gruppen unterschieden sich aber nicht signifikant ($F(1,48) = 1,02$).

Die Werte der Skala „Ruhe/Unruhe“ sanken in beiden Bedingungen in der Nachmessung ab, wobei diese Abnahme allerdings nicht signifikant war ($F(1,48) = 1,36$).

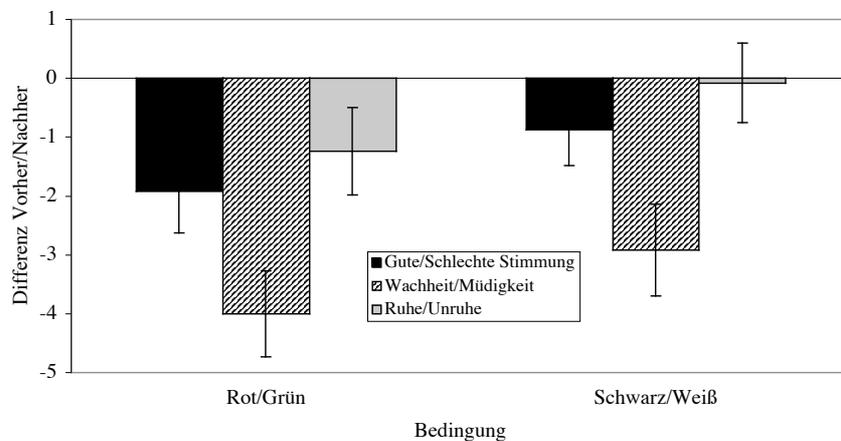


Abbildung 3.25: Mittlere Veränderung auf den Skalen „Gute/Schlechte Stimmung“, „Wachheit/Müdigkeit“ sowie „Ruhe/Unruhe“ nach Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

3.3 Interpretation und Diskussion

Die Ergebnisse werden getrennt für die Bereiche Leistung sowie objektive und subjektive Beanspruchung zusammengefasst und interpretiert.

3.3.1 Performanz

In Experiment 2 wurde bei einer niedrigen Umgebungsbeleuchtung von 5 Lux eine achromatische Textdarstellung, die positive Kontrastverhältnisse aufwies, mit einer äquivalenten chromatischen Textpräsentation – rote Schrift auf grünem Hintergrund – verglichen. Man hat vermutet, dass die Probanden in der Farbbedingung eine schlechtere Leistung aufweisen.

Die Ergebnisse zeigen, dass bei der farbigen Präsentation signifikant weniger Fehler entdeckt wurden als im Fall der Schwarz/Weiß-Darstellung. In der Farbbedingung war sowohl die Anzahl der gelesenen Wörter als auch die Fehlerentdeckungs-Genauigkeit relativ zur Menge des gelesenen Textes signifikant geringer als bei der Schwarz/Weiß-Darstellung. Darüber hinaus konnte festgestellt werden, dass in beiden Bedingungen die Leistung der Probanden über die gesamte Lesedauer von 45 Minuten konstant verlief. In Hinblick auf Genauigkeit, Lesetempo und die absolute Anzahl gefundener Fehler zeigten sich zwischen den Gruppen keine Unterschiede im Verlauf. Wie schon im ersten Experiment konnte kein Nachlassen der Leistung gegen Ende der Aufgabenbearbeitung festgestellt werden. Man kann also sagen, dass die Performanz in der Farbbedingung konsistent während des gesamten Versuchs schlechter war. Es gab auch keine Hinweise darauf, dass sich die beiden Gruppen bezüglich ihrer Rechtschreibkenntnisse unterschieden. Allerdings wurde von der Rot/Grün-Gruppe das sprachliche Textniveau als etwas schwieriger eingestuft, was aber durchaus eine Folge der schlechten Lesbarkeit sein konnte. Als mögliche Erklärung für den großen Leistungsunterschied zwischen den zwei Gruppen reicht diese Abweichung nicht aus.

Auch im zweiten Experiment wurde eine Analyse bezüglich der fünf verschiedenen Fehlerarten durchgeführt. Es zeigte sich, dass in der achromatischen Bedingung signifikant

weniger Grammatikfehler erkannt wurden in Relation zu den vier anderen Fehlertypen, was für die Farbbedingung allerdings nicht galt. Daraus kann man schließen, dass in den beiden Gruppen eine unterschiedliche Lesestrategie zur Anwendung kam. Die Schwarz/Weiß-Gruppe scheint die Texte während der gesamten Versuchsdauer flüchtiger gelesen zu haben. In beiden Gruppen nahm der Prozentsatz richtig erkannter Grammatikfehler gegen Ende des Versuchs allerdings nicht ab, was auf eine unveränderte Lesestrategie während der gesamten Aufgabenbearbeitung hindeutet.

Klammert man die Grammatikfehler von der weiteren Betrachtung aus, zeigte sich, dass auch die anderen vier Fehlerarten mit unterschiedlicher Genauigkeit erkannt wurden. Es lag genau das gleiche Bild wie in Experiment 1 vor. Wörter, die einen Buchstaben zuviel enthielten, wurden mit größter Genauigkeit erkannt, während Wörter, in denen ein Buchstabe fehlte, am wenigsten häufig entdeckt wurden.

Zusammenfassend kann man feststellen, dass in der chromatischen Bedingung die Performanz der Probanden deutlich schlechter war als in der Schwarz/Weiß-Darstellung. Dieses Ergebnis ist zum einen auf die Wahl der Farbkombination zurückzuführen. Bisherige Studien konnten zeigen, dass rote Schrift auf grünem Hintergrund von den Versuchsteilnehmern subjektiv als sehr unangenehm eingestuft wurde (Matthews et al., 1989; Pastoor, 1990; Shieh & Chen, 1997; Shieh et al., 1997). Auch in der vorliegenden Untersuchung zeigte sich, dass die Rot/Grün-Bedingung als signifikant unangenehmer bewertet wurde. Hauptsächlich kann man allerdings den nicht vorhandenen Luminanzkontrast in der Farbbedingung für das vorliegende Ergebnis verantwortlich machen (Legge & Rubin, 1986; Travis et al., 1990). Da sich die beiden Farben nicht in der Helligkeit unterschieden, ermöglichte es der vorliegende Chromatizitätskontrast, überhaupt etwas lesen zu können.

Experiment 2 wurde durchgeführt, um die Leistung der Versuchsteilnehmer aus Experiment 1 besser einstufen zu können. Es stellte sich die Frage, ob mit der schlechten Leistung unter negativen Kontrastverhältnissen bei Dunkelheit bereits die untere Grenze der Leistungsfähigkeit erreicht war oder die Möglichkeit bestand, durch Wahl einer sehr schlecht lesbaren Kombination von Text- und Hintergrundfarbe die Leistung der Probanden nochmals zu verschlechtern. In Tabelle 3.7 sind die absolute Anzahl entdeckter Fehler sowie die Fehlerentdeckungs-Genauigkeit in der chromatischen Bedingung aus Experiment 2 und der Negativ Dunkel-Bedingung aus Experiment 1 gegenübergestellt. Man sieht deskriptiv, dass die Leistung in der Negativ Dunkel-Gruppe nur geringfügig besser war als in

der sehr schlecht lesbaren Farbkombination aus Experiment 2. Dieses Ergebnis lässt einen Rückschluss darauf zu, wie ungeeignet eine Textpräsentation in negativen Kontrastverhältnissen bei Dunkelheit ist.

Tabelle 3.7

Gegenüberstellung der absoluten Anzahl entdeckter Fehler und der Fehlerentdeckungs-Genauigkeit. In Klammern sind die Standardfehler der Mittelwerte angegeben.

	Negativ Dunkel-Bedingung (Experiment 1)	Chromatische Bedingung (Experiment 2)
absolute Anzahl entdeckter Fehler	8,55 (0,63)	7,86 (0,78)
Fehlerentdeckungs-Genauigkeit (%)	56,75 (16,27)	46,36 (3,59)

3.3.2 Objektive Beanspruchung

Wie schon in Experiment 1 wurde die Vermutung aufgestellt, dass sich die Beanspruchung unter Kontrolle der motorischen Aktivität auch objektiv nachweisen lässt. Als Beanspruchungsindikatoren wurden erneut elektrodermale und kardiovaskuläre Parameter sowie die Atmung erfasst.

Die Ergebnisse zeigen, dass die Herzrate in beiden Gruppen während der Aufgabenbearbeitung signifikant anstieg im Vergleich zur ersten Ruhemessung und dann in der zweiten Ruhemessung wieder abfiel. Auch die Parameter der Herzratenvariabilität zeigten eine deutliche Beanspruchung während der Aufgabenbearbeitung. Der Prozentsatz der Intervalle mit mindestens 50ms Abweichung vom vorausgehenden Intervall (PNN50) sowie die Wurzel aus dem mittleren Quadrat sukzessiver Differenzen (RMSSD) sanken erwartungsgemäß in beiden Gruppen während der Aufgabenbearbeitung signifikant ab und stiegen dann in der zweiten Ruhemessung wieder an. Die spektralanalytische Auswertung der Herzratenvariabilität spiegelte ebenfalls die psychophysiologische Beanspruchung während der Bearbeitung der Lese-Korrektur-Aufgabe wider. Ebenso wie in Experiment 1 zeigte sich ein signifikanter Anstieg der normalisierten Leistung im mittleren Frequenz-

band in beiden Gruppen. Analog dazu fiel die Leistung in normalisierten Einheiten im oberen Band signifikant ab.

Als weiterer Beanspruchungsindikator stieg die mittlere tonische Hautleitfähigkeit in beiden Gruppen während der Aufgabenbearbeitung im Vergleich zur ersten Ruhemessung signifikant an. Diese Zunahme war in der Farbbedingung deskriptiv größer, allerdings statistisch nicht signifikant. Auch in früheren Studien konnte ein Anstieg der tonischen Hautleitfähigkeit bei der Bearbeitung kognitiver Anforderungen wie beispielsweise Reaktions- und Rechenaufgaben nachgewiesen werden (Albus et al., 1987; Fichera & Andreassi, 1998; Morell, 1989; Oshima et al., 2001). Betrachtet man die tonische Hautleitfähigkeit im Verlauf über die gesamten Versuchsabschnitte, zeigte sich ein ähnliches Bild wie in Experiment 1. Die Werte stiegen bis etwa zur Hälfte der Versuchsdauer an, um sich dann auf einem relativ konstanten Niveau einzupendeln, was möglicherweise eine Art Gewöhnung an die Versuchsdurchführung widerspiegelt.

Auch die Atemfrequenz stieg in beiden Gruppen während der Lese-Korrektur-Aufgabe signifikant an und kehrte in der zweiten Ruhemessung wieder auf das Ursprungsniveau zurück. Betrachtet man den Verlauf der mittleren Atemfrequenz über die insgesamt 15 Texte, zeigte sich mit zunehmender Versuchsdauer ein signifikantes Absinken der Frequenz, was wie schon in Experiment 1 als Müdigkeit oder Monotonie interpretiert werden kann.

Zusammenfassend lässt sich sagen, dass auch in Experiment 2 alle psychophysiologischen Indikatoren die Beanspruchung während der Bearbeitung der Lese-Korrektur-Aufgabe deutlich abbildeten. Es konnten zwischen den beiden Gruppen allerdings keine Unterschiede im Ausmaß der psychophysiologischen Beanspruchung festgestellt werden. Dieses Ergebnis ist – wie auch schon in Experiment 1 – mit der Annahme vereinbar, dass sich alle Versuchsteilnehmer gleichermaßen stark anstregten und Display-abhängige Unterschiede sich infolgedessen nur auf der Performanzebene manifestierten.

3.3.3 Subjektive Beanspruchung

Auf physischer Ebene zeigten sich am Ende der Aufgabenbearbeitung erwartungsgemäß deutliche Beschwerden. Vor allem konnte ein signifikanter Anstieg visueller Probleme wie Schmerzen, Tränen, Brennen und Jucken festgestellt werden. Die Zunahme der Augenbeschwerden war in der Farbbedingung deskriptiv größer, allerdings war dieser Unterschied nicht signifikant. Darüber hinaus kam es in beiden Gruppen zu einer signifikanten Zunahme von Kopfschmerzen, Benommenheit und Schwindel. Auch die Beschwerden im Nacken-, Schulter- und Rückenbereich waren in beiden Experimentalbedingungen am Ende der Aufgabenbearbeitung signifikant größer. An dieser Stelle muss allerdings festgestellt werden, dass dieses Ergebnis wie schon in Experiment 1 zumindest zum Teil auf die eingeschränkte Sitzhaltung aufgrund der physiologischen Ableitungen zurückzuführen ist.

Neben der physischen Beanspruchung zeigten sich bei allen Probanden auch Beschwerden im Bereich der subjektiven Befindlichkeit nach Beendigung der Lese-Korrektur-Aufgabe. Die Stimmung war in beiden Gruppen nach Aufgabenbeendigung schlechter. Auch das Müdigkeitsempfinden nahm deutlich zu. Zusammenfassend kann man feststellen, dass die Aufgabenbearbeitung bei allen Probanden zu deutlichen physischen und psychischen Beschwerden führte, wobei sich die Gruppen nicht signifikant im Ausmaß der Beschwerden unterschieden. Auch dieses Ergebnis deutet darauf hin, dass sich alle Probanden gleich stark anstrebten.

Die vorliegenden Daten zeigen, dass in der Farbbedingung eine deutlich schlechtere Leistung gezeigt wurde als im Fall der achromatischen Textdarstellung. Vergleicht man auf deskriptiver Ebene dieses Ergebnis mit dem Resultat des ersten Experiments, stellt man fest, dass sich die Performanz bei negativem Kontrast unter Dunkelheit nur geringfügig von der Leistung in der schlecht lesbaren Rot/Grün-Darstellung aus Experiment 2 unterscheidet.

Im ersten Experiment hatte sich auch bei hoher Umgebungsbeleuchtung eine schlechtere Leistung bei einer Textdarstellung in negativen Kontrastverhältnissen im Vergleich zur positiven Polarität gezeigt. Aus diesem Grund wäre es auch interessant gewesen, Experiment 2 unter Helligkeit durchzuführen, um die schlechte Performanz bei negativem Kontrast und Beleuchtung besser einzuordnen. Dieses war aber aus technischen Gesichtspunk-

ten nicht möglich, da bei Licht aufgrund von Reflektionen die Äquilibrium in der Farbbedingung nicht sichergestellt gewesen wäre.

Im Bereich der Textpräsentation spielt neben einer achromatischen Darstellung der Einsatz von Farben eine immer größer werdende Rolle, und die moderne Informationstechnik bietet die vielfältigsten Möglichkeiten, Publikationen farbig zu gestalten. Die Handhabung von Bildbearbeitungs- und Grafiksoftware wird immer einfacher, so dass diese Programme schon lange nicht mehr ausschließlich von professionellen Designern genutzt werden. Vor allem bei der Gestaltung von Web-Seiten findet man zahlreiche Schriftzeichen-Hintergrund-Farbkombinationen. Die Verwendung von Farben im Bereich der Bildschirmarbeit muss allerdings kritisch betrachtet werden, da sie mit erhöhten Anforderungen an den Sehapparat verbunden ist. Darüber hinaus besteht das Problem, dass bei farbiger Textpräsentation zwar zusätzliche Informationen im Farbkanal vorliegen, allerdings der Luminanzkontrast zwischen Text und Hintergrund reduziert wird. Für das Lesen von Texten ist allerdings ein ausreichender Helligkeitskontrast wichtig, da sich sonst die Lesbarkeit rapide verschlechtert. Ausgehend von den Resultaten der ersten beiden Experimente stellt sich an diesem Punkt nun die Frage, ob der Polaritätseffekt – das heißt ein Vorteil auf Performanzebene für eine Textdarstellung in positiver Polarität – ausschließlich für die typische Schwarz/Weiß-Darstellung gilt oder ob er sich auch auf eine chromatische Textpräsentation verallgemeinern lässt. Für das Präsentieren von Informationen – mittels Projektor oder Beamer – stellt beispielsweise gelbe Schrift auf blauem Hintergrund eine beliebte Farbkombination dar und wird sogar häufig einer achromatischen Darstellung vorgezogen. Es wird die Vermutung aufgestellt, dass sich auch in diesem Fall unabhängig von der Umgebungsbeleuchtung eine bessere Performanz bei einer Textpräsentation in positiven Kontrastverhältnissen zeigt.

Sowohl in Experiment 1 als auch in Experiment 2 konnte bereits mittels einer mehrdimensionalen psychophysiologischen Beanspruchungsanalyse eindeutig eine Beanspruchung bei Bildschirmarbeit auf objektiver Ebene nachgewiesen werden, wobei sich das Ausmaß der Beanspruchung aber nicht zwischen den Gruppen unterschied. Aus diesem Grund soll im dritten Teil der Untersuchung auf eine Erhebung der physiologischen Indikatoren verzichtet werden.

4. Experiment 3

4.1 Methode

In den folgenden Abschnitten erfolgt eine Beschreibung des dritten Experiments. Neben der untersuchten Stichprobe werden der Versuchsaufbau, die verwendeten Stimuli, das experimentelle Design sowie der Versuchsablauf beschrieben.

4.1.1 Stichprobenbeschreibung

Am dritten Experiment nahmen insgesamt 80 Versuchspersonen (52 weiblich, 28 männlich) im Alter von 19-53 Jahren ($M = 26,79$; $SD = 6,59$) teil. Alle Teilnehmer hatten Deutsch als Muttersprache. Die Zuordnung der Probanden zu den Versuchsbedingungen erfolgte randomisiert. Alle Teilnehmer wiesen eine normale oder korrigierte Sehfähigkeit auf. Die Teilnahme am Experiment wurde mit 5 Euro entlohnt.

4.1.2 Versuchsaufbau

Der Versuchsaufbau war nahezu identisch mit den ersten beiden Experimenten. Da sich in Experiment 1 und 2 gezeigt hatte, dass es aufgrund der Kinnstütze zu erheblichen Beschwerden im Schulter- und Nackenbereich gekommen war, wurde in dieser Untersuchung darauf verzichtet. An dem Versuch konnten maximal zwei Personen gleichzeitig teilnehmen. Die Arbeitsplätze waren parallel angeordnet und durch eine Stellwand getrennt. Die Darbietung der akustischen Signale erfolgte über Kopfhörer.

4.1.2.1 Subjektive Befindlichkeit und physische Beschwerden

Es wurden die gleichen Fragebögen wie in den ersten beiden Experimenten eingesetzt.

4.1.3 Beschreibung der Stimuli

Es wurde das gleiche Stimulusmaterial wie in den ersten beiden Experimenten verwendet und die Abfolge der Geschichten fand für jede Versuchsperson in randomisierter Form statt. Es gab zwei Arten der Textdarstellung. Der Text wurde sowohl bei einer hohen Raumbeleuchtung von 550 Lux als auch bei einer niedrigen Beleuchtung von 5 Lux entweder in einem positiven Kontrastverhältnis (blaue Schrift auf gelbem Hintergrund) oder in negativem Kontrast (gelbe Schrift auf blauem Hintergrund) präsentiert. In Tabelle 4.1 sind die Chromatizitäts-Koordinaten und Luminanzwerte der verwendeten Farben aufgeführt.

Tabelle 4.1

CIE Chromatizitäts-Koordinaten (x,y) und Luminanzwerte (L) der verwendeten Farben. Die Berechnung des Luminanzkontrastes erfolgte nach der Michelson Definition ($(L_{\max}-L_{\min})/(L_{\max}+L_{\min})$).

Farbe	x	y	Luminanz (cd/m ²)	Luminanzkontrast
Blau	.143	.098	22.3	0.79
Gelb	.420	.485	188	

4.1.4 Versuchsablauf

An dem Experiment konnten maximal zwei Versuchspersonen gleichzeitig teilnehmen. Da keine psychophysiologische Beanspruchungsanalyse durchgeführt wurde, verkürzte sich die gesamte Versuchsdauer auf circa 60 Minuten. Der Versuchsablauf ist in Abbildung 4.1 dargestellt. Vor Versuchsbeginn wurden die Probanden darauf hingewiesen, unbedingt den vorgegebenen Sehabstand einzuhalten.

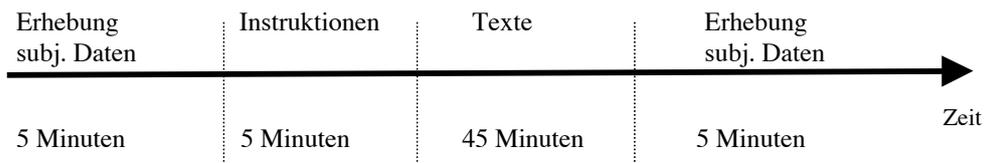


Abbildung 4.1: Schematische Darstellung des Versuchsablaufs

4.1.5 Design

Es handelte sich um ein 2×2 Design mit den Faktoren „Polarität“ und „Beleuchtung“. Der Faktor „Polarität“ lag in den Ausprägungen „Positiver Kontrast“ (= blaue Schrift auf gelbem Hintergrund) und „Negativer Kontrast“ (= gelbe Schrift auf blauem Hintergrund) vor. Der Faktor „Beleuchtung“ hatte die Abstufungen „Hell“ (künstliche Raumbeleuchtung von 550 Lux) und „Dunkel“ (5 Lux). Jeder Bedingung wurden per Zufall 20 Versuchspersonen zugeordnet.

Die abhängigen Variablen gliederten sich in die zwei Bereiche Leistung und subjektive Beanspruchung (Befindlichkeit und physische Beschwerden). Die einzelnen Variablen sind in Tabelle 4.2 dargestellt.

Tabelle 4.2

Abhängige Variablen getrennt für die Bereiche Performanz und subjektive Beanspruchung	
Performanz	Absolute Anzahl entdeckter Fehler pro Text
	Anzahl gelesener Wörter pro Text
	Genauigkeit (Anzahl entdeckter Fehler relativ zur Anzahl tatsächlich vorhandener Fehler im gelesenen Abschnitt)
	Anzahl falscher Alarme und falsch markierter Wörter
Subjektive Beanspruchung	Physische Beschwerden
	Befindlichkeit

4.1.6 Teststärkeüberlegungen

Unter praktischen Gesichtspunkten sollte der Unterschied zwischen den experimentellen Bedingungen sowie deren Interaktion mindestens „groß“ ($d = 0.8$) im Sinne der Konvention von Cohen (1977) sein. Eine A-priori Teststärkeanalyse mithilfe des Programms GPower (Erdfelder et al., 1996) ergab, dass damit bei $\alpha = \beta = .05$ insgesamt mindestens 84 Untersuchungsteilnehmer benötigt wurden. Es konnten allerdings nur 80 Probanden erhoben werden, so dass die Teststärke mit $(1 - \beta) = .942$ geringfügig unter der angestrebten Teststärke von .95 lag.

4.2 Ergebnisse

Der Ergebnisteil gliedert sich in die Bereiche Verhaltensdaten sowie subjektive Beanspruchung.

4.2.1 Verhaltensdaten

Im Folgenden wird die Leistung der Probanden in den vier Experimentalbedingungen dargestellt. Dazu zählt die absolute Anzahl entdeckter Fehler, die Genauigkeit bei der Fehlerentdeckung sowie die Anzahl gelesener Wörter.

4.2.1.1 Absolute Anzahl entdeckter Fehler

Abbildung 4.2 zeigt die Mittelwerte und Standardfehler der absoluten Anzahl entdeckter Fehler pro Textseite in den vier Gruppen. Es ist eindeutig zu erkennen, dass die mittlere Anzahl absolut entdeckter Fehler in der Negativ Hell-Gruppe ($M = 10,63$) und der Negativ Dunkel-Gruppe ($M = 9,93$) geringer war als in der Positiv Hell- ($M = 12,31$) und der Positiv Dunkel-Bedingung ($M = 11,61$). Eine $2 \times 2 \times 15$ MANOVA mit den Gruppenvariablen Beleuchtung (hell/dunkel) und Polarität (positiv/negativ) sowie der messwiederholten Variable Zeitverlauf (15 aufeinanderfolgende Texte) zeigte einen signifikanten Effekt der Polarität ($F(1,76) = 4,83, p = .031, R^2 = .06$). Der Faktor Beleuchtung hatte keinen Effekt ($F < 1$) und es gab auch keine Interaktion zwischen den beiden Variablen ($F < 1$).

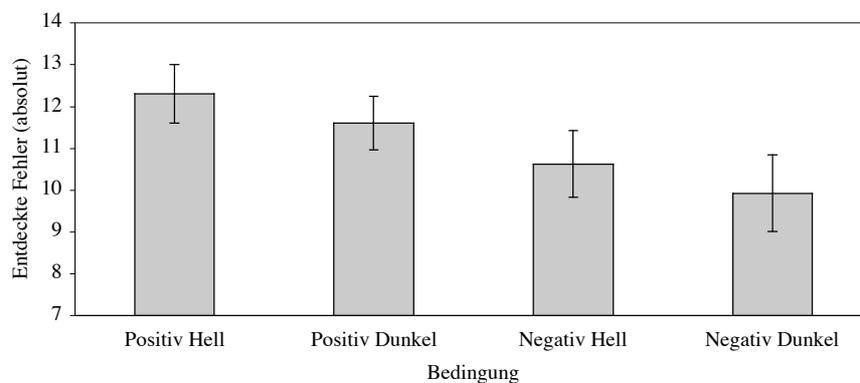


Abbildung 4.2: Durchschnittliche Anzahl richtig entdeckter Fehler pro Textseite. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Abbildung 4.3 zeigt den Verlauf der Anzahl richtig entdeckter Fehler über die 15 Texte. Es ist zu erkennen, dass die Anzahl der entdeckten Fehler pro Text relativ gleich bleibend war und die Leistung gegen Ende der Aufgabenbearbeitung nicht abnahm. Die statistische Analyse zeigte allerdings einen Effekt des Verlaufs ($F(14,63) = 2,69, p = .004, R^2 = .37$). Mithilfe orthogonaler Kontraste konnte gezeigt werden, dass dieser Effekt durch den Anstieg zu Beginn der Aufgabenbearbeitung bedingt war, da sich nur der erste ($F(1,76) = 22,82, p < .001, R^2 = .23$) und der zweite Text ($F(1,76) = 6,35, p = .014, R^2 = .08$) bezüglich der erkannten Fehler von den übrigen Texten unterschieden. Es wurde weder eine Interaktion zwischen der Polarität und dem Zeitverlauf ($F < 1$), noch zwischen der Beleuchtung und dem Zeitverlauf ($F(14,63) = 1,13$) festgestellt. Ebenso lag keine Interaktion zwischen diesen beiden Variablen und dem Faktor Zeitverlauf vor ($F(14,63) = 1,13$).

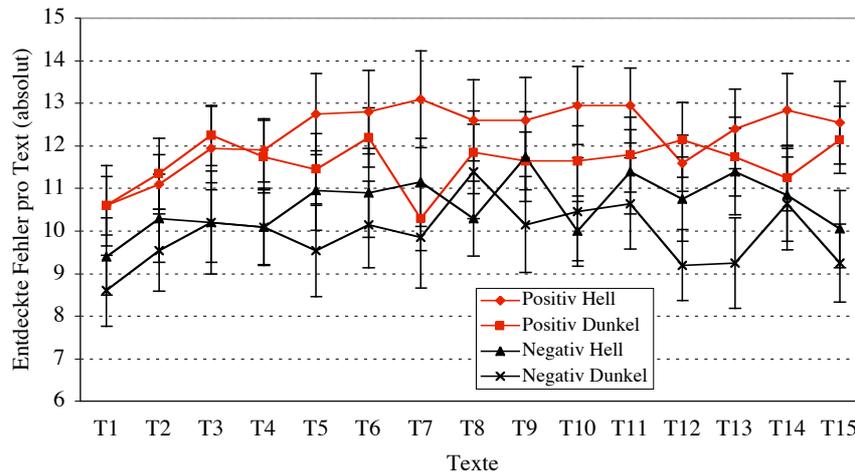


Abbildung 4.3: Die Abbildung zeigt den Verlauf der Anzahl richtig entdeckter Fehler über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

4.2.1.2 Anzahl gelesener Wörter

Die Tatsache, dass unter negativen Kontrastverhältnissen deutlich weniger Fehler entdeckt wurden als unter positiver Polarität, konnte zum einen darin begründet sein, dass bei negativer Textdarstellung weniger Wörter gelesen wurden. Zum anderen konnte eine geringere Genauigkeit bei der Fehlersuche die Ursache für dieses Ergebnis sein.

Abbildung 4.4 zeigt die Mittelwerte der Anzahl gelesener Wörter pro Textseite in den vier Gruppen. Es ist zu erkennen, dass die mittlere Anzahl gelesener Wörter sich in den vier Gruppen kaum unterschied. Die meisten Wörter wurden in der Positiv Hell-Gruppe ($M = 563$) gelesen, gefolgt von der Negativ Dunkel- ($M = 520$), der Negativ Hell- ($M = 505$) und der Positiv Dunkel-Bedingung ($M = 502$). Eine $2 \times 2 \times 15$ MANOVA zeigte entsprechend keinen signifikanten Effekt der Polarität ($F < 1$). Der Faktor Beleuchtung hatte ebenfalls keinen Effekt ($F < 1$) und es gab auch keine Interaktion zwischen den beiden Variablen ($F(1,76) = 1,27$).

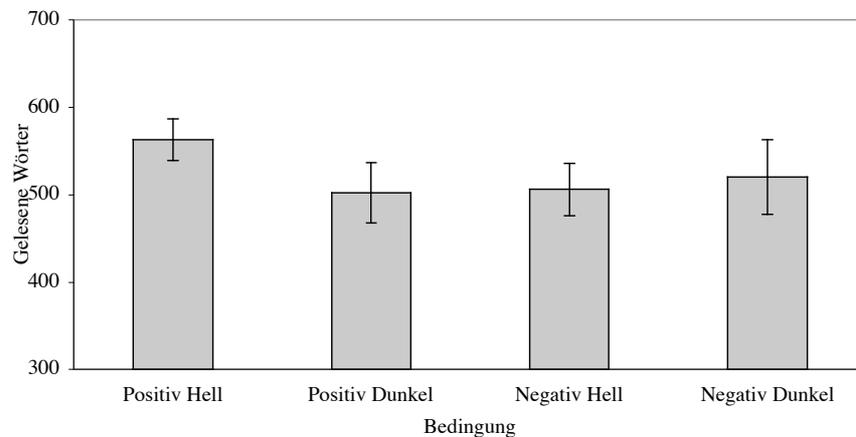


Abbildung 4.4: Illustration der durchschnittlichen Anzahl gelesener Wörter pro Textseite. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

Abbildung 4.5 zeigt den Verlauf der Anzahl gelesener Wörter über die 15 Texte. Es ist zu erkennen, dass die Anzahl der gelesenen Wörter pro Text relativ konstant war und die Leistung gegen Ende der Aufgabenbearbeitung nicht abnahm. Es zeigte sich allerdings ein Effekt des Zeitverlaufs ($F(14,63) = 5,89, p < .001, R^2 = .57$), der durch den Anstieg zu Beginn der Aufgabenbearbeitung verursacht wurde. Mittels orthogonaler Kontraste konnte gezeigt werden, dass sich nur der erste ($F(1,76) = 86,13, p < .001, R^2 = .53$) und der zweite Text ($F(1,76) = 11,07, p = .001, R^2 = .13$) bezüglich des Lesetempos von den übrigen Texten unterschieden. Es konnte weder eine Interaktion zwischen Polarität ($F < 1$), noch zwischen Beleuchtung und Zeitverlauf ($F < 1$) gefunden werden. Auch zeigte sich keine Interaktion zwischen dem Verlauf und diesen beiden Faktoren ($F(14,63) = 1,63$).

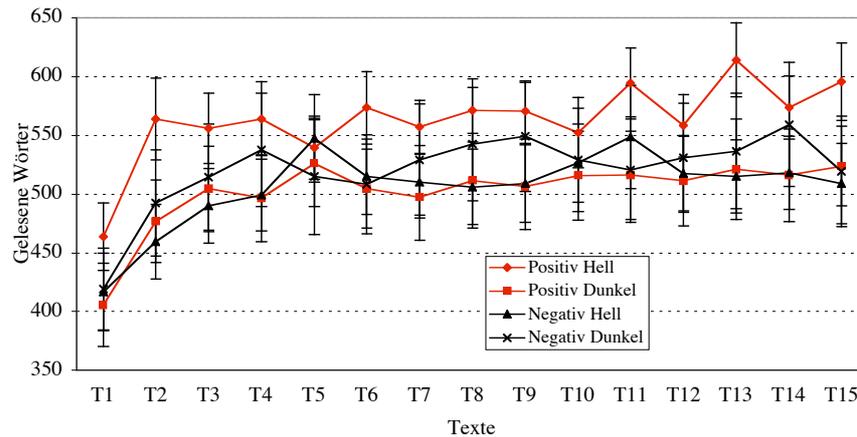


Abbildung 4.5: Die Abbildung zeigt den Verlauf der Anzahl gelesener Wörter über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

4.2.1.3 Relative Anzahl entdeckter Fehler (Genauigkeit)

Abbildung 4.6 zeigt die Fehlerentdeckungs-Genauigkeit gemittelt über alle 15 Texte in den vier Gruppen sowie die Standardfehler. Man sieht, dass sich die Genauigkeit bei positiver und negativer Kontrastwahl deutlich unterschied. Die höchste Genauigkeit bei der Fehlerentdeckung zeigte sich in der Positiv Dunkel-Gruppe ($M = 68\%$) gefolgt von der Positiv Hell- ($M = 65\%$), der Negativ Hell- ($M = 60\%$) und der Negativ Dunkel-Bedingung ($M = 55\%$). Das Ergebnis einer $2 \times 2 \times 15$ MANOVA konnte zeigen, dass der Effekt der Polarität statistisch signifikant war ($F(1,76) = 11,81, p = .001, R^2 = .13$). Der Faktor Beleuchtung hatte dagegen keinen Effekt ($F < 1$) und es gab auch keine Interaktion zwischen den beiden Variablen ($F(1,76) = 3,42$).

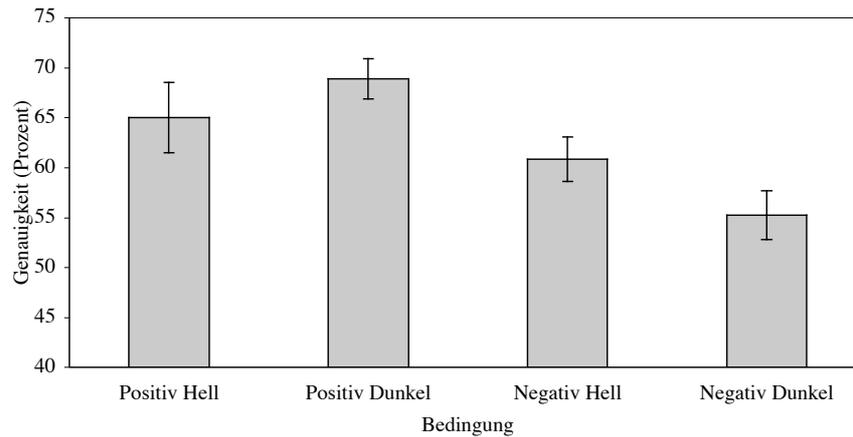


Abbildung 4.6: Durchschnittliche Genauigkeit bei der Fehlerentdeckung. Die Anzahl entdeckter Fehler wird hierbei in Relation gesetzt zur Anzahl tatsächlich vorhandener Fehler im gelesenen Abschnitt. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

In Abbildung 4.7 ist der Verlauf der Fehlerentdeckungs-Genauigkeit über die 15 Texte illustriert. Es ist zu erkennen, dass der Verlauf relativ konstant war und die Genauigkeit gegen Ende der Aufgabenbearbeitung nicht abnahm. Ein Effekt des Zeitverlaufs konnte nicht festgestellt werden ($F(14,63) = 1,26$). Es zeigte sich auch weder eine Interaktion zwischen Polarität und Zeitverlauf ($F < 1$), noch zwischen Beleuchtung und Zeitverlauf ($F(14,63) = 1,09$). Ebenso lag keine Interaktion des Zeitverlaufs mit diesen beiden Faktoren vor ($F < 1$).

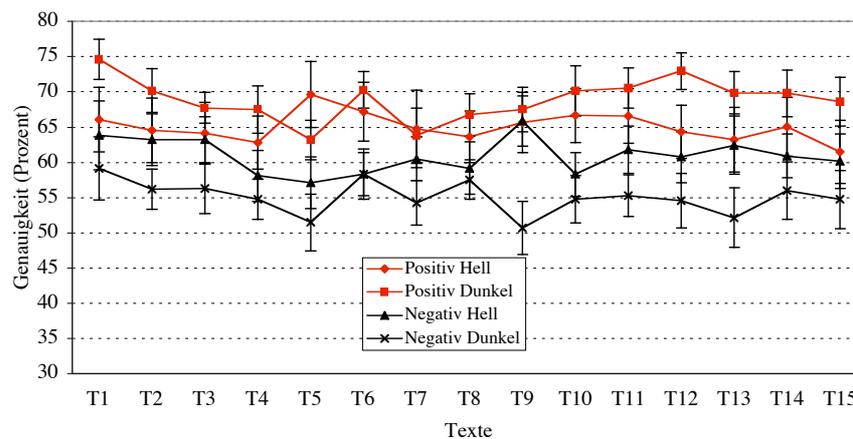


Abbildung 4.7: Verlauf der Fehlerentdeckungs-Genauigkeit über die 15 Texte. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

4.2.1.4 Fehlerarten

In Abbildung 4.8 wird die Genauigkeit bei der Fehlerentdeckung getrennt nach den fünf Fehlerarten dargestellt. Man kann erkennen, dass in allen vier Bedingungen Grammatikfehler weniger genau erkannt wurden als die vier anderen Fehlerarten. Mithilfe einer 2[2][5] MANOVA und orthogonaler Kontraste, die die Genauigkeit für Grammatikfehler mit dem Mittelwert der vier anderen Fehlerarten verglichen, konnte dies sowohl für negativen Kontrast ($F(1,38) = 39,95, p < .001, R^2 = .51$) als auch für positiven Kontrast ($F(1,38) = 35,17, p < .001, R^2 = .48$) gezeigt werden.

Klammert man die Grammatikfehler von der weiteren Betrachtung aus, so zeigte sich, dass auch die anderen vier Fehlerarten mit unterschiedlicher Genauigkeit erkannt wurden. Die Fehlerart „falscher Buchstabe“ wurde mit höherer Genauigkeit erkannt als die Fehlerarten „Buchstabe zuviel“, „Buchstabe zuwenig“ und „Buchstabe verdreht“. Dies konnte sowohl für negativen Kontrast ($F(1,38) = 52,87, p < .001, R^2 = .58$) als auch positiven Kontrast ($F(1,38) = 117,97, p < .001, R^2 = .76$) gezeigt werden. Die Fehlerart „Buchstabe zuwenig“ wurde mit geringster Genauigkeit entdeckt, was sowohl für negative ($F(1,38) = 45,45, p < .001, R^2 = .54$) als auch positive Kontrastwahl ($F(1,38) = 154,22, p < .001, R^2 = .8$) nachgewiesen wurde.

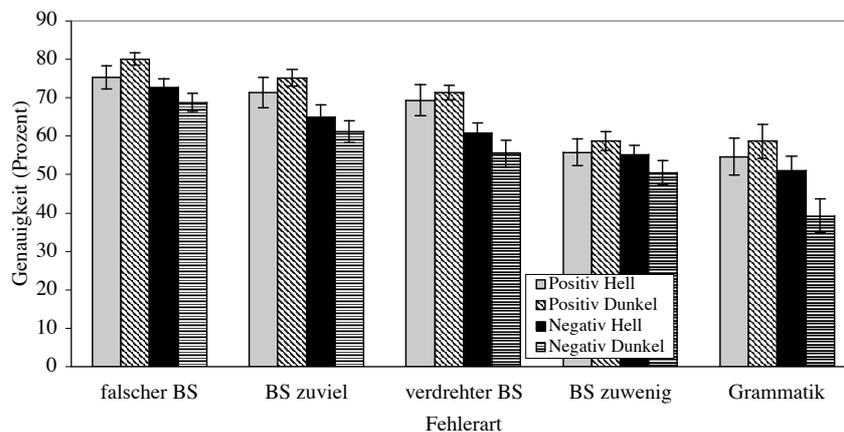


Abbildung 4.8: Die Abbildung zeigt die Fehlerentdeckungs-Genauigkeit getrennt nach Fehlerarten. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

4.2.1.5 Genauigkeit bei der Entdeckung von Grammatikfehlern

Abbildung 4.9 zeigt den Verlauf der Fehlerentdeckungs-Genauigkeit für Grammatikfehler über die 15 Texte. Es zeigte sich, dass für alle Gruppen die Kurve leicht schwankend verlief, allerdings die Genauigkeit gegen Ende der Aufgabenbearbeitung nicht abnahm. Ein Effekt des Zeitverlaufs konnte nicht festgestellt werden ($F(14,63) = 1,33$). Auch zeigte sich kein Effekt der Polarität ($F(14,63) = 1,07$) und der Beleuchtung ($F(14,63) = 1,98$) auf den Zeitverlauf, sowie auch keine Interaktion des Zeitverlaufs mit den anderen beiden Faktoren ($F < 1$).

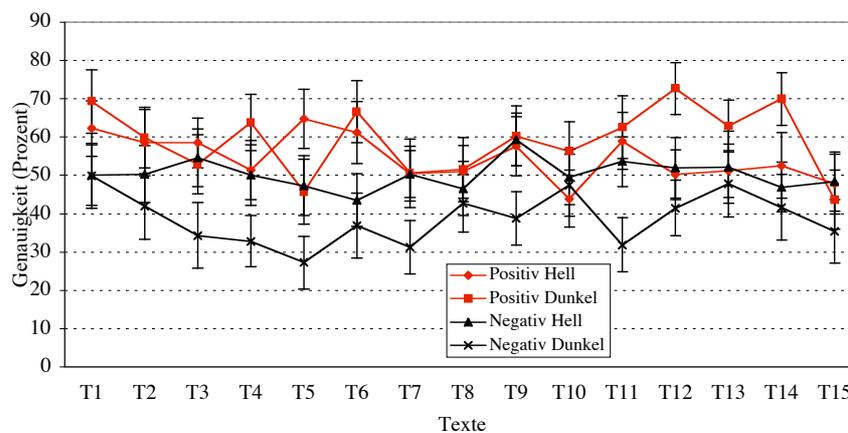


Abbildung 4.9: Dargestellt ist der Verlauf der Entdeckungs-Genauigkeit für Grammatikfehler. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

4.2.1.6 Falsche Alarmer und „Nicht übernehmen“

In Tabelle 4.3 ist die durchschnittliche Zahl falscher Alarmer und nicht übernommener Wörter pro Textseite für die verschiedenen Bedingungen dargestellt. Für die Anzahl der nicht übernommenen Wörter zeigte sich weder ein Effekt der Polarität ($F < 1$), noch der Beleuchtung ($F < 1$). Auch konnte keine Interaktion zwischen den Faktoren festgestellt werden ($F < 1$). Für die falsche Alarmrate bestand auch kein Effekt der Polarität ($F < 1$) und der Beleuchtung ($F < 1$). Ebenso lag auch keine Interaktion vor ($F < 1$).

Tabelle 4.3

Durchschnittliche Zahl falscher Alarme und nicht übernommener Wörter pro Seite. In Klammern ist die Standardabweichung angegeben.

	Falsche Alarme	„Nicht übernehmen“
Positiv Hell	0,53 (0,11)	0,35 (0,06)
Positiv Dunkel	0,39 (0,05)	0,28 (0,08)
Negativ Hell	0,49 (0,07)	0,36 (0,08)
Negativ Dunkel	0,45 (0,08)	0,35 (0,05)

4.2.1.7 Subjektive Angaben

Zum Abschluss des Versuchs wurden die Probanden gebeten, die Textschwierigkeit, die Textdarstellung, den empfundenen Zeitdruck sowie die eigene Sicherheit in der Rechtschreibung einzuschätzen. Die Ergebnisse sind in Tabelle 4.4 dargestellt. Bezüglich der Einschätzung der Textschwierigkeit konnte das Ergebnis einer Varianzanalyse mit Beleuchtung und Polarität als unabhängigen Variablen zeigen, dass der Effekt der Beleuchtung statistisch signifikant war ($F(1,76) = 4,65, p = .036, R^2 = .06$). Das heißt, in den beiden Dunkelbedingungen wurde das Textniveau als schwieriger eingestuft. Es zeigte sich allerdings kein Effekt der Polarität ($F < 1$) und auch keine Interaktion zwischen den beiden Variablen ($F < 1$). Für den subjektiv wahrgenommenen Zeitdruck ergab sich kein Effekt der Polarität ($F < 1$) und der Beleuchtung ($F < 1$), allerdings eine signifikante Interaktion der beiden Faktoren ($F(1,76) = 6,03, p = .016, R^2 = .07$).

Bei der Beurteilung der Textdarstellung zeigte sich kein Effekt der Polarität ($F < 1$), der Beleuchtung ($F < 1$) und auch keine Interaktion zwischen den beiden Variablen ($F < 1$). Für die Einschätzung der eigenen Rechtschreibsicherheit konnte kein Effekt der Polarität ($F(1,76) = 2,78$) und der Beleuchtung ($F < 1$) sowie auch keine Interaktion festgestellt werden ($F < 1$).

Tabelle 4.4

Subjektive Einschätzung des Textniveaus, der Textdarstellung, des empfundenen Zeitdrucks sowie der eigenen „Rechtschreibsicherheit“.

		Positiv Hell	Positiv Dunkel	Negativ Hell	Negativ Dunkel
Textschwierigkeit	1=sehr einfach bis 7=sehr schwierig	2,1 (0,14)	2,5 (0,24)	2,4 (0,24)	2,9 (0,24)
Textdarstellung	1=sehr angenehm bis 7=sehr unangenehm	4,9 (0,31)	4,7 (0,47)	5,2 (0,42)	4,6 (0,39)
Zeitdruck	1=gar nicht bis 7=sehr stark	4,6 (0,28)	3,5 (0,35)	4,0 (0,43)	4,65 (0,39)
Rechtschreibung	1=sehr sicher bis 7=sehr unsicher	2,4 (0,21)	2,5 (0,23)	2,8 (0,26)	2,9 (0,31)

4.2.2 Subjektive Beschwerden

Die subjektiven Beschwerden gliedern sich in die Bereiche physische Beschwerden und subjektive Befindlichkeit.

4.2.2.1 Physische Beschwerden

Der durchschnittliche Anstieg der Symptomkomplexe Augenbeschwerden, Kopfschmerzsymptome sowie Muskel- und Haltungsbeschwerden ist in Abbildung 4.10 dargestellt.

Die Zunahme der Augenbeschwerden war in allen Gruppen signifikant $F(1,76) = 150,05$, $p < .001$, $R^2 = .66$). Man kann erkennen, dass der Anstieg der visuellen Beschwerden in der Negativ Hell- ($M = 10,3$) und der Negativ Dunkel-Bedingung ($M = 9,85$) geringfügig größer war als in der Positiv Hell- ($M = 9,55$) und der Positiv Dunkel-Gruppe ($M = 9,75$). Das Ergebnis einer Varianzanalyse mit Beleuchtung und Polarität als unabhängigen Variablen zeigte allerdings weder einen signifikanten Effekt der Polarität ($F < 1$), noch der Helligkeit ($F < 1$). Es lag auch keine Interaktion zwischen den beiden Faktoren vor ($F < 1$).

Der durchschnittliche Anstieg der Kopfschmerzsymptome in allen vier Bedingungen war ebenfalls signifikant ($F(1,76) = 78,01, p < .001, R^2 = .51$). Das Ergebnis einer Varianzanalyse zeigte aber weder einen signifikanten Effekt der Beleuchtung ($F < 1$), noch der Polarität ($F(1,76) = 2,53$). Eine Interaktion zwischen den beiden Faktoren konnte nicht festgestellt werden ($F(1,76) = 3,02$).

Die Haltungs- und Muskelbeschwerden nahmen auch in allen vier Bedingungen signifikant zu ($F(1,76) = 47,18, p < .001, R^2 = .38$). Diese Zunahme war in den beiden Gruppen mit negativer Kontrastwahl größer. Das Ergebnis der Varianzanalyse konnte allerdings weder einen Effekt der Polarität ($F(1,76) = 2,07$), noch der Beleuchtung ($F < 1$) zeigen. Es lag auch keine Interaktion zwischen den Faktoren vor ($F < 1$).

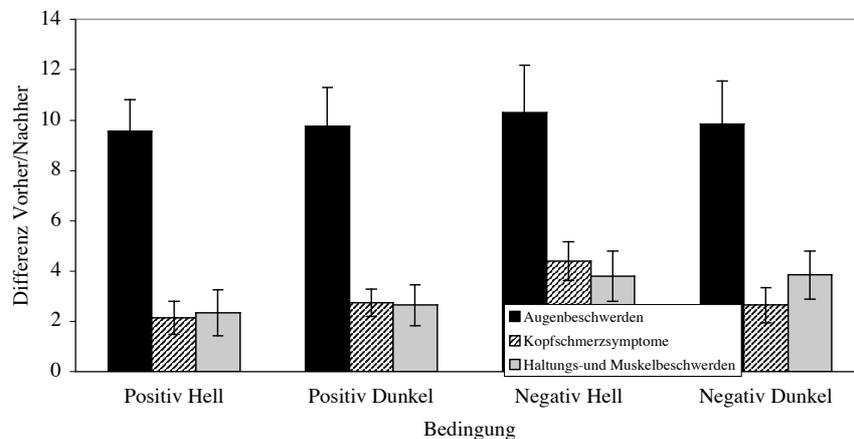


Abbildung 4.10: Durchschnittlicher Anstieg der Augenbeschwerden, Kopfschmerzsymptome sowie Haltungs- und Muskelbeschwerden nach Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

4.2.2.2 Befindlichkeit

Abbildung 4.11 zeigt die durchschnittliche Veränderung nach Aufgabenbearbeitung auf den Skalen „Gute/Schlechte Stimmung“, „Wachheit/Müdigkeit“ sowie „Ruhe/Unruhe“ des Mehrdimensionalen Befindlichkeitsfragebogens in allen Gruppen.

Man kann erkennen, dass der Skalenwert „Gute/Schlechte Stimmung“ in allen Bedingungen signifikant abnahm ($F(1,76) = 23,05, p < .001, R^2 = .23$), das heißt die Stimmung wur-

de schlechter. Es zeigte sich weder ein Effekt der Polarität ($F < 1$), der Beleuchtung ($F < 1$), noch eine Interaktion zwischen den beiden Faktoren ($F < 1$).

Der Wert der Skala „Wachheit/Müdigkeit“ nahm in allen Gruppen nach Aufgabenbearbeitung deutlich ab ($F(1,76) = 77,88, p < .001, R^2 = .51$), wobei diese Differenz in den Bedingungen mit negativer Kontrastwahl größer war. Das Ergebnis einer Varianzanalyse mit Beleuchtung und Polarität als unabhängigen Variablen konnte zeigen, dass der Effekt der Polarität signifikant war ($F(1,76) = 3,93, p = .051, R^2 = .05$). Die Beleuchtung hatte keinen Effekt ($F < 1$) und es gab auch keine Interaktion zwischen den Faktoren ($F < 1$).

Auch die Werte der Skala „Ruhe/Unruhe“ sanken in allen Bedingungen in der Nachmessung signifikant ab ($F(1,76) = 5,86, p = .018, R^2 = .07$). Die Probanden fühlten sich nach Aufgabenbearbeitung unruhiger und unausgeglichener als vorher. Es zeigte sich weder ein Effekt der Beleuchtung ($F < 1$), noch der Polarität ($F < 1$). Auch konnte keine Interaktion festgestellt werden zwischen den beiden Faktoren ($F(1,76) = 2,48$).

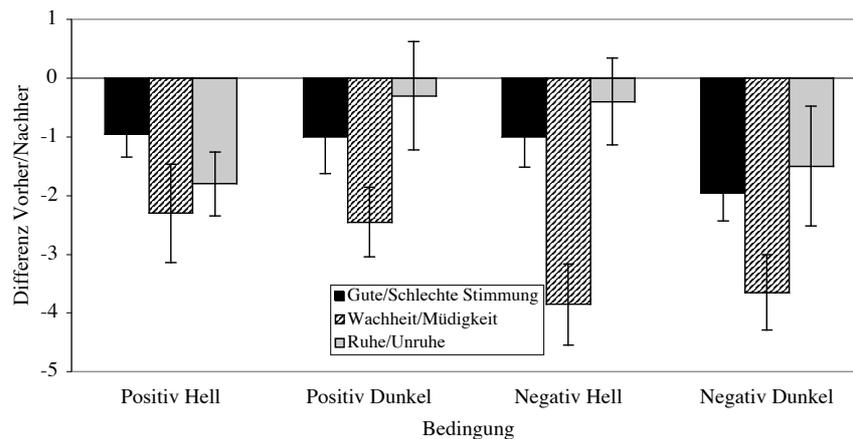


Abbildung 4.11: Mittlere Veränderung auf den Skalen „Gute/Schlechte Stimmung“, „Wachheit/Müdigkeit“ sowie „Ruhe/Unruhe“ nach Aufgabenbearbeitung. Die Fehlerbalken repräsentieren die Standardfehler der Mittelwerte.

4.3 Interpretation und Diskussion

Die Ergebnisse werden getrennt für die Bereiche Leistung und subjektive Beanspruchung zusammengefasst und interpretiert.

4.3.1 Performanz

Es wurde die Vermutung aufgestellt, dass sich der Polaritätseffekt aus Experiment 1 auch auf chromatische Bedingungen verallgemeinern lässt. Die Ergebnisse bestätigen die Hypothese und zeigen, dass auch bei einer chromatischen Textpräsentation unabhängig von der Beleuchtung unter negativen Kontrastverhältnissen die Leistung der Probanden schlechter war als im Rahmen einer Textpräsentation in positiver Polarität. Die Tatsache, dass unter negativen Kontrastverhältnissen durchschnittlich signifikant weniger Fehler pro Textseite entdeckt wurden, war primär in einer geringeren Genauigkeit bei der Fehlerentdeckung begründet. Bezüglich des Lesetempos unterschieden sich die vier Experimentalgruppen nicht, das heißt die Anzahl der gelesenen Wörter wies keinen Unterschied auf. In Experiment 1 hatte sich dagegen ein umgekehrter Effekt gezeigt. Hier konnte die bessere Performanz bei positiver Polarität in erster Linie auf ein schnelleres Lesetempo zurückgeführt werden. Darüber hinaus war von Interesse, ob sich in den vier Experimentalgruppen die Performanz im Verlauf der Aufgabenbearbeitung aufgrund von Ermüdung oder Motivationsdefiziten möglicherweise unterschied. Die Ergebnisse zeigen allerdings, dass in allen Bedingungen die Leistung der Probanden über die gesamte Lesedauer von 45 Minuten konstant verlief. In Hinblick auf Genauigkeit, Lesetempo und die absolute Anzahl gefundener Fehler zeigten sich zwischen den Gruppen keine Unterschiede im Verlauf und die Leistung nahm gegen Ende der Aufgabenbearbeitung nicht ab. Lediglich der erste Text wurde im Vergleich zu den restlichen Texten langsamer gelesen und infolgedessen wurden weniger Fehler gefunden. Man kann also sagen, dass – wie auch schon in der achromatischen Bedingung in Experiment 1 – die Performanz unter negativen Kontrastverhältnissen konsistent während des gesamten Versuchs schlechter war. Wie schon in den ersten beiden Experimenten gab es auch in diesem Fall keine Hinweise darauf, dass die beiden Gruppen

in Hinsicht auf ihre subjektive Leistungsfähigkeit differierten. Bezüglich der subjektiven Einschätzung der eigenen Rechtschreibsicherheit lag kein Unterschied vor. Allerdings wurde in den beiden Dunkel-Bedingungen das sprachliche Textniveau als etwas schwieriger eingestuft, was aber keine Erklärung für das vorliegende Ergebnis darstellt.

Auch im dritten Experiment wurde eine Analyse bezüglich der fünf verschiedenen Fehlerarten durchgeführt. Es zeigte sich, dass signifikant weniger Grammatikfehler erkannt wurden in Relation zu den vier anderen Fehlertypen, was für positive und negative Kontrastverhältnisse gleichermaßen galt. Somit gab es keine Hinweise darauf, dass sich die Experimentalgruppen hinsichtlich ihrer Lesestrategie unterschieden. Darüber hinaus deuten die Daten nicht darauf hin, dass mit zunehmender Versuchsdauer die Lesestrategie geändert wurde. Klammert man die Grammatikfehler von der weiteren Betrachtung aus, zeigte sich das gleiche Muster wie schon in den ersten beiden Experimenten. Wörter, die einen Buchstaben zuviel enthielten, wurden mit größter Genauigkeit erkannt, während Wörter, in denen ein Buchstabe fehlte, am wenigsten häufig entdeckt wurden.

Zusammenfassend kann man sagen, dass im dritten Teil der Untersuchung auf Performanzebene das Ergebnis aus Experiment 1 repliziert werden konnte. Auch im Fall einer chromatischen Textpräsentation ist die Leistung der Probanden unabhängig von der Beleuchtung besser unter positiven Kontrastverhältnissen. Bezüglich der subjektiven Einschätzung der Textdarstellung unterschieden sich die Gruppen – wie schon in Experiment 1 – nicht. Die Textpräsentation wurde von allen Probanden als gleichermaßen unangenehm eingestuft. Die Empfehlung, sowohl bei hoher als auch niedriger Umgebungsbeleuchtung auf negativen Kontrast zu verzichten, kann also auch auf eine farbige Textdarstellung verallgemeinert werden.

Vergleicht man die Leistung der Probanden in dieser Studie mit den Resultaten aus Experiment 1 kann man erkennen, dass unter Beleuchtung bei einer chromatischen Textdarstellung in positiver Polarität, das heißt blaue Schrift auf gelbem Hintergrund, deskriptiv mehr Fehler entdeckt wurden als im Fall einer Textdarbietung mit schwarzer Schrift auf weißem Hintergrund. An dieser Stelle kommt die Frage auf, wie dieses Ergebnis zu interpretieren ist. Ist die Blau/Gelb-Darstellung im Rahmen einer Lese-Korrektur-Aufgabe geeigneter für das Entdecken von Fehlern als schwarze Schrift auf weißem Hintergrund? Aufgrund theo-

retischer Erkenntnisse scheint diese Erklärung unwahrscheinlich zu sein. Es konnte gezeigt werden, dass sich in der Fovea verhältnismäßig wenige Rezeptoren für kurzwelliges Licht befinden (vgl. Karnath & Thier, 2003). Aus diesem Grund müsste die Schwarz/Weiß-Darstellung der Farbbedingung aus Experiment 3 auf Performanzebene deutlich überlegen sein. Für die bessere Leistung in der Blau/Gelb-Darstellung muss es eine andere Erklärung geben. Der experimentelle Ablauf in den beiden Experimenten war nicht direkt vergleichbar. Im Gegensatz zum letzten Experiment wurde in Experiment 1 neben der Leistung auch die psychophysiologische Beanspruchung untersucht. Die Probanden waren aufgrund der Ableitungen in ihrem Bewegungsfreiraum eingeschränkt und die Versuchsbedingungen waren dadurch unangenehmer. Dies konnte die Folge haben, dass die Leistung in Experiment 1 generell herabgesetzt war und somit nicht mit dem dritten Experiment verglichen werden konnte. In einer Folgeuntersuchung, die nicht Gegenstand dieser Arbeit ist, wurde die Blau/Gelb-Kombination der Schwarz/Weiß-Darstellung direkt gegenüber gestellt. Der Text wurde bei hoher Umgebungsbeleuchtung in vier verschiedenen Farbkombinationen präsentiert (Schwarz auf Weiß, Weiß auf Schwarz, Blau auf Gelb und Gelb auf Blau). Der Stichprobenumfang betrug insgesamt 84 Probanden. Auch in dieser Studie konnte ein deutlicher Vorteil für eine Textpräsentation in positiver Polarität festgestellt werden. Es zeigten sich allerdings keine Leistungsunterschiede zwischen der farbigen und der achromatischen Darstellung. In beiden Bedingungen wurden sowohl bei positiven als auch negativen Kontrastverhältnissen ungefähr gleich viele Fehler im Rahmen einer Lesekorrekture-Aufgabe entdeckt. Die entscheidende Variable war also auch in dieser Untersuchung die Wahl der Polarität und die Farbe hatte keinen Einfluss.

4.3.2 Subjektive Beanspruchung

Auch auf physischer Ebene wurden erwartungsgemäß bei fast allen Probanden am Ende der Aufgabenbearbeitung deutliche Beschwerden offensichtlich. In allen Gruppen konnte ein signifikanter Anstieg visueller Probleme sowie eine deutliche Zunahme von Kopfschmerzen, Benommenheit und Schwindel festgestellt werden. Auch die Beschwerden im Nacken-, Schulter- und Rückenbereich waren in allen Gruppen am Ende der Aufgabenbearbeitung deutlich stärker ausgeprägt. Im Vergleich zu den ersten beiden Experimenten

war das Ausmaß der Zunahme in allen Gruppen allerdings geringer, da die Bewegungsfreiheit der Probanden in Experiment 3 weniger eingeschränkt war.

Neben der physischen Beanspruchung zeigten sich bei allen Probanden auch Beschwerden im Bereich der subjektiven Befindlichkeit nach Beendigung der Lese-Korrektur-Aufgabe. Auf allen drei Skalen des mehrdimensionalen Befindlichkeitsfragebogens spiegelte sich die psychische Beanspruchung wider. Im Vergleich zur Vormessung kam es zu einer signifikanten Verschlechterung der Stimmung. Auch das Müdigkeitsempfinden nahm deutlich zu, wobei sich ein signifikanter Effekt der Polarität zeigte. Im Vergleich zum positiven Kontrast waren die Probanden bei negativer Polarität müder nach Beendigung der Aufgabe. Darüber hinaus fühlten sich alle Gruppen in der Nachmessung unruhiger und unausgeglichener als vorher.

5. Abschlussdiskussion

In den nächsten Abschnitten erfolgt eine abschließende Diskussion der drei Experimente. Anschließend werden offene Fragestellungen aufgezeigt, die in der vorliegenden Arbeit nicht geklärt werden konnten.

5.1 Polaritätseffekt

In der vorliegenden Untersuchung konnte erstmals nachgewiesen werden, dass unter achromatischen Bedingungen eine Textpräsentation in negativer Polarität sowohl bei einer üblichen Raumbeleuchtung von 550 Lux als auch bei Dunkelheit (5 Lux) zu einer deutlich schlechteren Leistung im Rahmen einer Lese-Korrektur-Aufgabe führte als positiver Kontrast, wofür in erster Linie ein geringeres Lesetempo verantwortlich war. Um die schlechte Leistung unter negativen Kontrastverhältnissen und Dunkelheit besser einstufen zu können, wurde in Experiment 2 die Performanz bei einer sehr schlecht lesbaren, äquilibriumanten Textpräsentation mit roter Schrift auf grünem Hintergrund untersucht. Die Ergebnisse zeigen, dass die Leistung bei dieser Textdarstellung unter ansonsten gleichen Bedingungen nur geringfügig schlechter war als in der Negativ Dunkel-Bedingung aus Experiment 1. Der Vorteil für die Positivdarstellung zeigte sich aber nicht nur im Rahmen einer achromatischen Textpräsentation, sondern konnte in Experiment 3 auch auf eine farbige Darstellung übertragen werden. Auch in diesem Fall war die Leistung der Probanden bei einer Textpräsentation in negativer Polarität unabhängig von den Lichtverhältnissen signifikant schlechter, wobei dafür in erster Linie eine geringe Genauigkeit bei der Fehlerentdeckung die Ursache war.

Die Tatsache, dass sich bei Helligkeit ein Vorteil für die Positivdarstellung zeigte, stellt ein erwartungsgemäßes Ergebnis dar. Frühere Studien, in denen der Einfluss der Polarität untersucht wurde, wiesen dagegen kein homogenes Bild auf. In einigen Studien zeigte sich eine bessere Leistung unter positiven Kontrastverhältnissen (Bauer & Cavonius, 1980; Gould et al., 1987). Andere Untersuchungen stellten dagegen keinen Unterschied auf der

Performanzebene fest (Creed et al., 1988; Cushman, 1986; Legge et al., 1987). Ein wichtiges methodisches Problem dieser Studien für die heutige Interpretation ist allerdings die Tatsache, dass in diesen Untersuchungen oftmals Röhrenmonitore mit geringer Bildwiederholrate verwendet wurden, die zu dieser Zeit Stand der Technik waren. Dies kann bei positiven Kontrastverhältnissen zu einer Wahrnehmung von Flimmern führen, was die Leseleistung möglicherweise stark herabgesetzt und infolgedessen den Vorteil der positiven Polarität verdeckt hat. Als Folge zeigte sich in diesen Studien kein Performanzunterschied zwischen einer Textdarstellung in positiver und negativer Polarität. Im Rahmen der vorliegenden Arbeit kam ein TFT-Monitor zum Einsatz, der gegenüber herkömmlichen CRT-Monitoren einige technische Vorteile aufweist und in der Regel flimmerfrei arbeitet. Es konnte erstmals gezeigt werden, dass im Rahmen einer Textpräsentation am TFT-Bildschirm die Leistung bei einer Textdarstellung in positiver Polarität signifikant besser ist. Das Ergebnis der vorliegenden Arbeit kann als Bestätigung dafür betrachtet werden, dass die Positivdarstellung heutzutage unter normalen Beleuchtungsbedingungen standardmäßig eingesetzt und auch nach den Sicherheitsregeln für Bildschirmarbeitsplätze empfohlen wird.

Es stellt sich natürlich die Frage, warum die Schlussfolgerungen bezüglich der Kontrastwahl bei dunkler Umgebung aus den Experimenten von Taptagaporn und Saito (1990), in denen Variationen im Pupillendurchmesser untersucht wurden, sich nicht auf der beobachtbaren Leistungsebene widerspiegeln. Ein Grund dafür kann sein, dass sich in der Regel im Rahmen von Textverarbeitung am Bildschirm die meisten Computernutzer für den positiven Kontrast entscheiden und man deswegen davon ausgehen kann, dass diese Art der Textpräsentation einen wesentlich höheren Vertrautheitsgrad aufweist. Viele Anwender haben vielleicht auch schon die Erfahrung gemacht, bei Dunkelheit am Bildschirm zu arbeiten und sind somit auch in dunklen Räumen an eine Positivdarstellung gewöhnt. Außerdem werden auch sämtliche Printmedien wie Bücher, Zeitungen etc. in positiver Polarität dargestellt, was zusätzlich die Vertrautheit erhöht. Darüber hinaus kann das vorliegende Ergebnis noch eine andere Ursache haben. Ein Argument für negativen Kontrast bei Dunkelheit war die Tatsache, dass dadurch der Leuchtdichteunterschied zwischen Bildschirm und Umgebung reduziert wird, da die Gesamtluminanz des Monitors geringer ist als bei einer Darstellung in positiver Polarität. Es gibt aber Hinweise darauf, dass nicht die gesamte Bildschirmoberfläche eine Rolle spielt bei Hell- und Dunkeladaptationsprozessen,

sondern nur die direkt fixierten Regionen des visuellen Feldes, in diesem Fall also die Schrift (vgl. Pawlak, 1986). Bei negativer Polarität ist die Schrift allerdings deutlich heller als der Hintergrund, wodurch bei Dunkelheit der Leuchtdichteunterschied zur Umgebung stark ansteigt. Dies könnte eine Erklärung dafür sein, warum in der vorliegenden Arbeit eine Textdarstellung in negativen Kontrastverhältnissen bei Dunkelheit entgegengesetzt der Empfehlung, die aus der Arbeit von Taptagaporn und Saito (1990) ableitbar ist, zu einer schlechteren Leistung führte. Die negative Polarität erhöhte sogar den Leuchtdichteunterschied zwischen Bildschirm und Umgebung anstatt ihn zu reduzieren. In der Studie von Taptagaporn und Saito (1990) bestand die Aufgabe der Probanden lediglich darin, auf den Monitor zu blicken. Es musste kein Text gelesen und somit auch keine bestimmten Bereiche des Bildschirms fixiert werden. Aus diesem Grund ergaben sich möglicherweise die geringeren Variationen im Pupillendurchmesser bei negativer Polarität und Dunkelheit, woraus ein Vorteil für den negativen Kontrast bei Dunkelheit gefolgert wurde. Das Ergebnis der vorliegenden Arbeit macht deutlich, dass es darüber hinaus generell als problematisch angesehen werden kann, aus Indikatoren – in diesem Fall Variationen im Pupillendurchmesser – auf ein Problem, das heißt die Auswirkungen auf Performanzebene, zu schließen, ohne den Einfluss der Kontrastwahl auf die Leistung selbst zu messen.

Das vorliegende Ergebnis kann aber auch noch einen anderen Grund haben. Bezüglich der Kontrastwahl in einer hellen Umgebung wurde argumentiert, dass bei positiver Polarität aus der höheren Leuchtdichte des Monitors eine Pupillenverengung resultiert, wodurch sich die Sehschärfe erhöht (Krüger, 1984). Man kann in diesem Zusammenhang davon ausgehen, dass bei dunkler Umgebung das Gleiche gilt und auch in diesem Fall positiver Kontrast die Qualität des retinalen Bildes steigert, wodurch die bessere Leistung bei positiver Polarität erklärt werden kann.

5.2 Objektive und subjektive Beanspruchung

Die psychophysiologischen Daten aus Experiment 1 und 2 zeigen, dass alle Indikatoren die Beanspruchung während der Bearbeitung der Lese-Korrektur-Aufgabe deutlich abbilden. Dieses Ergebnis stimmt mit früheren Studien überein, in denen die Effekte verschiedener

kognitiver Herausforderungen wie beispielsweise Rechen- oder Reaktionsaufgaben untersucht wurden. Wie sich dagegen typische Bildschirmarbeit wie Textverarbeitung oder Dateneingabe auf die objektive Beanspruchung auswirkt, konnte bisher nur unzureichend geklärt werden. Die Zahl der Studien auf diesem Gebiet ist zwar relativ umfangreich, allerdings lag der Fokus oftmals lediglich auf der Muskelaktivität, die durch die Bildschirmarbeit hervorgerufen wird (z.B. Birch et al., 2000; Finsen, Sogaard, Jensen, Borg & Christensen, 2001). In Untersuchungen, in denen zusätzlich zum Elektromyogramm noch andere psychophysiologische Indikatoren erfasst wurden, kann in fast allen Fällen das methodische Vorgehen kritisiert werden. Die motorische Aktivität der Versuchspersonen – bedingt durch Mausbewegungen oder Bedienung der Tastatur – fand keine oder nur unzureichende Berücksichtigung (Schleifer & Ley, 1994; Schleifer & Okogbaa, 1990). Da aber gezeigt werden konnte, dass allein durch physische Aktivität, ohne dass gleichzeitig kognitive Anforderungen vorliegen müssen, die Beanspruchungsindikatoren beeinflusst werden können, kann dies zu Konfundierungen führen (Garde et al., 2002). Aus diesem Grund ist es im Rahmen der Untersuchung objektiver Beanspruchung bei Bildschirmarbeit unabdingbar, die motorische Komponente zu kontrollieren. In der vorliegenden Arbeit wurden die Versuchspersonen aufgefordert, während der beiden Ruhemessungen Mausbewegungen zu simulieren. Veränderungen der psychophysiologischen Indikatoren während der Bearbeitung der Lese-Korrektur-Aufgabe konnten also nicht auf die motorische Aktivität zurückgeführt werden, sondern lediglich auf die kognitiven Aspekte der Aufgabe. Somit konnte erstmals gezeigt werden, dass typische Bildschirmarbeit wie Textverarbeitung unter Kontrolle der physischen Aktivität zu einer nachweisbaren, mehrdimensionalen objektiven Beanspruchung führte. An dieser Stelle muss allerdings erwähnt werden, dass nicht ausgeschlossen werden kann, dass es sich möglicherweise um einen unspezifischen Effekt handelte. Es könnte sein, dass die physiologische Beanspruchung durch die Aufgabenbearbeitung an sich und nicht spezifisch durch die Arbeit am Bildschirm hervorgerufen wurde. Um diese Frage zu klären, hätte man eigentlich eine zusätzliche Kontrollbedingung einführen müssen, in der die gleiche Aufgabe zum Beispiel in einer Papier-Version bearbeitet wird.

Die Ergebnisse dieser Studie zeigen aber auch, dass sich eine schlechte Leistung nicht gleichzeitig in einer signifikant höheren objektiven Beanspruchung widerspiegelte. Die wahrscheinlichste Erklärung für den nicht vorhandenen Gruppenunterschied auf physiologischer Ebene liegt darin, dass sich alle Versuchspersonen bei Bearbeitung der Aufgabe

gleichermaßen stark anstregten. Bei der Stichprobe handelte es sich in erster Linie um Studenten, die motiviert und ehrgeizig erschienen. Zudem wurde ein zusätzlicher Anreiz geschaffen, indem eine Prämie für eine besonders gute Leistung in Aussicht gestellt wurde. In welchem Ausmaß sich eine bestimmte Art der Textdarstellung für eine erfolgreiche Aufgabenbearbeitung eignete oder nicht, zeigte sich lediglich auf der Performanzebene. In diesem Zusammenhang muss aber noch eine Alternativerklärung für das vorliegende Ergebnis angeführt werden. Möglicherweise war die Diskriminationsgenauigkeit der physiologischen Beanspruchungsindikatoren zwischen den verschiedenen Belastungsbedingungen nicht groß genug. Jorna (1992) kam zu dem Schluss, dass beispielsweise die 0,1 Hz Komponente der Herzratenvariabilität sehr reliabel zwischen Ruhe und Belastung unterscheiden kann, deutlich weniger empfindlich aber Unterschiede zwischen verschiedenen Schwierigkeitsstufen einer Aufgabe abbildet. Unterschiede in der mentalen Beanspruchung spiegeln sich eventuell nur dann in der Messung dieses Frequenzbereichs wider, wenn sie sehr groß sind (z.B. Aasman et al., 1987). Allerdings kann argumentiert werden, dass sich auch in diesen Untersuchungen alle Teilnehmer gleichermaßen stark anstregten und sich deswegen keine Unterschiede in den physiologischen Beanspruchungsindikatoren zeigten. Neben dieser objektiv messbaren physiologischen Beanspruchung berichteten die Probanden nach Aufgabenbeendigung auch über eine deutliche subjektive Beanspruchung. Zusätzlich zu physischen Beschwerden wie visuelle Probleme, Kopfschmerzen etc. kam es auch zu Befindlichkeitsveränderungen wie Müdigkeit, Unruhe und schlechter Stimmung. Allerdings unterschieden sich in den drei Experimenten die Gruppen nicht sehr stark im Ausmaß der subjektiven Beanspruchung. Dieses Ergebnis kann zum einen auch darin begründet sein, dass sich alle Probanden gleich stark anstregten und sich diese Tatsache auf subjektiver Ebene widerspiegelt. Zum anderen besteht aber auch die Möglichkeit, dass die verwendeten Fragebögen nicht sensitiv genug waren, um die verschiedenen Abstufungen der Beanspruchung abzubilden.

5.3 Weiterführende Perspektiven

Die Ergebnisse zeigen, dass unabhängig von der Beleuchtung bei positiver Polarität eine bessere Leistung vorlag. An dieser Stelle kommt natürlich die Frage auf, inwieweit sich dieses Resultat verallgemeinern lässt. In der vorliegenden Arbeit wurde der Text in relativ kleiner Schriftgröße, langer Zeilenlänge sowie engem Zeilenabstand präsentiert. Interessant wäre es zu untersuchen, ob der Vorteil für die positive Polarität auch dann bestehen bleibt, wenn sich die Formatierung des Textes ändert. Zeigt sich der Polaritätseffekt auch bei einer Textpräsentation in Paragraphen oder bei größerer Schrift? Wie sieht es im Fall einer anderen Schriftart aus? Ein weiterer wichtiger Punkt betrifft die Lichtverhältnisse. In der vorliegenden Arbeit gab es nur zwei Abstufungen, und zwar wurde der Raum entweder komplett abgedunkelt oder es bestand eine konstante künstliche Beleuchtung von 550 Lux. Unter Alltagsbedingungen in üblichen Büroräumen sind die Lichtverhältnisse dagegen weitaus weniger konstant. Dies ist vor allem darin begründet, dass in den meisten Fällen eine Mischung aus Kunst- und Tageslicht vorliegt. Tageslicht unterliegt allerdings starken Schwankungen und kann zu Blendungen und Reflexionen führen. In diesem Zusammenhang stellt sich die Frage, ob sich die bessere Leistung bei positivem Kontrast auch dann zeigt, wenn die Leuchtdichten stärker variiert werden oder zusätzlich auf Tageslicht zurückgegriffen wird. Ebenso könnte untersucht werden, wie sich Variationen in der Helligkeitseinstellung des Monitors auswirken.

Auch bezüglich der subjektiven Beanspruchung – vor allem im Rahmen der psychischen Beschwerden – zeigten sich kaum signifikante Unterschiede zwischen den Gruppen. Dieses Ergebnis konnte zum einen darin begründet sein, dass sich alle Versuchspersonen gleich stark anstrebten. Das Resultat kann aber auch durch die Auswahl des Fragebogens zur Erfassung der psychischen Beanspruchung erklärt werden. Die Kurzform des mehrdimensionalen Befindlichkeitsfragebogens (MDBF) teilt nur sehr grob in die drei Skalen „Gute/Schlechte Stimmung“, „Wachheit/Müdigkeit“ und „Ruhe/Unruhe“ ein und die Ergebnisse der Befragung geben nur Hinweise auf die Befindlichkeit. In weiteren Untersuchungen sollte Wert darauf gelegt werden, das Ausmaß der psychischen Beanspruchung mittels anderer Fragebögen detaillierter zu erfassen. Darüber hinaus sollten die Versuchspersonen auch bezüglich ihrer subjektiven Anstrengungsbereitschaft befragt werden.

Um diese in der Diskussion angeführten offenen Fragen zu klären und die aus der vorliegenden Studie gewonnen Erkenntnisse weiter abzusichern, sind noch weitere Untersuchungen notwendig.

5.4 Schlussfolgerungen

Aus den Ergebnissen der drei Experimente kann die Schlussfolgerung gezogen werden, unabhängig von der Raumbeleuchtung auf eine Textpräsentation in negativer Polarität zu verzichten. Die Performanz im Rahmen einer Lese-Korrektur-Aufgabe ist sowohl bei Dunkelheit als auch künstlicher Raumbeleuchtung signifikant schlechter unter negativen Kontrastverhältnissen zwischen Schrift und Hintergrund. Dieses Resultat gilt nicht nur für eine Textdarstellung in Schwarz/Weiß, sondern lässt sich auch auf eine chromatische Form der Textpräsentation verallgemeinern.

Außerdem konnte in der vorliegenden Arbeit im Rahmen einer mehrdimensionalen psychophysiologischen Beanspruchungsanalyse gezeigt werden, dass typische Bildschirmarbeit wie die Durchführung einer Lese-Korrektur-Aufgabe unter Kontrolle der motorischen Aktivität zu einer nachweisbaren Beanspruchung führt. Unterschiede im Ausmaß der objektiven Beanspruchung zwischen den Gruppen konnten allerdings nicht nachgewiesen werden, was vereinbar ist mit der Annahme, dass sich alle Versuchsteilnehmer bei der Aufgabebearbeitung gleich stark anstrebten.

6. Literatur

- Aasman, J., Mulder, G. & Mulder, L.J. (1987). Operator effort and the measurement of heart-rate variability. *Human Factors*, 29, 161-70.
- Akselrod, S., Gordon, D., Madwed, J.B., Snidman, N.C., Shannon, D.C. & Cohen, R.J. (1985). Hemodynamic regulation: investigation by spectral analysis. *American journal of physiology*, 249, 867-875.
- Akselrod, S., Gordon, D., Ubel, F.A., Shannon, D.C., Berger, A.C. & Cohen, R.J. (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, 213, 220-222.
- Al' Absi, M., Everson, S.A. & Lovallo, W.R. (1995). Hypertension risk factors and cardiovascular reactivity to mental stress in young men. *International Journal of Psychophysiology*, 20, 155-160.
- Albus, M., Müller-Spahn, F., Ackenheil, M. & Engel, R.R. (1987). Psychiatric patients and their response to experimental stress. *Stress Medicine*, 3, 233-238.
- Althaus, M., Mulder, L.J., Mulder, G., Van Roon, A.M. & Minderaa, R.B. (1998). Influence of respiratory activity on the cardiac response pattern to mental effort. *Psychophysiology*, 35, 420-430.
- Appel, M.L., Berger, R.D., Saul, J.P., Smith, J.M. & Cohen, R.J. (1989). Beat to beat variability in cardiovascular variables: noise or music? *Journal of the American College of Cardiology*, 14, 1139-1148.
- Aslan, S., Nelson, L., Carruthers, M. & Lader, M.H. (1981). Stress and age effects on catecholamines in normal subjects. *Journal of Psychosomatic Research*, 25, 33-41.
- Backs, R.W. & Seljos, K.A. (1994). Metabolic and cardiorespiratory measures of mental effort: The effects of level of difficulty in a working memory task. *International Journal of Psychophysiology*, 16, 57-68.
- Bauer, D. & Cavonius, C.R. (1980). *Improving the legibility of visual display units through contrast reversal*. In: Grandjean, E. & Vigliani, E. (Eds). *Ergonomic Aspects of Visual Display Terminals* (S. 137-142). London: Taylor & Francis.
- Bildschirmarbeitsverordnung: Verordnung über Sicherheit und Gesundheitsschutz bei der Arbeit an Bildschirmgeräten. *BGBI*, I 1996, S. 1843-1844.

- Birch, L., Juul Kristensen, B., Jensen, C., Finsen, L. & Christensen, H. (2000). Acute response to precision, time pressure and mental demand during simulated computer work. *Scandinavian journal of work, environment and health*, 26, 299-305.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin, Heidelberg, New York, Barcelona, Hongkong, London, Mailand, Paris, Singapur, Tokio: Springer.
- Boucsein, W. (1992). *Electrodermal Activity*. New York: Plenum Press.
- Carroll, D., Turner, J.R., Hellowell, J.C. (1986). Heart rate and oxygen consumption during active psychological challenge: The effects of level of difficulty. *Psychophysiology* 23, 174-181.
- Casali, J.G. & Wierwille, W.W. (1983). A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human Factors*, 25, 623-41.
- Cohen, J. (1977). *Statistical Power Analysis for the behavioral sciences (revised edition)*. New York: Academic Press.
- Creed, A., Dennis, I. & Newstead, S. (1987). Proof-reading on VDUs. *Behaviour and Information Technology*, 6, 3-13.
- Creed, A., Dennis, I. & Newstead, S. (1988). Effects of display format on proof-reading with VDUs. *Behaviour and Information Technology*, 7, 467-478.
- Cushman, W.H. (1986). Reading from microfiche, a VDT, and the printed page: subjective fatigue and performance. *Human factors*, 28, 63-73.
- DBAG Forschungsteam MMI (1998). *Text*. Online im Internet: URL: http://www.inf.tu.dresden.de/ST2/pw/lv_bdt/hyperbase/buch/grundlagen/softwareergonomische_gestaltungsrichtlinien/text/text.htm (30.05.2003)
- Delaney, J.P.A. & Brodie, D.A. (2000). Effects of short-term psychological stress on the time and frequency domains of heart-rate variability. *Perceptual and Motor Skills*, 9, 515-524.
- DeValois, R.L. (1960) Color vision mechanisms in the monkey. *Journal of General Physiology*, 43 (2), 115-128.
- Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics*, 35, 1297-1326.
- DIN 33 405 (1987). *Psychische Belastung und Beanspruchung; Allgemeines, Begriffe*.
- Erdfelder, E., Faul, F. & Buchner, A. (1996). GPower: A general power analysis program. *Behavior Research Methods, Instruments & Computers*, 28, 1-11.

- Ergo/Gero (1998-2001). *Color for Text and Graph Legibility*. Online im Internet: URL: <http://www.ergogero.com/FAQ/Part6/cfaqPart6.html> (02.06.2003)
- Ertel, M., Junghanns, G., Pech, E. & Ullsperger, P. (1997). *Auswirkungen der Bildschirmarbeit auf Gesundheit und Wohlbefinden (Fb 762)*. Bremerhaven: Wirtschaftsverlag NW.
- Ertel, M., Junghanns, G. & Ullsperger, P. (1994). *Gesundheit am Bildschirmarbeitsplatz (GESBI)*. Bremerhaven: Wirtschaftsverlag NW.
- Fichera, L.V. & Andreassi, J.L. (1998). Stress and personality as factors in women's cardiovascular reactivity. *International Journal of Psychophysiology*, 28, 143-155.
- Finsen, L., Sogaard, K., Jensen, C., Borg, V. & Christensen, H. (2001). Muscle activity and cardiovascular response during computer-mouse work with and without memory demands. *Ergonomics*, 44, 1312-1329.
- Fournier, L.R., Wilson, G.F. & Swain, C.R. (1999). Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: Manipulations of task difficulty and training. *International Journal of Psychophysiology*, 31, 129-145.
- Garde, A.H., Laursen, B., Jorgensen, A.H. & Jensen, B.R. (2002). Effects of mental and physical demands on heart rate variability during computer work. *European journal of applied physiology*, 87, 456-461.
- Gould, J.D. (1987). Reading is slower from CRT displays than from paper: Attempts to isolate a single-variable explanation. *Human Factors*, 29, 269-299.
- Gould, J.D., Alfaro, L., Finn, R. & Haupt, B. (1987). Reading from CRT displays can be as fast as reading from paper. *Human Factors*, 29, 497-517.
- Gould, J.D. & Grischkowsky, N. (1984). Doing the same work with hard copy and with cathode-ray tube (CRT) computer terminals. *Human Factors*, 26, 323-337.
- Gouras, P. (1991) *Color Vision*. In: Kandel, E.R, Schwartz, J.H. & Jessell, T.M. *Principals of neural science* (S. 467-480). New York: Elsevier.
- Greif, S., Bamberg, E. & Semmer, N. (Hg.). (1991). *Psychischer Stress am Arbeitsplatz*. Göttingen: Hogrefe.
- Grossman, P. (1992). Respiratory and cardiac rhythms as windows to central and autonomic biobehavioral regulation: selection of window frames, keeping the panes clean and viewing the neural topography. *Biological psychology*, 34, 131-161.

- Heslegrave, R.J & Furedy, J.J. (1979). Sensitivities of HR and T-wave amplitude for detecting cognitive and anticipatory stress. *Physiology and Behavior*, 22, 17-23.
- Heuer, H., Hollendiek, G., Kröger, H. & Römer, T. (1989). Die Ruhelage der Augen und ihr Einfluss auf Beobachtungsabstand und visuelle Ermüdung bei Bildschirmarbeit. Resting position of the eyes and its influence on observation distance and visual fatigue during visual display work. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 36, 538-566.
- Hörhold, M. & Walschburger, P. (1997). Depressive Störung als Ausdruck misslingender Handlungskontrolle. Überprüfung einer psychophysiologischen Belastungsdiagnostik. Depressive disorders, state orientation and mental load: Examination of a psychophysiological diagnostic approach. *Zeitschrift für Klinische Psychologie Forschung und Praxis*, 26, 31-37.
- Ishihara, S. (1972). *Ishihara's Tests for Colour-Blindness*. Tokio: Kanehara Shuppan Co. Ltd.
- Jaschinski Kruza, W. (1991). Eyestrain in VDU users: viewing distance and the resting position of ocular muscles. *Human factors*, 33, 69-83.
- Johnson, L.C. & Lubin, A. (1966). Spontaneous electrodermal activity during waking and sleeping. *Psychophysiology*, 3, 8-17.
- Jorna, P.G. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological Psychology*, 34, 237-257.
- Junghanns, G., Ertel, M. & Ullsperger, P. (1998). *Anforderungsbewältigung und Gesundheit computergestützter Büroarbeit (Fb 787)*. Bremerhaven: Wirtschaftsverlag NW.
- Junghanns, G., Ertel, M. & Ullsperger, P. (1999). Zum Auftreten von Gesundheitsbeschwerden bei computergestützter Büroarbeit - eine multivariate Analyse auf der Grundlage einer fragebogengestützten Erhebung. *Zeitschrift für Arbeitswissenschaft*, 1, 18-24.
- Kalsbeek, J.W.H. & Ettema, J.H. (1963). Scored irregularity of the heart pattern and the measurement of perceptual or mental load. *Ergonomics*, 6, 306-307.
- Kahnemann, D. (1973). *Attentions and Effort*. Englewood Cliffs, USA: Prentice-Hall.
- Kamada, T., Sato, N., Miyake, S. & Kumashiro, M. (1992). Power spectral analysis of heart rate variability in Type As during solo and competitive mental arithmetic task. *Journal of Psychosomatic Research*, 36, 543-551.

- Kamath, M.V. & Fallen, E.L. (1993). Power spectral analysis of heart rate variability: a noninvasive signature of cardiac autonomic function. *Critical reviews in biomedical engineering*, 21, 245-311.
- Karnath, H.-O. & Thier, P. (2003). *Neuropsychologie*. Berlin, Heidelberg, New York, Hongkong, London, Mailand, Paris, Tokio: Springer.
- Kirby, L.D. (1999). Emotion or attention? The psychological significance of electrodermal activity. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 60, 3017.
- Kitney, R.I. (1975). Proceedings: Entrainment of the human RR interval by thermal stimuli. *Journal of physiology*, 252, 37-38.
- Knoblauch, K., Arditi, A. & Szlyk, J. (1988). Reading rate depends on luminance contrast but not chromatic contrast. *Invest. Ophthalmol. Vis. Sci.*, 29, 44.
- Kohlisch, O. & Schäfer, F. (1996). Physiological changes during computer tasks: responses to mental load or to motor demands? *Ergonomics*, 39, 213-224.
- Krüger, H. (1984). Lighting characteristics, legibility and visual comfort of displays. *Behaviour and Information Technology*, 3, 291-299.
- Kuhmann, W., Schäfer, F. & Boucsein, W. (1990). Effekte von Wartezeiten innerhalb einfacher Aufgaben: Eine Analogie zu Wartezeiten in der Mensch-Computer-Interaktion. *Zeitschrift für experimentelle und angewandte Psychologie*, 37, 242-265.
- Legge, G.E., Parish, D.H., Luebker, A. & Wurm, L.H. (1990). Psychophysics of reading: XI. Comparing color contrast and luminance contrast. *Journal of the Optical Society of America*, 7(10), 2002-2010.
- Legge, G.E., Rubin, G.S. & Luebker, A. (1987). Psychophysics of reading: V. The role of contrast in normal vision. *Vision Research*, 27, 1165-1177.
- Livingstone, M.S. & Hubel, D.H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of neuroscience the official journal of the Society for Neuroscience* 4 (1), 309-356.
- Lucini, D., Guzzetti, S., Casiraghi, S. & Pagani, M. (2002). Correlation between baroreflex gain and 24-h indices of heart rate variability. *Journal of hypertension*, 20, 1625-1631.
- Malliani, A., Pagani, M., Lombardi, F. & Cerutti, S. (1991). Cardiovascular neural regulation explored in the frequency domain. *Circulation*, 84, 482-492.

- Matthews, M.L. (1987). The influence of colour on CRT reading performance and subjective comfort under operational conditions. *Applied Ergonomics*, 18, 323-328.
- Matthews, M.L., Lovasik, J.V. & Mertins, K. (1989). Visual performance and subjective discomfort in prolonged viewing of chromatic displays. *Human Factors*, 31, 259-271.
- Matthews, M.L. (1989). Visual performance with coloured CRT displays: research update. *Applied Ergonomics*, 20, 58.
- Menozzi, M., Lang, F., Naepflin, U., Zeller, C. & Krueger, H. (2001). CRT versus LCD: Effects of refresh rate, display technology and background luminance in visual performance. *Displays*, 22, 79-85.
- Mills, C.B. & Weldon, L.J. (1987). Reading text from computer screens. *ACM Computing Surveys*, 19, 329-358.
- Milosevic, S. (1997). Drivers' fatigue studies. *Ergonomics*, 40, 381-389.
- Monk, A.F. & Hulme, C. (1983). Errors in proofreading: Evidence for the use of word shape in word recognition. *Memory and Cognition*, 11, 16-23.
- Montano, N., Ruscone, T.G., Porta, A., Lombardi, F., Pagani, M. & Malliani, A. (1994). Power spectrum analysis of heart rate variability to assess the changes in sympathovagal balance during graded orthostatic tilt. *Circulation*, 90, 1826-1831.
- Morell, M.A. (1989). Psychophysiological stress responsivity in Type A and B female college students and community women. *Psychophysiology*, 26, 359-368.
- Moya Albiol, L., Salvador, A., Costa, R., Martinez Sanchis, S., Gonzalez Bono, E., Ricarte, J. & Arnedo, M. (2001). Psychophysiological responses to the Stroop task after a maximal cycle ergometry in elite sportsmen and physically active subjects. *International Journal of Psychophysiology*, 40, 47-59.
- Mulder, G. & Mulder, L.J. (1981). Information processing and cardiovascular control. *Psychophysiology*, 18, 392-402.
- Mulder, G. (1985). *Attention, effort and sinus arrhythmia: how far are we?* In Orlebeke, J.F., Mulder, G. & van Doornen, L.P.J. (Eds.). *The psychophysiology of cardiovascular control* (S. 407-424). New York: Plenum Press.
- Muter, P., Latremouille, S.A., Treurniet, W.C. & Beam, P. (1982). Extended reading of continuous text on television screens. *Human Factors*, 24, 501-508.

- Muter, P. & Maurutto, P. (1991). Reading and skimming from computer screens and books: The paperless office revisited? *Behaviour and Information Technology*, 10, 257-266.
- Naesaenen, R., Karlsson, J. & Ojanpaeae, H. (2001). Display quality and the speed of visual letter search. *Displays*, 22, 107-113.
- Ohsuga, M., Shimono, F. & Genno, H. (2001). Assessment of phasic work stress using autonomic indices. *International Journal of Psychophysiology*, 40, 211-220.
- Oshima, A., Miyano, H., Yamashita, S., Owashi, T., Suzuki, S., Sakano, Y. & Higuchi, T. (2001). Psychological, autonomic and neuroendocrine responses to acute stressors in the combined dexamethasone/CRH test: a study in healthy subjects. *J Psychiatr Res*, 35, 95-104.
- Pagani, M., Furlan, R., Pizzinelli, P., Crivellaro, W., Cerutti, S. & Malliani, A. (1989). Spectral analysis of R-R and arterial pressure variabilities to assess sympatho-vagal interaction during mental stress in humans. *J Hypertens*, 7, 14-15.
- Pagani, M., Mazzuero, G., Ferrari, A., Liberati, D., Cerutti, S., Vaitl, D., Tavazzi, L. & Malliani, A. (1991). Sympathovagal Interaction During Mental Stress. A Study using Spectral Analysis of Heart Rate Variability in Healthy Control Subjects and Patients with a Prior Myocardial Infarction. *Circulation*, 83, 43-51.
- Pagani, M., Rimoldi, O., Pizzinelli, P., Furlan, R. & Crivellaro, W. (1991). Assessment of the neural control of the circulation during psychological stress. *Journal of the autonomic nervous system*, 35, 33-41.
- Pastoor, S. (1990). Legibility and subjective preference for color combinations in text. *Human Factors*, 32, 157-171.
- Pawlak, U. (1986). Ergonomic aspects of image polarity. *Behaviour and Information Technology*, 5, 335-348.
- Perini, R., Milesi, S., Fisher, N.M., Pendergast, D.R. & Veicsteinas, A. (2000). Heart rate variability during dynamic exercise in elderly males and females. *European journal of applied physiology*, 82, 8-15.
- Pomeranz, B., Macaulay, R.J., Caudill, M.A., Kutz, I. & Adam, D. (1985). Assessment of autonomic function in humans by heart rate spectral analysis. *American journal of physiology*, 248, 151-153.
- Redondo, M. (1992). Decrements in heart rate variability during memory search. *International Journal of Psychophysiology*, 13, 29-35.

- Renaud, P. & Blondin, J.P. (1997). The stress of Stroop performance: Physiological and emotional responses to color-word interference, task pacing, and pacing speed. *International Journal of Psychophysiology*, 27, 87-97.
- Richenhagen, G., Prümper, J., & Wagner, J. (1998). *Handbuch der Bildschirmarbeit*. Neuwied; Kriftel: Luchterhand.
- Rohmert, W. (1984). Das Belastungs-Beanspruchungs-Konzept. *Zeitschrift für Arbeitswissenschaft*, 38, 193-200.
- Schandry, R. (1988). *Lehrbuch Psychophysiologie: körperliche Indikatoren psychischen Geschehens*. Weinheim: Psychologie-Verl.-Union.
- Scher, H., Furedy, J.J. & Heslegrave, R. (1984). Phasic T-wave amplitude and heart rate changes as indices of mental effort and task incentive. *Psychophysiology*, 21, 326-333.
- Schleifer, L.M. & Ley, R. (1994). End-tidal PCO₂ as an index of psychophysiological activity during VDT data-entry work and relaxation. *Ergonomics*, 37, 245-254.
- Schleifer, L.M. & Okogbaa, O.G. (1990). System response time and method of pay: cardiovascular stress effects in computer-based tasks. *Ergonomics*, 33, 1495-1509.
- Shieh, K.K., Chen, M.T. & Chuang, J.H. (1997). Effects of color combination and typography on identification of characters briefly presented on VDTs. *International Journal of Human Computer Interaction*, 9, 169-181.
- Shieh, K. K. & Chen, M.T. (1997). Effects of screen color combination, work-break schedule, and workplace on VDT viewing distance. *International Journal of Industrial Ergonomics*, 20, 11-18.
- Sims, J., Carroll D., Turner, J.R. & Hewitt, J.K. (1988). Cardiac and metabolic activity in mild hypertensive and normotensive subjects. *Psychophysiology*, 25, 172-178.
- Snyder, H.L. (1988). *Image quality*. In: Helander, M. (Ed.). *Handbook of Human-Computer Interaction*. Amsterdam: Elsevier.
- Steyer, R., Schwenkmezger, P., Notz, P. & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF)*. Göttingen: Hogrefe.
- Svebak, S. 1982. The effect of task difficulty and threat of aversive electric shock upon tonic physiological changes. *Biological Psychology*, 14, 113-128.
- Taptagaporn, S. & Saito, S. (1990). How display polarity and lighting conditions affect the pupil size of VDT operators. *Ergonomics*, 33, 201-208.

- Task Force (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, 93, 1043-1065.
- Travis, D.S., Bowles, S., Seton, J. & Peppe, R. (1990). Reading from color displays: A psychophysical model. *Human Factors*, 32, 147-156.
- Tulen, J.H., Moleman, P., Van Steenis, H.G. & Boomsma, F. (1989). Characterization of stress reactions to the Stroop Color Word Test. *Pharmacology, Biochemistry and Behavior*, 32, 9-15.
- Turner, J.R. & Carroll, D. (1985). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded exercise: Further evidence of metabolically-exaggerated cardiac adjustments? *Psychophysiology*, 22, 261-267.
- Turner, J.R., Carroll, D. & Courtney, H. (1983). Cardiac and metabolic responses to "space invaders": An instance of metabolically-exaggerated cardiac adjustment? *Psychophysiology*, 20, 544-549.
- Udris, I. & Frese, M. (1988). *Belastung, Stress, Beanspruchung und ihre Folgen*. In Frey, D., Hoyos, Graf C. & Stahlberg, D. (Hg.). *Angewandte Psychologie. Ein Lehrbuch* (S.427-447). München: Psychologie Verlags Union.
- Udris, I., Kraft, M. & Mussmann, C. (1992). *Warum sind „gesunde“ Personen „gesund“? Untersuchungen zu Ressourcen von Gesundheit (Forschungsprojekt „Personale und organisationale Ressourcen der Salutogenese.“, Bericht Nr. 1)*. Zürich: ETH, Institut für Arbeitspsychologie.
- Vögele, C. (1998). Serum lipid concentrations, hostility and cardiovascular reactions to mental stress. *International Journal of Psychophysiology*, 28, 167-179.
- Wahlström, J., Hagberg, M., Johnson, P.W., Svensson, J. & Rempel, D. (2002). Influence of time pressure and verbal provocation on physiological and psychological reactions during work with a computer mouse. *European journal of applied physiology*, 87, 257-263.
- Wesseling, K.H. & Settels, I.J. (1985). *Baromodulation explains short term blood pressure variability*. In Orlebeke, J.F., Mulder, G. & van Doornen, L.P.J (Eds.). *The psychophysiology of cardiovascular control* (S. 333-352) New York: Plenum Press.

- Wieland-Eckelmann, R., Baggen, R., Saßmannshausen, A., Schmitz, U., Schwarz, R., Ademmer, C. & Rose, M.: *Gestaltung beanspruchungsoptimaler Bildschirmarbeit. Grundlagen und Verfahren für die Praxis*. Bremerhaven: Wirtschaftsverlag NW, 1996.
- Wilkinson, R.T. & Robinshaw, H.M. (1987). Proof-reading: VDU and paper text compared for speed, accuracy and fatigue. *Behaviour and Information Technology*, 6, 125-133.
- Wood, R., Maraj, B., Lee, C.M. & Reyes, R. (2002). Short-term heart rate variability during a cognitive challenge in young and older adults. *Age Ageing*, 31, 131-135.
- Wright, P. & Lickorish, A. (1983). Proof-reading texts on screen and paper. *Behaviour and Information Technology*, 2, 227-235.
- Yamamoto, Y., Hughson, R.L. & Peterson, J.C. (1991). Autonomic control of heart rate during exercise studied by heart rate variability spectral analysis. *Journal of applied physiology*, 71, 1136-1142.