

2 Empirischer Teil

2.1 Methode

Um sowohl die Nutzung medialer Stimuli als auch die Lern- und Leistungsergebnisse in ausreichender Weise standardisieren zu können und den Einfluss möglicher Störvariablen gering zu halten, wurde einem laborexperimentellen Zugang der Vorzug gegenüber einem feldexperimentellen Ansatz gegeben. Aufgrund der beabsichtigten Erfassung irreversibler Lerneffekte und der zeitlichen Beanspruchung der Versuchsteilnehmer durch die einzusetzende Testdiagnostik kam hinsichtlich der zu realisierenden Treatmentbedingungen nur ein Between-Subject-Design in Frage. Wie dargestellt wurde, lässt die bisherige Forschungslage den Schluss zu, dass sich unterschiedliche mediale Darbietungsformen und Inhalte in unterschiedlicher Weise auf kognitive Prozesse auswirken können. Daher berücksichtigt das Untersuchungsdesign sowohl einen **Faktor A (Gewalt)** als einen **Faktor B (Medium)**. Der Faktor A (Gewalt) wurde zweifachgestuft variiert, indem gewalthaltige und gewaltneutrale Medien eingesetzt wurden. Der Faktor B (Medium) wurde zweifachgestuft variiert, indem als interaktives Medium Computerspiele und als nicht-interaktives Medium Filme eingesetzt wurden. Durch den Einbezug beider Faktoren können Interaktionen zwischen der Art des Mediums und dem dargebotenen Inhalt ermittelt werden. Zusätzlich wurde eine fünfte nicht-mediale Explorationsbedingung realisiert, um veranschaulichen zu können, wie sich die vier hinsichtlich Gewalt und Interaktivität variierten Medienbedingungen in ihren Auswirkungen jeweils von einer nicht-medialen Freizeitbeschäftigung unterscheiden (vgl. Tabelle 8).

Tabelle 8. Übersicht über das realisierte zweifaktorielle Untersuchungsdesign mit zusätzlicher nicht-medialer Explorationsbedingung

A ₁ neutral B ₁ Spiele X11	A ₂ gewalthaltig B ₁ Spiele X21	A ₁ neutral B ₂ Filme X12	A ₂ gewalthaltig B ₂ Filme X22	Nicht-mediale Explorations- bedingung
--	---	---	--	---

2.1.1 Stichprobe

Die Stichprobenauswahl orientierte sich zum einen am Ziel einer möglichst hohen Generalisierbarkeit der Befunde, zum anderen an zentralen forschungsethischen Rahmenbedingungen. Hinsichtlich der Merkmale Geschlecht, Bildungsgrad und Häufigkeit des Computerspielens sollte eine möglichst heterogene Stichprobe gewonnen werden. Wenngleich eine repräsentative Auswahl von Mitgliedern einer Population in Laborexperimenten im Sinne einer Zufallsstichprobe nicht zu realisieren ist, sollte durch Berücksichtigung dieser Merkmale zumindest sichergestellt werden, dass die einbezogenen Versuchspersonen nicht einer sehr homogenen Personengruppe entstammen.

Besondere Vorüberlegungen betrafen das Alter der einzubeziehenden Versuchspersonen. Zunächst ist festzuhalten, dass das im Hintergrund der Untersuchung stehende Erkenntnisinteresse nach dem Erklärungswert kognitiver Minderungshypothesen für den negativen Zusammenhang von Gewaltmediennutzung und Schulleistung den Einbezug kindlicher oder jugendlicher Teilnehmer nahelegt. Es war jedoch intendierte Zielsetzung, in dem Experiment Medien einzusetzen, die explizite Gewaltdarstellungen enthalten und die damit zum überwiegenden Teil nur für ein volljähriges Publikum freigegeben sind. Zwar ist bekannt, dass auch unter Heranwachsenden bereits Erwachsenenangebote genutzt werden (vgl. Kapitel 1.1.2).⁵⁶ Dieser Sachverhalt ändert jedoch nichts an der Tatsache, dass es forschungsethisch hochgradig problematisch wäre, minderjährige Probanden im Rahmen einer experimentellen Untersuchung Angeboten auszusetzen, bei denen von einem entwicklungsbeeinträchtigenden Potential auszugehen ist. Hieraus ergibt sich ein forschungsethisches und methodologisches Dilemma, für dessen Auflösung grundsätzlich zwei verschiedene Lösungsmöglichkeiten existieren: Die erste Lösung besteht darin, bei den gewalthaltigen Medienbedingungen in der Weise gewaltreduzierte Angebote zu wählen, dass keine jugendbeeinträchtigenden Wirkungen mehr befürchtet werden müssen, indem beispielsweise ausschließlich Angebote mit einer Altersfreigabe ab 12 Jahren eingesetzt werden. In diesem Fall wäre jedoch die Konzeption eines Treatments nach dem intendierten Gewaltgrad nicht zu erreichen gewesen, da Angebote ab 12 Jahren in der Regel genau jene Merkmale nicht aufweisen, die im Experiment auf ihr Wirkpotential hin untersucht werden sollten. Eine zweite Lösung, die im Rahmen der vorliegenden Konzeption gewählt wurde, besteht darin, das Alter der Versuchspersonen soweit heraufzusetzen, dass jugendschutzrechtlichen und damit forschungsethischen Bedenken Rechnung ge-

56 Dies zeigt sich insbesondere bei männlichen Jugendlichen: So nutzen unter 15-Jährigen 26 Prozent regelmäßig Spiele (Mädchen: 2,2 %) und rund 23 Prozent häufig Filme ab 18 Jahren (Mädchen: 6,1 %). Unter 10-jährigen Jungen schauen 30 Prozent regelmäßig Fernsehinhalte ab 16 Jahren (Mädchen: 14 %); rund 21 Prozent der Jungen spielen regelmäßig Spiele ab 16 oder 18 Jahren (Mädchen: 3 %, vgl. Mößle, et al., 2007).

tragen wird. Es wurde daher entschieden, ein Mindestalter von 18 Jahren für die Teilnahme am Versuch vorauszusetzen. Im Interesse einer altershomogenen Untersuchungspopulation junger Erwachsener wurde die Altersobergrenze auf 25 Jahre festgelegt. Die Entscheidung zugunsten einer erwachsenen Stichprobe bedingt, dass die Teilnehmer mit größerer Wahrscheinlichkeit bereits über eingehende Erfahrungen mit gewalthaltigen Medien verfügen und damit auch mit geringerer emotionaler Beteiligung auf diese reagieren könnten. Damit besteht grundsätzlich die Möglichkeit, dass emotionale und kognitive Wirkungen in jüngeren Alterspopulationen mittels einer erwachsenen Stichprobe nicht zuverlässig aufgedeckt werden können. Dieser Einschränkung, die eine konservative Testung des emotionalen und kognitiven Wirkpotentials gewalthaltiger Medien erwarten lässt, sollte jedoch damit begegnet werden, dass Moderatorvariablen des Distresserlebens berücksichtigt werden und damit auch Aussagen zu Versuchspersonen möglich sind, die in besonders hohem Maße auf gewalthaltige Medien reagieren (vgl. Kapitel 2.1.6).

Neben den genannten Einschlusskriterien wurden die folgenden Ausschlusskriterien definiert:

- Geringe deutsche Sprachkenntnisse
- Nicht durch Sehhilfen zu korrigierende Sehbeeinträchtigung
- Aktuell vorliegende Suchterkrankung
- Neigung zu lichtreizbedingten Krampfanfällen (fotosensitive Epilepsie)
- Starke Ängste und Abneigungen gegenüber medialen Gewaltdarstellungen

Eine eingeschränkte Kenntnis der deutschen Sprache hätte ein mangelndes Instruktionsverständnis, eine starke Sehbeeinträchtigung keine zuverlässige diagnostische Urteilsbildung und eine aktuell vorliegende Suchterkrankung unvorhersehbare Schwankungen in der kognitiven Leistungsfähigkeit erwarten lassen. Eine Neigung zu lichtreizbedingten Krampfanfällen wurde als gesundheitlicher Risikofaktor gewertet, der eine Partizipation am Versuch aus forschungsethischen Gründen nicht vertretbar erscheinen ließ, da die Versuchspersonen über einen vergleichsweise langen Zeitraum Medienangeboten mit schnellen Bildwechseln ausgesetzt werden sollten. Bei Personen mit besonders starken Ängsten bzw. gravierenden Abneigungen gegenüber medialen Gewaltdarstellungen hätte befürchtet werden müssen, dass die Angebote auf derartig starken Widerstand stoßen würden, dass die Medienrezeption auf inadäquate Weise verfälscht würde und in der Folge artifizielle Ergebnisse begünstigt würden.⁵⁷ Dies hätte zudem den Ablauf der Untersuchungen maßgeblich stören können.

⁵⁷ Hierbei galt es selbstverständlich, nur Probanden mit besonders extremer Abneigung auszuschließen und nicht etwa Personen, die einfach nur "ungern" Gewaltmedien nutzen. Ein Ausschluss dieser Extremgruppe wurde in Hinblick auf die vorliegenden Fragestellungen als

Es wurde eine Gesamtstichprobengröße von $N = 360$ Versuchspersonen angestrebt. Die dieser Entscheidung vorausgegangenen Überlegungen betrafen forschungsökonomische Gesichtspunkte sowie die zu erwartenden Effektgrößen. In Ermangelung vergleichbarer Studien konnte ausgehend von Daten zur Wirkung medialer Gewalt auf andere affektive und kognitive Parameter zumindest davon ausgegangen werden, dass hinsichtlich eventuell vorliegender kurzfristiger Wirkungen von Mediengewalt auf kognitive Leistungsparameter allenfalls mittlere Effekte erwartet werden können.⁵⁸ In Hinblick auf den Untersuchungsgegenstand wurde jedoch auch kleinen Effekten eine praktische Relevanz zugesprochen, da experimentell nur kurzfristige Ausschnitte von kognitiven Prozessen beobachtet werden können, die in der Alltagswirklichkeit im Rahmen langfristiger Lernprozesse kumulativ aufeinander aufbauen (vgl. Kapitel 1.4.3). In Erwägung der Stichprobengrößen, die für die Aufdeckung unterschiedlicher Populationseffekte notwendig geworden wären, wurde schließlich entschieden, dass Effekte ab $f \geq .15$ mit einer Teststärke von $1 - \beta \geq .90$ aufgedeckt werden sollen.⁵⁹

Die Akquise der $N = 360$ Versuchspersonen orientierte sich an den bereits genannten Merkmalen Geschlecht, Bildungsgrad (hoch, niedrig) und Häufigkeit des Computerspielens (nie: 0h/Woche, wenig: $< 6h/Woche$, viel: $\geq 6h/Woche$). Durch die Parallelisierung wurden in der Endauswahl jeweils 30 bezüglich dieser Variablen homogene Versuchspersonen angestrebt, die den Untersuchungsgruppen randomisiert zugewiesen wurden (vgl. Tabelle 9).

unproblematisch erachtet, weil von diesen Personen ohnehin nicht zu erwarten ist, dass sie sich im Alltag überhaupt gewalthaltigen Medien zuwenden.

- 58 Per Konvention gelten in Varianzanalysen Effektgrößen ab $f = .10$ als klein, Effektgrößen ab $f = .25$ als mittelgroß und Effektgrößen ab $f = .40$ als groß (vgl. Bortz & Döring, 2003; Cohen, 1988).
- 59 In Orientierung an gängige Konventionen wurde hierbei ein Alphafehlerniveau von $\alpha = .05$ festgelegt. Testpoweranalysen mit dem Programm GPOWER 2 (Erdfelder, Faul, & Buchner, 1996) führten zu der Entscheidung, die Wahl der Stichprobengröße an einer kleinen bis mittleren Effektgröße von $f \geq .15$ auszurichten. Da jedoch einige Parameter wie insbesondere die Interkorrelation zwischen den Messzeitpunkten der Testverfahren vorab nicht bekannt waren, wurde auf Basis grober Schätzungen entschieden, pro Treatmentgruppe mindestens 70 Versuchspersonen einzubeziehen (entspricht $N = 280$ für die vier Mediengruppen und weiteren $N = 70$ für die Explorationsbedingung). Dieses Vorgehen ermöglicht zudem die Berücksichtigung eines ergänzenden dichotomen Gruppierungsfaktors, womit sich pro Gruppe immer noch ca. $n = 35$ Versuchspersonen ergeben würden. Um gleichmäßige Zellbesetzungen zu ermöglichen (vgl. Tabelle 9) wurde eine Gesamtstichprobengröße von $N = 360$ festgelegt. Da die a priori Bestimmung der Stichprobe nicht anhand exakter Kriterien vorgenommen werden konnte wird bei der Darstellung der Ergebnisse eine ergänzende Post-Hoc-Poweranalyse durchgeführt.

Tabelle 9. Zielvorgabe für die Untersuchungsstichprobe (N = 360)

180 Männer						180 Frauen					
Matching nach Häufigkeit des Computerspielens (trichotomisiert)											
nie (60)		wenig (60)		viel (60)		nie (60)		wenig (60)		viel (60)	
Matching nach Bildungsgrad (dichotomisiert)											
HB (30)	NB (30)	HB (30)	NB (30)	HB (30)	NB (30)	HB (30)	NB (30)	HB (30)	NB (30)	HB (30)	NB (30)

Anmerkung. HB = hohe Bildung (Abitur, Fachhochschulabschluss oder Hochschulabschluss vorliegend oder angestrebt), NB = niedrige Bildung (kein Abitur, Fachhochschulabschluss oder Hochschulabschluss vorliegend oder angestrebt). Wenigspieler < 6 h/Woche, Vielspieler ≥ 6h/Woche.

Die Untersuchung wurde über Flyer, Plakate, Zeitungsannoncen, Mitteilungen im Internet (z. B. über die Social-Networking Plattform www.Lokalisten.de) angekündigt (Anwerbeflyer vgl. Anhang C). Jedem Teilnehmer wurde für die vollständige Teilnahme an beiden Untersuchungstagen eine pauschale Versuchspersonenentschädigung in Höhe von 40,- Euro angeboten. Als zusätzliches Incentive wurden die Interessenten auf eine Verlosung von Preisen unter allen Teilnehmern (fünf Geschenkgutscheine von Amazon im Gesamtwert von 210,- Euro) hingewiesen.

Die Kontaktaufnahme verlief primär über eine eigens dafür eingerichteten Homepage, auf der sich Interessenten über die Studie informieren, per einfachem Mausklick ihr Interesse an einer Teilnahme bekunden und unmittelbar einen kurzen Onlinefragebogen ausfüllen konnten. Dieser enthielt allgemeine soziodemographische Fragen sowie Fragen zum Computerspielverhalten. Hier konnten die Interessenten auch ihre Telefonnummer für einen Rückruf hinterlassen. Interessenten, die die Kriterien für die Studie erfüllten, wurden anschließend von einem geschulten Interviewer zurückgerufen. Alternativ konnten Interessenten die Studienleitung auch direkt über eine eigens dafür eingerichtete Hotline erreichen. In beiden Fällen erfolgte im Rahmen des telefonischen Erstkontaktes eine Befragung über ein standardisiertes Interview, in welchem die Einschluss- und Ausschlusskriterien für die Teilnahme abgeklärt wurden. Die Angaben der betreffenden Personen wurden dabei unmittelbar in eine Datenbank aufgenommen. Diese Datenbank wurde so eingerichtet, dass dem Interviewer die Eignung oder Nichteignung eines Teilnehmers von der Software in Echtzeit zurückgemeldet wurde. Gleichfalls stellte die Software eine zufällige Zuweisung zu einer der Haupttreatmentbedingungen sicher (Filme, Computerspiele, oder nicht-mediale Explorationsbedingung). Geeigneten Personen wurde unmittelbar ein Wunschtermin in einem der für

die entsprechende Haupttreatmentbedingung vorab definierten Untersuchungsblöcke angeboten.⁶⁰

Im Anschluss an die Terminvergabe wurden die Kandidaten zunächst telefonisch, und einige Tage später zusätzlich schriftlich (vgl. Anhang D), über den weiteren Ablauf der Untersuchung und die Bedingungen, die mit einer Auszahlung der Versuchspersonenentschädigung verknüpft sind, informiert.⁶¹ Zudem wurde eine Checkliste mit Verhaltensregeln beigelegt, sowie eine Einverständniserklärung, die von der Versuchsperson zu unterschreiben und zur Untersuchung mitzubringen war. Hierin erklärten die Versuchspersonen unter anderem, dass Mindestalter von 18 Jahren erreicht zu haben. Das Anschreiben wies zudem erneut auf den vereinbarten Untersuchungstermin und den Ort der Untersuchung hin und enthielt zudem eine Anfahrsbeschreibung. Lag zwischen Terminvereinbarung und Termin ein längerer Zeitraum (länger als einige Tage), wurden am Vortag der Untersuchung zusätzliche Reminder (Anrufe oder SMS) eingesetzt, um die betreffenden Personen aktuell an ihren anstehenden Untersuchungstermin zu erinnern.

Die Datenerhebung fand in einem Viermonatszeitraum von Juli bis Oktober 2006 an insgesamt 41 Datenerhebungstagen statt. Dabei wurden im Schnitt an jedem Versuchstag 9 Versuchspersonen untersucht. Pro Untersuchungstermin ergab sich eine durchschnittliche Auslastung von 4,7 der insgesamt 6 Teilnehmerplätze, womit eine ökonomische Versuchsdurchführung erreicht wurde. Insgesamt wurden Datensätze zu $N = 369$ Versuchspersonen gewonnen (vgl. Tabelle 10).

Tabelle 10. Realisierte Stichprobe ($N = 369$)

185 Männer						184 Frauen					
Matching nach Häufigkeit des Computerspielens (trichotomisiert)											
nie (60)		wenig (59)		viel (66)		nie (62)		wenig (61)		viel (61)	
Matching nach Bildungsgrad (dichotomisiert)											
HB (32)	NB (28)	HB (29)	NB (30)	HB (34)	NB (32)	HB (30)	NB (32)	HB (31)	NB (30)	HB (31)	NB (30)

Anmerkung. HB = hohe Bildung (Abitur, Fachhochschulabschluss oder Hochschulabschluss vorliegend oder angestrebt), NB = niedrige Bildung (kein Abitur, Fachhochschulabschluss oder Hochschulabschluss vorliegend oder angestrebt). Wenigspieler < 6 h/Woche, Vielspieler $\geq$ 6h/Woche.

60 Die Zufallszuweisung der Versuchspersonen zur gewalthaltigen oder gewaltneutralen Bedingung erfolgte immer erst unmittelbar vor Versuchsbeginn mittels des Verfahrens "Ziehen ohne Zurücklegen" durch die Versuchsassistenten. Damit ist die Zuweisung der Probanden als Parallelisierung mit anschließender Randomisierung zu verstehen.

61 Es wurde darauf hingewiesen dass die Versuchspersonenentschädigung nur bei vollständiger Teilnahme an beiden Untersuchungsblöcken ausbezahlt wird.

50 Prozent der Teilnehmer sind männlichen Geschlechts. Das Alter der Probanden beträgt 18 bis 25 Jahre ($M = 20.65$, $SD = 2.20$). Rund 51 Prozent weisen ein hohes Bildungsniveau auf. 33 Prozent der Probanden sind Nichtspieler, 33 Prozent Gelegenheitsspieler und 34 Prozent Vielspieler. Durch die Rekrutierung nach Parallelisierungsmerkmalen konnten die angezielten Zellbesetzungen annähernd erreicht werden, wenngleich sich insbesondere hinsichtlich der Rekrutierung niedriggebildeter und gleichzeitig nichtspielender Männer Schwierigkeiten in der Anwerbung ergaben und hier zwei Versuchspersonen weniger als vorgesehen für die Teilnahme gewonnen werden konnten. Tabelle 11 ist die Anzahl von Versuchspersonen pro Treatmentbedingung zu entnehmen.

Tabelle 11. Zuordnung der Versuchspersonen zu den Versuchsbedingungen

		Faktor A: Gewalt		
		gewalthaltig	gewaltneutral	
Faktor B: Medium	Spiele	n = 75	n = 75	n = 150
	Filme	n = 73	n = 74	n = 147
	nicht-medial	X	X	n = 72
		n = 148	n = 149	N = 369

2.1.2 Operationale Definitionen der unabhängigen Variablen (UV)

Bei der Konzeption des Untersuchungsdesigns sind in besonderer Weise Erkenntnisse berücksichtigt worden, die sich aus zwei wesentlichen Problembereichen zum Einsatz von Unterhaltungsmedien als experimentelles Stimulusmaterial ergeben. Der erste Bereich betrifft den Umgang mit motivationalen Aspekten der Nutzung von Unterhaltungsmedien und der zweite Bereich grundsätzliche Probleme, die sich beim Einsatz interaktiven Stimulusmaterials innerhalb einer laborexperimentellen Versuchsanordnung ergeben.

Hinsichtlich motivationaler Aspekte besteht bei Medienwirkungsstudien grundsätzlich die Gefahr, dass im Labor artifizielle Rezeptionsprozesse nachgebildet

werden, die der Lebenswirklichkeit und Interessenlage des Probanden nicht entsprechen. Dieses Problem besteht bei Unterhaltungsmedien in besonderer Weise, da diese im privaten Leben freiwillig, stimmungsabhängig und den allgemeinen Ansprüchen entsprechend eingesetzt werden. So werden sich einige Probanden möglicherweise niemals freiwillig im privaten Leben gewalthaltigen medialen Stimuli aussetzen. Um dies konsequent zu berücksichtigen, müssten im Rahmen von experimentellen Untersuchungen die Probanden frei entscheiden dürfen, welche Angebote sie während der Treatmentphase nutzen möchten, womit jedoch keine Randomisierung der Versuchspersonen zu bestimmten Versuchsbedingungen, sondern eine Selbstselektion des Stimulusmaterials gegeben wäre (vgl. Reinecke & Trepte, 2008). Ein solches Design wäre kein experimentelles Forschungsdesign mehr und ließe damit auch keine Aussagen über die Kausalität der Wirkbeziehungen zu.

Ein zweiter zentraler Problemkomplex betrifft den Einsatz von Computerspielen als interaktives Stimulusmaterial, der spezielle Ansprüche an das Forschungsdesign stellt. So muss zum einen bedacht werden, dass bei der Realisierung gewalthaltiger Computerspielbedingungen eine Vergabe des Gewaltstimulus nicht selbstverständlich garantiert ist, da sich unterschiedliche Probanden je nach Vorlieben und eigenem Kompetenzlevel deutlich darin unterscheiden könnten, wie sie das Spiel nutzen und damit auch inwieweit sie überhaupt gewalthaltige Handlungen ausführen (Klimmt, Hartmann, & Vorderer, 2005; Klimmt & Trepte, 2003, S. 118).⁶² Zudem muss beim Einsatz gewalthaltigen und gewaltneutralen Stimulusmaterials unabhängig vom Medium darauf geachtet werden, dass sich die gewaltneutralen Bedingungen neben der Gewalt nicht durch weitere zentrale Merkmale von den gewalthaltigen Bedingungen unterscheiden, indem diese beispielsweise deutlich weniger unterhaltsam sind (Klimmt & Trepte, 2003, S. 118; Klimmt, Vorderer, & Ritterfeld, 2004).

Nachfolgend wird vor dem Hintergrund dieser Überlegungen die Operationalisierung der unabhängigen Variablen beschrieben. Für die Identifikation gewalthaltiger vs. gewaltneutraler Medien wurde unter Rückgriff auf die Literatur eine für diese Untersuchung geeignete Operationalisierung von Mediengewalt abgeleitet (vgl. Kapitel 2.1.2.1). Der Auswahl interaktiver (Computerspiele) vs. nicht-interaktiver Medien (Filme) gingen Überlegungen zum geeigneten Interaktions-

62 Ein Beispiel hierfür ist das Spiel *Grand Theft Auto IV*. Hier kann sich der Spieler z. B. entscheiden, durch die Stadt zu fahren, Billard oder Bowling zu spielen oder Kleidung einzukaufen. Ebenso ist es jedoch möglich, sich die Zeit mit einem Amoklauf in der Innenstadt oder der Eliminierung von Polizisten zu vertreiben (Rehbein, Kleimann, & Mößle, 2008). Die hochgradige Interaktivität lässt bezüglich anderer unabhängiger Variablen kein einheitliches Rezeptionsgeschehen zu, so dass in diesem Fall nicht mehr zu garantieren wäre, dass die Spieler sich einem gewalthaltigen Spielszenario aussetzen.

grad der Computerspiele voraus (vgl. Kapitel 2.1.2.2). An beiden Kriterien orientierten sich Auswahl und Präsentation des Stimulusmaterials (vgl. Kapitel 2.1.3).

2.1.2.1 Faktor A: Gewalt

Die Identifizierung geeigneten Stimulusmaterials setzte zunächst ein Verständnis darüber voraus, was genau unter einem gewalthaltigen und einem gewaltneutralen Medium zu verstehen ist. Zugunsten einer erhöhten Konstruktvalidität wurde das Ziel verfolgt, diese Unterscheidung unter Anwendung explizit formulierter Kriterien vorzunehmen, die sowohl für Computerspiele als auch Filme anwendbar sind. Wie bereits in Kapitel 1.1.2 ausgeführt, operieren die inhaltsanalytischen Studien zur Mediengewalt mit unterschiedlichen Definitionen, die in der Regel eine recht weite Auslegung des Gewaltbegriffs erkennen lassen. So werden auch regelgeleitete, sportliche Wettkämpfe, die Zerstörung von Sachgegenständen, psychische Gewalt, die Beeinträchtigung nichtmenschlicher Lebewesen oder ausschließlich verbale Androhungen von Gewalt als gewalthaltige Interaktionen behandelt und in differenzierter Weise eingeordnet und bewertet (vgl. P. Grimm, et al., 2005; *National Television Violence Study 1*, 1997; *National Television Violence Study 3*, 1998). Zudem unterscheiden die meisten vorhandenen Definitionen nicht nach verschiedenen Darstellungsformen wie comichaft-verfremdeter vs. realitätsnaher Darstellung. In Hinblick auf die Zielsetzung, Medienangebote mit explizit visualisierten und akustisch untermalten Gewaltdarstellungen einzusetzen, erschienen diese Zugänge für den vorliegenden Anwendungszweck zu weit gefasst. Auch gängige Dichotomisierungen zur Beschreibung von Gewalt wie etwa real/fiktional (z. B. Nachrichten vs. Actionfilm), natürlich/künstlich (Schauspiel vs. Animation), „sauber“/„schmutzig“ (bezogen auf die Darstellung von Gewaltfolgen) oder auch legitim/illegitim („Der gute Aggressor“ vs. „Der böse Aggressor“) reichen für den vorliegenden Forschungsgegenstand nicht aus. Für das Experiment wurde daher die folgende Gewaltdefinition entwickelt:

<< Unter medialer Gewalt wird die fiktional-mediale Inszenierung einer intentionalen, ernsthaften körperlichen Schädigung anderer Menschen oder menschenähnlicher Wesen verstanden, die von einem oder mehreren menschlichen Angreifern ausgeht und deren Folgen durch akustische und visuelle Effekte explizit dargestellt und dabei nicht in besonderer Weise stilistisch verfremdet werden >>

Diese Definition rekuriert damit auf körperliche Angriffe, die von Menschen gegenüber anderen Menschen ausgehen und schließt dabei auch ein, dass die Konsequenzen der Gewalthandlungen explizit dargestellt werden müssen. Damit gelten

Medien dann als gewalthaltig, wenn die in ihnen vorhandenen Gewaltereignisse die Merkmale des folgenden Kriterienkatalogs vollständig aufweisen:

1. Es handelt sich um fiktionale Gewalt
2. Die Gewalt geht von menschlichen oder menschenähnlichen Wesen aus (personalisierte Gewalt)
3. Die Gewalt richtet sich gegen menschliche oder menschenähnliche Wesen (personale Gewalt)
4. Die Gewalt ist nicht gegen sich selbst gerichtet (keine Autoaggressivität)
5. Die Gewalt ist intentional (willentlich durch den Angreifer herbeigeführt)
6. Die Gewalt zielt auf eine ernsthafte physische Schädigung des Opfers ab oder wird in dieser Weise wirksam (schwere Verletzung, physische Folter oder Tod)
7. Die Gewaltfolgen werden durch visuelle und akustische Effekte explizit hervorgehoben
8. Die Darstellung der Gewalt ist naturalistisch ⁶³

Die Auswahl geeigneten Stimulusmaterials orientierte sich neben dieser Gewaltdefinition an weiteren Gesichtspunkten. So wurden nur Medien mit einer **hohen Quantität von Gewaltereignissen** aufgenommen. Eine hohe Gewaltquantität wurde dann als vorliegend erachtet, wenn der Anteil der Gewalt gemessen an der Gesamtlänge eines Films bzw. am Anteil gewaltfreier Handlungen innerhalb eines Computerspiels dominiert. Die eingesetzten Medien mussten somit eine Vielzahl von Gewalthandlungen im Sinne der Gewaltdefinition aufweisen (Filme) oder für den Spielerfolg notwendig machen (Computerspiele). Zudem wurde für gewalthaltige Medien vorausgesetzt, dass sie eine **Alterseinstufungen ab 16 oder ab 18 Jahren** aufweisen. Zugrundegelegt wurden hier die Bewertungen der Freiwilligen Selbstkontrolle der Filmwirtschaft (FSK) bzw. für die Computerspiele die Kennzeichen der Unterhaltungssoftware Selbstkontrolle (USK). Ferner wurde als wesentliches Merkmal definiert, dass der Darstellung von Gewalt eine eindeutige **Unterhaltungsentention** zugrundeliegen muss und diese **keine moralische Problematisierung** implizieren darf. Medien, die vorwiegend einen edukativen Anspruch geltend machen und Gewalt (sozial-)kritisch thematisieren, wurden als ungeeignet bewertet. Für eine Einlösung dieses Unterhaltungsanspruchs mussten die Gewalthandlungen in stimmiger Weise in eine **narrative Rahmenhandlung** eingebettet sein. Dies bedeutet, dass das Angebot nicht aus einer simplen Aneinanderreihung von Gewaltszenen besteht, sondern dass diese innerhalb einer autonomen, kompakten und für ein breites Publikum verständlichen Narration dargeboten

63 Keine naturalistische Darstellung liefern Trick- und Animationsfilme und sind damit ungeeignet. Bei interaktiven Medien muss eine größtmöglichen Annäherung an eine realitätstreue Umsetzung erkennbar sein.

werden. Zudem wurde festgelegt, dass die Darstellung der Gewalt auf eine **realistische Vermittlung von Gewalthandlungen und Folgen** abzielt, die für den Rezipienten nicht losgelöst von der physikalischen Realität (Naturgesetze) existiert. So durfte die technisch-ästhetische Umsetzung von Gewalt nicht so gestaltet sein, dass sie per se lächerlich, unglaubwürdig oder parodistisch wirkt. Für die Computerspiele wurde zudem festgelegt dass **der Spieler primär in der Rolle des Angreifers** wirksam werden muss und nicht in erster Linie Opfer von Gewaltereignissen wird.⁶⁴ Eine vergleichsweise hohe Offenheit in der Auswahl bestand hinsichtlich des Gewaltlegitimationskontextes. So konnte es sich um ungerechtfertigte Gewalt oder auch scheinbar oder offensichtlich legitimierte Gewalt (z. B. Soldat im Kriegseinsatz) handeln, solange diese in ein glaubwürdiges und stimmiges Szenario eingebettet ist.

Weiterhin wurde für die gewaltneutralen Medien eine sehr enge Auslegung angesetzt, indem diese keinerlei Formen physischer Gewalt gegen andere Menschen, menschenähnliche Wesen oder Tiere/Monster, belebte Maschinen (Roboter, Androiden) oder Fahrzeuge, in denen sich Menschen aufhalten könnten, beinhalten durfte.⁶⁵

2.1.2.2 Faktor B: Medium

Eine systematische Manipulation von Interaktivität innerhalb experimenteller Treatmentbedingungen mag zunächst unproblematisch erscheinen: Computerspiele sind immer interaktive Medien und klassische Filmformate sind immer nicht-interaktive Medien. Zum Zwecke dieser Studie war die Auswahl des interaktiven Stimulusmaterials jedoch hinsichtlich des Grades der angezielten Interaktivität zu präzisieren. So kann sich die Interaktivität von Computerspielen von einfachen, sehr beschränkten Spielszenarien bis zu hochgradig offenen Spielwelten mit einer Vielzahl von Spieloptionen erstrecken. Die Wahl eines hochgradig offenen Spielszenarios hätte mit hoher Wahrscheinlichkeit auch sehr differenzierte Rezeptionsgeschehen zwischen verschiedenen Versuchspersonen zur Folge. Auch wenn es richtig ist, dass die naturgemäße Interaktivität von Computerspielen generell kein einheitliches Rezeptionsgeschehen ermöglicht (vgl. Klimmt, et al., 2004), sollte das Verhalten der Spieler für dieses Experiment zumindest in der Weise kontrolliert

64 Ein Beispiel hierfür sind Spiele der Kategorie "Survival-Horror", bei dem der Spieler häufig unter stark begrenzten Angriffsmöglichkeiten versuchen muss, in einem feindlichen Szenario zu überleben.

65 Das Vorgehen orientierte sich damit an dem in der experimentellen Versuchsplanung anerkannten Max-Kon-Min-Prinzip, indem die Primärvarianz zwischen den Faktoren (gegenüber der Fehlervarianz) möglichst hoch ausfallen soll, ohne dass jedoch artifizielle und wenig alltagsnahe Versuchsbedingungen realisiert werden (Sarris & Reiß, 2005, S. 206).

werden, dass der Stimulus „mediale Gewalt“ in jedem Fall in angemessener Intensität garantiert werden kann und gleichfalls sichergestellt ist, dass gewaltneutrale Spiele keine versteckten violenten Handlungsmöglichkeiten beinhalten.⁶⁶ Das interaktive Medium sollte daher so ausgelegt sein, dass dem Spieler keine Alternative zur Ausübung von Gewalt bzw. zum Verzicht auf Gewalt zur Erreichung seiner Spielziele zur Verfügung steht. Deshalb erschienen gewalthaltige Spiele, die dem Spieler auch gewaltreduzierte, der Gewaltdefinition nicht entsprechende Lösungsmöglichkeiten wie eine Verhaftung oder Betäubung der Gegner offen lassen (z. B. Taktikshooter wie die *SWAT*- oder *Rainbow-Six*-Reihe), nicht geeignet. Auch Spiele, die ein längerfristiges Ausweichen oder Verstecken vor Gegnern erlauben, erschienen ungeeignet, denn hierdurch könnten Spieler durchgängig gewaltfreien Lösungsstrategien den Vorzug geben und dadurch nicht wie intendiert Mediengewalt rezipieren.

Ein weiterer Aspekt der Interaktivität, welcher eine einheitliche Stimulusvergabe im Experiment erschweren kann, betrifft den Erfolgsausgang der Computerspielnutzung. So wird darauf hingewiesen, dass Versuchspersonen sehr unterschiedliche Erfolge in Computerspielen erzielen können (vgl. Klimmt, et al., 2005). Wenngleich hierzu noch überzeugende empirische Belege ausstehen, wird vom Erfolgsausgang und der Beherrschung des Spielgeschehens vermutet, dass diese Faktoren in komplexer Weise mit dem Unterhaltungserleben verknüpft sind, indem sowohl Unter- als auch Überforderung das Unterhaltungserleben negativ beeinflussen können (vgl. Klimmt, 2006; Klimmt, Hartmann, & Frey, 2007). Zudem gilt bei der Rezeption gewalthaltiger Computerspiele, dass auch die Quantität und Qualität der expliziten Gewaltdarstellungen unmittelbar vom Spielerfolg des Spielers abhängen können. Wenn ein Spiel in einer Weise komplex oder so schwierig ist, dass einige Versuchspersonen gar nicht in der Lage sind, virtuelle Gewalt auszuüben, dann entspräche das Treatment für diese Personen auch keinem Gewalt-treatment. Damit würden Versuchspersonen mit hoher Spielexpertise selektiv einer besonders hohen Dosis von Mediengewalt ausgesetzt sein, was hinsichtlich der Zielsetzung, differenzielle Wirkungen von Mediengewalt auf unterschiedliche Subpopulationen zu ermitteln, als problematisch erachtet wurde.

Den genannten Problemen wurde durch unterschiedliche Maßnahmen Rechnung getragen. Zugunsten einer Annäherung der einzelnen Rezeptionsverläufe wurden Spiele ausgewählt, die ein abgegrenztes Spielszenario beinhalteten, welches vom Spieler nicht verlassen werden konnte. Dieses Spielszenario durfte ausschließlich dem intendierten Stimulus entsprechende Spielhandlungen ermögli-

66 So könnten zum Beispiel bei einem Rennspiel mit anderen beteiligten Fahrzeugen einige Spieler dazu neigen, andere Fahrzeuge von der Straße zu drängen oder in Unfälle zu verwickeln, auch wenn das Spiel eigentlich einen fairen und gewaltfreien Rennwettbewerb vorsieht.

chen und keine optionale Vielfalt sehr differenzieller Spielhandlungen aufweisen. Zusätzlich wurde allen Spielern ein einheitlicher Level vorgegeben (z. B. eine einheitliche Rennstrecke oder ein bestimmtes Kampfszenario), welcher vom Spieler nicht verändert bzw. verlassen werden durfte. Dieses Vorgehen entspricht damit der von Klimmt vorgeschlagenen experimentellen Strategie einer Realisierung invarianter Inseln innerhalb des komplexen Angebotsraums des interaktiven Mediums (Klimmt, et al., 2004, S. 149). Ein einheitlicherer Spielverlauf zwischen den Probanden wurde zusätzlich dadurch begünstigt, dass die einzusetzenden Spiele einen mittleren Schwierigkeitsgrad und eine schnelle Zugänglichkeit aufweisen sollten, um auch unerfahrenen Personen einen schnellen Einstieg zu ermöglichen. Auf Spiele mit einer sehr komplexen Steuerung wurde grundsätzlich verzichtet.⁶⁷ Ganz im Gegenteil wurden Angebote gewählt, die eine leicht zugängliche Steuerung mit besonders hohem Involvementpotential aufweisen, indem z. B. bei einem Rennspiel ein authentisches Lenkrad mit Fußpedalen eingesetzt wurde und zwei Shooterspiele mit Lightgunkompatibilität gewählt wurden, bei denen der Spieler mit einer Lichtpistole direkt auf den Bildschirm zielen kann. Bei der Auswahl der Spiele wurde jedoch ebenso Wert darauf gelegt, dass sich erfahrene Spieler durch die angebotenen Spiele ebenfalls angesprochen und herausgefordert fühlen. Daher sollten die ausgewählten Spiele trotz schneller Zugänglichkeit Spielraum für den Einsatz besonderer Fertigkeiten lassen. Um die Vorkenntnisse der Versuchspersonen anzugleichen, wurden großflächige Hinweistafeln vorgesehen, die Auskunft über das Spielziel und die Steuerung liefern sollten. Damit lässt sich der Kriterienkatalog für den angezielten Interaktivitätsgrad der Computerspielbedingungen wie folgt zusammenfassen:

- Der Interaktionsrahmen lässt in der Gewaltbedingung ausschließlich violente Handlungen zu
- Der Interaktionsrahmen lässt in der Neutralbedingung ausschließlich nicht-violente Handlungen zu
- Der Interaktionsrahmen bewegt sich innerhalb eines klar definierten Spielszenarios mit einer überschaubaren Anzahl von Handlungsmöglichkeiten
- Ein mittlerer Schwierigkeitsgrad, eine leichte und involvierende Steuerung sowie anschauliche Instruktionen ermöglichen auch bei geringen Vorkenntnissen einen sofortigen Einstieg in das Spiel

Trotz dieser Maßnahmen sind unterschiedliche Erfolgsausgänge beim experimentellen Einsatz von Computerspielen unvermeidbar. Dies wurde jedoch kontrolliert, indem die Versuchspersonen in Bewertungsheften im Anschluss an das Treatment dazu aufgefordert wurden, die Schwierigkeit von Spiel und Steuerung sowie ihren eigenen Spielerfolg zu beurteilen (vgl. Kapitel 2.1.8).

⁶⁷ Aus diesem Grunde wurde auch von einem Einsatz von First-Person-Shootern abgesehen, die eine sehr komplexe Steuerung aufweisen und bei unerfahrenen Personen ein mehrstündiges Training erfordern.

2.1.3 Auswahl und Präsentation des medialen Stimulusmaterials

Realisiert wurde eine Treatmentdauer von 60 Minuten. Diese in Hinblick auf experimentelle Medienwirkungsstudien vergleichsweise hohe Dimensionierung des Stimulusmaterials erschien wünschenswert, damit die Dauer des im Labor nachgestellten Treatments auch der freizeitlichen Beschäftigung mit Medien nahekommt. Zudem sollte in jedem Fall vermieden werden, dass evtl. vorhandene und im Alltag auftretende Effekte in Folge experimenteller Ökonomieverwägungen und zu niedrig dosierter Stimuli nicht aufgedeckt werden.

Das insgesamt 60-minütige Treatment wurde auf drei Treatmentblöcke (3 x 20 Minuten) aufgeteilt, und hier jeweils mit unterschiedlichen Medienangeboten realisiert. 20 Minuten wurden als hinreichend erachtet, um innerhalb der Mediengewaltbedingungen eine kompakte und in sich abgeschlossene Narration zu realisieren und den Versuchspersonen der Spielbedingungen eine Adaption an die Spielsteuerung und damit ein Eintauchen in das Spielerlebnis zu ermöglichen.⁶⁸ Mit der größeren Bandbreite verschiedener medialer Stimuli sollte die Wahrscheinlichkeit dafür erhöht werden, dass divergierende Vorlieben auf Seiten der Probanden durch unterschiedliche Angebote abgedeckt werden und diese damit einen optimalen Unterhaltungsgrad erreichen. Dieses Vorgehen hat zudem den Vorteil, dass auch die externe Validität der Untersuchung erhöht wird, indem die Daten nicht nur Aussagen über die kognitiven Wirkungen eines einzigen medialen Stimulus pro Treatmentbedingung zulassen, sondern eventuelle Wirkungen auf ein breiteres Spektrum unterschiedlicher Stimuli zurückgeführt werden können. Schließlich wird damit auch die Wahrscheinlichkeit dafür erhöht, dass sich unbekannte medienimmanente Störgrößen, die nicht Teil der intendierten Manipulation sind, zwischen den Gruppen ausmitteln. Um Reihenfolgeeffekte zu vermeiden, wurde die Darbietungsreihenfolge der drei Treatmentblöcke innerhalb aller medialen Bedingungen variiert. Dies wurde dadurch erreicht, dass den Versuchspersonen vor Versuchsbeginn im Rahmen der randomisierten Zuweisung zur gewalt- oder gewaltneutralen Bedingung zufällig eines von drei Bewertungsheften zugewiesen wurde, welches jeweils eine andere Reihenfolge in der Nutzung der Medienstationen vorgibt:

Bewertungsheft 1: Stimulus 1 --> Stimulus 2 --> Stimulus 3 Bewertungsheft 2: Stimulus 2 --> Stimulus 3 --> Stimulus 1 Bewertungsheft 3: Stimulus 3 --> Stimulus 1 --> Stimulus 2

68 Gerade unerfahrene Spieler könnten bei einer nur kurzen Spielphase auf einer sehr basalen Ebene der Interaktion mit dem Spiel verbleiben, in der sie in erster Linie damit beschäftigt wären, die Folgen ihrer ersten Steuerungsversuche mit einem Gemisch aus Faszination, Verwirrung und Verzweiflung zur Kenntnis zu nehmen, und damit die Rezeption in inadäquater und hinsichtlich gängiger Spielerlebnisse alltagsferner Weise erleben.

Die Auswahl des konkreten Stimulusmaterials orientierte sich neben den Faktoren Gewalt und Interaktivität daran, den Vorlieben eines bildungsheterogenen Publikums männlicher und weiblicher Probanden mittels einer breiten Palette von Mainstream Angeboten entgegenzukommen. Hierfür erschien es wünschenswert, dass die Angebote bekannten Schemata erfolgreicher Blockbuster und beliebter Video- und Computerspielen folgen. Es wurde daher entschieden, die sowohl hinsichtlich gewalthaltiger Computerspiele als auch gewalthaltiger Filme populären Genre **Action**, **Horror** und **Martial Arts (Kampf)** zu verwenden.⁶⁹ Damit wurde es auch möglich, hinsichtlich der Gewaltbedingungen jeweils ähnliche Inhalte auf Seiten des Computerspiels und der Filmbedingung (z. B. Horror-Zombiespiel vs. Horror-Zombiefilm) bereitzustellen, um hier eine Annäherung der Inhalte (Art der Gewalt) unter Manipulation des Mediums (interaktiv vs. nicht-interaktiv) zu ermöglichen (vgl. Anhang E).⁷⁰

Hinsichtlich der gewaltneutralen Bedingungen bestand der Anspruch hingegen darin, gegenüber der jeweiligen Gewalt-Referenzbedingung Angebote mit einem vergleichbaren Unterhaltungspotential bereitzustellen. Ein systematisch geringeres Unterhaltungserleben in den gewaltneutralen Bedingungen würde Verzerrungen befürchten lassen, indem sich Probanden dem Material weniger engagiert zuwenden oder sich möglicherweise sogar vollständig aus dem Rezeptionserleben zurückziehen könnten. Damit könnten eventuelle Unterschiede in den abhängigen Variablen nicht mehr in eindeutiger Weise auf die Gewaltdarstellungen zurückgeführt werden. Eine Angleichung des Unterhaltungserlebens ist jedoch insbesondere bei Computerspielen nicht einfach zu realisieren: Wie bereits gezeigt werden konnte, weisen die meisten Computerspiele in irgendeiner Form gewaltbezogene Hand-

69 Gewalt im Fernsehen ist typisches Charakteristikum von Spielfilmen bzw. fiktionalen Formaten (Groebel & Gleich, 1993, S. 67). Im US-Fernsehen zeigt sich der höchste Anteil von Gewalt sowie der höchste Anteil von Blut und Verletzungsdarstellungen unter den Gewaltsequenzen im Genre "Movies" (Groebel & Gleich, 1993, S. 67; *National Television Violence Study 3*, 1998, S. 60; 82). Hinsichtlich Computerspielen findet sich die höchste Rate getöteter Spielfiguren in Shooter- und Actionspielen (Thompson & Haninger, 2001). Beliebte Shooterspiele behandeln häufig Action (z. B. Half Life, Call of Duty), Horror (z. B. Doom) oder Kampfszenarien (z. B. Mortal Kombat, Street Fighter). Diese Genre stellen auch im Bereich der Filme ein populäres Format dar (Actionfilme, Horrorfilme und Martial-Arts-Filme).

70 Hierbei ist selbstverständlich immer nur von einer Annäherung, keinesfalls aber von einer Gleichsetzung der Inhalte zu sprechen, da sich Computerspiele und Filme durch eine Vielzahl von Konstruktionsprinzipien unterscheiden, die unter anderem die Art und Geschwindigkeit der Darstellung sowie die Häufigkeit von Gewaltereignissen umfassen. Inwieweit sich hier charakteristische Merkmale von Spielen und Filmen in unterschiedlicher Weise auf die betrachteten AV auswirken, war ein zentrales Erkenntnisinteresse dieser Untersuchung, weshalb eine weiterführende Angleichung auch gar nicht von Interesse war.

lungen auf.⁷¹ Eine unbedachte Auswahl gewaltneutraler Spiele könnte sich für erwachsene Spieler als ungeeignet erweisen oder ein in weit geringerem Maße involvierendes Spielgeschehen bieten. Vor diesem Hintergrund erschienen einfache Brettspielumsetzungen oder Geschicklichkeitsspiele (z. B. *Tetris*) als ungeeignet. Stattdessen boten sich gewaltfreie Sport- und Rennspiele an, die ebenfalls actionreich und grafisch ansprechend gestaltet sind, und zudem eine große Beliebtheit sowohl unter Männern als auch Frauen aufweisen. Hinsichtlich der gewaltneutralen Filmbedingung wurden die ebenfalls sehr populären Genre Komödie und Drama angezielt, die ebenfalls häufig keinerlei Gewaltereignisse im Sinne der hier verwendeten Definition aufweisen.

Der Unterhaltungswert der eingesetzten Filme und Spiele wurde vorab im Rahmen kleinerer Pilotexperimente untersucht. Diese wurden in Zusammenarbeit mit dem Institut für Journalistik und Kommunikationsforschung Hannover (IJK) innerhalb eines anwendungsbezogenen Experimentalseminars für Masterstudenten des Studiengangs Medienmanagement im Zeitraum Oktober 2005 bis Januar 2006 realisiert. Zentrale Fragestellungen zum Unterhaltungsrating von Filmen und Computerspielen wurden in studentischen Arbeitsgruppen unter Beratung und Anleitung des Projektteams bearbeitet. Die Pilotexperimente fanden in den Räumen des IJK statt und wurden von den Studierenden selbständig ausgewertet. In schriftlichen Berichten wurden die Konzeptionen und deren Bedeutung für das Gesamtvorhaben festgehalten. Die Ergebnisse flossen unmittelbar in die Stimulusauswahl des Hauptexperimentes ein und führten dazu, dass einzelne bereits vorgesehene Stimuli aufgrund eines zu geringen Unterhaltungspotential ausgeschlossen und durch alternative Angebote ersetzt wurden. So wurde zum Beispiel ein in der gewaltneutralen Bedingung zunächst vorgesehenes Flipperspiel (*Ultimate Pinball Challenge*, PC) aufgrund sehr geringer Unterhaltungswerte nicht in das Hauptexperiment übernommen. Einen weiteren maßgeblichen Einfluss hatte die Pilotisierung auf die Bedingung "gewalthaltige Filme". Hierfür waren zunächst, wie in der gewaltneutrale Bedingung, Kurzfilme vorgesehen (*American Showdown* [Action], *Kampfansage* [Martial Arts] und *I'll see you in my dreams* [Zombie-Horror]). Die Filme erzielten zwar, wie intendiert, hohe Gewalt-, jedoch sehr geringe Unterhaltungsratings, und wurden daher nicht in das Hauptexperiment übernommen. In diesem wurde alternativ auf in sich geschlossene Ausschnitte beliebter gewalthaltiger Blockbuster-Filme zurückgegriffen. Die Endauswahl ist Tabelle 12 zu entnehmen.

71 Aufgrund der für diese Untersuchung vergleichsweise streng definierten Kriterien gewaltneutraler Medien (vgl. Kapitel 2.1.2.1) sind so zum Beispiel auch Spiele für die gewaltneutrale Bedingung ungeeignet, in denen der Spieler physisch gegen comichaft gezeichnete Tierwesen vorgehen kann (Beispiel Super Mario) oder in denen Roboter oder Androiden bekämpft werden.

Tabelle 12. Gesamtübersicht über das Stimulusmaterial der medialen Treatmentbedingungen

	Spiele gewalthaltig	Filme gewalthaltig	Spiele gewaltneutral	Filme gewaltneutral
Material 1 (20 Min.)	Action / Krieg TIMECRISIS II	Action / Krieg WINDTALKERS	Rennen XPAND-RALLY	Drama C. BRUNNER + SCHWARZ- FAHRER
Material 2 (20 Min.)	Horror HOUSE OF THE DEAD III	Horror DAWN OF THE DEAD	Tennis TOP SPIN	Milieustudie LOTTO NORMAL + SILVERSTAR
Stimulus 3 (20 Min.)	Martial Arts / Kampf MORTAL KOMBAT DECEPTION	Martial Arts / Kampf KILL BILL VOL.1	Sport XLS ON TOUR	Komödie DANGLE + BJÖRN UND DIE HÜRDEN DER BEHÖRDEN

2.1.3.1 Konzeption der nicht-medialen Explorationsbedingung

Für die Wahl der nicht-medialen Explorationsbedingung galten im wesentlichen die beiden Vorgaben, dass hierbei keine Bildschirmmedien zum Einsatz kommen sollten und dass hierbei natürlich auch der Stimulus "Gewalt" nicht variiert bzw. appliziert werden sollte. Stattdessen wurden für die Explorationsbedingung beliebte Freizeitaktivitäten vorgesehen, bei denen die Versuchspersonen unter Verzicht auf Bildschirmmedien spielerisch tätig werden. Hierfür wurde zunächst erwogen, diesen Anspruch in unmittelbarer Anlehnung an die Medienbedingungen zu realisieren, indem die nicht-mediale Spielaktivität alleine und in sitzender beziehungsweise motorisch gering beanspruchender Weise auszuführen sein sollte. Die hierfür möglichen Konzeptionen (z. B. Puzzlespiel, Geschicklichkeitsspiel, Denkspiel) erschienen jedoch in der Weise als ungeeignet, dass ihnen in Hinblick auf die anzuzielende erwachsene und bildungsheterogene Stichprobe kaum ein ausreichendes Unterhaltungspotential zugesprochen werden konnte. Aus den gleichen Gründen wurde auch eine reine Ruhe- oder Entspannungsbedingung als problematisch erachtet. Eine Warteraumsituation mit Zeitschriften zum Lesen hätte artifizielle Interferenzeffekte mit dem verbalen Gedächtnismaterial erwarten lassen. In Abwägung der Vor- und Nachteile verschiedener Konzeptionen wurde schließlich entschieden, dem Faktor Unterhaltung und Involvement einen besonders hohen Stellenwert einzuräumen. Gewählt wurden die Freizeitaktivitäten Tischfußball, Tischtennis und Darts. Die Versuchspersonen konnten während der 20-minütigen Treatmentphasen diese Angebote frei nutzen. Dadurch weicht die Explorationsbedingung nicht nur hinsichtlich des Medienfaktors, sondern auch

weiterer Gesichtspunkte von den vier medialen Bedingungen ab. Dies erschien jedoch vertretbar, da die Hauptintension des Experimentes in einer differenziellen Beschreibung der Wirkungen interaktiver Mediengewalt auf kognitive Leistungen bestand, und hinsichtlich dieses Faktors bereits gewaltneutrale Vergleichsbedingungen vorgesehen waren. Für die Interpretation der Daten bedingt dies jedoch, dass Unterschiede zwischen medialen Bedingungen und der nicht-medialen Freizeitbedingung nicht ausschließlich auf den Medienfaktor zurückgeführt werden können. Die Explorationsbedingung ist daher auch nicht Teil der in dieser Arbeit behandelten Wirkhypothesen, sondern dient ausschließlich der zusätzlichen Veranschaulichung und Einordnung der Befunde.

2.1.4 Operationale Definitionen der abhängigen Variablen (AV)

Erfasst werden sollten die räumlich-visuelle und verbale Gedächtnisleistung, die verbale Lernleistung und die Konzentrationsleistung. Während hinsichtlich der Konzentrationsleistung eine Messung unmittelbar nach dem Treatment vorgesehen wurde, bestand hinsichtlich der Gedächtnis- und Lernleistung der Anspruch einer um 24 Stunden verzögerten Nachmessung.⁷² Dieses Vorgehen knüpft an Erkenntnisse zur Konsolidierung expliziter Gedächtnisinhalte an, wonach davon auszugehen ist, dass emotionale Wirkungen auf initial gelernte Informationen nur schwer nachweisbar sind, da die eigentliche Konsolidierungsphase einige Zeit in Anspruch nimmt (vgl. Kapitel 1.4.3). Eine um 24 Stunden verzögerte Nachmessung erschien darüber hinaus notwendig, um die externe Validität der Befunde zu erhöhen, indem Gedächtniseffekte über einen längeren Zeitraum nachgewiesen werden können. Eine Messung direkt nach dem Treatment hätte zudem befürchten lassen, dass mögliche Leistungsunterschiede auf die (aktuelle) Fähigkeit zum Abruf der Wissensbestände zurückzuführen sein könnten und nicht auf die Konsolidierung.⁷³ Das Lerntestverfahren sollte in besonderer Weise die Aneignung neuer Wissensinhalte über mehrere Lernphasen beobachtbar machen. Um natürliche Leistungsschwankungen zwischen den Gruppen, die auch bei einer Randomisierung auftreten können, kontrollieren zu können, sollte die Messung der abhängigen Variablen jeweils eine Prämessung vor dem Treatment mit einschließen (vgl. Tabelle 13).

72 Durch das Sampling, die Randomisierung und die verhältnismäßig großen Treatmentgruppen besteht aus statistischen Erwägungen heraus kaum die Gefahr, dass Versuchspersonen unterschiedlicher Gruppen sich innerhalb dieser 24 Stunden ihres Lebensalltags systematisch unterschiedlich verhalten.

73 Die Untersuchung von Wirkungen auf die aktuelle Fähigkeit zum Abruf semantischer Gedächtnisinhalten steht nicht im Fokus dieser Untersuchung (vgl. Kapitel 1.4.3).

Tabelle 13. Übersicht über das Messdesign der abhängigen Variablen

	Prä-messung	Messungen zwischen Treatment-phasen	Post-messung Tag 1	Post-messung 24 h verzögert
Verbale Gedächtnisleistung	X			X
Visuelle Gedächtnisleistung	X			X
Lernleistung	X	X		X
Konzentrationsleistung	X		X	

Die abhängigen Variablen wurden mithilfe psychometrischer Testdiagnostik erfasst. Für Gedächtnis- und Konzentrationsleistung kamen Papier-Bleistift-Verfahren einschlägiger Tests zum Einsatz, für die Lernleistung wurde ein mehrstufiges, computergestütztes Lerntestverfahren angewandt. Im Folgenden wird die Operationalisierung der berücksichtigten Leistungskonstrukte beschrieben.

2.1.4.1 Räumlich-visuelle und verbale Gedächtnisleistung

In der Untersuchung soll überprüft werden, ob gewalthaltige Medien einen Einfluss auf die Konsolidierung zuvor encodierter semantischer Lerninhalte nehmen können. Neben der testdiagnostischen Gütekriterien orientierte sich die Auswahl geeigneter Testdiagnostik primär an den folgenden drei Kriterien:

- Das Verfahren beansprucht die Messung expliziter (semantischer) Gedächtnisleistung.
- Es sind Messwiederholungen vorgesehen, die eine Beurteilung der Gedächtnisleistung nach 24 Stunden erlauben.
- Eine Einsetzbarkeit als Gruppentest ist gegeben.

Gewählt wurde der *VVM (Visueller und Verbaler Merkfähigkeitstest)*. Dieser erfasst kurzfristige und langfristige Gedächtnisleistung für explizite räumlich-visuelle und semantische Inhalte (Schelling & Schächtele, 2001, S. 36). Das Testverfahren sieht einen unmittelbaren sowie einen verzögerten Abruf vor, der sowohl 2 Stunden als auch 24 Stunden später erfolgen kann. Im räumlich-visuellen Teil müssen die Probanden einen Weg innerhalb eines Stadtplanes auswendig lernen, der in der verwendeten Testform „Theater“ zu einem neu gebauten Theater führt. Nach der Einprägephase von zwei Minuten muss der Weg erneut in einen Stadtplan eingezeichnet werden, auf dem der vorgegebene Weg nicht mehr zu sehen ist (vgl.

Abbildung 15). Für die Bewältigung dieser Aufgabe stehen dem Probanden abermals zwei Minuten zur Verfügung.

Stadtplan: THEATER

Zeichnen Sie bitte den eingetragten Weg in diesen Stadtplan.

Sie haben 2 Minuten Zeit.

Theaterbau

Beantworten Sie bitte die folgenden Fragen zum Theaterbau. Tragen Sie Ihre Antworten in die freie Zeile hinter der jeweiligen Frage ein.

Sie haben 4 Minuten Zeit.

1. Vorname und Name des Schauspielers?			
2. Vorname und Name der Gärtnerin?			
3. Name der Straße, an der der Haupteingang liegt?			
4. Der Nachname des Ehepaares, das das Theater leitet?			
5. Die Vornamen dieses Ehepaares?			
6. Wie viele Sitzplätze hat das Theater?			
7. Vor wie viel Jahren ist der Schauspieler gestorben?			
8. Wie viel Millionen Euro kostete das Theater?			
9. Wie viel Quadratmeter hat der Park?			
10. Welche Hausnummer hat das Theater?			
11. Wie hoch ist der Quadratmeterpreis des Grundstückes?			
12. Wie viele Räume hat das Theater?			
13. Wie viele Abteilungen hat das Theater?			
Welche 8 Unternehmen sind an der Stiftung beteiligt?			
14.	18.		
15.	19.		
16.	20.		
17.	21.		

Abbildung 15. Reproduktionsaufgabe räumlich-visuelle (links) und verbale (rechts) Gedächtnisleistung des VVM (Visueller und verbaler Merkfähigkeitstest)

Im verbalen Teil muss eine aus syntaktisch einfachen Sätzen bestehende Beschreibung des „angefahrenen“ Theaters eingetragt werden. Dabei werden verbale Zusammenhänge in Form von Zahlen, Namen und Begriffen erlernt. Die Beschreibung des Testverfahrens ist fiktiv und knüpft damit nicht an bereits bestehende Wissensbestände der Probanden an:

"Das neue Theater verfügt über X Sitzplätze. Es trägt den Namen des vor X Jahren verstorbenen Schauspielers X. Die Gesamtkosten für das Theater betrugen X Millionen Euro. (...)"

Im Anschluss findet eine schriftliche Abfrage der zuvor dargebotenen Theaterinformationen statt (vgl. Abbildung 15). Gruppe 1 („Namen“) beinhaltet fünf Fragen, die sich auf Namen von Schauspielern, Angestellten, Erbauern etc. beziehen. Gruppe 2 („Zahlen“) umfasst insgesamt acht Fragen, die sich auf numerische Informationen wie Anzahl der Sitzplätze, Kosten, Quadratmeter etc. beziehen. Gruppe 3 („Begriffe“) besteht aus acht Fragen, die die Geschäftsfelder derjenigen Unternehmen kennzeichnen, die den Theaterbau finanziert haben.

Dieselben Aufgaben werden den Probanden nach einem Zeitintervall von 24 h zum zweiten Messzeitpunkt erneut präsentiert. Zu diesem Zeitpunkt werden den Versuchspersonen die zuvor gelernten Informationen nicht erneut dargeboten. Stattdessen müssen alle Informationen erneut aus dem Gedächtnis reproduziert werden. Durch den Einbezug sowohl verbaler als auch visueller Informationen wird

die Inanspruchnahme unterschiedlicher Gedächtnissysteme und damit auch eine Vielschichtigkeit der zugrunde liegenden kortikalen Projektionsareale erreicht.

2.1.4.2 Lernleistung

Um die Auswirkung unterschiedlicher Treatmentbedingungen auf die Lernleistung abbilden zu können, sollte zusätzlich ein Verfahren verwendet werden, welches diese als State-Konstrukt operationalisiert. Zwar ermöglicht auch der *VVM* die Ermittlung von Vergessensraten über einen Zeitraum von 24 Stunden, legt jedoch wie nahezu alle Gedächtnistests ein Trait-Diagnostik-Verständnis zugrunde. Damit bestand hinsichtlich dieses Verfahrens die Gefahr, dass dieses für situative Leistungsschwankungen nicht sensibel genug sein könnte und demnach treatmentspezifische Auswirkungen auf die Gedächtnisleistung nicht zuverlässig aufgedeckt werden können. Zudem sollte ein Verfahren zum Einsatz kommen, welches in Anlehnung an schulische Lernprozesse im Rahmen eines mehrstufigen Lernprozesses den Erwerb umfassenderer semantischer Wissensbestände abbildet.

Eingesetzt wurde *WERNICKO-45*, ein computergestütztes Lerntestverfahren zum fiktiven Fremdspracherwerb, welches am Universitätsklinik der Universität Münster entwickelt wurde (Breitenstein & Knecht, 2002; Knecht, et al., 2004). Innerhalb von fünf ca. zehnminütigen Lernphasen wird ein Vokabular von 45 emotional neutralen Pseudowörtern einer fiktiven Sprache (jeweils aus 4 Buchstaben bestehend) erlernt. Die Worte werden vom Programm auditiv dargeboten und jeweils mit dem Bild eines Alltagsgegenstandes gekoppelt. 50 Prozent der präsentierten Wort-Bild-Kopplungen sind korrekt, die anderen 50 Prozent stellen unterschiedliche Falschkopplungen dieses Bildes mit den übrigen Wörtern dar. Die Reihenfolge der Kopplungen ist über alle Trials randomisiert. In jedem Block wird jede Vokabel zweimal mit dem richtigen Bild gekoppelt, und zweimal falsch mit jeweils einem anderen Bild.⁷⁴ Das Programm wird von der Versuchsperson über die Computertastatur gesteuert, indem bei jeder Wort-Bild-Präsentation entweder eine grüne Taste für „richtig“ oder eine rote Taste für „falsch“ gedrückt wird. Um das Verfahren zeitlich zu standardisieren, sind für die Eingabe nach der akustischen Präsentation des Pseudowortes jeweils max. 1,2 Sekunden vorgesehen. Gibt die Versuchsperson ihre Antwort nicht schnell genug ein, erscheint die Meldung „Zeit überschritten“, womit die Versuchsperson dazu aufgefordert wird, schneller zu antworten.

74 Auf alle fünf Lernphasen bezogen bedeutet dies, dass im gesamten Lerntest jedes Pseudowort zehnmal mit der richtigen Bildzuordnung und zehnmal mit unterschiedlichen falschen Bildzuordnungen dargeboten wird.

WERNICKO-45 ist in seiner Grundkonzeption ein assoziativer Lerntest und spricht damit implizites Lernen im Rahmen statistischer Lernprozesse an. Die eingesetzte Version wurde jedoch modifiziert, um explizites Lernen abzubilden. Beurteilt die Versuchsperson eine präsentierte Wort-Bild-Kopplung richtig, wird ein lachendes Gesicht präsentiert um anzuzeigen, dass die Eingabe korrekt war. Bei falscher Beurteilung erscheint hingegen ein trauriges Gesicht. Damit ist es den Probanden möglich, bewusste Lernstrategien einzusetzen, indem korrekte Zuordnungen eingeprägt werden. Dies wird auch dadurch ermöglicht, dass zwischen den einzelnen Items aufgrund der Präsentationszeit des Feedbacks insgesamt zwei Sekunden liegen (vgl. Abbildung 16). Damit ist die Phase zwischen den Stimuli insgesamt doppelt so lang wie in der impliziten *WERNICKO*-Fassung, bei der diese nur eine Sekunde beträgt (Breitenstein & Knecht, 2002; Knecht, et al., 2004).

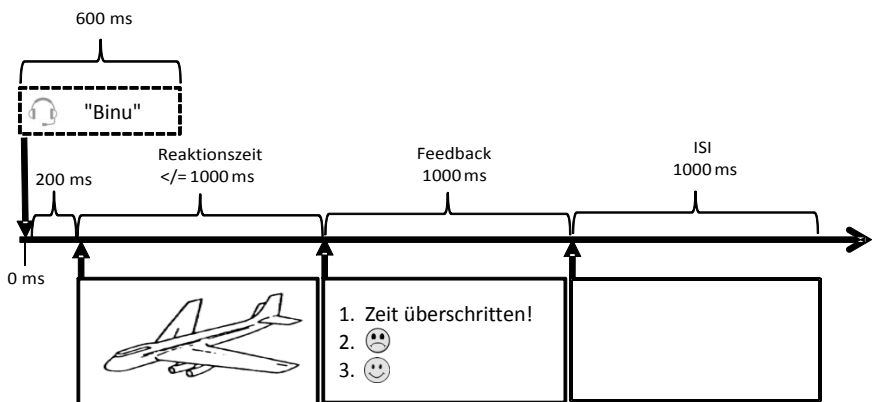


Abbildung 16. Beispielvorgabe eines einzelnen Testitems aus dem für explizites Lernen abgewandelten Lerntestverfahren *WERNICKO* (für die implizite Ursprungsfassung vgl. Breitenstein & Knecht, 2002; Knecht, et al., 2004). ISI = Interstimulusintervall.

Im Verlaufe der Lernphasen steigern Probanden Ihre Trefferquote je nach individueller Leistungsfähigkeit von anfänglich 50 Prozent (Zufallstrefferwahrscheinlichkeit) auf ca. 80 bis 90 Prozent. Der Lernerfolg in den einzelnen Phasen wird von dem Programm aufgrund seiner Implementierung in das Programm *SuperLab Pro* automatisch protokolliert. Im Anschluss an die fünf Lernphasen ist ein Transfertest vorgesehen. Hier werden die Vokabelpaare nur noch akustisch präsentiert. Die Versuchsperson müssen also beispielsweise entscheiden, ob das deutsche Wort "Baum" dem Pseudowort "Nalf" entspricht. In der Transferphase erhalten die Versuchspersonen über die Richtigkeit ihrer Angaben kein Feedback.

WERNICKO hat sich als sensibel für situative Schwankungen in der Lernleistung herausgestellt, was sowohl anhand medikamentöser (vgl. Knecht, et al., 2004) als auch verhaltensbezogener Treatmentbedingungen (vgl. Winter, et al., 2007) nachgewiesen werden konnte. Um Deckeneffekte auszuschließen und den Lernzuwachs über die Lernphasen einschätzen zu können, wurde vor dem Hauptexperiment eine Pilottestung des Verfahrens mit $N = 15$ Versuchspersonen durchgeführt. Dabei zeigte sich, dass zwei Lernphasen notwendig sind, um einen ersten bedeutsamen Lernzuwachs zu erzielen. Während nach der ersten Lernphase die Trefferquote nur unwesentlich über der Zufallswahrscheinlichkeit (50 %) lag, bestand nach der zweiten Lernphase eine Trefferquote von 60 bis 70 Prozent. Aufgrund dieses Ergebnisses wurde entschieden, vor der ersten Treatmentphase zwei Lernphasen zu realisieren. Nach der vierten Lernphase wurde ein Lernniveau von 75 bis 85 Prozent erreicht. Damit war kein Deckeneffekt hinsichtlich der Realisierung einer fünften Lernphase zu befürchten, zumal diese erst 24 Stunden verzögert am Folgetag stattfinden sollte. Es wurde deshalb entschieden, alle fünf vorgesehenen Lernphasen von *WERNICKO* sowie einen Transfertest zu realisieren.

2.1.4.3 Konzentrationsleistung

Bei der Auswahl eines geeigneten Instrumentes zur Erfassung der Konzentrationsleistung wurde besonderer Wert darauf gelegt, dass dessen Operationalisierung eine inhaltliche Abgrenzung zum Konstrukt Aufmerksamkeitsleistung erlaubt, die als wahrnehmungsnahe, selektive Beachtung relevanter Reize oder Informationen verstanden wird (vgl. Schmidt-Atzert, Büttner, & Bühner, 2004; Schmidt-Atzert, Krumm, & Bühner, 2008, S. 61). Konzentration kann in Abgrenzung zur Aufmerksamkeit als unter Anspannung erlebte kognitive Zuwendung zu einer definierten Arbeitsaufgabe ohne aufgabenimmanente ablenkende Störreize bzw. konkurrierende Informationen verstanden werden (Schmidt-Atzert, et al., 2004). Damit fokussiert der Begriff Aufmerksamkeit auf Wahrnehmungsleistungen und der Begriff Konzentration auf den gesamten Prozess der Informationsverarbeitung und den damit verbundenen Kontrollanforderungen (vgl. Schmidt-Atzert, et al., 2004, S. 9; Schmidt-Atzert, et al., 2008, S. 61). Die theoretische Trennung dieser Konzepte findet sich in gängigen Testverfahren nicht immer wieder und die Begriffe Aufmerksamkeits- und Konzentrationsleistung werden häufig synonym gebraucht. Die Erfassung von Konzentrationsleistung orientierte sich an folgenden Kriterien:

- Erfasst werden in geringerem Maße Funktionen der Wahrnehmung und Selektion von Reizen, sondern primär Funktionen der Weiterverarbeitung von Informationen, die in hohem Maße Kontrollanforderungen und exekutive Pro-

zesse beanspruchen und damit als arbeitsgedächtnisnahe Konzentrationsleistungen definiert werden können.

- Eine Einsetzbarkeit als Gruppentest ist gegeben.
- Zeitökonomische Mehrfachmessungen sind möglich.

Nach Recherche möglicher testdiagnostischer Verfahren, die den genannten Ansprüchen genügen, wurde der *Konzentrations-Leistungs-Test (KLT-R; Mayrhofer, 2004)* ausgewählt. Die Befunde einer faktorenanalytischen Studie von Schmidt-Atzert, Bühner und Enders (2006) lassen sich dahingehend interpretieren, dass der *KLT-R* im Vergleich zu anderen Aufmerksamkeits- und Konzentrationstests besonders hoch auf den Faktor Gedächtnisleistung lädt und damit in besonderer Weise arbeitsgedächtnisnahe Konzentrationsleistung abbildet.⁷⁵ In dem Verfahren müssen die Probanden unter Zeitdruck Rechenaufgaben lösen. Jede Rechenaufgabe besteht dabei aus drei einzelnen Rechenschritten (vgl. Abbildung 17):

1. Der Proband löst die erste aus drei einstelligen Ziffern bestehende Rechenaufgabe (Addition und Subtraktion) und merkt sich das Zwischenergebnis.
2. Der Proband löst eine zweite aus drei einstelligen Ziffern bestehende Rechenaufgabe (Addition und Subtraktion) und merkt sich auch dieses Zwischenergebnis.
3. Der Proband entscheidet, welches der beiden Zwischenergebnisse größer ausfällt.
4. Ist die obere Zahl größer, so muss der Proband das untere Zwischenergebnis von dem oberen abziehen. Ist die obere Zahl hingegen kleiner, so müssen beide Zwischenergebnisse aufaddiert werden.

Hierbei dürfen sich die Probanden keine Zwischenergebnisse notieren. Der Test beansprucht damit in besonderer Weise das Arbeitsgedächtnis (vgl. Kapitel 1.4.1.1), indem die meisten Personen für derartige Aufgabenstellungen sowohl räumlich-visuelle Anteile (räumlich-visueller Notizblock) als auch verbale Anteile (artikulatorische Schleife) des Arbeitsgedächtnis in Anspruch nehmen (J. R. Anderson, 1996, S. 174 ff.). Diamond und Kollegen (2007) weisen zudem auf die Bedeutsamkeit aufgabenimmanenter Auswertungsregeln hin, wie zum Beispiel die Anweisung, unter bestimmten Voraussetzungen eine bestimmte Rechenprozedur anzuwenden. Derartige Entscheidungsregeln können den Autoren zufolge eine

75 Hierbei ist auf die Unschärfe des Konstrukts Konzentrationsleistung hinzuweisen. So weisen Schmidt-Atzert und Kollegen (2006) darauf hin, dass in nahezu allen Konzentrationstests die Erfassung des Primärkonstruktes mit weiteren Fähigkeiten und Fertigkeiten konfundiert ist. Ein von den Autoren aus vielfältigen Testverfahren abgeleiteter Konzentrationsfaktor (Studie 2) umfasst ein heterogenes Leistungsspektrum, das auch Aspekte der Wahrnehmungsgeschwindigkeit und Informationsverarbeitungsgeschwindigkeit einschließt (vgl. Schmidt-Atzert, et al., 2006, S. 41 ff.). Hinsichtlich des *KLT-R* ist zudem davon auszugehen, dass das Verfahren neben arbeitsgedächtnisnaher Konzentrationsleistung auch allgemeine Rechenfertigkeiten abbildet.

simple Aufgabe zu einer komplexen machen, indem sie die Inanspruchnahme des präfrontalen Cortex erforderlich machen (Diamond, et al., 2007, S. 18).

Block 1		Block 2		Block 3	
8 - 3 + 7		9 + 5 + 8		5 - 2 + 9	
5 + 6 - 3		6 - 3 + 4		4 + 7 - 2	
3 + 6 - 5		5 - 2 + 6		9 + 2 - 5	
5 + 7 - 4		4 + 7 + 5		7 - 4 + 6	
9 + 8 - 6		8 + 5 - 7		6 - 3 + 8	
5 - 3 + 5		4 + 9 - 5		4 + 9 - 6	
7 - 4 + 9		2 + 6 - 5		7 + 6 - 5	
2 + 5 - 3		7 + 4 - 3		4 + 9 + 3	
3 + 6 - 5		5 - 2 + 6		9 + 2 - 5	
5 + 7 - 4		4 + 7 + 5		7 - 4 + 6	
8 - 3 + 7		9 + 5 + 8		5 - 2 + 9	
5 + 6 - 3		6 - 3 + 4		4 + 7 - 2	

Abbildung 17. Testbogenausschnitt des Konzentrations-Leistungs-Tests (Aufgabenstellung verfremdet)

Da die Versuchspersonen zwischen 18 und 25 Jahren alt sind, kam die für den Erwachsenenbereich geeignete Version *KLT-R(6-13)* zum Einsatz. Das vollständige Testverfahren enthält neun Aufgabenblöcke, für die separat Indikatoren der Leistungsmenge und der Leistungsgüte gebildet werden können. Aus diesem Grunde erschien es für den experimentellen Anwendungszweck dieser Untersuchung angemessen, dass Testverfahren auf zwei Messzeitpunkte mit jeweils drei Aufgabenblöcken aufzuteilen. Hinsichtlich der Gütekriterien entspricht das Verfahren mit einer gegebenen Objektivität und einer hohen Reliabilität und internen Konsistenz allen testdiagnostischen Anforderungen. Als Leistungsindikator wird die Anzahl korrekt bearbeiteter Aufgaben (Gesamtleistungsmenge abzüglich der Fehler) zugrundegelegt.

2.1.5 Kovariate: Verbale Intelligenz

Die verbale Intelligenz der Probanden wurde erhoben, um diese als Kovariate in den Auswertungen berücksichtigen zu können und damit die Fehlervarianz der

Unterschiede zwischen den Treatmentbedingungen zu reduzieren. Erfasst wurde diese mit der Aufgabengruppe Analogien (AN) aus dem *Intelligenz-Struktur-Test 2000 R* (I-S-T 2000 R; Amthauer, Brocke, Liepmann, & Beauducel, 2001). Dieser Untertest beansprucht innerhalb des *I-S-T 2000 R* zusammen mit zwei weiteren Untertests (Satzergänzung [SE], Gemeinsamkeiten [GE]) die Messung verbaler Leistungsaspekte von Intelligenz. Unter verbaler Intelligenz verstehen die Testautoren "die Fähigkeit zum Umgang mit sprachlichem Material im Rahmen des schlußfolgernden Denkens" (Amthauer, et al., 2001, S. 92). Die durch den *I-S-T 2000 R* erfasste verbale Intelligenz korreliert von allen Untertests am stärksten mit der Deutschnote und weist auch Korrelationen mit den Schulfächern Mathematik, Englisch, Chemie und Physik auf.

Die Aufgabengruppe Analogien besteht aus 20 Testitems, in denen den Probanden jeweils drei Wörter vorgegeben werden. Für jedes korrekt bearbeitete Item wird ein Punkt vergeben, womit insgesamt maximal 20 Punkte erreicht werden können. Zwischen den ersten beiden Wörtern besteht eine bestimmte Beziehung. Die Aufgabe des Probanden besteht darin, unter fünf Auswahlwörtern das Wort zu identifizieren, mit welchem sich zum dritten vorgegebenen Wort eine ähnliche Beziehung herstellen lässt (vgl. Abbildung 18). Für die Bearbeitung sind sieben Minuten vorgegeben (vgl. Amthauer, et al., 2001). Der Ablauf der Testdiagnostik orientierte sich eng an den Vorgaben im Testmanual (vgl. Anhang I).

Wald : Bäume = Wiese : ?				
a) Gräser	b) Heu	c) Futter	d) Grün	e) Weide

Abbildung 18. Beispielaufgabe aus dem Untertest Analogien (AN) des Intelligenz-Struktur-Test 2000 R

Die Rohwerte der Versuchspersonen wurden nach der Normtabelle des Testmanuals (Altersbereich 21 – 25 Jahre) in Standardwerte transformiert (Amthauer, et al., 2001). Die Transformation bewirkt insbesondere eine Reduktion der Schiefe der Rohwertverteilung von -.49 auf .05 und eine Reduktion der positiven Kurtosis von 3.36 auf 3.02. Während die Rohwertverteilung signifikant von einer Normalverteilung abweicht ($p < .01$) trifft dies auf die Standardwertverteilung nicht zu (ns). Für die Kovarianzanalysen wird ausschließlich auf die Standardwertverteilung verbaler Intelligenz zurückgegriffen (vgl. Abbildung 19).

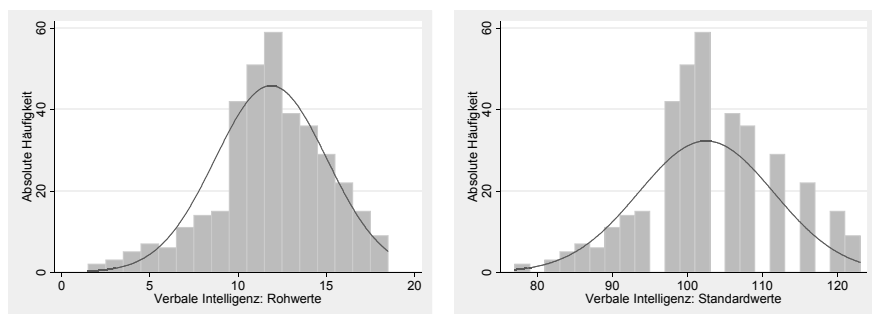


Abbildung 19. Verteilung der Rohwerte (links) und Standardwerte (rechts) der Aufgabengruppe Wortanalogien des Intelligenz-Struktur-Test 2000 R (n = 365).

2.1.6 Moderatorvariablen

2.1.6.1 Neurotizismus

Für die Erfassung von Neurotizismus wurde die Skala *Nervosität, Sensibilität und emotionale Labilität (N)* des *Hamburger Persönlichkeitsinventars* eingesetzt (Andresen, 2002). Die Skala besteht aus insgesamt 14 unipolaren viergestuften Items (vgl. Tabelle 14) und gehört zum Subsystem Emotion, Befindlichkeit und Stimmungen des Persönlichkeitsinventars. Das Instrument ist dabei nicht in erster Linie klinisch ausgerichtet, sondern erfasst Neurotizismus als "psychosoziale Streßreagibilität und Bereitschaft zu valenznegativen Affekten vor allem in sozial-interaktiven und -evaluativen Kontexten" (Andresen, 2002, S. 75). Zu allen n = 365 Versuchspersonen ließen sich gültige Skalenwerte ermitteln. Die Skalenrohwerte wurden anhand der Grobnormentabellen des Manuals (Andresen, 2002, S. 166) in rangnormalisierte Standardwerte transformiert (vgl. Abbildung 20).

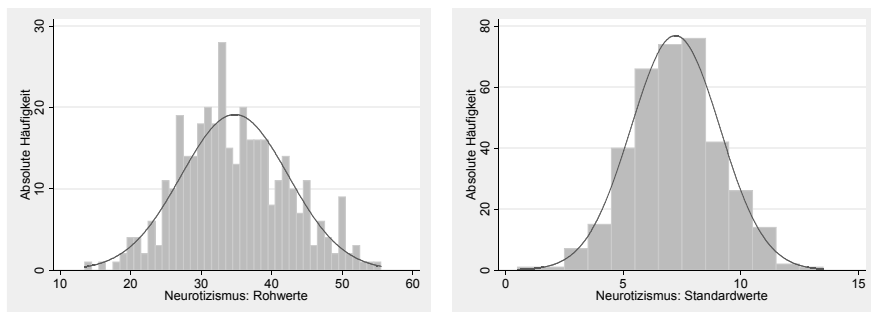


Abbildung 20. Verteilung der Rohwerte (links) und Standardwerte (rechts) der Skala Nervosität, Sensibilität und Labilität (N) des Hamburger Persönlichkeitsinventars ($n = 365$).

Tabelle 14. Itemkennwerte der Skala Nervosität, Sensibilität und emotionale Labilität (N) des Hamburger Persönlichkeitsinventars (n = 365)

Item	<i>M</i>	<i>SD</i>	<i>r</i> _{it}	<i>α</i>
Ich neige zu Stimmungsschwankungen.	2.65	.85	.54	.85
Ich fühle mich manchmal ohne Grund einfach miserabel.	2.48	1.03	.53	
Ich neige zu völlig fruchtlosen Grübeleien.	2.34	1.00	.40	
Ich reagiere empfindlich auf kritisch gemeinte Andeutungen.	2.46	.86	.34	
Ich reagiere oft genervt auf schlechte Angelegenheiten anderer Leute.	2.83	.82	.39	
Ich reagiere oft übertrieben emotional auf kleinere Vorkommnisse.	2.38	.98	.55	
Ich fühle mich leicht abgelenkt.	2.57	.92	.40	
Ich fühle mich oft einsam.	2.13	1.00	.55	
Manchmal kann es mir niemand recht machen.	2.89	.89	.45	
Ich nehme viele alltägliche Rückschläge zu schwer.	2.32	.91	.55	
Ich erlebe viele negative Gefühle wie Schreck, Wut oder Schmerz.	2.17	.99	.57	
Meine Alltagsbelastungen machen mich oft müde und angespannt.	2.76	.91	.45	
Mir wachsen viele Verpflichtungen über den Kopf.	2.36	.87	.52	
Ich wünschte, ich wäre seelisch stabiler.	2.44	1.09	.57	

Anmerkung. *r*_{it} = Trennschärfe (Part-whole-Korrektur). *α* = Cronbachs Alpha. Skalenwerte von 1 bis 4 (völlig unzutreffend, eher unzutreffend, eher zutreffend, völlig zutreffend)

2.1.6.2 Allgemeine Selbstwirksamkeitserwartung

Eingesetzt wurde die Skala zur *Allgemeinen Selbstwirksamkeitserwartung (SWE)* von Schwarzer und Jerusalem (1999). Das Instrument umfasst zehn vierstufige unipolare Items, die allgemeine optimistische Kompetenzerwartungen erfassen (vgl. Tabelle 15). Selbstwirksamkeitserwartung kann als allgemeiner Schutzfaktor aufgefasst werden: „Die Zuversicht in die eigenen Fähigkeiten und Fertigkeiten ist im Sinne einer stabilen personalen Coping-Ressource zu verstehen“ (Schwarzer, 1993, S. 189). Bei Aufsummierung der Items sind Werte zwischen 10 und 40

Punkten zu erreichen. Dem Manual sind keine Hinweise für den Umgang mit fehlenden Werten zu entnehmen. Es wurde entschieden, bei einem einzigen individuell vorliegenden Missingwert den Itemmodalwert (3) als Ersatzwert zu nutzen. Bei mehr als einem fehlenden Wert wurde der Skalenrohwert auf Missing gesetzt. Letzteres traf auf insgesamt sieben Personen zu, so dass zu $n = 358$ Personen gültige Skalenrohwerte vorliegen.

Tabelle 15. Itemkennwerte der Skala Selbstwirksamkeitserwartung ($n = 358$)

Item	<i>M</i>	<i>SD</i>	<i>r</i> _{it}	<i>α</i>
Wenn sich Widerstände auftun, finde ich Mittel und Wege, mich durchzusetzen.	3.14	.68	.64	
Die Lösung schwieriger Probleme gelingt mir immer, wenn ich mich darum bemühe.	3.32	.68	.71	
Es bereitet mir keine Schwierigkeiten, meine Absichten und Ziele zu verwirklichen.	2.92	.79	.63	
In unerwarteten Situationen weiß ich immer, wie ich mich verhalten soll.	2.58	.66	.64	
Auch bei überraschenden Ereignissen glaube ich, dass ich gut mit ihnen zurechtkommen kann.	2.94	.57	.68	
Schwierigkeiten sehe ich gelassen entgegen, weil ich meinen Fähigkeiten immer vertrauen kann.	2.82	.73	.74	.87
Was auch immer passiert, ich werde schon klarkommen.	3.17	.71	.69	
Für jedes Problem kann ich eine Lösung finden.	3.01	.76	.71	
Wenn eine neue Sache auf mich zukommt, weiß ich, wie ich damit umgehen kann.	2.85	.66	.72	
Wenn ein Problem auftaucht, kann ich es aus eigener Kraft meistern.	2.91	.65	.68	

Anmerkung. r_{it} = Trennschärfe (Part-whole-Korrektur). $α$ = Cronbachs Alpha. Skalenwerte von 1 bis 4 (stimmt nicht, stimmt eher, stimmt kaum, stimmt genau).

Die Skalenrohwerte wurden nach geschlechtsspezifischer und für Deutschland bevölkerungsrepräsentativer Normwerttabelle (Altersgruppe 14-30 Jahre) in Standardwerte transformiert (vgl. Schumacher, Klaiberg, & Brähler, 2001). Die Standardwertverteilung weist jedoch keine verbesserten Verteilungseigenschaften auf als die Rohwertverteilung (vgl. Abbildung 21). Es wurde daher entschieden, für die nachfolgenden Auswertungen die Rohwertverteilung von Selbstwirksamkeitserwartung zu verwenden.

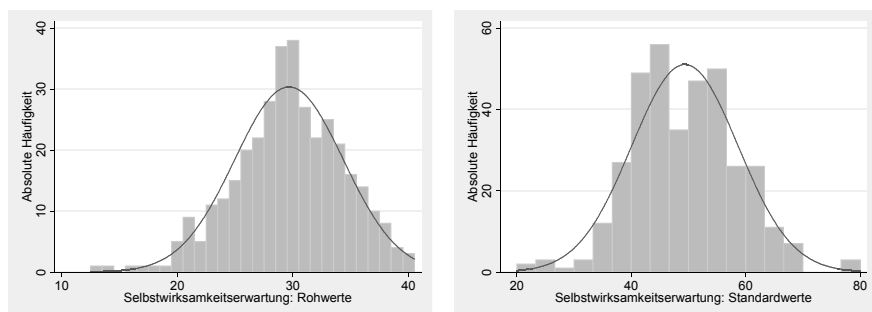


Abbildung 21. Verteilung der Rohwerte (links) und Standardwerte (rechts) der Skala Selbstwirksamkeitserwartung ($n = 358$).

2.1.6.3 Prosozialität

Für die Erfassung von Prosozialität wurde die Skala *Altruismus, Fürsorglichkeit und Hilfsbereitschaft (A)* des *Hamburger Persönlichkeitsinventars* verwendet (Andresen, 2002). Die Skala erfasst verschiedene Aspekte einer globalen prosozialen Orientierung wie Zärtlichkeitsbedürfnisse, Friedfertigkeit, Harmoniesuche, Hingabefähigkeit, Selbstaufopferungsbereitschaft, Mitleidsfähigkeit und Empathie (vgl. Andresen, 2002, S. 76) und besteht aus insgesamt 14 unipolaren viergestuften Items. Von den $n = 365$ Versuchspersonen hat eine Versuchsperson mehr als ein Item der Skala nicht beantwortet. Bei dieser Person wurde die gesamte Skala auf Missing gesetzt. Damit liegen zu $n = 364$ Personen gültige Skalenwerte vor (vgl. Tabelle 16).

Tabelle 16. Itemkennwerte der Skala Altruismus, Fürsorglichkeit und Hilfsbereitschaft (Prosozialität) des Hamburger Persönlichkeitsinventars (n = 364)

Item	<i>M</i>	<i>SD</i>	<i>r</i> _{it}	<i>α</i>
Ich Sorge sehr gern für andere.	3.05	.78	.48	.79
Die Nächstenliebe ist für mich die schönste menschliche Tugend.	2.76	.88	.50	
Ich habe ein sehr sanftes, anschmiegsames Wesen.	2.86	.81	.41	
Friedfertigkeit ist vielleicht mein auffälliger Charakterzug.	2.73	.85	.29	
Ich mag alles Zarte und Weiche sehr.	2.86	.86	.38	
Ich reagiere sehr empfindsam auf das Leid anderer Menschen.	3.06	.83	.47	
Ich organisiere gern Nachbarschaftshilfe.	1.70	.81	.32	
Ich spüre oft das Bedürfnis, ausländischen Waisenkindern zu helfen.	2.12	1.00	.39	
Zärtlichkeit bedeutet mir viel mehr als leibenschaftliche Sexualität.	3.06	.82	.30	
Babys finde ich "schrecklich süß".	2.90	1.04	.45	
Ich beobachte und beaufsichtige gerne Kinder beim Spielen.	2.48	1.08	.44	
Man sagt, ich habe ein zu weiches Herz.	2.62	.96	.49	
Ich mache mir häufig ernsthafte Sorgen um andere Menschen.	2.94	.85	.47	
Ich lese geliebten Menschen alle Wünsche von den Augen ab.	2.73	.88	.33	

Anmerkung. r_{it} = Trennschärfe (Part-whole-Korrektur). α = Cronbachs Alpha. Skalenwerte von 1 bis 4 (völlig unzutreffend, eher unzutreffend, eher zutreffend, völlig zutreffend).

Die Skalenrohwerte wurden anhand der Grobnormentabellen des Manuals (Andresen, 2002, S. 166) in rangnormalisierte Standardwerte transformiert (vgl. Abbildung 22).

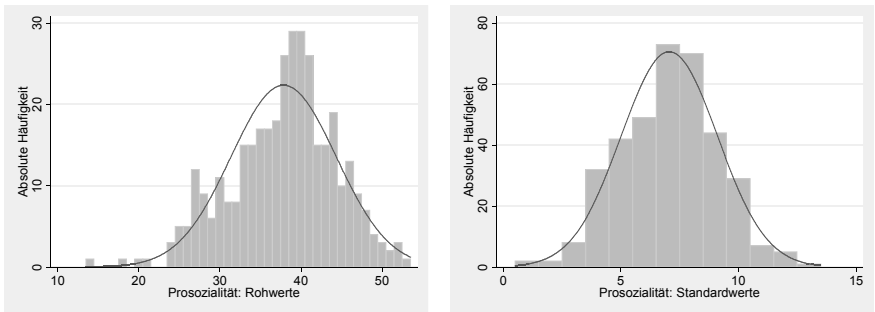


Abbildung 22. Verteilung der Rohwerte (links) und Standardwerte (rechts) der Skala Altruismus, Fürsorglichkeit und Hilfsbereitschaft (Prosozialität) des Hamburger Persönlichkeitsinventars (n = 364).

2.1.6.4 Habituelle Gewaltexposition

Alle Versuchspersonen wurden schon bei der Akquise um eine Angabe darüber gebeten, ob sie Video- und Computerspiele nutzen und wenn ja, wie viele Stunden in der Woche sie üblicherweise damit verbringen. Diese Einschätzung wurde im Rahmen des Telefoninterviews (vgl. Kapitel 2.1.11) mittels einer zwölfstufigen Skala festgehalten, die Werte zwischen 0 und > 20 Stunden (pro Woche) einnehmen konnte.⁷⁶ Alle Versuchspersonen schätzten zudem während der schriftlichen Befragung am zweiten Versuchstag per Freifeldeingabe ein, wie viele Stunden in der Woche sie üblicherweise mit nicht-interaktiven Medien (Fernsehen, Video und DVD) verbringen. Um die Angaben zur Computerspiel- und Filmnutzungszeit miteinander verrechnen zu können, wurden die Freifeldangaben der Versuchspersonen aus der schriftlichen Befragung auf das zwölfstufige Format der telefonischen Abfrage der Computerspielzeit umgerechnet.

Zusätzlich wurden die Versuchspersonen darum gebeten, für beide Medien einzuschätzen, welcher Anteil der von ihnen innerhalb der letzten 6 Monate genutzten Angebote mit einer Altersfreigabe ab 16 Jahren oder höher eingestuft wurde (vgl. Abbildung 23). Aus den Angaben zur Nutzungszeit und den Angaben zur relationalen Bedeutung von Gewaltinhalten im jeweiligen Medienmenu⁷⁷ wurde für beide

76 Skalenintervalle: "0 h" "0-2 h" "2-4 h" "4-6 h" "6-8 h" "8-10 h" "10-12 h" "12-14 h" "14-16 h" "16-18 h" "18-20 h" "> 20 h".

77 Hierbei kann eingewendet werden, dass zumindest bei Filmmedien die Nutzung von Angeboten ab 18 Jahren auch pornografische Inhalte umfassen kann. Damit kann es bei Personen mit einer hohen Nutzung von pornografischen Inhalten zu einer Überschätzung der Gewaltmedienexposition kommen. Diese Ungenauigkeit in der Messung wurde jedoch zugunsten

Medien ein Gewaltexpositionsindex errechnet, indem der Anteil der Nutzung von Angeboten mit einer Einstufung ab 16 Jahren und höher mit der jeweiligen Nutzungsdauer des Mediums verrechnet wurde.⁷⁸ Anschließend wurde die Gewaltexposition basierend auf der Nutzung beider Medien aufsummiert.⁷⁹

Wenn Sie an die Video-/Computerspiele denken, die Sie in den letzten sechs Monaten gespielt haben: Ab welchem Alter waren diese Spiele freigegeben?

_____ (%) waren unter 16 Jahren freigegeben (z.B. ab 6 oder ab 12 Jahren)

_____ (%) waren ab 16 Jahren freigegeben

_____ (%) waren ab 18 Jahren freigegeben

_____ (%) hatten keine Altersfreigabe (z.B. indiziert / in Deutschland verboten)

= 100%

Wenn Sie an die Filme denken, die Sie in den letzten sechs Monaten gesehen haben: Ab welchem Alter waren diese Filme freigegeben?

_____ (%) waren unter 16 Jahren freigegeben (z.B. ab 6 oder ab 12 Jahren)

_____ (%) waren ab 16 Jahren freigegeben

_____ (%) waren ab 18 Jahren freigegeben

_____ (%) hatten keine Altersfreigabe (z.B. indiziert / in Deutschland verboten)

= 100%

Abbildung 23. Erfassung der Altersfreigaben der in den letzten sechs Monaten genutzten Angebote hinsichtlich interaktiver (links) und nicht-interaktiver Unterhaltungsmedien

Bei einer Versuchsperson konnte kein Gewaltexpositionsindex errechnet werden, weil diese Person die Freifeldangabe zur Nutzungszeit nicht-interaktiver Medien nicht beantwortet hat. Bei zwei Nichtspielern liegen Angaben zur Fernsehzeit, nicht jedoch zum Anteil von Gewaltinhalten innerhalb dieses Mediums vor, weshalb auch zu diesen Versuchspersonen keine Ermittlung der Gewaltexposition möglich ist. Damit verbleiben n = 362 Versuchspersonen mit gültigen Skalenwerten. Die Verteilung der Gewaltexpositionswerte erweist sich als linksschief verteilt. Im Falle linksschief verteilter Variablen empfehlen Tabachnick und Fidell (2007) im Rahmen multivariater Analyse eine Quadratwurzel- oder logarithmische Trans-

einer kompakten und die Befragungspersonen nicht überfordernden Abfragemethode in Kauf genommen.

78 Gesamtmedienzeit abzüglich des Anteils von Angeboten mit einer Alterseinstufung unter 16 Jahren = Gewaltmedienzeit. Beispiel: 15 Stunden in der Woche fernsehen, 40 % unter 16 Jahren eingestuft = 9 Stunden Gewaltfilmmutzung pro Woche.

79 So wurde davon ausgegangen, dass eine einstündige Nutzung von gewalthaltigen Filmen plus eine einstündige Nutzung von gewalthaltigen Spielen eine höhere Gewaltmedienexposition charakterisiert als eine einstündige Nutzung gewalthaltiger Filme unter Verzicht auf Computerspiele.

formation der Daten vorzunehmen. Im Falle der vorliegenden Verteilung erweist sich eine Quadratwurzeltransformation als besonders geeignet und verbessert die Verteilungseigenschaften nachhaltig (vgl. Abbildung 24). Die Verteilung nähert sich dabei an eine Normalverteilung an, weist jedoch nach wie vor eine positive Kurtosis (2.5) und damit einen zu steilen Verteilungsverlauf auf.

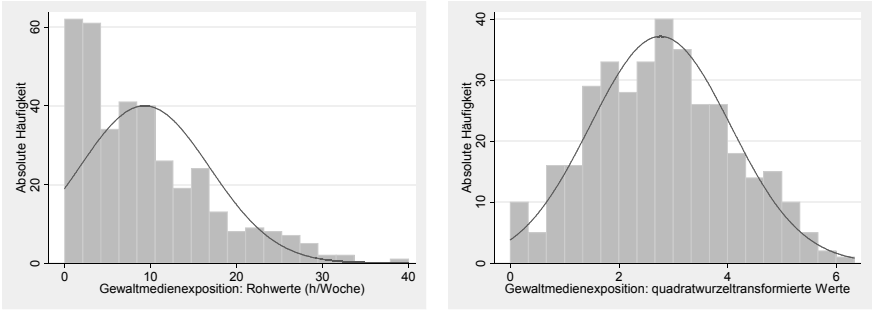


Abbildung 24. Verteilung der Rohwerte (links) und quadratwurzeltransformierten Werte (rechts) des Gewaltexpositionsindexes (n = 362)

2.1.6.5 Gewaltpräferenz

Die Vorliebe für mediale Gewaltinhalte wurde in Anlehnung an Gentile et al. (2004, S. 10) mittels einer zehnstufigen Selbsteinschätzungsskala erhoben. Um eine möglichst akkurate Einschätzung der Präferenz von Gewalt zu erhalten, wurde diese zu beiden infrage stehenden Bildschirmmedien abgefragt (vgl. Abbildung 25).

Bei actionreichen Video-/Computerspielen (z.B. Shootern / „Ballerspielen“):
Wie hoch ist das Gewaltlevel, das Sie üblicherweise bevorzugen?

Niedriges Gewaltlevel	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hohes Gewaltlevel
	1	2	3	4	5	6	7	8	9	10	

Bei actionreichen Filmen (z.B. Action- oder Kriegsfilme):
Wie hoch ist das Gewaltlevel, das Sie üblicherweise bevorzugen?

Niedriges Gewaltlevel	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hohes Gewaltlevel
	1	2	3	4	5	6	7	8	9	10

Abbildung 25. Erfassung der Gewaltpräferenz hinsichtlich actionreicher interaktiver (oben) und actionreicher nicht-interaktiver Unterhaltungsmedien (unten)

Zwischen beiden Angaben besteht mit $r = .75$ eine hohe Korrelation. Es wurde daher entschieden, beide Angaben innerhalb eines Gewaltpräferenzindexes zusammenzufassen. Da bei den Nichtspielern nicht nach der Gewaltpräferenz in Video-/Computerspielen gefragt wurde, wird für diese nur die Filmgewaltpräferenz zugrundegelegt. Bei Normalspielern und Vielspielern, denen beide Fragen vorgelegt wurden, wird hingegen die durchschnittliche Angabe beider Präferenzstatements zugrundegelegt. Hierfür wurden beide Angaben aufsummiert und das Ergebnis durch zwei geteilt. Um die Verteilungseigenschaften des Indexes zu verbessern, wurde im Falle nicht ganzzahliger Ergebnisse auf die nächsthöhere Stufe des Indexes aufgerundet.⁸⁰ Hieraus ergibt sich die in Abbildung 26 dargestellte Verteilung, die mit Ausnahme der extremeren Ausprägungen 9 und 10 auf eine recht gleichmäßige Verteilung zwischen den Stufen des Indexes hindeutet.

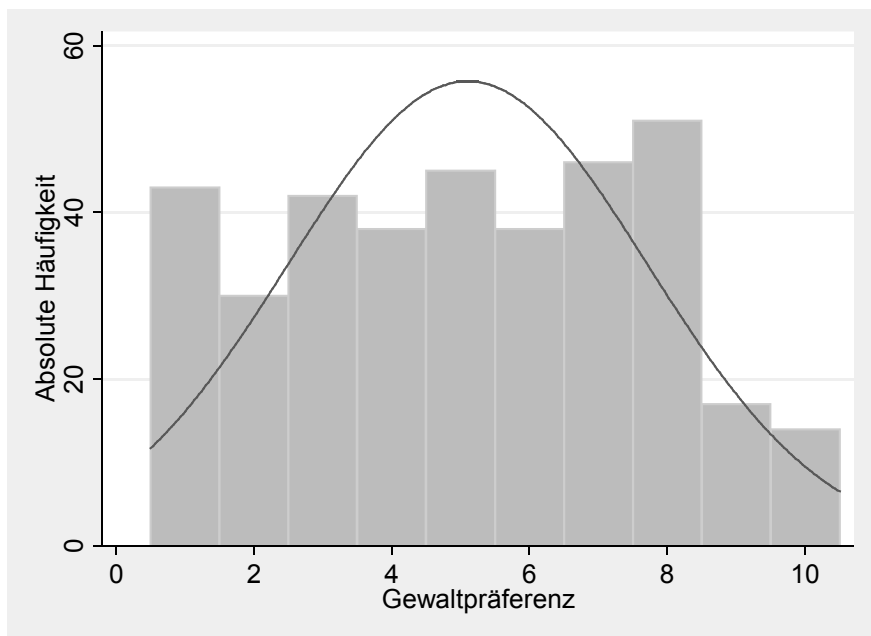


Abbildung 26. Verteilung der Gewaltpräferenzwerte der Versuchspersonen (n = 365)

80 So wurde z. B. bei einer Person, welche als Computerspielgewaltpräferenz 8 und als Filmgewaltpräferenz 9 angibt nicht der Wert 8.5 sondern 9 vergeben. Von einer Max-Kodierung, d.h. der Kodierung des höheren Ausprägungswertes beider Items, wurde abgesehen, da das Maß allgemeine Gewaltpräferenz anzeigen sollte.

2.1.7 Distresserleben

Zur Erfassung des Distresserlebens wurden zeitökonomisch anwendbare Selbstbeurteilungssitems eingesetzt. Auf die Erhebung physiologischer Daten wurde verzichtet, da der Anspruch, negative Gefühlsreaktionen mittels biopsychologischer Erlebniskorrelate erfassen zu wollen, breitbandige Messansätze notwendig gemacht hätte, die unter forschungsökonomischen Gesichtspunkten als nicht angemessen bewertet wurden. Zudem wurde befürchtet, dass die Abnahme von Speichelproben sowie die Anwendung elektronischer Messmethoden der Untersuchung einen "medizinischen Charakter" verliehen hätten, der das Versuchspersonenverhalten in nicht intendierter Weise hätte beeinflussen können. In verschiedenen aktuelleren Untersuchungen konnte inzwischen demonstriert werden, dass Selbstberichtsinstrumente sowohl zur Messung der allgemeinen Arousalwirkung von Computerspielen (vgl. Reinecke & Trepte, 2008) als auch zur Ausdifferenzierung von Eustress und Distress (Maass, et al., 2010) erfolgreich eingesetzt werden können.

Zur Erfassung negativer Gefühle im Zusammenhang mit den Treatment wurden die Probanden unmittelbar nach jeder Treatmentphase um ihre Einschätzung darüber gebeten, inwieweit sie in dieser Stress, Angst und Ekel empfunden hatten (vgl. Abbildung 27). Zu jedem emotionalen Zustand liegen somit zu jedem Probanden drei Messzeitpunkte vor.

Versetzen Sie sich bitte noch mal innerlich in das Fernseherlebnis zurück und geben Sie an, inwieweit die folgenden Aussagen auf Sie zutreffen.				
	völlig unzutreffend	eher unzutreffend	eher zutreffend	völlig zutreffend
Ich empfand Stress.	☐	☐	☐	☐
Ich empfand Ekel.	☐	☐	☐	☐
Ich empfand Angst.	☐	☐	☐	☐

Abbildung 27. Operationalisierung des Distresserlebens am Beispiel der Filmbedingungen

Eine Itemanalyse ergibt, dass die drei negativen Emotionen zu jedem Messzeitpunkt in ausreichendem Maße miteinander korrelieren. Das Cronbachs Alpha bewegt sich dabei je nach Messzeitpunkt zwischen .68 (MZIP2) und .73 (MZIP1), was in Hinblick auf eine Anzahl von drei Items als akzeptabel gelten kann. Zum Zwecke der Auswertung wurde auf Basis aller Items für jeden Probanden ein Rohpunktwert "Distresserleben" errechnet, der einen Wert zwischen 0 (alle Items verneint) und 27 (alle Items bejaht) annehmen kann. Der Gesamtskalenwert erreicht mit $\alpha = .86$ einen guten Konsistenzwert (vgl. Tabelle 17).

Tabelle 17. Itemkennwerte der Skala Distresserleben (n = 365)

MZP	Item	M	SD	r _{it}	α	r _{it}	α
1	Ich empfand Angst.	1.26	.62	.63		.62	
1	Ich empfand Ekel.	1.53	.93	.61		.56	.73
1	Ich empfand Stress.	1.64	.89	.59		.53	
2	Ich empfand Angst.	1.22	.60	.60		.59	
2	Ich empfand Ekel.	1.47	.85	.56	.86	.50	.68
2	Ich empfand Stress.	1.61	.88	.59		.45	
3	Ich empfand Angst.	1.23	.60	.61		.64	
3	Ich empfand Ekel.	1.50	.92	.70		.59	.72
3	Ich empfand Stress.	1.66	.88	.56		.47	

Anmerkung. r_{it} = Trennschärfe (Part-whole-Korrektur). α = Cronbachs Alpha. Skalenwerte von 1 bis 4.

Die Skala ist auch unter den für die nachfolgenden Auswertungen maßgeblichen Probanden der Gewaltbedingungen (n = 147) linksschief verteilt. Um die Verteilungseigenschaften der Skala zu verbessern, wurde eine Quadratwurzeltransformation der Daten vorgenommen, wodurch die Schiefe der Verteilung deutlich reduziert wird (vgl. Abbildung 28).

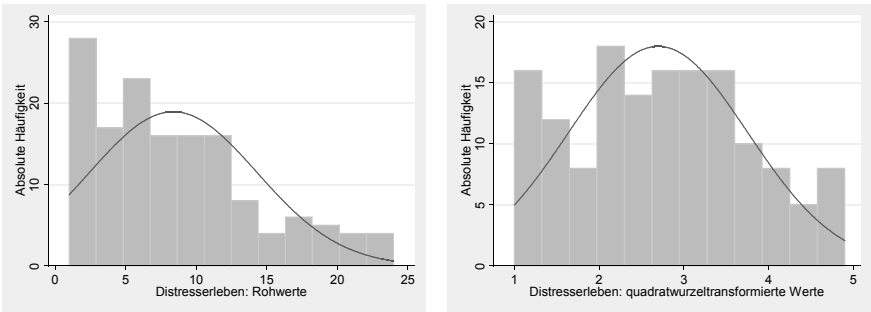


Abbildung 28. Verteilung und Normalverteilungsplot der Rohwerte (links) und quadrtwurzeltransformierten Werte (rechts) der Skala Distresserleben (nur Versuchspersonen der Gewaltbedingungen, n = 147)

2.1.8 Instrumente für den Manipulationscheck

Die Erhebung der Kontrollvariablen für den Manipulationscheck erfolgte im Rahmen der Bewertungshefte, die von den Versuchspersonen im Anschluss an jedes Treatment bearbeitet wurden (vgl. Kapitel 2.1.11). Um die Gewaltmanipulation der eingesetzten Medien auf ihre Wirksamkeit hin überprüfen zu können, wurden die Versuchspersonen um eine Einschätzung zum Gewaltgehalt der dargebotenen Angebote gebeten (vgl. Abbildung 29).

Wie <u>gewalthaltig</u> fanden Sie HOUSE OF THE DEAD III ?											
überhaupt nicht gewalthaltig	1	2	3	4	5	6	7	8	9	10	sehr gewalthaltig

Abbildung 29. Rating-Item zum Gewaltgehalt am Beispiel eines gewalthaltigen Computerspiels

Weiterhin sollte kontrolliert werden, ob unter den Treatmentbedingungen ein vergleichbares Unterhaltungserlebens erreicht wurde (vgl. Abbildung 30).

Wie <u>unterhaltsam</u> fanden Sie HOUSE OF THE DEAD III ?											
überhaupt nicht unterhaltsam	1	2	3	4	5	6	7	8	9	10	sehr unterhaltsam

Abbildung 30. Rating-Item zum Unterhaltungserleben am Beispiel eines gewalthaltigen Computerspiels

Zusätzlich sollte überprüft werden, ob die gewalthaltigen und gewaltneutralen Spiele, wie intendiert, einen ähnlichen Schwierigkeitsgrad aufweisen. Deshalb wurden die Versuchspersonen im Rahmen der interaktiven Bedingungen dazu aufgefordert, den Schwierigkeitsgrad von Spiel und Steuerung einzuschätzen. Zusätzlich wurden die Versuchspersonen darum gebeten, ihren eigenen Spielerfolg einzuschätzen und sich hierfür selbst eine Schulnote zu geben (vgl. Abbildung 31). Die Abfragemethoden zum Unterhaltungserleben, zum Gewaltgehalt und zum Schwierigkeitsgrad wurden je nach Treatmentbedingung im Wortlaut minimal variiert, um sie an den jeweiligen Gegenstandsbereich (Computerspiel, Film, nicht-mediale Aktivität) anzupassen.

Die Steuerung und das Spielen fand ich insgesamt...				
sehr leicht	eher leicht	mittelmäßig	eher schwer	sehr schwer
☐	☐	☐	☐	☐

Meine Leistungen im Spielen würde ich mit der folgenden Schulnote bewerten:					
Note	Note	Note	Note	Note	Note
1	2	3	4	5	6
☐	☐	☐	☐	☐	☐

Abbildung 31. Rating-Items zum Schwierigkeitsgrad und Spielerfolg in den interaktiven Medienbedingungen

2.1.9 Kontrolltechniken: Störvariablen

Aufgrund des laborexperimentellen Zugangs sowie der nahezu vollständigen Standardisierung des Untersuchungsablaufs kann von einer Kontrolle vielfältiger möglicher Störeinflüsse ausgegangen werden. Die Parallelisierung hinsichtlich Geschlecht, Spielhäufigkeit und Bildungsniveau mit anschließender Randomisierung der homogenen Versuchspersonen zu den Untersuchungsgruppen stellt eine besonders effektive experimentelle Kontrolltechnik dar. Mit einer Begrenzung des Altersspektrums einbezogener Versuchsteilnehmer wurde zudem die Altersheterogenität verringert.

Tabelle 18. Antizipierte Störeinflüsse im Untersuchungsdesign und angewandte Kontrolltechniken

Störvariable	Kontrolltechnik
Fitness vor Versuchsbeginn	Abfrage der Fitness vor Versuchsbeginn. Test auf Gruppenunterschiede.
Arbeitsmotivation vor Versuchsbeginn	Abfrage der Arbeitsmotivation vor dem ersten Leistungstest. Test auf Gruppenunterschiede.
Manipulationsversuche	Offene Abfrage von Manipulationsversuchen am Ende der Befragung. Bei Vorliegen: Versuchspersonenausschluss.
Raumtemperatur	Erfassung der Raumtemperatur in den Versuchsräumen. Test auf Gruppenunterschiede.
Gruppengröße	Erfassung der Gruppengröße. Test auf Gruppenunterschiede

Trotz dieser Maßnahmen können sich Versuchsgruppen zufällig hinsichtlich bestimmter Merkmale unterscheiden, die einen Einfluss auf die abhängigen Variablen aufweisen. In Tabelle 18 werden vorab definierte Störvariablen aufgeführt, die im Experiment nicht mit Sicherheit konstant zu halten waren.

Die im Rahmen von Selbstbeurteilungsinstrumenten⁸¹ zu kontrollierenden Störvariablen wurden mittels einfacher Single-Items und Skalen direkt in die Testhefte integriert. So wurde zu Beginn jedes Versuchstages die aktuelle Fitness der Versuchsperson mittels eines einfachen Single-Items erhoben (vgl. Abbildung 32).

Bevor es losgeht: Bitte schätzen Sie kurz ein, wie fit (vital, munter) Sie sich heute fühlen.

Sehr fit

☐

+++

☐

++

☐

+

☒

-

☐

--

☐

Überhaupt nicht fit

Abbildung 32. Kontrollitem zum aktuellen Fitnesszustand

Um die Arbeitsmotivation der Versuchspersonen zu erfassen wurden diese vor dem ersten Leistungstest darum gebeten, einzuschätzen, inwieweit die Aussagen "Ich bin gespannt, wie gut ich hier abschneiden werde" und "Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen" auf sie zutreffen (7-stufig; 1 = trifft nicht zu, 7 = trifft zu). Aus beiden Angaben wurde ein Mittelwertindex gebildet. Um manipulatives Verhalten der Versuchspersonen zu erfassen, wurden die Versuchspersonen zudem am Ende der Untersuchung um Angaben darüber gebeten, ob sie sich bei den Tests absichtlich weniger angestrengt haben, um die Ergebnisse der Untersuchung zu beeinflussen und ob sie manche Fragen, die ihnen unangenehm waren, nicht wahrheitsgemäß beantwortet haben. Versuchspersonen, die diesen Fragen zustimmen, sollten aus den Datenanalysen ausgeschlossen werden.

Ein weiterer wichtiger Teil der Störvariablenkontrolle wurde durch Versuchsleiter und Versuchsassistentz durchgeführt. So sah das Versuchsleiterprotokoll (vgl. Anhang F) eine Kurzbewertung des Untersuchungsablaufs vor, in welcher auch mögliche Störungen zu protokollieren waren. Hier wurde auch festgehalten, wie viele Versuchspersonen gleichzeitig am Versuch teilgenommen haben. Da die Raumtemperatur in Ermangelung einer Klimaanlage in den Laborräumen nicht konstant gehalten werden konnte und die Untersuchung in den Sommermonaten

81 Selbstberichtsinstrumente können in diesem Zusammenhang zwar nicht dafür herangezogen werden, Störeinflüsse im Einzelfalls auszuschließen, da hier eine vermehrte Beantwortung nach sozialer Erwünschtheit angenommen werden muss. Sie können jedoch auf gravierende Störeinflüsse, die im Einzelfall einen Versuchspersonenausschluss notwendig machen, hinweisen. Zusätzlich ermöglichen sie auf der Gruppenebene einen Vergleich der Treatmentbedingung nach der Häufigkeit des Auftretens bestimmter Störeinflüsse.

vorgenommen wurde, musste auch dieser Parameter durch Versuchsleitung und Assistenz festgehalten werden.

2.1.10 Einrichtung der Laborräume

Die Laborräume bestanden aus einem Eingangsbereich mit Garderobe (vgl. Abbildung 33), einem Lernraum sowie einem medialen und einem nicht-medialen Treatmentraum. Da entsprechende institutseigene Räumlichkeiten nicht zur Verfügung standen, wurden diese für die Zeitdauer des Versuchs zentral im Stadtgebiet Hannover angemietet und eingerichtet. Um zu einer ungezwungenen Atmosphäre beizutragen, wurden die Laborräume durch Dekorationsgegenstände wie Bilder und Pflanzen wohnlich gestaltet. Auf eine Videoüberwachung wurde verzichtet.



Abbildung 33. Einrichtung des Eingangsbereichs der Laborräume

Der Lernraum bestand aus sechs Arbeitsplätzen mit Lerncomputer sowie einem Schreibtisch für den Versuchsleiter (vgl. Abbildung 34). Der VL sitzt, wie in testdiagnostischen Situationen üblich, den Versuchspersonen gegenüber. Der gesamte Raum wurde hell beleuchtet und schlicht eingerichtet, um die Ernsthaftigkeit einer Lern- bzw. Prüfungssituation zu unterstreichen. Bei den Lerncomputern handelte es sich um einfache Computer mit 14-Zoll-Monitor, auf denen die Experimentalssoftware *Super Lab* sowie das Lerntestverfahren *WERNICKO* eingerichtet waren. Rechtsseitig auf den Tischen wurde die Tastatur für die Bedienung der Lernsoftware angeordnet. Auf der linken Seite befand sich ausreichend Platz zum Bearbeiten der Test- und Befragungsinstrumente. Die Stühle wiesen ausreichend Komfort für kürzere Büroarbeitszeiten auf und waren höhenverstellbar.

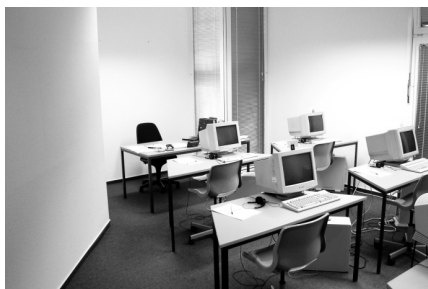


Abbildung 34. Einrichtung des Lernraums

Die geschützte Atmosphäre des eigenen Wohnraums ist im Labor nur eingeschränkt reproduzierbar. Damit bestand für die Untersuchung die Gefahr, dass sich die Versuchspersonen nicht intensiv genug auf die dargebotenen Freizeitbedingungen einlassen würden und sich hierbei möglicherweise unwohl fühlen. Deshalb wurde bei den Freizeitbedingungen Wert auf hochwertige Ausstattungsmerkmale gelegt. So wurden hinsichtlich der Medienbedingungen auf großformatige Anzeigegeräte (Fernseher, Monitor, ca. 70 cm Bildschirmdiagonale), eine gute Soundtechnik mittels Stereokopfhörer sowie authentische Eingabemedien für die Computerspiele (z. B. authentische Lightgun, Lenkrad mit motorischem Feedback und Fußpedalen) zurückgegriffen, um die Faszinationskraft der Medienangebote zu erhöhen. Darüber hinaus sollte die gesamte Einrichtung im Freizeitraum einen geschützten Rahmen vermitteln. Dies wurde u.a. durch Abtrennung der Rezeptionsbereiche durch Trennwände erreicht (vgl. Abbildung 35).



Abbildung 35. Einrichtung der Medienbedingungen

Das Licht im Raum wurde leicht abgedunkelt, um auch an sonnigen Nachmittagen eine geeignete Stimmung aufkommen zu lassen. Eine klare Rollentrennung

zwischen Versuchsleiter (Lernen) und Versuchsassistentz (Freizeit) sollte es den Versuchspersonen zudem erleichtern, sich in Gegenwart der Versuchsassistentz entspannt auf die Freizeitbedingungen einzulassen (vgl. Anhang G, H, I). Im Raum befanden sich insgesamt sechs Medienstationen bestehend aus Anzeigergerät (Fernseher oder Monitor), Abspielgerät (XBOX, Playstation 2, PC oder DVD-Player) und den für die Spielbedingung notwendigen Eingabegeräten. Die Fernsehgeräte wiesen von der Raummitte nach außen. Auf der einen Seite des Raumes befanden sich die gewaltneutralen, auf der anderen Seite des Raumes die gewalthaltigen Bedingungen. Vor jeder Rezeptionsphase wurde eine einminütige Orientierungsphase vorgesehen, in der sich die Versuchspersonen mittels großformatiger Anzeigetafeln über das zu nutzende Medium informieren konnten. Jeweils im Wochenwechsel fanden im Medientreatmentraum die Filmbedingungen oder die Computerspielbedingungen statt, um evtl. vorhandenen saisonalen Störeinflüssen entgegenzuwirken. Durch die parallele Darbietung gewalthaltiger und gewaltneutraler Medien konnten weitere denkbare Störeinflüsse, wie etwa mögliche Versuchsleitereffekte, effektiv kontrolliert werden.

Die fünfte, explorative Bedingung fand in einem zweiten Treatmentraum statt (vgl. Abbildung 36). Dieser enthielt die drei frei zugänglichen Freizeitbereiche Tischtennis, Tischfußball und Dart. Der Raum bot ausreichend Platz für bis zu sechs Versuchspersonen. Für die Versuchsassistentz war ein Platz im Raum vorgesehen, bei dem diese alle Bereiche überblicken konnte.



Abbildung 36. Einrichtung der explorativen Zusatzbedingung

2.1.11 Ablauf der Untersuchungen

Jeder Versuch wurde an beiden Untersuchungstagen von einem Versuchsleiter und am ersten Tag zusätzlich von einer Versuchsleiterassistentz betreut. Der gesamte Versuch wurde von der Begrüßung der Versuchspersonen zu Beginn des ersten

Versuchstag bis zu ihrer Verabschiedung am Ende des zweiten Versuchstages mithilfe wortwörtlicher Instruktionsmanuale vollständig standardisiert (vgl. Anhang G, H, I). Die Aufgaben des Versuchsleiters bestanden maßgeblich in der Testdiagnostik und Befragung der Versuchspersonen. Die Versuchsassistentin war hingegen für die Treatmentbedingung sowie die Vorbereitung des ersten Versuchstages zuständig.

Die Studie wurde als *Freizeit-Medien-Studie* angekündigt. Den Versuchspersonen wurde zu den Hintergründen der Untersuchung nur mitgeteilt, dass die Untersuchung einen Beitrag zur Lern- und Freizeitforschung leisten soll und dass einige Aktivitäten die Versuchspersonen möglicherweise an ihre Schulzeit erinnern werden (vgl. Anhang G). Eine Geheimhaltung des Erkenntnisinteresses erschien notwendig, um Versuchspersoneneffekte zu minimieren. Alle Teilnehmer wurden zudem explizit auf die Freiwilligkeit der Untersuchung und die Sicherstellung des Datenschutzes hingewiesen. Ihnen wurde auch mitgeteilt, dass sie den Versuch jederzeit abbrechen können. Die Versuchspersonen wurden darüber hinaus dazu angehalten, während des Versuchs nicht miteinander über die Lerninhalte des Versuchs zu sprechen und ihre Handys für die Dauer des Versuchs auszuschalten. Durch die Konzipierung eines Gruppenversuchs konnten bis zu sechs Versuchspersonen gleichzeitig an dem Experiment teilnehmen (Minimal: 2 Personen).

Jeder Versuch gliederte sich in einen ersten und einen zweiten Untersuchungstag. Dieses Vorgehen wurde zum einen aufgrund der Zielsetzung notwendig, für den Gedächtnis- und Lernleistungstest eine um 24-Stunden verzögerte Nachmessung zu realisieren (vgl. Kapitel 2.1.4). Zum anderen erschien eine Aufteilung der insgesamt rund 4,5 Stunden in Anspruch nehmenden Datenerhebung notwendig, um die Versuchspersonen nicht zu überfordern. Durch das Sampling, die Randomisierung und die verhältnismäßig großen Treatmentgruppen konnte erwartet werden, dass Versuchspersonen unterschiedlicher Gruppen sich innerhalb dieses Zeitraumes nicht systematisch unterschiedlich verhalten. Zudem wurden die Versuchspersonen darum gebeten, zwischen den Untersuchungstagen auf den Genuss von Alkohol zu verzichten (vgl. Anhang G).

Am ersten Untersuchungstermin wurden mit Prä- und Postmessung die vollständige Konzentrationsmessung (*KLT-R*), die ersten vier der insgesamt fünf Lern-testphasen des computergestützten Fremdsprachtests (*WERNICKO*) sowie Einprä-gephase und Prämessung des räumlich-visuellen und verbalen Gedächtnismaterials (*IWM*) durchgeführt. Die Konzentrationsmessung schloss die Abfrage zur Arbeits-motivation mit ein (vgl. Kapitel 2.1.9). An diesem ersten Versuchstag fanden auch alle drei Treatmentblöcke statt. Der erste Untersuchungstermin dauerte insgesamt etwa 3 Stunden (vgl. Abbildung 37, Anhang G, H).

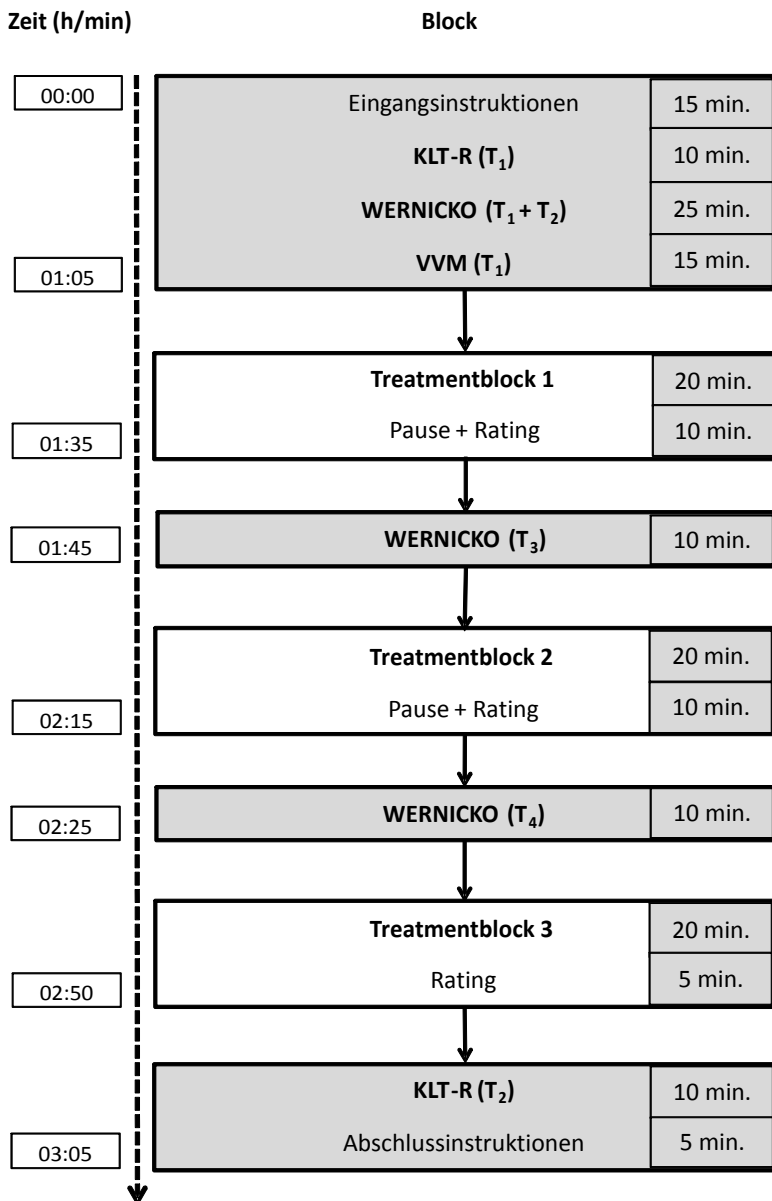


Abbildung 37. Versuchsablauf am ersten Untersuchungstag (Gesamtdauer: 185 Minuten). KLT-R = Konzentrations-Leistungs-Test. WERNICKO = Fremdsprachenlernetest. VVM = Visueller und Verbaler Merkfähigkeitstest.

Vierundzwanzig Stunden verzögert nach Beginn der ersten Untersuchung fand der zweite Untersuchungstermin statt. Hier wurde der computergestützte Fremdsprachentests (*WERNICKO*) mittels einer fünften Lernphase sowie einer Transferaufgabe abgeschlossen. Den Abschluss des *Visuellen und Verbalen Merkfähigkeitstests (VVM)* bildete eine Nachmessung der räumlich-visuellen und verbalen Gedächtnisleistung zu den am Vortag erlernten Inhalten. Zusätzlich wurde am zweiten Versuchstag die Intelligenzdiagnostik durchgeführt sowie die Daten zu allen weiteren Befragungsinstrumenten mittels eines Fragebogens erhoben. Nichtspieler erhielten dabei eine leicht gekürzte Variante, die keine Fragen zur Computernutzung beinhaltete. Am Ende des zweiten Versuchstages wurden die Versuchspersonen verabschiedet. Dabei wurde ihnen für ihre Teilnahme und Unterstützung gedankt und die Versuchspersonenentschädigung direkt in bar ausgezahlt. Die Versuchspersonen wurden zudem darüber in Kenntnis gesetzt, dass sie nach Abschluss der Feldphase der Untersuchung bei Interesse genauere Informationen über die Hintergründe der Untersuchung erhalten können.⁸² Der zweite Versuchstag nahm weitere 1,5 Stunden pro Versuchsperson in Anspruch (vgl. Abbildung 38, Anhang I).

82 Hierfür konnten sich alle Interessenten in eine Emailliste eintragen. Das nach der Feldphase versandte Informationsschreiben enthielt Hinweise zum wissenschaftlichen Hintergrund und Erkenntnisinteresse der Untersuchung und legte dar, welche Methoden mit welcher Zielsetzung eingesetzt wurden. Nach Veröffentlichung dieser Arbeit ist eine Ergebnispräsentation der Daten unter den Interessenten geplant.

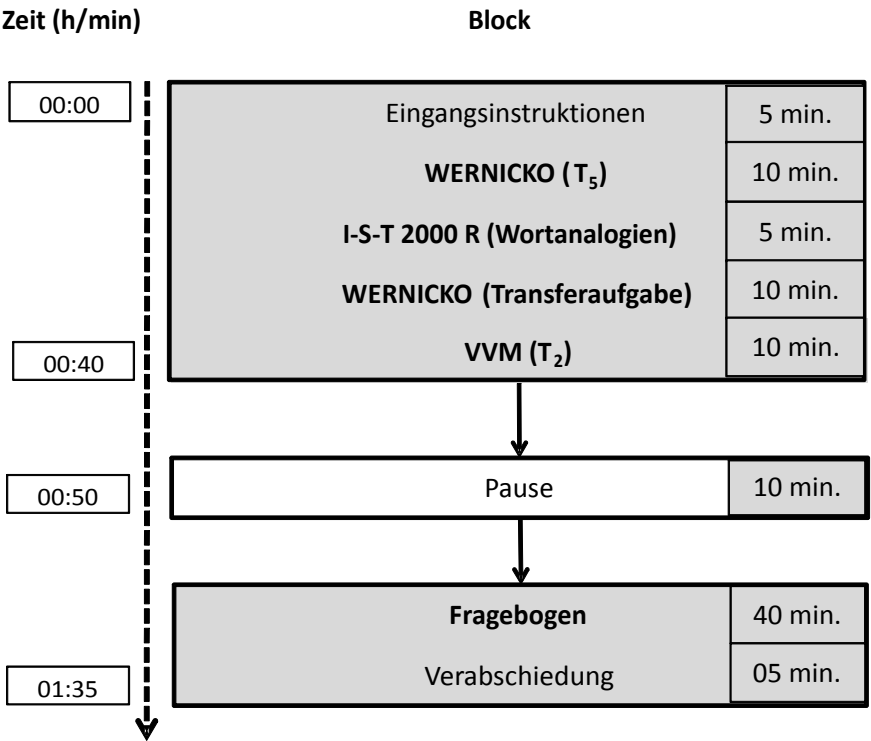


Abbildung 38. Versuchsablauf am zweiten Untersuchungstag (Gesamtdauer: 95 Minuten). KLT-R = Konzentrations-Leistungs-Test. WERNICKO = Fremdsprachenlernetest. VVM = Visueller und Verbaler Merkfähigkeitstest. I-S-T 2000 R = Intelligenz-Struktur-Test 2000 R, Untertest Analogien der Dimension verbale Intelligenz.

2.2 Ergebnisse

2.2.1 Elektronische Datenverarbeitung

Die Daten wurden in eine SPSS-Datenmatrix eingegeben und nach einer Kontrolle von Eingabefehlern zusätzlich in das STATA-Format exportiert. Alle nachfolgenden Datenauswertungen fanden mit den Programmen STATA (Version SE 10) und SPSS (Version 15) statt. Zum Zwecke der Post-Hoc-Poweranalysen wurde die Freeware G*Power (Version 3) eingesetzt. Visualisierung von Ergebnissen wurde zum Teil in Office Excel (Version 2007) vorgenommen.