

**Einfluss eines Benzodiazepines  
sowie eines Phytopharmakons  
auf emotionale Befindlichkeit  
und kognitive Leistungsfähigkeit**

- eine monozentrische, randomisierte, placebokontrollierte,  
pharmakopsychologische Doppelblindstudie –

**als Dissertation vorgelegt von:**

**Wolfgang Berger**

**Om Rodde 12**

**D – 51427 Bergisch-Gladbach**

- Die veröffentlichte Fassung ist von der Philosophischen Fakultät der Universität Köln angenommen -

## **Gesamtübersicht zur Arbeit**

### **Studientitel**

Einfluss eines Benzodiazepines sowie eines Phytopharmakons auf emotionale Befindlichkeit und kognitive Leistungsfähigkeit - eine monozentrische, randomisierte, placebokontrollierte, pharmakopsychologische Doppelblindstudie -

### **Rationale und wissenschaftliche Zielsetzungen**

Der psychologischen Farbforschung mangelt es an aktuellen und methodisch anspruchsvollen Studienergebnissen. Die vorliegende pharmakopsychologische Studie wurde als explorative Kontrolluntersuchung zur Studie „Influence of Valerian Treatment on Reaction Time, Alertness and Concentration in Volunteers“ (Kuhlmann, Berger, Podzuweit und Schmidt, 1999) konzipiert und knüpft an die Tradition der psychologischen Farbforschung an. Folgende Zielsetzungen wurden verfolgt:

- Exploration der individuellen Farbwahrnehmung der computergenerierten Farbpalette des Kölner Farbpräferenztests durch den methodischen Einsatz eines Semantischen Eindrucksdifferentials (ED)
- Definition der Zielvariablen und Entwicklung einer speziellen KFPT-Auswertestrategie
- Veränderungsmessung der KFPT-Zielvariablen unter medikamentösem Einfluss durch ein Benzodiazepin und ein Phytopharmakon
- Gegenüberstellung der KFPT-Befunde zum Farbwahlverhalten (als Indikator für die emotionale Befindlichkeit) mit den Ergebnissen zur kognitiven Leistungsfähigkeit im Determinationstest (DT) zwecks einer ersten Beurteilung der Pharmakosensitivität des KFPT und seiner Eignung als Verlaufsparemeter.

### **Studiendesign**

Zur Beantwortung der o.g. wissenschaftlichen Fragestellungen wurde der doppelblinde, placebo-kontrollierte Parallelgruppenvergleich als die geeignetste und allgemein akzeptierte Methode gewählt. Da psychologische Effekte nach einmaliger Sedativa-Einnahme als auch nach 14-tägiger medikamentöser Einnahme untersucht werden sollten, wurde ein zweigeteiltes Studiendesign entwickelt.

*Studienteil 1* (einmalige abendliche Medikationseinnahme)

Unizentrischer, dreiarmiger, doppelblinder, kontrollierter, randomisierter Parallelgruppenvergleich Baldrian-Extrakt vs. Benzodiazepin vs. Placebo

*Washout-Phase* (7 Tage ohne Medikationseinnahme)

*Studienteil 2* (14tägige abendliche Medikationseinnahme)

Unizentrischer, zweiarmiger, doppelblinder, kontrollierter, randomisierter Parallelgruppenvergleich Baldrian-Extrakt vs. Placebo

### **Studiendauer/Proband**

23 Tage (Studienteil 1: 2 Tage / Washout-Phase: 7 Tage / Studienteil 2: 14 Tage)

## **Studienteilnehmer**

102 weibliche und männliche Probanden im Alter zwischen 30 und 60 Jahren

## **Studiendurchführung**

Die praktische Umsetzung der Studie erfolgte in den Räumlichkeiten des „Instituts für interdisziplinäre medizinische Forschung“ (IMF) in Köln.

Die psychometrischen Untersuchungen wurden vom Doktoranden, die medizinischen Untersuchungen von einem beim IMF beschäftigten approbierten Arzt durchgeführt.

## **Studienmedikationen**

### **Baldrian-Trockenextrakt**

Darreichungsform: 300 mg Kapseln (Studienteil 1) / Dragées (Studienteil 2)  
Applikationsart: oral, zum Einnehmen  
Dosierungsschema: 2 Kapseln / Dragées abends vor dem Schlafengehen  
Behandlungsdauer: 1 Tag (Studienteil 1) bzw. 14 Tage (Studienteil 2)

### **Flunitrazepam**

in Form, Größe und Farbe mit dem Baldrianpräparat und der Placebo-Zubereitung identische Kapseln

Darreichungsform: 1 mg Filmtablette, verkapselt  
Applikationsart: oral, zum Einnehmen  
Dosierungsschema: 1 Kapsel + 1 Placebo-Kapsel abends vor dem Schlafengehen  
Behandlungsdauer: 1 Tag (Studienteil 1)

### **Placebo**

In Form, Größe und Farbe mit dem Baldrian- bzw. Flunitrazepampräparat identische Kapsel bzw. identisches Dragée

Darreichungsform: Kapseln (Studienteil 1) / Dragées (Studienteil 2)  
Applikationsart: oral, zum Einnehmen  
Dosierungsschema: 2 Kapseln / Dragées abends vor dem Schlafengehen  
Behandlungsdauer: 1 Tag (Studienteil 1) bzw. 14 Tage (Studienteil 2)

## **Einschlusskriterien**

- schriftliches Einverständnis nach mündlicher und schriftlicher Aufklärung
- 30-60jährige weibliche und männliche Probanden
- Farbtüchtigkeit/normale Tagessehschärfe mit oder ohne Sehhilfe mindestens 0.7/0.7

- normales Körpergewicht ( $\pm 10\%$  nach BROCA)
- nach DIPS-Befund keine psychiatrische Erkrankung (ICD-10)
- keine anamnestisch bekannte organische Schlafstörungen (ICD-10/G 47)
- soziales Trinkverhalten (evaluiert durch den ATV)
- Atemalkoholkonzentrations-Messwert (AAK) 0.00 mg/l  
(keine Alkoholisierung bei der Aufnahme bzw. den Folgeuntersuchungen)

### **Ausschlusskriterien**

- Einnahme von Psychopharmaka, Hypnotika, Sedativa, Analgetika oder anderen Pharmaka, die das Schlafverhalten oder die kognitive Leistungsfähigkeit beeinflussen, in den vier Wochen vor Beginn oder während der Teilnahme an der Studie
- psychiatrische Erkrankungen
- Alkohol-, Drogen- und Medikamentenabhängigkeit
- akute Vergiftung mit Alkohol, Schlaf-/ Schmerzmitteln, Neuroleptika, Antidepressiva, Lithium
- Hypersomnien, Störungen des Schlaf-Wach-Rhythmus, Parasomnien
- jede Erkrankung, die bekanntlich zu sekundären Schlafstörungen führt, wie z.B. Neurodermitis, chronischer Juckreiz (Lebererkrankungen) u.a.
- Myasthenia gravis
- spinale und/oder zerebrale Ataxie
- eingeschränkte Leber- oder Nierenfunktion
- respiratorische Insuffizienz (anamnestisch)
- nicht kompensierte Herzinsuffizienz
- Morbus Parkinson
- arterielle Hypotonie
- Engwinkelglaukom
- hirnorganisches Psychosyndrom
- erhöhte Krampfbereitschaft
- schwere konsumierende Erkrankungen (z.B. Malignom)
- schwere Blutbildveränderungen  
(Hb, Erythro-, Leuko- und Thrombozyten  $> 10\%$  außerhalb der Norm)
- bekannte Überempfindlichkeit gegen Flunitrazepam
- bekannte Überempfindlichkeit gegen Baldrian-Extrakt
- Schwangerschaft und Stillzeit bei Frauen im gebärfähigen Alter
- fehlender Konzeptionsschutz bei Frauen im gebärfähigen Alter

## **Methodologie**

### **Psychometrische Zielvariablen**

Kölner Farbpräferenztest (KFPT):

Die KFPT-Zielvariablen wurden durch eine Clusteranalyse der mittleren Dimensionswerte (Baseline) für Erregung, Valenz, Potenz des semantischen Eindrucksdifferentials bestimmt.

- Cluster (Farbverbände)
- Einzelfarbtöne

Semantisches Eindrucksdifferential (ED):

- Dimensionswerte für Erregung, Valenz, Potenz der 14 KFPT-Farbtöne

Determinationstest (DT):

- Median der Reaktionszeit
- Anzahl richtiger Reaktionen
- Anzahl falscher Reaktionen
- Anzahl ausgelassener Reaktionen
- Anzahl der bearbeiteten Reize.

### **Sicherheitsvariablen (Drug Safety)**

- Nebenwirkungen/unerwünschte Ereignisse (UE)
- klinische Laborbefunde
- Tolerabilität/Verträglichkeit

### **Statistische Methoden**

Bei dieser explorativen Studie wurden vorab keine Hypothesen hinsichtlich des Farbwahlverhaltens der Probanden im KFPT formuliert. Alle Analysen wurden deskriptiv durchgeführt. Für die Interpretation der KFPT-Daten wurden zusätzlich p-Werte für die Cluster bzw. Einzelfarben berechnet. Innerhalb der Medikationsgruppen wurden für jeden Parameter die Zeitpunkte Baseline, Retest 1, Retest 2 mittels des Wilcoxon-Tests miteinander verglichen. Die Medikationsgruppen wurden anhand der Parameter zur Baseline mit dem Mann-Whitney-U-Test, zu Retest 1 und Retest 2 mit dem Paardifferenzen-U-Test (Differenz zur Baseline) verglichen; p-Werte  $\geq 0.2$  wurden als nicht signifikant (n.s.) wiedergegeben. Weiterhin wurden Spearman-Rangkorrelationskoeffizienten zwischen Valenzwert im semantischen Eindrucksdifferential und der Wahlfähigkeit des Clusters bzw. des Einzelfarbtönen im KFPT berechnet.

## **Auswertekollektive**

Insgesamt lagen 111 Dokumentationen zur Auswertung vor. Neun Probanden wurden lediglich voruntersucht (gescreent). Aufgrund von Verstößen bezüglich der Ein- oder Ausschlusskriterien wurden diese neun Probanden von der Studienteilnahme ausgeschlossen.

102 Probanden wurden in die Studie aufgenommen (ITT-Kollektiv). Keiner dieser Probanden wies bei der Aufnahme eine unerlaubte Begleiterkrankung oder –medikation auf. Sämtliche Einschlusskriterien wurden erfüllt und keines der Ausschlusskriterien wurde verletzt.

Insgesamt vier Probanden brachen die Studie vorzeitig ab (ST1: 3 Probanden/ST2: 1 Proband). Da nicht für alle Kontrollzeitpunkte eines Studienteils Messwerte vorlagen, wurden die Dokumentationen dieser vier Probanden für die jeweiligen ATP-Analysen nicht berücksichtigt.

### *Kollektiv Studienteil 1*

Für die ATP-Auswertung des Studienteils 1 (N=99) konnten 33 Probanden der Gruppe Baldrian-Extrakt, 34 Probanden der Placebo- und 32 Probanden der Flunitrazepam-Gruppe herangezogen werden. Drei Probanden, (Baldrian-Extrakt N=1, Flunitrazepam N=2) wurden von der ATP-Analyse ausgeschlossen.

### *Kollektiv Studienteil 2*

Im ATP-Kollektiv des Studienteils 2 (N=91) wurden die Daten von 47 Probanden der Baldrian-Extrakt-Gruppe und 44 Probanden der Placebo-Gruppe berücksichtigt. Die Daten von insgesamt 11 Probanden wurden aufgrund von Protokollverstößen (10 Probanden) bzw. vorzeitigem Behandlungsende (1 Proband) von der ATP-Analyse ausgeschlossen.

Grundsätzlich wurde das ATP-Kollektiv (According-To-Protocol=Kollektiv derjenigen Probanden, welche die Studie gemäß dem Untersuchungsplan komplett durchlaufen hatten) zugrunde gelegt, da dies - nach den Planungsvorgaben der Studie - das relevante Kollektiv für die Analyse der Hauptzielvariablen darstellte. Bei der zusätzlich durchgeführten ITT-Analyse (Intention-To-Treat) wurden auch diejenigen Probanden berücksichtigt, die lediglich einen Aufnahme- und Folgewert aufwiesen.

## **Ergebnisse und Schlussfolgerungen**

### **Kölner Farbpräferenztest (KFPT)**

#### *Studienteil 1*

Die Baselinebefunde des Gesamtkollektivs waren durch eine große Range der Messwerte geprägt; Streuung und Extrema lagen zum Teil über dem arithmetischen Mittel (Cluster 1, 1b, 5b, 10c), was auf große interindividuelle Varianzen im Farbwahlverhalten hinweist, welche die Generalisierbarkeit der erhobenen Ergebnisse (Mittelwerte) einschränkt. Trotzdem bleibt als Befund festzuhalten, dass bei den schönen Pyramiden das Cluster 2 und der Farbton 1b (rot) bevorzugt wurden, während bei den hässlichen Pyramiden der Farbton 10c (schwarz) eindeutig präferiert wurde.

In der Placebo-Kontrollgruppe wurden bei der Folgetestung keine psychologisch relevanten Messwertveränderungen gegenüber der Ausgangslage objektiviert. Auch in den beiden Verumgruppen wurden nur wenige ausgeprägte Veränderungen der Farbpräferenzen beim Retest evaluiert. Unter Flunitrazepam zeigte sich bei der Wahl der schönen Pyramiden im Cluster 2 eine bei der Folgetestung signifikante Veränderung ( $p=0.016^*$ ) und bei Farbton 10c ein Trend ( $p=0.061+$ ). Unter Baldrian-Extrakt wurde bei der Wahl der hässlichen Pyramiden im Cluster 1

signifikante Veränderungen bei Retest 1 ( $p=0.019^*$ ) und bei Farbton 1b ein Trend ( $p=0.063+$ ) gemessen.

Vor dem Hintergrund dieser Ergebnisse erscheint die mögliche Beeinflussung des Farbwahlverhaltens durch die beiden aktiven pharmakologischen Substanzen unwahrscheinlich. Die im DT objektivierte Beeinträchtigung der kognitiven Leistungsfähigkeit durch das Benzodiazepin korrespondiert nicht mit dem KFPT-Befund. Farbwahlverhalten und emotionale Befindlichkeit wurden durch Flunitrazepam offensichtlich nicht verändert.

Im Gruppenvergleich Placebo vs. Baldrian-Extrakt ergaben sich insgesamt keine bedeutsamen Messwertunterschiede. Die Gegenüberstellung von Flunitrazepam und Baldrian-Extrakt zeigte nur bei den schönen Pyramiden für das Cluster 2 bei Retest 1 ( $p=0.023^*$ ) und bei den hässlichen Pyramiden für das Cluster 1 in der Ausgangslage ( $p=0.024^*$ ) signifikante Unterschiede. Für Flunitrazepam und Placebo ergab sich bei den schönen Pyramiden ein signifikanter Unterschied bei Retest 1 im Cluster 2 ( $p=0.40^*$ ) und für den Farbton 5b ein Trend ( $p=0.074+$ ). Insgesamt wurden nach einmaliger abendlicher Einnahme der Medikation bei der Folgetestung am darauf folgenden Morgen weder innerhalb der drei Behandlungsgruppen noch im Gruppenvergleich gravierende Veränderungen des Farbwahlverhaltens deutlich.

#### *Studienteil 2*

Medikationsunterschiede (Baldrian-Extrakt/Placebo) hinsichtlich des Farbwahlverhaltens ließen sich nicht belegen. Ausgeprägte Veränderungen der Farbpräferenzen nach 2wöchiger Therapie waren nur bei Cluster 3 - bei schönen und hässlichen Pyramiden – in beiden Medikationsgruppen festzustellen. Diese Befunde wurden nicht durch die Ergebnisse anderer Cluster bzw. Einzelfarben bestätigt. Insgesamt kam es innerhalb der Medikationsgruppen nur zu geringen Messwertveränderungen.

In der Gegenüberstellung von Baldrian-Extrakt und Placebo zeigten sich nach 14tägiger Therapie keine relevanten Unterschiede im Farbwahlverhalten. Ein medikamentöser Einfluss durch das Baldrianpräparat auf die emotionale Befindlichkeit ist auszuschließen.

### **Semantisches Eindrucksdifferential (ED)**

#### *Studienteil 1*

Gegenüber der Ausgangslage im Gesamtkollektiv (Baseline) veränderte sich nach einmaliger Einnahme der Medikationen (Retest 1) das Farberleben – unabhängig von der Behandlungsgruppe - hinsichtlich der E/V/P-Dimensionen nur unwesentlich.

Im Vergleich der drei Medikationsgruppen ergaben sich nur bei Cluster 3 für den Vergleich Baldrian-Extrakt vs. Placebo in der Dimension „Potenz“ ( $p=0.020^*$ ) und beim Farbton 5c für den Vergleich Baldrian-Extrakt vs. Flunitrazepam ( $p=0.015^*$ ) signifikante Gruppenunterschiede. Diese beiden Befunde alleine sind jedoch kein Indiz für ein generell verändertes Farberleben. Insgesamt zeigten die Gruppenvergleiche keine statistisch signifikanten und klinisch relevanten Veränderungen in der Farbwahrnehmung der Probanden.

#### *Studienteil 2*

In Relation zu den Ausgangsbefunden im Gesamtkollektiv veränderte sich nach 14tägiger Einnahme der Medikationen das Farberleben hinsichtlich der E/V/P-Dimensionen nur unwesentlich. Klinisch relevante Unterschiede zwischen den Behandlungsgruppen waren nicht objektivierbar. Eine Beeinträchtigung des Farberlebens unter Baldriantherapie ist damit unwahrscheinlich.

## **Determinationstest (DT)**

### *Studienteil 1*

In drei Variablen des Determinationstests – Reaktionszeit (Median), Anzahl der richtigen Reaktionen, Anzahl der bearbeiteten Reize - war der Leistungszuwachs der Flunitrazepam-Gruppe signifikant geringer als in den beiden Vergleichsgruppen. Dabei waren die unter Flunitrazepam aufgetretenen Einschränkungen der Leistungsfähigkeit psychologisch relevant. Im DT zeigte sich bei keiner Variablen ein statistisch auffälliger Unterschied zwischen Placebo und Baldrian-Extrakt. Nach einmaliger vorabendlicher Einnahme waren beim Retest am Morgen des Folgetages Reaktionsgüte und Leistungsfähigkeit der Probanden beider Medikationsgruppen in der Hauptzielvariablen (Median der Reaktionszeit) gleichwertig.

Die einmalige Einnahme eines Baldrian-Extraktes führte zu keiner Beeinträchtigung des Reaktionsvermögens, der Aufmerksamkeit oder der Koordinationsleistung. Ein vigilanzmindernder Hangover-Effekt wurde nicht objektiviert. Hinsichtlich des erzielten Übungsfortschrittes waren Baldrian-Extrakt und Placebo-Kontrollgruppe in ihrer Leistungsfähigkeit vergleichbar.

Für das Phytopharmakon ließen sich keine sedativen Residualwirkungen belegen, wohl aber für die chemisch definierte aktive Substanz. Unter Flunitrazepam war das Reaktionsvermögen der Probanden in der DT-Hauptzielvariablen (Median der Reaktionszeit) beeinträchtigt. Auch bei der niedrigen Dosierung von 1 mg/die Flunitrazepam konnte ein Hangover-Effekt bei 25% der Probanden nachgewiesen werden, auch war der Übungseffekt signifikant geringer als in den Vergleichsgruppen.

Verkehrsrelevante kognitive Leistungsminderungen am Folgetag sind auch nach nur einmaliger vorabendlicher Flunitrazepam-Einnahme nicht auszuschließen.

Die DT-Ergebnisse für Placebo und Baldrian-Extrakt korrespondieren mit den farbpsychologischen Resultaten von KFPT und ED. In allen Verfahren zeigten sich in beiden Behandlungsgruppen keine Hinweise auf kognitive Beeinträchtigungen bzw. emotionale Veränderungen. Die im DT objektivierete Beeinträchtigung des Reaktionsvermögens unter Flunitrazepam wurde durch die farbpsychologischen Ergebnisse nicht gestützt. Im ED und KFPT zeigten sich keine bedeutsamen Veränderungen des Farberlebens oder der Farbpräferenzen.

### *Studienteil 2*

Die Studienergebnisse belegten keine Vigilanzminderung nach einer 14tägigen Baldriantherapie. Im DT wurden in keiner Variablen signifikante Unterschiede zu Placebo objektiviert. Das Konfidenzintervall schließt für die variable Reaktionszeit (Median) eine relevante Beeinträchtigung unter Baldrian-Extrakt mit 97.5%iger Sicherheit aus. Die Äquivalenz zwischen Baldrian-Extrakt und Placebo konnte bestätigt werden. Das Ausmaß der gemessenen Leistungssteigerungen im Abschlusstest nach zweiwöchiger Behandlung war in beiden Untersuchungsgruppen vergleichbar.

DT-Ergebnisse und farbpsychologische Resultate korrespondieren miteinander. Sowohl kognitive Leistungsfähigkeit als auch emotionale Befindlichkeit werden durch das Baldrian-Präparat offensichtlich nicht beeinträchtigt.

## **Arzneimittelsicherheit**

Insgesamt erwiesen sich sowohl das Phytopharmakon als auch das Benzodiazepin als sichere, gut verträgliche und tolerable Medikamente, wobei die Quote unerwünschter Ereignisse im

ersten Studienteil höher war als im zweiten, was durch die Mitführung des synthetischen Arzneimittels begründet ist (Suggestiveffekt bei den Probanden).

Schwerwiegende Nebenwirkungen traten in beiden Studienteil nicht auf. Der Schweregrad der meisten Nebenwirkungen wurde als "gering" oder "mäßig" eingestuft. Bis auf wenige Ausnahmen waren die dokumentierten Beschwerden bis zum Abschlussbesuch abgeklungen.

Zu einem Therapieabbruch wegen Nebenwirkungen kam es lediglich einmal - nach Placebo-Einnahme. Ausschlaggebend war das Auftreten von Quaddeln an Hals und Beinen sowie unerträglicher Juckreiz.

## **Diskussion**

Das gewählte Studiendesign der randomisierten, placebokontrollierten Doppelblindstudie entsprach hinsichtlich Planung, Durchführung und Auswertung den aktuellen wissenschaftlichen Anforderungen und ist „state of the art“. Die für pharmakopsychologische Studien gültigen ethischen und rechtlichen Anforderungen des deutschen Arzneimittelgesetzes (AMG) und die Regeln der „Good Clinical Practice“ (GCP) wurden erfüllt. Methodischen Forderungen an farbpsychologische Studien, wie z.B. Spezifizierung der Farben und mehrdimensionale Erfassung der Gefühlserebnisse, wurde gleichfalls entsprochen.

Sämtliche Untersuchungen wurden gemäß den Planungsvorgaben durchgeführt, fehlende Angaben bzw. Messwerte waren aufgrund der sorgfältigen Dokumentation und des engmaschigen Monitorings selten. Protokollverstöße wurden nur in wenigen Ausnahmefällen dokumentiert und waren ohne Relevanz für die Aussagekraft der Studie. Die hohe Qualität bei Planung, Durchführung und Datenerhebung tragen maßgeblich dazu bei, die wissenschaftliche Aussagekraft der evaluierten Ergebnisse zu erhöhen.

Im Rahmen der Studiendurchführung wurde zunächst eine farbpsychologische Exploration der computergenerierten Farbpalette des KFPT mittels semantischen Eindrucksdifferentials (ED) durchgeführt. Die farbpsychologische Exploration des aus vierzehn Farben bestehenden KFPT-Farbensatzes erfolgte mittels Baselinewerten des semantischen Eindrucksdifferentials. Die anschließende Clusteranalyse führte zur Einteilung der KFPT-Farbpalette in drei Cluster und drei Einzelfarbtöne, die sich wie folgt charakterisieren: Die Cluster 1 und 2 sowie der Farbton 5b (gelb) zeigen eine vorrangige Valenzwirkung; Cluster 3 und insbesondere der Farbton 10c (schwarz) werden als primär kraftvoll eingeschätzt. Der Farbton 1b (rot) nimmt eine Sonderstellung ein, da er fast gleichermaßen als erregend, angenehm und kraftvoll erlebt wird („Superfarbe rot“).

Die drei explorierten Cluster bzw. Einzelfarbtöne bilden das Farberleben mit unterschiedlichen Akzenten ab – so fokussieren die Einzelfarbtöne auf die Dimensionen Potenz (1b/10c) und Erregung (5b), die Cluster auf die Dimensionen Valenz (Cluster 1 und 2) und Potenz (Cluster 3). Cluster und Einzelfarben ergänzen sich dabei und spiegeln die Dreidimensionalität des EVP-Modells wider. Hervorhebung oder Verzicht auf ein Cluster bzw. einen speziellen Farbton erscheinen deshalb wenig sinnvoll – auch ginge die Affinität zum Farbensatz des Farbpolyeders (FPT) verloren.

Unter Berücksichtigung der farbpsychologischen Besonderheiten des achromatischen Farbtons 10c (schwarz) sowie der beiden Farbtöne mit maximaler Sättigung und mittlerer Helligkeit 1b (rot) und 5b (gelb) wurden die drei Cluster bzw. verbleibenden Einzelfarbtöne als KFPT-Zielvariablen für die statistische Auswertung festgelegt. Im Rahmen des Studienverlaufs wurden Messwertveränderungen der definierten KFPT-Zielvariablen unter medikamentösem Einfluss durch ein Benzodiazepin, ein Phytopharmakon sowie eines Scheinmedikamentes (Placebo) objektiviert.

Aussagekraft und Interpretation der KFPT-Befunde basieren auf den Ergebnissen der Clusteranalyse und dem anschließend entwickelten Auswertekonzept mit drei Clustern und drei Einzelfarbtönen. Die in der Ausgangslage vorherrschenden Dimensionsausprägungen charakterisieren die Cluster/Einzelfarbtöne auch nach Medikationseinnahme (Studienteile 1 und 2) – relevante Veränderungen des Farberlebens lassen sich nicht feststellen und die Akzentuierung der Dimensionen bleibt bei Clustern und Einzelfarben im gesamten Studienverlauf konstant.

Gemäß dem hypothesenbildenden Pilotcharakter der Studie wurden a priori keine Hypothesen zum Farbwahlverhalten formuliert, die ex-post confirmatorisch geprüft werden sollten. Sämtliche statistische Analysen wurden deskriptiv durchgeführt, ggf. wurden zusätzlich p-Werte berechnet. Abschließend wurden die Befunde zur Farbwahrnehmung mit den DT-Ergebnissen zur kognitiven Leistungsfähigkeit der „Valerian-Treatment Study“ (Kuhlmann et al., 1999) verglichen.

Eine erste Beurteilung der Pharmakosensitivität des KFPT erfolgte durch den Vergleich mit den DT-Ergebnissen zur kognitiven Leistungsfähigkeit der parallel durchgeführten „Valerian-Treatment-Studie“. Korrespondierend fanden sich in beiden Studien weder Hinweise auf eine Veränderung der Farbwahrnehmung bzw. des Farbwahlverhaltens, noch konnte eine Beeinträchtigung der kognitiven Leistungsfähigkeit durch die Einnahme eines Baldrianpräparates festgestellt werden.

Während sich unter Baldrian-Extrakt keine Leistungsminderungen zeigten, wurde unter Flunitrazepam in der „Valerian-Treatment“-Studie bei drei Variablen im DT ein signifikant zurückbleibender Übungsfortschritt objektiviert. So konnte z.B. in der Hauptzielvariablen des DT (Median der Reaktionszeit) bei 25% der Probanden der für Flunitrazepam typische Hangover-Effekt (Dreßing, 1992; Gerhard, Linnenbrink, Georghiadou und Hobi, 1996) mit einer Vigilanzminderung objektiviert werden. In der farbpsychologischen Kontrollstudie wurde dieses Ergebnis jedoch nicht bestätigt – Veränderungen der Farbwahrnehmung und des Farbwahlverhaltens durch das Benzodiazepin wurden nicht beobachtet.

Die Resultate zur Pharmakovigilanz belegen insgesamt eine gute Verträglichkeit in Verbindung mit einer geringen Nebenwirkungsrate. Die Anzahl der Nebenwirkungen war erwartungsgemäß im ersten Studienteil höher als im zweiten Studienteil, was durch die Mitführung des synthetischen Arzneimittels begründet ist (Suggestiveffekt beim Benzodiazepin).

Diese Befunde stehen im Einklang mit der vorliegenden verkehrspsychologischen Literatur. Eine Gegenüberstellung der KFPT-Befunde mit den Ergebnissen anderer farbpsychologischer Untersuchungen ist nicht möglich, da der KFPT in der vorliegenden Studie erstmalig eingesetzt wurde und somit keine weiteren Daten vorliegen. Ein Vergleich mit historischen FPT-Daten aus pharmakopsychologischen Studien (Fahrenberg & Prystav, 1966; Lienert, 1961) ist aufgrund methodischer Unterschiede hinsichtlich Testmaterial, Farbpalette, Zielvariablen und Auswertung unzulässig. Eine Beurteilung der Pharmakosensitivität des KFPT ist deshalb auch nur eingeschränkt – auf Basis der aktuell vorliegenden Ergebnisse - möglich. Erst in weiteren pharmakopsychologischen Studien kann letztlich die Eignung des KFPT als therapiesensitiver Parameter für den Einsatz in klinischen Arzneimittelstudien geklärt werden.

Eine abschließende Beurteilung der psychometrischen Gütekriterien des KFPT ist momentan gleichfalls nicht möglich. Spezielle Test-Retest-Reliabilitäten wurden nicht berechnet. Die geringen Messwertveränderungen in der Placebo-Kontrollgruppe verdeutlichen jedoch Konstanz und geringe Varianz der KFPT Messergebnisse bei allen drei Testabnahmen – was für die Reliabilität des Verfahrens spricht.

Hinsichtlich der Beurteilung von Gültigkeit und Zuverlässigkeit liefert die vorliegende Arbeit nur einen ersten Beitrag auf dem langen Weg der Validierung - der KFPT befindet sich noch im Stadium der Erprobung. Die Validierung eines projektiven Verfahrens wie dem KFPT ist ein jahrelanger Prozess und erfordert eine Vielzahl weiterer Einzeluntersuchungen. Nur so kann die ausreichende Stabilität der Farbvariablen als notwendige Voraussetzung der Validität empirisch gesichert werden.

Wie jeder andere Persönlichkeitstest erlaubt auch der KFPT nur Aussagen, die mit einer relativen Unsicherheit behaftet sind. Als Einzeltest sind z.B. die diagnostischen Möglichkeiten des KFPT beschränkt. Nur im Rahmen eines umfassenden psychodiagnostischen Settings erlangen seine Befunde klinische Bedeutsamkeit. Die Befunde der einzelnen Verfahren einer Testbatterie ergänzen und korrigieren sich; erst dadurch erzeugen sie eine übergreifende Validität. Vor diesem Hintergrund ist die Gegenüberstellung der KFPT-Befunde mit den DT-Ergebnissen von besonderem Interesse, erhält man doch Aufschluss über Korrespondenz oder Divergenz.

Um die Einsatz- und Anwendungsmöglichkeiten des KFPT auf eine breite Basis zu stellen, sind weitere farbpsychologische Untersuchungen notwendig. So sind z.B. individuelle Farbpräferenzen ohne Berücksichtigung des sozio-kulturellen Hintergrundes und geschlechtsspezifischer Unterschiede nicht generalisierbar. Für die Konstruktion farbpsychologischer Tests empfiehlt daher Choungourian (1968) die Entwicklung kulturspezifischer Normwerte für Frauen und Männer – spezifische Normwerte für den KFPT müssten erst zukünftig evaluiert werden. Die Ergebnisse der vorliegenden Untersuchung sind im Hinblick auf mögliche Geschlechtsunterschiede im Farberleben inkonsistent – Generalisierungen mithin unzulässig. Gleichfalls ließ sich das weit verbreitete Vorurteil, dass Männer generell den Farbton „blau“ und Frauen „rot“ bevorzugen (Cheskin, 1948), auf Basis des vorliegenden Datenmaterials nicht bestätigen.

Eine der wichtigsten Aufgaben für die Weiterentwicklung des KFPT ist sicherlich die Erstellung eines speziellen Testmanuals. Dieses Manual sollte Informationen über das zu messende Konstrukt, Zusammensetzung der Testform, Durchführungshinweise für den Versuchsleiter sowie Auswertungs- und Interpretationsbeispiele enthalten. Darüber hinaus sollten eine Eichstichprobe und relevante Teilpopulationen (z.B. nach Alter und Geschlecht) festgelegt und Normwerttabellen, Angaben zur Reliabilität und Validität für das eigentliche Testverfahren sowie für die gesamte Eich- und Kontrollstichprobe gemacht werden.

Wie erleben Menschen Farben und welche Gefühle lösen sie aus? Farben gelten als Indikatoren des emotionalen Geschehens. Allen Farbtests liegt das Konstrukt zugrunde, „dass Farben in einer besonderen, hervorgehobenen und markanten Weise Beziehungen zur Affektivität zeigen“ und „Farben der nächste und offenbar adäquate Ausdruck und Repräsentant von Affekten sind“ (Heiss, Halder & Höger, 1975). Der psychologischen Farbforschung mangelte es bisher an aktuellen und methodisch sauberen Studienergebnissen. Mit der vorliegenden Arbeit ist ein erster Schritt getan hier Abhilfe zu schaffen. Es sind jedoch weitere intensive Forschungsanstrengungen erforderlich, um der Farbpsychologie in Zukunft wieder mehr wissenschaftliches Gewicht zu verleihen – und somit einen qualifizierten Beitrag zur Erforschung von Emotionen und Affekten zu leisten.

Experimentelle und empirische Farbforschung ist zugleich auch immer experimentelle und empirische Emotionsforschung. Vor diesem Hintergrund wurden abschließend die vorliegenden Ergebnisse in einen größeren wissenschaftlichen Zusammenhang gestellt und mit aktuelleren Befunden zur Emotionsforschung sowie deren Methodik in Verbindung gesetzt.

Darüber was eine Emotion ist, worin ihre Funktion besteht und in welchem Verhältnis sie zur Kognition steht, wird in der psychologischen Forschung seit Jahrzehnten kontrovers diskutiert (Bischof, 1989; Dörner, 1989; Schneider, 1989). So betonen Kleinginna und Kleinginna (1981), dass in der Varianz von Definitionen eines der Hauptprobleme der Emotionsforschung besteht. Durch das Fehlen einer Definition besteht auch kein Konsens darüber, was die primären Dimensionen emotionalen Erlebens sind und wie diese adäquat erfasst werden können. Standardisierte Verfahren fehlen bisher weitgehend – deshalb ist der KFPT mit seinem standardisierten Farbensatz und der computergestützten Generierung visueller Reize zur Auslösung emotionaler Reaktionen und der Beurteilung von Farbpräferenzen von besonderem wissenschaftlichen Interesse – zumal die farbpsychologische Exploration der KFPT-Farbpalette durch das semantische Eindrucksdifferential erfolgte.

Zugrunde liegende Annahmen Ertels zur dreidimensionalen Organisation des emotionalen Erlebens (Erregung, Valenz, Potenz) im semantischen Raum werden durch Studienergebnisse von Mehrabian und Russel (1974) bestätigt. Deren Resultate belegen, dass das semantische Eindrucksdifferential ein breites Spektrum der affektiven Befindlichkeit zuverlässig und zeitökonomisch erfasst. Der im Rahmen der vorliegenden Arbeit gewählte methodische Ansatz, die farbpsychologische Exploration des KFPT durch das ED vorzunehmen, wird durch die oben angeführten Autoren bestätigt – die Übertragung von ED-Befunden auf die Erfassung emotionalen Geschehens und Farbwahrnehmung erscheint mithin methodisch abgesichert. Das semantische Eindrucksdifferential beschreibt auch unter Berücksichtigung aktuellerer Forschungsergebnisse subjektive Erlebniswelten adäquat. Dieser generelle Anspruch der Dreidimensionalität des semantischen Eindrucksdifferentials wird später jedoch wiederum von Russel (1979) relativiert. Seine neueren Befunde belegen eine große Varianz der Dominanzdimension, welche diese als wenig geeignet zur aussagekräftigen Beschreibung des emotionalen Geschehens erscheinen lassen. Russel beurteilt deshalb die Relevanz der Dimension „Dominanz“ eher kritisch, was ihn dann zukünftig von nur zwei Dimensionen emotionalen Geschehens ausgehen lässt (Russel, 1980). Die Schwierigkeit der konsistenten Identifizierung einer dritten Basisdimension – neben Valenz und Erregung - zur Verteilung der Reize im affektiven Raum betonen auch Storm und Storm (1987). Watson und Tellegen (1985) postulieren gleichfalls ein zweidimensionales System und differenzieren eine positive vs. eine negative Affektdimension. Demgegenüber belegt eine empirische Arbeit von Feist (2000) mittels Faktorenanalyse die Dreidimensionalität des emotionalen Raumes. Auch die Ergebnisse von Valdez und Mehrabian (1994) zu Effekten von Farben auf Emotionen basieren auf einem dreidimensionalen Modell (Pleasure-Arousal-Dominance).

Ein leicht einzusetzendes non-verbales Verfahren zur direkten Erfassung des subjektiven Erlebens auf den drei Dimensionen Valenz (Pleasure), Erregung (Arousal) und Dominanz (Dominance) mittels standardisierter Bildreize ist das von Lang (1980) konstruierte Self-Assessment Manikin (SAM). In Ergänzung zum semantischen Eindrucksdifferential erscheint insbesondere die von Cook, Atkinson und Lang (1987) entwickelte Computerversion des SAM als ein geeignetes Instrument zur Evaluierung der Dimensionsausprägungen des KFPT-Farbensatzes. Dabei wäre zu prüfen, ob und inwieweit ED-Befunde und SAM-Resultate korrelieren.

Neben dem KFPT ist die Zahl geeigneter Skalen und Testverfahren zur Erfassung von Gefühlen stark limitiert. Deshalb werden abschließend die von Behr und Becker (2004) vorgestellten SEE (Skalen zum Erleben von Emotionen) kurz referiert, auch wenn diese keinen direkten Bezug zur Farbpsychologie bzw. zur Exploration des Farberlebens aufweisen.

Abschließend möchte ich mich ausdrücklich bei meinem Doktorvater, Herrn Prof. Egon Stephan, sowie Herrn Prof. Udo Undeutsch für Ihre freundliche Unterstützung und kompetente Beratung bei der Anfertigung der Dissertation bedanken.

## Inhaltsverzeichnis

GESAMTÜBERSICHT ZUR ARBEIT.....	2
INHALTSVERZEICHNIS .....	13
ABKÜRZUNGEN UND FACHBEGRIFFE.....	15
ETHIK.....	16
1. EINLEITUNG .....	17
2. RATIONALE UND WISSENSCHAFTLICHE ZIELSETZUNG .....	20
3. UNTERSUCHUNGSPLAN .....	22
3.1 Studiendesign .....	22
3.1.1 Studienablauf und Untersuchungen.....	23
3.1.2 Diskussion des Studiendesigns .....	24
3.2 Studienpopulation .....	24
3.2.1 Einschlusskriterien .....	25
3.2.2 Ausschlusskriterien .....	25
3.2.3 Abbruchkriterien.....	26
3.3 Medikationen .....	26
3.3.1 Verabreichte Medikationen .....	26
3.3.2 Dosis und Dosierungsschema .....	29
3.3.3 Randomisierung.....	30
3.3.4 Verblindung.....	30
3.3.5 Begleittherapie.....	31
3.3.6 Behandlungscompliance.....	31
3.3.7 Arzneimittelsicherheit.....	31
3.3.7.1 Unerwünschte Ereignisse und medizinische Sicherheitsvariablen ..	31
3.3.7.2 Klinisches Labor.....	32
4. PSYCHOMETRISCHE MESS- UND RATINGVERFAHREN .....	33
4.1 Kölner Farbpräferenztest (KFPT).....	33
4.2 Semantisches Einrucksdifferential (ED).....	35
4.3 Determinationstest (DT).....	37
5. DATENQUALITÄT .....	38
5.1 Quelldaten-Überprüfung (Source Data Verification) .....	38
5.2 Dokumentation der Daten .....	38
6. DATENANALYSE UND STATISTISCHE METHODEN .....	39
6.1 Stichprobenumfang.....	39
6.2 Definition der Auswertekollektive .....	40
6.3 Generelle Überlegungen zur Datenanalyse .....	41
6.3.1 Abweichungen vom Untersuchungsablauf .....	41
6.3.2 Behandlungscompliance .....	41
6.4 Interimanalyse .....	41
6.5 Analyse der Primärvariablen .....	42
6.5.1 Kölner Farbpräferenztests (KFPT).....	42
6.6 Analyse der Sekundärvariablen .....	43
6.6.1 Semantisches Eindrucksdifferential (ED).....	43
6.6.2 Determinationstest (DT).....	45

6.7	Analyse der Arzneimittelsicherheit.....	46
6.7.1	Analyse unerwünschter Ereignisse (Adverse Events) .....	46
6.7.2	Analyse des klinischen Labors .....	46
7.	ERGEBNISSE ZUR STUDIENPOPULATION .....	47
7.1	Abweichungen vom Untersuchungsablauf.....	47
7.2	Fehlende Angaben .....	47
7.3	Randomisierung und Behandlungscompliance.....	47
7.4	Verifizierung der Ein- und Ausschlusskriterien .....	48
7.5	Analysierte Populationen .....	49
7.6	Soziodemographische Charakteristika .....	52
8.	ERGEBNISSE DER PSYCHOMETRISCHEN VERFAHREN UND SCHLUSSFOLGERUNGEN .....	53
8.1	Primärer Endpunkt.....	54
8.1.1	Kölner Farbpräferenztest (KFPT).....	54
8.2	Sekundäre Endpunkte .....	66
8.2.1	Semantisches Eindrucksdifferential (ED).....	66
8.2.2	Determinationstest (DT).....	74
8.3	Schlussfolgerungen aus den Ergebnissen psychometrischer Verfahren .....	88
9.	ERGEBNISSE ZUR ARZNEIMITTELSICHERHEIT UND SCHLUSSFOLGERUNGEN .....	92
9.1	Unerwünschte Ereignisse (Adverse Events).....	92
9.2	Klinisches Labor .....	98
9.3	Schlussfolgerungen zur Arzneimittelsicherheit .....	98
10.	DISKUSSION.....	99
11.	LITERATURVERZEICHNIS .....	111

## ABKÜRZUNGEN UND FACHBEGRIFFE

AAK	Atem-Alkohol-Konzentration
AMG	Arzneimittelgesetz
ART	Adverse Reaction Terminology
ATP	According-To-Protocol
ATV	Alkoholiker-Trenn-Verfahren
Compliance	Vergleich verabreichter/verbrauchter Medikation
CRF	Case Report Form (Studienordner/Proband)
DIPS	Diagnostisches Interview bei psychischen Störungen
Drug Safety	Arzneimittelsicherheit
DT	Determinationstest
ED	Semantisches Eindrucksdifferential
EK	Ethik-Kommission
FAS	Full-Analysis-Set
FPT	Farbpyramidentest
GABA	Gamma-Aminobuttersäure
$\gamma$ GT	Glutamyl-Transferase
GOT	Glutamat-Oxalazetat-Transaminase
GPT	Glutamat-Pyruvat-Transaminase
Hangover-Effekt	Unerwünschte sedierende Nebenwirkung von Medikamenten
Hb	Hämoglobin
ICD-10	International Classification of Psychiatric Diseases (10. Revision)
ICH	International Conference on Harmonization
IMF	Institut für interdisziplinäre medizinische Forschung
ITT	Intent-To-Treat
KFPT	Kölner Farbpräferenztest
LÄK	Landesärztekammer
LOCF	Last Observation Carried Forward
PPS	Per-Protocol-Set
PR	Prozentrang
Randomisierung	Zufallszuteilung der Population auf Untersuchungsgruppen
RR	Symbol für den mit dem Riva-Rocci-Apparat gemessenen Blutdruck
Screening	Voruntersuchung zwecks Überprüfung der Eignung zur Studienteilnahme
SDV	Source Data Verification
ST	Studienteil
UE	Unerwünschtes Ereignis
Wash-Out-Phase	Auswaschphase (medikationsfrei Zeit)
ZNS	Zentrales Nervensystem

## **ETHIK**

### **Unabhängige Ethik-Kommission**

Aufgrund der vorgesehenen Einnahme von zugelassenen und vermarkteten Arzneimitteln mussten gemäß deutschem Arzneimittelgesetz (AMG) der Untersuchungsplan, Probandeninformation und Einverständniserklärung vor Studienbeginn einer unabhängigen Ethik-Kommission (EK) zur Beurteilung möglicher medizinischer Risiken für die Teilnehmer vorgelegt werden.

Gegen das Studienprotokoll (Studienablauf, Methodik, Einnahmendauer und Dosierung beider Medikationen etc.) wurden von der zuständigen EK der Landesärztekammer (LÄK) Nordrhein keine Einwände gegen die Durchführung der Doppelblindstudie erhoben.

### **Ethische Durchführung der Studie**

Die Studie wurde in Übereinstimmung mit den Richtlinien der aktuell gültigen Version zur "Guten klinischen Praxis" (Good Clinical Practice) sowie länderspezifischen gesetzlichen Bestimmungen gemäß den Prinzipien der Deklaration von Helsinki und ICH (International Conference on Harmonization (ICH) Tripartite Guideline, Guideline for GCP, May 1996) durchgeführt.

### **Information und Aufklärung der Studienteilnehmer**

Vor Studienbeginn wurde jeder Proband über Wesen, Bedeutung und Tragweite der Untersuchung in verständlicher Form aufgeklärt. Die Bereitschaft zur freiwilligen Studienteilnahme durch den Probanden erfolgte in schriftlicher Form und konnte jederzeit – ohne Begründung – widerrufen werden.

### **Planung und administrative Bedingungen**

Studienplanung, -organisation und -durchführung erfolgten durch den Doktoranden in enger Kooperation mit den Universitäten Köln und Bielefeld.

Die Universität Köln stellte für die experimentellen farbpsychologischen Untersuchungen den von Stephan entwickelten computergestützten Farbpräferenztest (KFPT) zur Verfügung und Höger (Universität Bielefeld) den für die Evaluierung des subjektiven Farberlebens mittels semantischen Eindrucksdifferentials (ED) notwendigen Farbensatz. Die praktische Umsetzung der Studie erfolgte in den Räumlichkeiten des „Instituts für interdisziplinäre medizinische Forschung“ (IMF) in Köln. Die psychometrischen Untersuchungen wurden vom Doktoranden, die medizinischen Untersuchungen von einer beim IMF beschäftigten approbierten Ärztin durchgeführt.

# 1 EINLEITUNG

Der Psychiater Emil Kraepelin (1856-1926), ein Schüler des Begründers der experimentellen Psychologie Wilhelm Wundt, etablierte 1899 eine Klassifikation von Psychosen, die bis heute in wesentlichen Teilen unverändert geblieben ist. Neben seiner Arbeit als Psychiater blieb Kraepelin der experimentellen Psychologie Wundts weiterhin eng verbunden und leistete einen wesentlichen Beitrag zur Etablierung der Psychologie in der Klinik (z.B. 1895: Der psychologische Versuch in der Psychiatrie).

Im Jahre 1892 veröffentlichte Kraepelin dann ein Buch mit dem Titel „Über die Beeinflussung einfacher psychischer Vorgänge durch einige Arzneimittel – Experimentelle Untersuchungen“. In dieser Publikation zeigte Kraepelin auf, dass man die Wirkung von Arzneimitteln mit psychologischen Methoden untersuchen und beurteilen kann. Kraepelin kann deshalb zu Recht als der Begründer der experimentellen klinischen Psychologie und der Pharmakopsychologie betrachtet werden (Wittling, 1976).

Für Heinrich Düker war jegliche Erfassung von Psychopharmakawirkungen mit psychologischen Methoden Pharmakopsychologie (Janke, 1988). Das Hauptanliegen der experimentellen Pharmakopsychologie ist demnach der objektive Nachweis der Zielwirkung psychotroper Substanzen und die Profilierung ihrer Wirkaspekte mittels psychometrischer Testverfahren und Ratingskalen. Die Befunde der experimentellen Pharmakopsychologie leisten einen wesentlichen Beitrag zur Charakterisierung von Arzneimitteln und zur medizinischen Risiko-Nutzen-Analyse. Die pharmakopsychologischen Methoden ermöglichen es, Präparateunterschiede aufzuzeigen und liefern für den behandelnden Arzt eine Entscheidungshilfe bei der Therapiewahl. Bei ZNS-wirksamen Substanzen ist die Klärung ihres möglichen Einflusses auf kognitive Prozesse, psychomotorische Funktionen und Emotion erforderlich, um valide Informationen über Reaktions- und Funktionstüchtigkeit sowie Befindlichkeit des Patienten nach Einnahme des Medikaments zu besitzen. Durch den Einsatz spezieller psychologischer Testverfahren lassen sich darüber hinaus differentielle Aspekte der Pharmakotherapie aufklären, z.B. führt die Depression selbst zu einer Beeinträchtigung von Aufmerksamkeit, Wahrnehmung, Gedächtnis und intellektueller Leistungsfähigkeit (Kraepelin, 1927; Raskin, 1982; Weingarter, 1982). Bei älteren Patienten können zusätzlich altersbedingte kognitive Defizite hinzukommen. Antidepressiva mit sedierenden und anticholinergen Eigenschaften können diese Beeinträchtigungen noch verstärken und die – aufgrund der antidepressiven Therapie - zu erwartende Besserung der depressiven Symptomatik verhindern (Hindmarch, 1987; Siegried & O’Conolly, 1986). Eine der wesentlichen Aufgaben der experimentellen Pharmakopsychologie ist daher die Entwicklung, Prüfung und Bereitstellung neuer psychologischer Testverfahren, die für die Untersuchung chemischer Wirkstoffe und Pharmaka erforderlich sind (Wegener & Leonard, 1992).

In der Pharmakopsychologie steigt das Interesse an Testverfahren, mit denen pharmakologische Effekte auf die Informationsverarbeitung und kognitive Leistungsfähigkeit sowie Persönlichkeitsmerkmale evaluiert werden können. Häufig eingesetzt werden bisher Ratingskalen, Selbst- und Fremdbeurteilungsverfahren, z.B. zur Diagnosesicherung, Schweregradbestimmung und Therapiekontrolle psychiatrischer Erkrankungen.

Computergestützte psychometrische Leistungs- und Persönlichkeitstests kommen im Bereich der Pharmakopsychologie vergleichsweise selten zum Einsatz. Pharmakosensitivität, Validierung, Reliabilität und Objektivität vieler Verfahren sind noch nicht hinreichend geklärt. Erschwerend ist, dass z.B. viele Reaktionszeitexperimente aus dem Bereich der Kognitionspsychologie einen relativ aufwendigen apparativen Aufbau bedingen. Der Gebrauch von Computern bei der Durchführung psychologischer Experimente beinhaltet weiterhin die Entwicklung validierter Versuchssteuerungs- und Auswerteprogramme. Dies erfordert bei experimentellen Psychologen vertiefte Programmier- und Hardwarekenntnisse, die leider häufig nicht vorhanden sind. Damit bleibt die wissenschaftliche Neuentwicklung computergestützter Testverfahren nur auf wenige deutsche Universitäten (wie z.B. Köln) konzentriert.

Der begrenzte Einsatz computergestützter Testverfahren im Rahmen der Pharmakopsychologie ist jedoch letztlich durch den Mangel an wissenschaftlich fundierten Untersuchungsergebnissen zur Evaluierung von Arzneimittelwirkungen mittels Computertests zu erklären. Die Überprüfung der therapeutischen Sensitivität und der Eignung als Verlaufspareter zur Objektivierung des Behandlungserfolges stehen bei den meisten computergestützten Testverfahren noch aus.

Frühere Studien zum Farberleben und insbesondere zur Eignung des Farbpyramidentests (FPT) als Verlaufspareter pharmakopsychologischer Untersuchungen wurden nicht weitergeführt. Den Untersuchungen von Höger (1980) zum Farberleben älterer Menschen sind nur wenige gleichgerichtete Projekte gefolgt. In jüngster Zeit werden die Forschungsbemühungen wieder intensiviert. In Anlehnung an den FPT wurde an der Universität Köln von Stephan die experimentelle Version eines computergestützten Farbwahlverfahrens entwickelt. Eine Beurteilung hinsichtlich der Gewährleistung testtheoretischer Gütekriterien und eine Umschreibung der Einsatzmöglichkeiten des Kölner Farbpräferenztests (KFPT) – explizit in der Pharmakopsychologie – steht noch aus. Ob das Verfahren in diesem Anwendungsbereich (insbesondere zur Evaluierung medikamentös bedingter Veränderungen der Farbwahrnehmung und der emotionalen Befindlichkeit) valide Testergebnisse generiert, und ob diese Ergebnisse auch mit Resultaten etablierter kognitiver Leistungstests korrespondieren, ist bis dato ungeklärt.

Unabhängig von den mit ihnen verbundenen Gegenständen und ihrer Kennzeichnungsfunktion erzielen Farben eine eigenständige psychologische Wirkung,

die sich vor allem in der emotionalen Befindlichkeit des Betrachters niederschlägt. Deshalb wurden bereits seit Beginn des Jahrhunderts Untersuchungen zur Farbwahrnehmung und emotionaler Wirkung bestimmter Farben durchgeführt. Die psychologische Farbforschung besitzt somit zwar eine langjährige Tradition, doch führen die Ergebnisse – zumindest bezüglich systematischer Beziehungen zwischen Farben und durch sie ausgelöste Gefühlserlebnisse – zu keinem einheitlichen Erkenntnisstand. Konsens besteht lediglich dahingehend, dass unterschiedliche Farben auch unterschiedliche Gefühle auslösen. So lassen sich z.B. unter dem bipolaren Aspekt „erregend-beruhigend“, einheitlich und konsistent, folgende mit Farben verbundene Gefühlserlebnisse beschreiben: Erregend wirken Rot, Orange und Gelb; beruhigend wirken Grün und Blau. Die Frage, ob die Gefühlswirkung einer speziellen Farbe anhand regelhafter Beziehungen vorhergesagt werden kann, bleibt bis heute unbeantwortet. Als Ergebnis jahrzehntelanger wissenschaftlicher Farbforschung sind diese Aussagen trivial und unbefriedigend.

Methodische Defizite erklären die geringe Aussagekraft und mangelnde Generalisierbarkeit der Ergebnisse vieler farbpsychologischer Studien. Eine genaue Spezifizierung der verwendeten Farbtöne wurde bei den meisten Untersuchungen nicht vorgenommen. Weiterhin lässt der oft nur geringe Stichprobenumfang keine allgemeinen Schlussfolgerungen zu. Letztlich erfolgt die Beschreibung der Gefühlserlebnisse zumeist nur eindimensional (z.B. angenehm – unangenehm); weitere emotionale Dimensionen bleiben unberücksichtigt. Die Reproduzierbarkeit vorliegender Ergebnisse ist mithin nicht möglich.

Kaum eine farbpsychologische Untersuchung genügt den oben genannten methodischen Anforderungen: Spezifizierung der Farben, Repräsentativität der Stichprobe, mehrdimensionale Erfassung der Gefühlserlebnisse. So lassen sich z.B. die Untersuchungen von Eysenck (1941) insbesondere die beiden letzten Punkte unberücksichtigt. Guilford und Smith (1959) erfüllen zwar die beiden ersten methodischen Forderungen, doch wird lediglich die Annehmlichkeit des Gefühlseindrucks als einzige Dimension berücksichtigt. Oyama, Tanaka und Chiba (1962) evaluieren mehrere Gefühlsdimensionen und spezifizieren auch die verwendeten Farben, doch erlaubt die geringe Fallzahl keine Verallgemeinerung der Ergebnisse. Ohne einheitliche Methodik und vergleichbare Versuchsbedingungen sind die bisweilen widersprüchlichen Befunde nicht aufzuklären. Eine Gegenüberstellung vorliegender Forschungsergebnisse, z.B. im Rahmen einer Meta-Analyse, ist methodisch fragwürdig und wenig weiterführend. Da die Farbforschung zunehmend aus dem wissenschaftlichen Fokus getreten ist, fehlen aktuelle und methodisch gesicherte Studienergebnisse. Seit den Untersuchungen von Höger (1980) zur Psychologie des Farberlebens Älterer sind nur wenige neuere Ergebnisse publiziert worden. In einer Revision vorliegender Literatur zur Farbpsychologie kommen Whitfield und Wiltshire (1990) zur Conclusion, dass „Farbe und Farbton nicht synonym“ und „weitere Forschung notwendig ist“.

## 2 RATIONALE UND WISSENSCHAFTLICHE ZIELSETZUNG

Der psychologischen Farbforschung mangelt es an aktuellen und methodisch einwandfreien Studienergebnissen. Die vorliegende Untersuchung verfolgte das Anliegen, an die Tradition der psychologischen Farbforschung anzuknüpfen und methodologisch fortzuentwickeln. Dabei sollten die drei methodischen Grundforderungen berücksichtigt sowie die pharmakopsychologischen Einsatzmöglichkeiten eines an der Universität Köln von Stephan entwickelten computergestützten farbpsychologischen Testverfahrens untersucht werden.

Für den Bereich der Pharmakopsychologie sind neben Leistungstests Verfahren bedeutsam, welche Hinweise auf die emotionale Verfassung des Patienten geben (z.B. bei depressiven Erkrankungen). Auch hier werden meist Fremd- oder Selbstbeurteilungsverfahren eingesetzt, da validierte psychometrische Tests kaum zur Verfügung stehen. Eine Alternative zu sprachgebundenen Verfahren stellt der nonverbale Farbpyramidentest dar. Der FPT baut auf Erkenntnissen der Farbpsychologie auf, die einen engen Zusammenhang von Farbpräferenzen und Affekten, Gefühlen und Stimmungen belegen. Der FPT erlaubt Rückschlüsse auf die aktuelle Stimmung einer Persönlichkeit. Deshalb ist der Test auch als begleitendes Diagnostikum bei Verlaufsuntersuchungen geeignet. Der FPT hat ebenfalls Verwendung bei der Kontrolle von Psychotherapien oder psychopharmakologischer Untersuchungen gefunden. Trotzdem ist der Farbpyramidentest in den letzten Jahrzehnten als experimentelles Untersuchungsinstrument fast in Vergessenheit geraten. Die ersten psychopharmakologischen Untersuchungen mit dem FPT führte Hiltmann (1951) mit dem Stimulans „Perhedrin-Roche“ durch. Lienert (1961) untersuchte das Farbwahlverhalten mittels FPT unter dem Einfluss von Lysergsäurediätylamid (LSD), Fahrenberg und Prystav (1966) prüften die Wirkung des chemisch definierten Tranquilizers „Adumbran“, Müller und Romeyke (1970) evaluierten in einer Langzeitstudie die Wirkung eines Thymoleptikums, Heiss, Fahrenberg und Seuss (1971) untersuchten den Einfluss des zentral vasodilatierenden Medikaments „Hydergin“. Die erzielten Ergebnisse lassen es als sinnvoll erscheinen, die empirischen Einsatzmöglichkeiten des FPT weiter zu verfolgen. Der aktuelle Stand der pharmakopsychologischen Farbforschung ist jedoch durch den Mangel an neueren Studienergebnissen gekennzeichnet.

Die vorliegende Studie wurde als explorative Kontrolluntersuchung zur Studie „Influence of Valerian Treatment on Reaction Time, Alertness and Concentration in Volunteers“ konzipiert. Die von Kuhlmann et al. (1999) publizierten Ergebnisse zur kognitiven Leistungsfähigkeit sollten anhand des unabhängigen Außenkriteriums „Farbwahlverhalten“ in einer parallel durchgeführten Untersuchung verifiziert werden. Da das Farbwahlverhalten als ein „unmittelbarer Spiegel des aktuellen emotionalen Zustandes,“ (Höger, 1975) zu betrachten ist, sollte weiterhin die Eignung des KFPT als diagnostisches Verfahren zur Erfassung der

Affektstruktur, insbesondere ihre Labilisierung oder Störbarkeit unter medikamentösen Einfluss, untersucht werden.

Folgende wissenschaftliche Zielsetzungen wurden mit der vorliegenden pharmakopsychologischen Studie verfolgt:

- Exploration der individuellen Farbwahrnehmung der computergenerierten Farbpalette des Kölner Farbpräferenztests durch den methodischen Einsatz eines „semantischen Eindrucksdifferentials“ (ED)
- Definition der KFPT-Zielvariablen und Entwicklung einer testspezifischen Auswertestrategie
- Veränderungsmessung der KFPT-Zielvariablen unter medikamentösem Einfluss durch ein Benzodiazepin und ein Phytopharmakon
- Gegenüberstellung der KFPT-Befunde zum Farbwahlverhalten (als Indikator für die emotionale Befindlichkeit) mit den Ergebnissen zur kognitiven Leistungsfähigkeit des Determinationstests (DT) zwecks einer ersten Beurteilung der Pharmakosensitivität des KFPT und seiner Eignung als Verlaufsparemeter.

### 3      **UNTERSUCHUNGSPLAN**

#### 3.1      **Studiendesign**

Zur Beantwortung der wissenschaftlichen Fragestellungen wurde der doppelblinde, placebokontrollierte Parallelgruppenvergleich als die geeignetste und allgemein akzeptierte Methode gewählt.

Da psychologische Effekte nach einmaliger Sedativa-Einnahme als auch nach 14-tägiger medikamentöser Einnahme untersucht werden sollten, wurde ein zweigeteiltes Studiendesign entwickelt.

<b>Studienteil 1</b> (Tage 1-2)	Unizentrischer, <u>dreiarmiger</u> , doppelblinder, kontrollierter, randomisierter Parallelgruppenvergleich Baldrian-Extrakt vs. Benzodiazepin vs. Placebo
<b>Auswaschphase</b>	7 Tage ohne Medikationseinnahme (drug-free status)
<b>Studienteil 2</b> (Tage 9-23)	Unizentrischer, <u>zweiarmiger</u> , doppelblinder, kontrollierter, randomisierter Parallelgruppenvergleich Baldrian-Extrakt vs. Placebo

Tab. 1: Studiendesign

Für den *Studienteil 1* wurde ein dreiarmiges Untersuchungsdesign konzipiert. Es sollte einerseits geklärt werden, ob der Baldrian-Extrakt am Morgen nach der vorherigen abendlichen Einnahme emotionale Befindlichkeit und kognitive Leistungsfähigkeit nicht mehr beeinträchtigt als das Placebo. Durch die einmalige Einnahme von Flunitrazepam sollte andererseits die Eignung bzw. Pharmakosensitivität des kognitiven Leistungstests DT zur Erfassung einer - für das Benzodiazepin zu erwartenden - Beeinträchtigung der kognitiven Leistungsfähigkeit belegt werden, indem ein Unterschied zwischen der Flunitrazepam- und der Placebo-Gruppe objektiviert wird. Weiterhin sollten bisher nicht untersuchte mögliche Auswirkungen der einmaligen Einnahme eines Benzodiazepins auf Farbwahrnehmung und Farbwahlverhalten evaluiert werden.

Im *Studienteil 2* wurde – nach einwöchiger Auswaschphase - das Studiendesign auf eine zweiarmige Randomisierung (Baldrian-Extrakt/Placebo) zurückgeführt. Hier sollte der Effekt einer zweiwöchigen Einnahme des Baldrianpräparates untersucht werden. Ein Vergleich mit Placebo war ausreichend, um evtl. Beeinträchtigungen der Affektstruktur und der kognitiven Leistungsfähigkeit zu erfassen.

Medikationsgruppen	Medikamentöse Behandlungen		
	Studienteil 1 (dreiarmliges Design)	Auswaschphase	Studienteil 2 (zweiarmiges Design)
	Baldrian-Extrakt	N=34	ohne Medikation N=102
Flunitrazepam	N=34	N=51	
Placebo	N=34		

Tab. 2: Medikationsgruppen und Behandlungen im Studienverlauf

### 3.1.1 Studienablauf und Untersuchungen

UNTERSUCHUNGEN	STUDIEN TEIL 1 (einmalige Med.Einnahme)		STUDIEN TEIL 2 (14tägige Med.Einnahme)	
	TAG 1	TAG 2	TAG 9	TAG 23
Information/Einverständniserklärung	x			
Verifizierung der Ein-/Ausschlusskriterien	x	x	x	x
Randomisierung	x			
<i>Psychometrische Tests</i> <i>- KFPT und DT</i>	x	x		x
<i>Ratingverfahren</i> <i>- ED</i>	x	x		x
Ausgabe der Medikation	x		x	
Rücknahme der Medikation		x		x
Behandlungs-Compliance		x		x
Beurteilung der Verträglichkeit		x		x
Labor	x			x
Erfassung unerwünschter Ereignisse (UE)		x		x
Dokumentation evtl. Studienabbrüche		x	x	x

Tab 3: Untersuchungsplan

### **3.1.2 Diskussion des Studiendesigns**

Das Studiendesign ist so konzipiert, dass die evaluierten Ergebnisse Rückschlüsse bezüglich der Pharmakosensitivität des KFPT ermöglichen und eine etwaige Konvergenz mit den DT-Ergebnissen der „Valerian-Treatment Study“ (Kuhlmann et al. 1999) untersucht werden kann.

Um die Vergleichbarkeit der farbpsychologischen Resultate mit den DT-Ergebnissen der „Valerian-Treatment Study“ zu gewährleisten, wurde auf die Entwicklung einer eigenständigen Prüfanlage bewusst verzichtet. Design, Studienpopulation, Ein- und Ausschlusskriterien, Ablauf- und Durchführungsbedingungen der beiden – ansonsten voneinander unabhängigen Studien – entsprechen einander und erlauben somit einen direkten Vergleich der Ergebnisse beider Untersuchungen. Es bestanden lediglich methodische Unterschiede bezüglich der Wahl der eingesetzten psychometrischen Tests und Ratingverfahren sowie der statistischen Auswertung. Ansonsten divergierten die beiden Studienkonzeptionen nicht. Die KFPT-Testabnahmen und die Datenerhebungen für das semantische Eindrucksdifferential (ED) erfolgten separat und in Ergänzung zu den in der „Valerian-Treatment Study“ vorgesehenen Testdurchführungen.

Das Design erfüllt sämtliche methodologischen Standards einer pharmakopsychologischen Studie und ist „state-of-the-art“: Randomisierte Zuordnung der Probanden auf die Untersuchungsgruppen, doppelblinde Verabreichung der Medikationen und Placebokontrolle, psychometrische Testdurchführungen unter Doppelblindbedingungen für Proband und Untersucher. Die Stichprobengröße basiert auf vorheriger statistischer Fallzahlberechnung und war ausreichend, um auf breiter Datenbasis zuverlässige und aussagekräftige Resultate zu generieren. Eine Spezifizierung der Farben des KFPT wurde vorab vorgenommen und eine mehrdimensionale Erfassung der Gefühlserlebnisse mittels semantischen Eindrucksdifferentials war vorgesehen. Damit genügt das Design allen methodischen Anforderungen einer experimentellen farbpsychologischen Untersuchung.

### **3.2 Studienpopulation**

Insgesamt wurden 102 weibliche und männliche gesunde Teilnehmer im Alter zwischen 30 und 60 Jahren in die Untersuchung einbezogen. Zum Verständnis der Testinstruktionen und der sich anschließenden Testdurchführungen mittels PC (über Probandenpanel und Monitor) war eine normale Tagesseshschärfe und Farbtüchtigkeit bei sämtlichen Probanden zwingend erforderlich.

Um auswertbare Ergebnisse zu generieren und um mögliche Interaktionen der verabreichten Arzneimittel mit exogenen Wirkstoffen wie z.B. Alkohol auszuschließen, wurden die Probanden vor Aufnahme in die Studie hinsichtlich ihres

Trinkverhaltens untersucht. Nur Probanden mit unauffälligem Testergebnis im Alkoholiker-Trenn-Verfahren (ATV) wurden eingeschlossen. Vor jeder Testabnahme wurden die Probanden auf eine mögliche Alkoholisierung hin überprüft. Probanden mit einem Atemalkoholkonzentrations-Messwert > 0.00 mg/l (AAK) wurden von der weiteren Studienteilnahme ausgeschlossen; die psychometrischen Tests wurden nicht mehr durchgeführt.

### **3.2.1 Einschlusskriterien**

- schriftliches Einverständnis nach mündlicher und schriftlicher Aufklärung
- 30-60jährige weibliche und männliche Probanden
- Farbtüchtigkeit/normale Tagessehschärfe mindestens 0.7/0.7
- normales Körpergewicht ( $\pm 10\%$  nach Broca)
- nach DIPS-Befund keine psychiatrische Erkrankung (ICD-10)
- keine anamnestisch bekannte organische Schlafstörungen (ICD-10/G 47)
- soziales Trinkverhalten (evaluiert durch den ATV)
- Atemalkoholkonzentrations-Messwert (AAK) 0.00 mg/l  
(keine Alkoholisierung bei der Aufnahme bzw. den Folgeuntersuchungen)

### **3.2.2 Ausschlusskriterien**

- Einnahme von Psychopharmaka, Hypnotika, Sedativa, Analgetika oder anderen Pharmaka, die das Schlafverhalten oder die kognitive Leistungsfähigkeit beeinflussen, in den vier Wochen vor Beginn oder während der Teilnahme an der Studie
- psychiatrische Erkrankungen
- Alkohol-, Drogen- und Medikamentenabhängigkeit
- akute Vergiftung mit Alkohol, Schlaf-/ Schmerzmitteln, Neuroleptika, Antidepressiva, Lithium
- Hypersomnien, Störungen des Schlaf-Wach-Rhythmus, Parasomnien
- jede Erkrankung, die bekanntlich zu sekundären Schlafstörungen führt, wie z.B. Neurodermitis, chronischer Juckreiz (Lebererkrankungen) u.a.
- Myasthenia gravis
- spinale und/oder zerebrale Ataxie
- eingeschränkte Leber- oder Nierenfunktion

- respiratorische Insuffizienz (anamnestisch)
- nicht kompensierte Herzinsuffizienz
- Morbus Parkinson
- arterielle Hypotonie
- Engwinkelglaukom
- hirnorganisches Psychosyndrom
- erhöhte Krampfbereitschaft
- schwere konsumierende Erkrankungen (z.B. Malignom)
- schwere Blutbildveränderungen  
(Hb, Erythro-, Leuko- und Thrombozyten > 10% außerhalb der Norm)
- bekannte Überempfindlichkeit gegen Flunitrazepam
- bekannte Überempfindlichkeit gegen Baldrian-Extrakt
- Schwangerschaft und Stillzeit bei Frauen im gebärfähigen Alter
- fehlender Konzeptionsschutz bei Frauen im gebärfähigen Alter

### **3.2.3 Abbruchkriterien**

Der Proband konnte die Studie jederzeit auf eigenen Wunsch – ohne Angabe von Gründen – abbrechen. Weiterhin wurde beim Auftreten nicht tolerabler Nebenwirkungen bzw. unerwünschter Ereignisse (UE), bei einer deutlichen Verschlechterung des Allgemeinzustandes, bei Einnahme nicht-erlaubter Zusatzmedikation oder mangelnder Compliance die Studienteilnahme abgebrochen.

## **3.3 Medikationen**

### **3.3.1 Verabreichte Medikationen**

Als Psychopharmaka bezeichnet man chemisch-synthetische oder organische Wirkstoffe, die direkt oder über das autonome Nervensystem das ZNS beeinflussen und dadurch temporär das Erleben und Verhalten beeinflussen.

Wegen ihrer „psychotropen“ (-trop: auf etwas gerichtet) Wirkung können Psychopharmaka gezielt zur Behandlung psychischer Störungen und zur Modulierung von psychischen Zuständen eingesetzt werden.

Kognitive Veränderungen und mögliche mentale Beeinträchtigungen lassen sich besonders gut bei sedierenden Psychopharmaka wie Beruhigungs- und Schlafmitteln objektivieren. Sowohl bei chemisch definierten wie auch pflanzlichen

psychotropen Arzneimitteln ist potentiell eine Beeinträchtigung der psychophysischen Leistungsfähigkeit nicht auszuschließen (Ammon, 1982). Gerade bei zentralnervös wirksamen Substanzen können in interindividuell unterschiedlichem Ausmaß psychophysische Leistungen beeinträchtigt sein. Arbeits- und verkehrsmedizinische/-psychologische Untersuchungen haben wiederholt gezeigt, dass die meisten Psychopharmaka - in Abhängigkeit von der Dosierung - mentale Leistung und Vigilanz mehr oder weniger stark herabsetzen können.

Für die Auswahl der im Rahmen der Studie zu verabreichenden Arzneimittel war entscheidend, dass a) entweder psychometrisch nachweisbare Beeinträchtigungen der psychophysischen Leistungsfähigkeit sicher anzunehmen, oder im umgekehrten Fall, extrem unwahrscheinlich sind und b), dass die Einnahme über einen längeren Zeitraum ethisch und medizinisch unbedenklich ist.

Da sich die Behandlung von Probanden mit indikationsgebundenen Psychopharmaka verbietet, wurden im Rahmen der Studie ein mildes pflanzliches Sedativum (Baldrian) - bei dem kognitive Beeinträchtigungen eher unwahrscheinlich sind - und *einmalig* ein chemisch definiertes Sedativum (Flunitrazepam) - mit bekanntem Hangover-Effekt - als Arzneimittel verabreicht. Da eine nicht indizierte längerfristige Behandlung von gesunden Personen mit einem Benzodiazepin wie Flunitrazepam ethisch nicht vertretbar ist, wurde für die Verlaufsuntersuchung lediglich der Baldrian-Extrakt von den Probanden eingenommen. Gegen Einnahmedauer und Dosierung beider Medikationen wurden von der zuständigen Ethik-Kommission der Landesärztekammer Nordrhein (LÄK) keine medizinischen Einwände erhoben.

Baldrian ist eine seit Jahrhunderten eingesetzte Heilpflanze. Die Anwendung von Baldrian in kurativen Medizin reicht bis ins Altertum zurück. Bereits griechische und römische Ärzte setzten Baldrian als Spasmolytikum, Analgetikum und Diuretikum ein. Doch erst im Verlauf des 18. Jahrhunderts wurde Baldrian vom englischen Arzt John Hull als Sedativum in die neuere Medizin eingeführt (Kautz, 1994). An der Anwendung des Baldrians als Beruhigungs- und Schlafmittel hielt die Volksmedizin bis heute fest. Die weit verbreitete Skepsis bei Patienten gegenüber chemisch definierten Arzneimitteln hat dazu geführt, dass Baldrianzubereitungen zunehmend an Bedeutung von Schlafstörungen gewinnen (Krieglstein und Grusla, 1988).

Die wichtigsten Inhaltsstoffe der Baldrianwurzel sind im Wesentlichen ätherisches Öl (Valeren- und Acetoxyvalerensäure sowie Valeriansäure), Valepotriate, Alkaloide, Zucker, Stärke, Amino- und Phenolcarbonsäuren. Durch seine zentral dämpfende Wirkung und spasmolytischen Eigenschaften normalisiert Baldrian das Schlafverhalten. Als ein Ansatzpunkt des Baldrians gelten die GABA-Rezeptoren. GABA ist der bedeutendste inhibitorische Neurotransmitter, der eine Schlüssel-funktion bei Stress- und Angstreaktionen einnimmt. Weiterhin dämpfen GABAerge Neurone im Neocortex und Hypocampus sekundär auch noradrenerge und dopaminerge vigilanzfördernde Effekte und wirken damit schlafunterstützend. Medizinische Indikationsgebiete für Baldrian sind deshalb nervös bedingte Einschlafstörungen und Unruhezustände.

Vorliegende Studienergebnisse belegen eine beruhigende vegetative Wirkung des Baldrians, ohne dass die kognitive Leistungsfähigkeit beeinträchtigt wird (Herberg, 1994; Kautz, 1994; Krieglstein & Grusla, 1988; Stephan, 1980). Resultate, welche den Einfluss einer Baldriantherapie auf den emotionalen und affektiven Bereich der Persönlichkeit untersuchen, liegen bisher nicht vor.

Mit ihrer geringen Toxizität, der fast ausschließlich auf das zentrale Nervensystem gerichteten Wirkung und nur wenigen metabolischen Interaktionen sind Benzodiazepine häufig eingesetzte Therapeutika zur Behandlung von Schlafstörungen. Gerade die gute Verträglichkeit, das Fehlen auch subjektiver unangenehmer Nebenwirkungen, können jedoch auch zu Problemen im Umgang mit Benzodiazepinen führen – es besteht ein ausgesprochen hohes Missbrauchs-Risiko. Benzodiazepine wirken entspannend und beruhigend, die Welt erscheint freundlicher und harmonischer und die eigenen realen Probleme verlieren ihre Dringlichkeit. Länger anhaltender, regelmäßiger Gebrauch von Benzodiazepinen, auch in den empfohlenen Dosen, führt zu „psychischer Abhängigkeit“ (Spiegel, 1991). Andererseits kehren Unruhe, Ängstlichkeit und Bedrängnis beim Absetzen des Benzodiazepinpräparates rasch und oft sogar verstärkt wieder. Ein abruptes Absetzen des Benzodiazepines sollte deshalb unbedingt vermieden werden. Zu empfehlen ist eine stufenweise De-Eskalation (Ausschleichen) der therapeutischen Dosis, um „Rebound“-Phänomene zu vermeiden.

Die beruhigende, entspannende Wirkung der Benzodiazepine wird auf die Fähigkeit zurückgeführt, die Effekte des Neurotransmitters GABA (Gamma-Aminobuttersäure) zu verstärken. GABA ist der mengenmäßig vorherrschende und weitestverbreitete erregungshemmende Neurotransmitter. Angriffspunkte der Benzodiazepine sind die GABAergen Synapsen, deren Empfindlichkeit durch das Benzodiazepin-Molekül gesteigert wird, so dass die gleiche GABA-Menge eine stärkere bzw. länger anhaltende postsynaptische Hemmung auslöst. Die besondere Anordnung der GABAergen Neuronen in praktisch allen Teilen des zentralen Nervensystems soll dafür verantwortlich sein, dass Benzodiazepine vor allem exzessive Erregung dämpfen und dass die Dämpfung ein bestimmtes Maß nicht übersteigen kann. Da sich GABA praktisch nur im zentralen Nervensystem findet, wird auch verständlich, dass Benzodiazepine keine peripheren Effekte, die klinisch als vegetative Nebenwirkungen in Erscheinung treten, auslösen.

Bei der Behandlung von Einschlafstörungen mit kurz wirksamen Benzodiazepinpräparaten (Wirkstoff z.B. Flunitrazepam) kann außerdem eine unerwünschte Tagessedation auftreten (Hangover-Effekt), die zur Beeinträchtigung der Konzentrations- und Leistungsfähigkeit führen kann (Dreßing, 1992). Zu den potentiellen Nebenwirkungen von Flunitrazepam gehört - neben dem für Benzodiazepine bekannten Gewöhnungs- und Suchteffekt - auch die Vigilanzminderung (Gerhard et al. 1996). Weiterhin sind Interaktionen mit Alkohol bei Benzodiazepinen vielfach wissenschaftlich belegt, jedoch für Flunitrazepam nicht explizit beschrieben. Auswirkungen von Flunitrazepam auf die Affektstruktur und kognitive Leistung sind bisher nicht untersucht worden, doch können die für Benzodiazepine

bekannten Nebenwirkungen aufgrund des bekannten Wirkprofils dieser Substanzklasse auch bei Flunitrazepam erwartet werden.

In der Studie wurden folgende Medikationen verabreicht:

#### Baldrian-Trockenextrakt

Darreichungsform: 300 mg Kapseln (Studienteil 1) / Dragées (Studienteil 2)

Applikationsart: oral, zum Einnehmen

Dosierungsschema: 2 Kapseln / Dragées abends vor dem Schlafengehen

Behandlungsdauer: 1 Tag (Studienteil 1) bzw. 14 Tage (Studienteil 2)

#### Flunitrazepam

in Form, Größe und Farbe mit dem Baldrianpräparat und der Placebo-Zubereitung identische Kapseln

Darreichungsform: 1 mg Filmtablette, verkapselt

Applikationsart: oral, zum Einnehmen

Dosierungsschema: 1 Kapsel/1 Placebo-Kapsel abends vor dem Schlafengehen

Behandlungsdauer: 1 Tag (Studienteil 1)

#### Placebo

In Form, Größe und Farbe mit dem Baldrian-Extrakt bzw. Flunitrazepampräparat identische Kapsel bzw. identisches Dragée

Darreichungsform: Kapseln (Studienteil 1) / Dragées (Studienteil 2)

Applikationsart: oral, zum Einnehmen

Dosierungsschema: 2 Kapseln/ Dragées abends vor dem Schlafengehen

Behandlungsdauer: 1 Tag (Studienteil 1) bzw. 14 Tage (Studienteil 2)

### **3.3.2 Dosis und Dosierungsschema**

Da eine mögliche Beeinträchtigung alltagsüblicher Dosierungen der Arzneimittel erfasst werden sollte, entsprachen Dosierungen und abendliche Einnahme den Herstellerempfehlungen.

### 3.3.3 Randomisierung

Unter Randomisierung ist die zufällige Verteilung der Studienteilnehmer auf die verschiedenen Untersuchungs- bzw. Behandlungsgruppen zu verstehen. Mit diesem methodischen Vorgehen sollte eine positive oder negative Selektion der Probanden vermieden und eine gleichmäßige und zufällige Verteilung der Probanden auf die Behandlungsgruppen sichergestellt werden. Dies war für die spätere wissenschaftliche Aussagekraft der explorierten Studienergebnisse von grundlegender Bedeutung.

Die Randomisierung wurde vom Institut für interdisziplinäre medizinische Forschung (IMF) mit dem Programm „SAS PROC PLAN“ durchgeführt. Das Ergebnis wurde in einer Randomliste festgehalten. Die Zuteilung der Probanden zu den Behandlungsgruppen erfolgte randomisiert in 4er Blöcken (Studienteil 1) bzw. 6er Blöcken (Studienteil 2).

Die Randomisierung für beide Studienteile wurde *einmalig* durchgeführt. Der erste in die Studie aufgenommene Proband erhielt die Nummer 1, der zweite die Nummer 2 usw. Die Probanden der Flunitrazepam-Gruppe aus Studienteil 1 wurden für Studienteil 2 den beiden anderen Gruppen (Baldrian-Extrakt und Placebo) gleichmäßig zugeteilt. Diese Verfahrensweise war biometrisch zulässig, da eine siebentägige Auswaschphase zwischen den Studienteilen vorgesehen war. Um eine gleichmäßige Verteilung der Teilnehmer des dreiarmligen Studienteils 1 auf die zwei Gruppen des Studienteils 2 zu ermöglichen, sollten 102 Probanden in die Studie eingeschlossen werden. Die Gruppenstärke für Studienteil 1 sollte 34, für Studienteil 2 bei 51 Probanden betragen.

### 3.3.4 Verblindung

Die höchste wissenschaftliche Aussagekraft haben Studien, die unter doppelblinden Bedingungen durchgeführt werden. Damit soll der Einfluss der Erwartungshaltung des Studienteilnehmers wie auch der Einfluss des Untersuchers auf die Studienergebnisse minimiert werden. Die Behandlungszuordnung der Studienteilnehmer ist nicht offensichtlich, da die Medikationen verblindet abgegeben werden, d.h. man sieht der Kapsel oder Tablette nicht an, ob sie aktiven Wirkstoff (Verum) oder Traubenzucker (Placebo) enthält. Eine perfekte Verblindung stellt hohe Ansprüche an den Arzneimittelhersteller, da die Studienmedikationen nicht nur gleich aussehen, sondern auch gleich riechen und schmecken müssen.

In der vorliegenden Studie unterlagen sowohl Untersucher als auch Probanden der Verblindung (Doppelblindstudie). Die Verblindung war sichergestellt durch:

- äußerliche Identität (Form, Größe, Farbe) der Medikationen
- gleiche Anzahl von Kapseln (Studienteil 1) bzw. Dragées (Studienteil 2).

Der Behandlungscode für den einzelnen Probanden wurde unter Verschluss gehalten (Notfallcouverts) und durfte während der Studie nur bei Eintritt eines medizinischen Notfall gebrochen werden, um Kenntnis von der Art der Medikation zu erlangen und um geeignete Behandlungsmaßnahmen einleiten zu können. Im Falle einer Entblindung war der Studienteilnehmer sofort von der weiteren Studienteilnahme auszuschließen. Probanden, Untersucher und sämtliche in die Studie involvierten Personen unterlagen bis nach Beendigung des experimentellen Teils der Studie der Verblindung. Erst nach Schließung der Datenbank (database closure) wurde der Random-Code für die statistische Auswertung gebrochen.

### **3.3.5 Begleittherapie**

Die gleichzeitige Einnahme von Psychopharmaka, Hypnotika, Sedativa, Analgetika oder anderen Pharmaka, welche bekanntermaßen die kognitive Leistungsfähigkeit beeinflussen, vier Wochen vor Beginn oder während der Teilnahme an der Studie, führten zum Ausschluss des Probanden. Alle anderen Begleittherapien mussten seit mindestens vier Wochen bestehen und durften während der Studienteilnahme nicht verändert werden. Andernfalls wurde im Rahmen der statistischen Auswertung über den Ausschluss des Probanden von der According-To-Protocol-Analyse im Einzelfall entschieden.

### **3.3.6 Behandlungcompliance**

Die Probanden wurden angehalten, Restmedikation (angebrochene und leere Packungen) beim nächsten Besuch mitzubringen und zurückzugeben.

Die Compliance wurde bei Besuch 2 (Tag 2) und Besuch 4 (Tag 23) durch Feststellung der nicht verbrauchten Kapseln bzw. Dragées übergeprüft (Pill-Counting) und im CRF (Case Report Form) eines jeden Probanden dokumentiert.

### **3.3.7 Arzneimittelsicherheit**

#### **3.3.7.1 Unerwünschte Ereignisse und medizinische Sicherheitsvariablen**

Eine Nebenwirkung bzw. ein unerwünschtes Ereignis (UE) ist jede unerwünschte, schädigende oder pathologische Veränderung (z.B. Laborwerte) nach Einnahme der Studienmedikation oder eines Placebos, unabhängig davon, ob der Verdacht eines Zusammenhangs zwischen den eingetretenen Veränderungen und der verabreichten Medikation besteht oder nicht. Unabhängig vom kausalen Zusammenhang musste jedes im Rahmen der Studie aufgetretene UE, das der

Proband bei den Kontrolluntersuchungen spontan oder auf Befragung angab, erfasst und dokumentiert werden. Die Kausalitätsbewertung sollte nach Möglichkeit unter Blindbedingungen vorgenommen werden - d.h. die Behandlungszuteilung des betreffenden Probanden wurde nicht dekodiert, ausgenommen im Falle einer notwendigen Decodierung zur Einleitung einer Notfallbehandlung.

Alle während des Studienverlaufs aufgetretenen und dokumentierten unerwünschten Ereignisse wurden in die Analyse einbezogen. Ziel der statistischen Auswertung der unerwünschten Ereignisse war die tabellarische Darstellung von Häufigkeiten und Inzidenzraten (probandenbezogen) aller im Studienverlauf aufgetretener UE.

#### *Medizinische Sicherheitsvariablen*

Das Gewicht in kg Körpergewicht wurde auf immer derselben Waage gewogen. Nach mindestens fünfminütigem ruhigen Sitzen wurden Blutdruck und Pulsfrequenz im Sitzen gemessen. Die Messung des Blutdruckes erfolgte mit immer demselben Blutdruckmessgerät, die Pulsfrequenz wurde durch einminütiges Zählen des Radialispulses erfasst. Körpergewicht, Blutdruck und Pulsfrequenz wurden zu allen Untersuchungszeitpunkten gemessen.

An Tag 2 und Tag 23 sowie bei einem evtl. vorzeitigem Studienabbruch wurde die Verträglichkeit vom Arzt und vom Probanden folgendermaßen eingestuft: „sehr gut“, „gut“, „mäßig“, „unbefriedigend“.

#### **3.3.7.2 Klinisches Labor**

Bei der Aufnahme- und der Abschlussuntersuchung (Tag 1/Tag 23) wurden folgende Parameter erhoben: Hämoglobin, Erythrozyten, Leukozyten, Thrombozyten, Kreatinin, GOT, GPT,  $\gamma$ GT. Neben der Erfassung evt. unerwünschter Nebenwirkungen (Abweichungen von den Referenzbereichen des untersuchten Labors) dienen die Parameter gleichzeitig zur Verifizierung der Ausschlusskriterien bei der Aufnahme.

## 4 PSYCHOMETRISCHE MESS-UND RATINGVERFAHREN

### 4.1 Kölner Farbpräferenztest (KFPT)

Der an der Universität Köln entwickelte KFPT ist ein wissenschaftliches Routineverfahren zur Untersuchung persönlicher Farbpräferenzen, mit dem Ziel einer möglichst quantitativen Aussage hinsichtlich des individuellen Farbwahlverhaltens.

Das Verfahren ist eine computergestützte Weiterentwicklung des Farbpyramidentests (FPT) von Heiss und Halder. Beide Verfahren basieren auf den Erkenntnissen der Farbpsychologie über den engen Zusammenhang von Farben und Affekten, Gefühlen und Stimmungen und zielen auf die Erfassung der aktuellen und der allgemeinen Affektstruktur der Persönlichkeit ab. Der Geltungsbereich ist damit auf den affektiven und emotionalen Bereich der Persönlichkeit festgelegt.

Beim „klassischen“ FPT besteht das Testmaterial aus 14 verschiedenen 2,5 cm<sup>2</sup> großen „Farbplättchen“ und zwei „Vorlagen“. Die Vorlage zeigt ein Muster mit 15 Feldern, die von der gleichen Größe sind wie die Farbplättchen. Die Felder sind so angeordnet, dass sie die Form eines Dreiecks bilden. Durch den treppenförmigen Anstieg von der fünf Felder umfassenden Basis bis zur, aus einem Feld bestehenden „Spitze“ entsteht der Eindruck einer Pyramide. Der KFPT ist - wie der FPT- ein nonverbaler Wahltest, in dem die Farbe als Testmaterial verwendet wird. Im Unterschied zum FPT wird das Testmaterial beim KFPT über den Computer generiert.

In einer (auf der linken Seite des VGA-Monitors) vorgegeben Farbpalette stehen 14 verschiedene Farben, einschließlich unbunter Farbtöne, zur Auswahl. Die KFPT-Farbpalette entspricht *nicht* dem Farbangebot des klassischen FPT. Auf Grundlage des empirischen Materials von Diester und Stapf (1969) stellte Höger für den KFPT einen modifizierten Farbsatz zusammen, der die Gefühlswirkung, Potenz und Valenz in ihrer Abhängigkeit von Helligkeit, Sättigung und farbtongleicher Wellenlänge zeigt. Von jeder farbtongleichen Wellenlänge wurden insgesamt 3 Nuancen ausgewählt - eine helle Nuance (a), eine mit maximaler Sättigung und etwa mittlerer Helligkeit (b) und eine dunkle Nuance (c) und auf elf Farbtöne plus drei achromatische Farben verteilt. Die von Höger zusammengestellte KFPT-Farbpalette umfasst drei helle Farbtöne (1a, 4a, 6a), fünf gesättigte Farbtöne (1b, 5b, 6b, 8b, 9b), drei dunkle Farbtöne (1c, 5c, 8c) sowie drei unbunte Farbtöne (weiß 10a, grau 10b, schwarz 10c).

Die Farbcodierung der einzelnen KFPT-Farbtöne erfolgte nach dem Munsell-System (1915). Das „Munsell-book of colors“ ermöglicht es jede gegebene Farbe zu codieren. Dieses Farbsystem beruht auf den Attributen „hue“ (Farbton im Sinne der Position im sichtbaren Spektrum bzw. auf dem Farbkreis), „value“ (Helligkeit) und chroma (Farbstärke oder Sättigung), die als Beschreibungsdimensionen verwendet werden. Sie sind dabei in von der Wahrnehmung her

gleich groß erscheinende Intervalle eingeteilt. Es handelt sich daher um ein wahrnehmungspsychologisches Farbsystem, welches jedoch in die physikalische Farbcodierung übertragen werden kann (Höger, 1980).

Jede Farbnuance der KFPT-Farbpalette ist definiert durch die Angabe von Wellenlänge (hue), Helligkeit (value) und Sättigung (chroma), und einheitlich gestaltet nach der konventionellen Formel h-v/c. So codiert z.B. 5R-7/6 den Farbton 1a: die Zahl 5 vor dem Buchstaben bezeichnet die wellenlängebestimmte Nuance des Farbtones; der Buchstabe R steht für den englischen Farbnamen (red); die 7 dient als Kennwert der Helligkeitsstufe, und die 6 nach dem Querstrich ist der Index für den Sättigungswert.

FARBTON	MUNSELLWERT
1a	5R-7/6
4a	2,5Y-8/6
6a	5G-7/6
1b	5R- 5/14
5b	2,5GY-5/12
6b	2,5G-5/12
8b	5PB-5/12
9b	10P-5/12
1c	5R-3/6
5c	2,5GY-5/6
8c	5PB-3/6
10a	N9.5
10b	5Y7.2
10c	N1

Tab 4: KFPT-Farbpalette und Farbcodierung nach MUNSELL

Vom Probanden wird nicht nur eine Wahl gefordert, sondern eine kontinuierliche Wahlfolge. Je 15 Wahlen werden zu einer Pyramide zusammengefasst. Aufgabe des Probanden ist es drei „schöne“ und drei „hässliche“ Pyramiden zu wählen.

Bei einer kompletten Testdurchführung sind insgesamt 90 Wahlvorgänge notwendig. Der Proband kann die Farbtöne nach Belieben aus der Palette auswählen, bestimmte Farben bevorzugen und andere vermeiden. So kann der Proband z.B. eine Farbe kontinuierlich wählen und einfarbige Pyramiden

zusammenstellen oder aber zwischen den Farben wechseln. Per „Mausklick“ wird ein Farbton der Palette ausgewählt und durch „Herüberziehen“ mit der „Maus“ in der Pyramide placiert (rechte Seite des VGA-Monitors). Weiterhin besteht die Möglichkeit, einen zuvor gewählten Farbton der Pyramide durch eine Korrektur (die Wahl einer anderen Farbe) zu überdecken. Dem Probanden ist ein breiter Spielraum gegeben, innerhalb dessen die Wahlen hin und her pendeln und sich verschiedene affektive Strukturen zeigen können. Die Protokollierung der Wahlen erfolgt mittels Computerprogramm.

Angaben zu den testtheoretischen Gütekriterien des KFPT, Normwerten und Einsatzmöglichkeiten liegen bisher nicht vor. Insbesondere Ergebnisse zur Pharmakosensitivität des Verfahrens und die Frage, ob der KFPT in der Lage ist, durch Pharmaka induzierte psychologische Veränderungen abzubilden, stehen noch aus. Aufgrund seiner starken Affinität zum FPT sind dessen bekannten methodischen Probleme zur Reliabilitätsbestimmung (Michel, 1962) sowie die belegte Validität (Michel, 1959) und Objektivität auf den KFPT übertragbar. Die Zielvariablen des FPT (Häufigkeit der Farbtöne und Syndrome, Formungen der Pyramiden, Verlaufsformeln, Kippindex) wurden für den KFPT bewusst nicht übernommen, da sich die Farbpaletten beider Verfahren unterscheiden und sich Farbverbände bei neueren farbpsychologischen Untersuchungen als die stabilsten und aussagekräftigsten Merkmale erwiesen haben (Heiss, Halder & Höger 1975). Stattdessen wurden Cluster (Farbverbände) und verbleibende Einzelfarbtöne und deren Veränderungen über den Untersuchungszeitraum als Zielvariablen für den KFPT definiert.

#### **4.2 Semantisches Eindrucksdifferential (ED)**

Die Interpretation und Gewichtung der mit dem KFPT gewonnenen Testergebnisse zum Farbwahlverhalten setzt die Kenntnis von Beziehungen zwischen den Dimensionen der Farbwahrnehmung und den Dimensionen der Gefühlsenerlebnisse voraus. Angesichts des Fehlens von Daten zur Überprüfung des Farberlebens der computergenerierten Farben, war eine Exploration der speziellen Farbpalette des KFPT erforderlich. Auf die Erkenntnisse zur Farbwirkung der FPT-„Farbplättchen“ konnte nicht zurückgegriffen werden, da es sich hier um eine gänzliche andere Darbietung des Testmaterials handelt, die eine Übertragung der Befunde ausschließt.

Methodisch wurde das Farberleben durch den Einsatz eines semantischen Farbdifferentials evaluiert. Bei Gefühlen handelt es sich um Erlebnisqualitäten, die sich mit Wahrnehmungen bzw. Vorstellungen verbinden und sprachlich in differenzierter Weise bezeichnet werden können. Die Sprache schafft ein abbildendes System für Gefühlsenerlebnisse, bei welchem zudem die Reduktion auf grundlegende Beschreibungsdimensionen besteht.

Eine Untersuchung von Ertel (1964) belegt, dass die von Osgood, Suci, und Tannenbaum (1957) ermittelten Faktoren des sogenannten „semantischen Raumes“ als emotionale Dimensionalität zu interpretieren sind. Ertel beschreibt

die Faktoren als:

- *Erregung*

(im Sinne von Erregung, Spannung verursachend, Unruhe verbreitend vs. Erregung mindernd, Ruhe verbreitend)

- *Valenz*

(im Sinne von angenehm, anziehend vs. unangenehm, abstoßend wirkend)

- *Potenz*

(im Sinne von kraftvoll wirkend, dominierend vs. kraftlos, unterlegen wirkend).

Mit der "Methode des Semantischen Differentials" entwickelte Ertel ein Eindrucksdifferential, welches die relativ einfache Bestimmung des von einem Gegenstand für eine Person ausgehenden Gefühlseindrucks ermöglicht, und zwar getrennt nach den Dimensionen "Erregung", "Valenz" und "Potenz".

Bei der Instruktion werden die Versuchspersonen gebeten, vom unmittelbaren und gefühlsmäßigen Eindruck her, den eine Farbe auf sie ausübt, zu beurteilen, ob diese z.B. mehr dem Begriff "Ruhe" oder "Bewegung" entspricht. Die Intensität wurde durch die angekreuzte Skalenstufe ausgedrückt. Der Nullpunkt der Skalen ist geteilt. Falls sich der Proband für keinen der beiden Begriffe des Gegensatzpaares entscheidet, sollte zumindest eine tendenzielle Wertung vorgenommen werden.

Das Eindrucksdifferential besteht aus insgesamt 18 gegensätzlichen Begriffspaaren, welche jeweils die Pole einer achtstufigen Skala bezeichnen. Je sechs sind den drei Dimensionen zugeordnet. Die Gegensatzpaare des Eindrucksdifferentials sind weitgehend rein und ermöglichen eine direkte Bestimmung der Dimensionswerte unter Umgehung einer Faktorenanalyse aus dem arithmetischen Mittel der Einstufungen auf den zur gleichen Dimension gehörenden Skalen. Durch die sechs zu einer Dimension gehörenden Begriffspaare bzw. Skalen erfolgt eine sechsfach wiederholte Einschätzung des jeweiligen Gegenstandes auf der entsprechenden Dimension, wodurch sich die Reliabilität der Dimensionswerte entsprechend erhöht.

Das Berechnungsverfahren besteht darin, zunächst den Skalenpunkten die Zahlenwerte 1 bis 8 zuzuweisen. Anschließend werden je Versuchsperson und Farbnuance aus dem arithmetischen Mittel der jeweils drei zu einer Dimension gehörenden Skalen die Dimensionswerte für Erregung, Valenz und Potenz bestimmt. Aus den individuellen Mittelwerten wird wiederum das arithmetische Mittel für die Gesamtgruppe gebildet. Da die Beziehungen zwischen Gefühls- und Farbwahrnehmungsvariablen nicht linear sind, entfallen Korrelationskoeffizienten als Beschreibungsgrößen. Der Einsatz des Eindrucksdifferentials hat sich in Farbuntersuchungen bewährt (Höger, 1980).

### 4.3 Determinationstest (DT)

Der Determinationstest (DT) ist eine computergestützte Weiterentwicklung des Wiener Determinationsgerätes (WDG). Im Gegensatz zum WDG werden beim DT die Reize über den Bildschirm (Standard VGA-Monitor) ausgegeben. Die Reaktion erfolgt durch Betätigung der entsprechenden Tasten am Probandenpanel. Der DT entspricht in seinem apparativen Aufbau nicht dem traditionellen Vorläufer WDG. Trotz dieser Unterschiede besteht nach den Angaben im Testmanual „eine große Übereinstimmung der Testergebnisse von DT und WDG“ (Schuhfried, 1996).

Wie beim WDG sind beim DT optische Signale und Töne mit Tasten oder Fußpedalen zu beantworten. Die Reizdarbietung erfolgt auf dem Bildschirm. Zehn optische Reize in den Farben weiss, gelb, rot, grün und blau bilden eine obere und eine untere Reihe. Die fünf diesen Farben zugeordneten Reaktionstasten sind auf dem Probandenpanel so angebracht, dass mit beiden Händen reagiert werden kann. Zwei weitere Reize (weiße rechteckige optisch abgehobene Felder) erfordern die Betätigung des linken oder rechten Fußpedals. Zwei akustische Reize (hoher bzw. tiefer Ton) sind den beiden „Ton“-Tasten in der Mitte des Panels zugeordnet. Die untere, rechteckige schwarze Taste ist immer der tiefe Ton, die obere, rechteckig graue Taste immer dem hohen Ton zugeordnet.

In der vorliegenden Studie wurde der Test im Modus „Adaptiv“ durchgeführt. Dabei wurde das Tempo laufend an die Reaktionszeit des Probanden angepasst. Die Reizdauer ergab sich aus dem Mittelwert der letzten 8 Reaktionen. Wurde bei einem Reiz nicht richtig (zeitgerecht oder verspätet) geantwortet, wurde statt der Reaktionszeit die doppelte Reizdarbietungszeit verwendet. Das Tempo der Reizausgabe wurde laufend an die Arbeitsgeschwindigkeit des Probanden angepasst, dass etwa 70% der Reize richtig beantwortet wurden. Bei dieser Form der Reizdarbietung bewegte sich der Proband immer an der Grenze seiner Leistungsfähigkeit. Beim ersten Reiz wurde angenommen, dass zuvor 8mal mit einer Reaktionszeit von drei Sekunden reagiert wurde. Die Testdauer betrug acht Minuten. Sämtliche oben beschriebenen Stimuli wurden vorgegeben. Hauptzielvariable im DT war der Median der Reaktionszeit, d.h. die Zeitspanne zwischen der Darbietung eines Reizes und der Betätigung der adäquaten Taste des Panels.

## **5 DATENQUALITÄT**

### **5.1 Quelldaten-Überprüfung (Source Data Verification)**

Source Data Verification (SDV) ist ein essentieller Bestandteil zur Sicherung der Glaubwürdigkeit der erhobenen wissenschaftlichen Daten. Die Datenverifizierung dient dem Ziel nachzuweisen, dass die erhobenen und in der Case Report Form (CRF) dokumentierten Daten auf real existierende Probanden und an ihnen vorgenommene Untersuchungen basieren, nachvollziehbar und korrekt sowie vollständig sind.

Während der Studiendurchführung wurden schriftliche Einverständniserklärungen der Studienteilnehmer und sämtliche Untersuchungsbefunde (z.B. automatisierte Mess- und Testergebnisse, Laborbefunde, Dokumentationen unerwünschter Ereignisse) sowie die Behandlungs-Compliance mit den für die Studie vorgesehenen Prozeduren (Einhaltung der Kontrollzeitpunkte, Ausgabe und Rücknahme der Medikation incl. Pill-Counting) von einem unabhängigen Monitor des IMF, der selbst nicht in die Studiendurchführung involviert war, hinsichtlich Korrektheit und Vollständigkeit der Daten (missing data) überprüft.

### **5.2 Dokumentation der Daten**

Sämtliche studienbezogenen Daten eines Probanden wurden in einem separaten Probanden-Ordner (CRF=Case Report Form) dokumentiert und während des Studienverlaufs archiviert. Das CRF spiegelte den Untersuchungsplan wider und enthielt sämtliche Testergebnisse bzw. medizinischen Befunde in chronologischer Reihenfolge. Der Proband war lediglich anhand der Random-Nr. und Initialen zu identifizieren. Damit erfüllte die Anonymisierung vollumfänglich geltende datenschutzrechtliche Bestimmungen. Erhobene Daten wurden direkt im Anschluss an die Untersuchungen im Probanden-Ordner dokumentiert bzw. als Ausdrucke (print-outs) abgelegt (z.B. Messergebnisse des Drogenscreenings oder Rohdaten der psychometrischen Tests). Erforderliche Korrekturen im CRF waren mit Datum, Unterschrift des Untersuchers und einer Erklärung zu versehen. Neben der Ablage im Probanden-Ordner verblieben die Resultate der computergenerierten Testergebnisse auf dem jeweiligen elektronischen Datenträger als back-up. Diese Daten wurden nach Abschluss der experimentellen Untersuchungen – und nach erfolgtem Abgleich mit den print-outs – direkt in die Datenbank für die Studienauswertung transferiert, die CRF-Daten wurden nach durchgeführtem data cleaning per doppelter Dateneingabe (double data entry) durch zwei verschiedene Personen eingegeben und ein Cross-Check vorgenommen.

## 6 DATENANALYSE UND STATISTISCHE METHODEN

Der prospektive Charakter einer kontrollierten Arzneimittelstudie verlangt die vorherige Festsetzung aller biometrischen Details im Rahmen der Studienplanung. Ziel der kontrollierten klinischen Studie ist die Ableitung einer allgemeingültigen Aussage über das Nutzen-Risiko-Verhältnis zweier oder mehrerer medikamentöser Behandlungen bzw. deren Einfluss auf bestimmte psychologische Konstrukte aufgrund eines mit einer vorgegebenen Wahrscheinlichkeit reproduzierbaren Studienergebnisses.

Voraussetzungen für die Generalisierbarkeit der Ergebnisse sind die Anlage der Studie als „blind randomisierter Versuch“ sowie die Anwendung interferenzstatistischer Verfahren (statistische Testverfahren) auf die Studienergebnisse. Reproduzierbarkeit der Studienergebnisse setzt ferner voraus, dass die Studienhypothese, die Wahl der statistischen Verfahren, die Fehlerwahrscheinlichkeit und die zur Absicherung der Reproduzierbarkeit des Studienergebnisses statistisch erforderliche Fallzahl vor Studienbeginn festgelegt worden sind.

### 6.1 Stichprobenumfang

Um valide und möglichst generalisierbare Aussagen zu generieren, geht jeder wissenschaftlichen Studie eine Festlegung der sogenannten Fallzahl voran, also der Anzahl von Probanden, die im Rahmen der Studie rekrutiert werden sollten. Ziel einer Fallzahlschätzung ist es immer, über genügend Datensätze zur Auswertung der Studie zu verfügen, um eine statistisch abgesicherte Aussage über die Studienergebnisse treffen zu können. Sich hinsichtlich der Fallzahl allein von praktischen Erwägungen leiten zu lassen birgt das Risiko, möglicherweise keine Signifikanz der Studienergebnisse nachweisen zu können. Meist werden Fallzahlberechnungen anhand von bekannten klinischen und statistischen Kenngrößen durchgeführt. Für die Fallzahlberechnung wurden publizierte Daten einer klinischen Studie von Albrecht et al. (1995) zugrunde gelegt, wo das Wiener Determinationsgerät als Hauptzielkriterium diente. Da für die confirmatorische Statistik der parallel durchgeführten „Valerian-Treatment Study“ (Kuhlmann et al., 1999) ein 95%-Konfidenzintervall berechnet werden sollte, lag der Fallzahlberechnung folgende Formel (Hartung, 1982, Kap. IIIV S. 511-513) zugrunde:

$$n \geq 2 \left( \frac{(u_{1-\alpha/2} + u_{1-\beta}) \cdot \sigma}{(d_1 - d)} \right)^2 \text{ (Konfidenzintervalles)}$$

wobei

n: Fallzahl jeder Medikationsgruppe

$\alpha$ : Signifikanz, hier 5%

$\beta$ : Wahrscheinlichkeit für den Fehler 2. Art, hier 20%

(daraus ergibt sich eine POWER von 80%)

$\sigma$ : Streuung der Gesamtstichprobe, hier 47

d1: Stelle, an welcher der Fehler abgesichert werden soll, hier = 0

(kein Unterschied zwischen Baldrian-Extrakt und Placebo)

d: maximal noch klinisch irrelevanter Unterschied, hier = 30 (DELTA)

u: Quantile der Standardnormalverteilung, hier für  $\alpha = 5\%$

d.h.  $u_{1-\alpha/2} = u_{0.975} = 1.9600$

$u_{1-\beta} = u_{0.800} = 0.8416 \Rightarrow N=39$

Anhand dieser Vorgaben der Fallzahlberechnung ergab sich eine Gruppenstärke von N=39. Unter Berücksichtigung einer Dropout-Rate für Studienabbrecher von 20% wurde eine Fallzahl von mindestens 98 zu randomisierenden Probanden festgelegt. Aufgrund der Randomisierung in 6er Blöcken (Studienteil 1) bzw. 4er Blöcken (Studienteil 2) wurde eine Fallzahl von N=102 festgelegt:

Studienteil 1: 102 (34/34/34) randomisierte Probanden

Studienteil 2: 102 (51/51) randomisierte Probanden.

## 6.2 Definition der Auswertekollektive

Insbesondere bei Untersuchungen mit größeren Fallzahlen ist davon auszugehen, dass nie alle Teilnehmer die Studie genauso durchlaufen wie es die Planung vorsieht. Deshalb wurde im Vorhinein festgelegt, wie mit Protokollverletzungen (z.B. fehlende Daten, mangelnde Compliance oder divergierende Untersuchungszeitpunkte) umzugehen ist. Nach Studienabschluss wurden zwei Auswertungen - basierend auf unterschiedlichen Kollektiven - durchgeführt: Grundsätzlich wurde das ATP-Kollektiv (According-To-Protocol = Kollektiv derjenigen Probanden, welche die Studie gemäß dem Untersuchungsplan komplett durchlaufen hatten) zugrunde gelegt, da dies - nach den Vorgaben in der Planung der Studie - das relevante Kollektiv für die Analyse der Hauptzielvariablen darstellte. Bei der zusätzlich durchgeführten ITT-Analyse (Intention-To-Treat) wurden auch diejenigen Probanden berücksichtigt, die lediglich einen Aufnahme- und Folgewert aufwiesen. Die ITT-Analyse wertet Daten von allen Probanden aus, die randomisiert und einer Behandlung zugeführt wurden – unabhängig davon, ob die Planungsvorgaben exakt eingehalten wurden. Das Ziel der ITT-Analyse war es, den „realen“ Durchführungsbedingungen bei der Auswertung Rechnung zu tragen.

## **6.3 Generelle Überlegungen zur Datenanalyse**

### **6.3.1 Abweichungen vom Untersuchungsverlauf**

Die Zuordnung der Daten in das jeweilige Auswertekollektiv (Safety Analysis, FAS, PPS) wurde vor der statistischen Analyse für jeden Probanden im Einzelfall im Rahmen eines „Blind Data Review Meetings“ zwischen Statistiker und Versuchsleiter diskutiert und entschieden. Hierbei wurden Abweichungen vom vorgesehenen Untersuchungsverlauf in geringfügig (minor violation) und bedeutsam (major violation) differenziert. Bedeutsame Verletzungen der Ein- und Ausschlusskriterien, des Untersuchungsverlaufs oder der Compliance führten zum Ausschluss des Probanden von der PPS-Datenanalyse. Für die verbleibenden Probanden mit geringfügigen Abweichungen vom Untersuchungsplan wurden Listen angefertigt, um die Zuordnung der Probanden für die einzelnen Analysekollektive festzulegen und die numerische Zahl der Auswertekollektive zu fixieren.

Falls ein Proband die Studie vorzeitig abgebrochen haben sollte, wurden - gemäß der „Last Observation Carried Forward“ -Methode (LOCF) - die zuletzt erhobenen Daten fortgeschrieben.

### **6.3.2 Behandlungcompliance**

Für den Studienteil 1 wurden diejenigen Probanden, welche die Medikation am Tag 1 nicht eingenommen hatten, von der ATP- Analyse ausgeschlossen.

Probanden, die weniger als 24 Dragées im Studienteil 2 verbraucht hatten, wurden in der According-To-Protocol-Auswertung des Studienteils 2 nicht berücksichtigt.

## **6.4 Interimanalyse**

Vor allem in Langzeitstudien mit großen Fallzahlen werden meist geplante Zwischenauswertungen durchgeführt. Das Zwischenergebnis kann zur vorzeitigen Beendigung der Studie führen, wenn z.B. eine medikamentöse Behandlung mit höheren Sicherheitsrisiken behaftet ist als die Behandlung mit einer Vergleichssubstanz oder die Wirksamkeit der Therapie fraglich erscheint. Die vorliegende Studie erstreckte sich lediglich über einen kurzen Therapiezeitraum, so dass die Durchführung einer statistischen Interimanalyse als nicht notwendig erachtet wurde.

## 6.5 Analyse der Primärvariablen

### 6.5.1 Kölner Farbpräferenztest (KFPT)

Laut Michel (1962) sollte die Reliabilität des FPT nicht auf die Wahlhäufigkeit einzelner Farbtöne reduziert werden. Die Auswertung des FPT dürfe nicht von den Wahlhäufigkeiten der einzelnen, isolierten Farbtöne ausgehen, sondern sollte auf der Analyse von Farbverbänden basieren. Explizit weist er darauf hin, dass die einzelnen FPT-Variablen nicht wie bei psychometrischen Tests jeweils für sich interpretiert werden sollten, sondern stets nur in ihrer Stellung zum Merkmalsgesamt.

Michels Forderung korrespondiert mit den Befunden aktuellerer farbpsychologischer Untersuchungen, wonach sich Farbverbände (Cluster) als die stabilsten und aussagekräftigsten Merkmale erwiesen haben. Die Ergebnisse von Heiss, Halder und Höger (1975) berücksichtigend, wurde die 1951 entwickelte Auswertestrategie des klassischen FPT für den neu entwickelten KFPT *nicht* übernommen. Man sollte sich vor dem Trugschluss hüten, je differenzierter die die Auswertung (Verlaufsformel, Kippindex, Formungen, Syndrome), desto ergiebiger die Interpretation. Verfeinerung der Auswertung geht stets mit einem Verlust an Reliabilität und damit im Allgemeinen auch mit Validität einher (Michel, 1962).

Deshalb sollte für den KFPT auf Grundlage des im Rahmen der Studie erhobenen Datenmaterials von Kölner Farbpräferenztest und semantischen Eindrucksdifferential (ED) ein eigenständiges Auswertekonzept unter Berücksichtigung von Farbverbänden als Zielvariablen entwickelt werden. Die Zusammenfassung mehrerer Einzelfarben zu Clustern hat methodisch den Vorzug, dass sich hierdurch die Reliabilität des KFPT erhöht – die Clusterbildung entspricht formal gesehen einer Testverlängerung (Lienert, 1969), die bei weniger reliablen Einzelfarben durch eine Gruppenbildung von Farben gleichgerichteter Eindrucksqualität zu einer relativ hohen Korrelation zwischen Leer- und Retest führt (Höger, 1968).

Eine adäquate Analyse und Interpretation der KFPT-Daten setzt die Kenntnis der Beziehungen zwischen den Dimensionen der Farbwahrnehmung und den Dimensionen der Gefühlserlebnisse der computergenerierten Farben voraus. Die hierzu notwendige farbpsychologische Exploration der speziellen KFPT-Farbpalette wurde mittels des ED durchgeführt. Basierend auf den Dimensionswerten des semantischen Eindrucksdifferentials für Erregung, Valenz und Potenz, wurde für den 14-Farbensatz des KFPT im Anschluss eine Clusteranalyse durchgeführt (Details siehe Kapitel 6.6.1).

Nach eingehender Diskussion der Ergebnisse der hierarchischen Clusteranalyse mit Höger und Stephan wurden – unter Beachtung der farbpsychologischen Besonderheiten des achromatischen Farbtönen 10c (schwarz) sowie der beiden Farbtöne mit maximaler Sättigung und mittlerer Helligkeit 1b (rot) und 5b (gelb) -

folgende drei Cluster bzw. Einzelfarbtöne als KFPT-Zielvariablen für die statistische Analyse definiert.

Cluster 1	1a/4a/6a/10b
Cluster 2	6b/8b/8c/9b/10a
Cluster 3	1c/5c
Einzelfarbe 1	1b
Einzelfarbe 2	5b
Einzelfarbe 3	10c

Tab. 5: KFPT-Zielvariablen (14-Farbensatz)

Bei dieser explorativen Studie wurden vorab keine Hypothesen hinsichtlich des Farbwahlverhaltens formuliert. Alle Analysen wurden deskriptiv durchgeführt. Für die Interpretation der KFPT-Daten wurden zusätzlich p-Werte für die Cluster bzw. Einzelfarben berechnet. Innerhalb der Medikationsgruppen wurden für jeden Parameter die Zeitpunkte Baseline, Retest 1, Retest 2 mittels des Wilcoxon-Tests miteinander verglichen. Die Medikationsgruppen wurden anhand der Parameter zur Baseline mit dem Mann-Whitney-U-Test, zu Retest 1/Retest 2 mit dem Paardifferenzen-U-Test (Differenz zur Baseline) verglichen; p-Werte  $\geq 0.2$  wurden als nicht signifikant (n.s.) wiedergegeben. Weiterhin wurden Spearman-Rangkorrelationskoeffizienten zwischen Valenzwert im semantischen Eindrucksdifferential und der Wahlhäufigkeit des Clusters bzw. des Einzelfarbtönen im KFPT berechnet.

## 6.6 Analyse der Sekundärvariablen

### 6.6.1 Semantisches Eindrucksdifferential (ED)

Für die Auswertung des semantischen Eindrucksdifferentials wurden zunächst Scores für die Dimensionswerte Erregung, Valenz, Potenz berechnet. Dabei war die Skalenpolung – im Gegensatz zur Vorgehensweise von Ertel - so gewählt, dass hohe numerische Werte jeweils eine hohe Dimensionsausprägungen wiedergeben und umgekehrt. Aus den sechs transformierten Rohwerten einer Dimension wurde aus dem arithmetischen Mittel der jeweilige Dimensionswert für Erregung, Valenz und Potenz des speziellen Farbtönen der KFPT-Farbpalette bestimmt.

Auf Basis der mittleren E/V/P-Dimensionswerte (Baseline) der vierzehn KFPT-Farbtöne wurde anschließend eine Clusteranalyse mit der SPSS-Version 8.0 durchgeführt. Dabei wurden die Farbtöne und die Merkmale bzgl. derer sie sich ähneln sollten (die – über die Probanden – gemittelten ED-Dimensionswerte für

Erregung, Valenz, Potenz) zu Clustern zusammengefasst. Die Cluster sollten in sich möglichst homogen und untereinander heterogen sein. Die Clusterbildung erfolgte sukzessive in Schritten (hierarchisch), wobei z.B. im 1. Schritt die am besten zueinanderpassenden Farbtöne zu einem Cluster zusammengefasst wurden usw. Je später eine Fusion der Farbtöne erfolgte, desto schlechter passten die fusionierten Farbtöne des Clusters zusammen. Wie gut die Items (Farbtöne) zusammenpassen, wurde durch die „Fehlerquadratsumme“ (Summe über die Streuung der einzelnen Farbtöne im Cluster) festgestellt. Die Clusterbildung wurde abgebrochen, wenn die Fehlerquadratsumme „sprunghaft anstieg“ (z.B. um Faktor 10). Am Ende der Clusterbildung sollten möglichst alle Farbtöne einem Cluster zugeordnet werden können - und nur möglichst wenige einzelne Farbtöne übrig bleiben.

### **Ergebnisse der Clusteranalyse (14-Farbensatz KFPT)**

Cluster 1	▶	Farbtöne 1a/4a/6a/10b
Cluster 2	▶	Farbtöne 6b/8b/8c/9b/10a
Cluster 3	▶	Farbtöne 1c/5c
Einzelfarbe 1	▶	Farbton 1b
Einzelfarbe 2	▶	Farbton 5b
Einzelfarbe 3	▶	Farbton 10c

Die Ergebnisse der Clusteranalyse bestätigen die Einteilung der Farbpalette in a) helle, b) gesättigte und c) dunkle Farbtöne nur bedingt. So ist z.B. der Farbton 10b (Cluster 1) eher den hellen als den gesättigten Farbtönen zuzuordnen. Das aus fünf Farbtönen bestehende Cluster 2 setzt sich heterogen aus einem hellen (10a), drei gesättigten (6b, 8b und 9b) und einem dunklen Farbton (8c) zusammen. Lediglich das Cluster 3 besteht aus zwei einheitlichen Farbtönen – den dunklen Farbtönen 1c, 5c. Hinsichtlich der verbleibenden Einzelfarben (1b, 5b, 10c) ist anzumerken, dass kein heller Farbton als prägnante Einzelfarbe isoliert wurde.

Ergänzend zur Clusterbildung wurden Clusterprofile erstellt. Das Clusterprofil spiegelt die für die Clusteranalyse relevanten Merkmale (mittlere Dimensionswerte für Erregung, Valenz, Potenz im Eindrucksdifferential) innerhalb eines Clusters wider. So ist z.B. die Standardabweichung neben den Extrema ein Maß dafür, wie gut die Farbtöne des Clusters beim jeweiligen Merkmal übereinstimmen. Ein Beispiel für eine gelungene Clusterbildung ist Cluster 2: In allen ED-Dimensionen zeigt sich eine geringe Streuung und Minima/Maxima liegen eng um den Mittelwert.

Cluster / Farbtöne	ED-Dimension	M	S	Min	Max	N
Einzelfarbe 10c	Erregung	3,6140	,	3,614	3,614	1
	Valenz	4,7590	,	4,759	4,759	1
	Potenz	6,5166	,	6,517	6,517	1
Cluster 3 1c, 5c	Erregung	4,1623	,4264	3,671	4,428	2
	Valenz	3,3744	,4422	2,864	3,638	2
	Potenz	4,6700	1,1416	3,354	5,394	2
Cluster 1 1a, 4a, 6a, 10b	Erregung	3,1365	,4357	2,853	3,638	4
	Valenz	5,0504	,7828	4,412	5,924	4
	Potenz	2,9047	,4438	2,601	3,414	4
Cluster 2 6b, 8b, 9b, 8c, 10a	Erregung	4,1558	,3530	3,742	4,572	5
	Valenz	5,6922	,3141	5,315	6,037	5
	Potenz	4,6330	,5365	4,325	5,581	5
Einzelfarbe 1b	Erregung	6,0623	,	6,062	6,062	1
	Valenz	5,9525	,	5,953	5,953	1
	Potenz	6,1224	,	6,122	6,122	1
Einzelfarbe 5b	Erregung	5,6401	,	5,640	5,640	1
	Valenz	4,4019	,	4,402	4,402	1
	Potenz	5,3248	,	5,325	5,325	1

Tab. 6: Profil der Clusteranalyse

Die Profile charakterisieren die Cluster bzw. Einzelfarbtöne wie folgt: Die Cluster 1 und 2 sowie der Farbton 5b (gelb) zeigen eine vorrangige Valenzwirkung; Cluster 3 und insbesondere der Farbton 10c (schwarz) werden als primär kraftvoll eingeschätzt. Der Farbton 1b (rot) nimmt eine Sonderstellung ein, da er fast gleichermaßen als erregend, angenehm und kraftvoll erlebt wird („Superfarbe rot“).

### 6.6.2 Determinationstest (DT)

Als Zielvariablen im DT wurden definiert:

- Median der Reaktionszeit (d.h. die Zeitspanne zwischen der Darbietung eines Reizes und der Reaktion durch den Probanden)
- Anzahl richtiger Reaktionen
- Anzahl falscher Reaktionen

- Anzahl ausgelassener Reaktionen
- Anzahl der bearbeiteten Reize.

## **6.7 Analyse der Arzneimittelsicherheit**

### **6.7.1 Analyse unerwünschter Ereignisse (Adverse Events)**

Im Verlauf der Studie aufgetretene bzw. sich verschlechternde medizinische Befunde wurden bei jedem Untersuchungstermin erfasst und als unerwünschte Ereignisse (Adverse Events) im CRF (Case Report Form) dokumentiert. Berichtete unerwünschte Ereignisse wurden im Rahmen der statistischen Analyse gelistet. Absolute und relative Häufigkeiten der Symptome, Zuordnung zu Organsystemklassen, Schweregrad und mögliche Beziehung zur Studienmedikation, Vergleiche zwischen den drei verabreichten Medikationen bzgl. Nebenwirkungen, wurden tabellarisch dargestellt.

### **6.7.2 Analyse des klinischen Labors**

Laborwerte wurden beschrieben und interpretiert als Abweichungen vom Referenzbereich (Normalwerte eines Parameters) vor und während der Einnahme der Medikation. Bezogen auf die jeweilige Behandlungsgruppe wurden sämtliche untersuchten Laborparameter tabellarisch dargestellt.

## **7 ERGEBNISSE ZUR STUDIENPOPULATION**

### **7.1 Abweichungen vom Untersuchungsverlauf**

Die Studie wurde gemäß den Vorgaben des Untersuchungsplans durchgeführt. Modifikationen des geplanten Untersuchungsverlaufs waren nicht erforderlich.

### **7.2 Fehlende Angaben**

Die KFPT-Testergebnisse von drei Probanden (Nr. 1, 2, 3) fehlten bei allen Untersuchungszeitpunkten. Aufgrund technischer Probleme zu Studienbeginn war eine Speicherung der Baseline-Testergebnisse bei diesen ersten drei Probanden nicht möglich. Die Durchführung der Folgetestungen erübrigte sich deshalb. Lediglich das Eindrucksdifferential wurde bei diesen drei Probanden durchgängig erhoben. Für die KFPT-Analyse lagen somit Daten von 96 Probanden (Studienteil 1) bzw. 88 Probanden (Studienteil 2) vor.

Auswertbare Daten für das semantische Eindrucksdifferential (ED) lagen für beide Studienteile komplett vor (ST1:N=99/ST2:N=91). Vor Rückgabe der bearbeiteten Erfassungsbögen wurde vom Versuchsleiter darauf geachtet, dass die Bögen vollständig ausgefüllt waren – unbeantwortete Items wurden sofort nachträglich bearbeitet.

Die Testergebnisse des Determinationstests (DT) sowie die individuelle Studierendokumentation (CRF=Case Report Form) waren für jeden Probanden gleichfalls vollständig verfügbar (ST1:N=99/ST2:N=91).

### **7.3 Randomisierung und Behandlungscompliance**

Die Randomisierung wurde durchgehend eingehalten. Kein Notfallcouvert wurde geöffnet. Die Verblindung war für Proband und Untersucher während des gesamten Untersuchungszeitraumes sichergestellt. Für die Auswertung wurde die Probanden-Nummer beibehalten, als Medikations-Code wurde derjenige der erhaltenen Medikation eingegeben. Die Restmedikation wurde kontrolliert und gezählt (Pill-Counting).

## 7.4 Verifizierung der Ein- und Ausschlußkriterien

Bei der Aufnahme wurden neben einer ausführlichen Anamnese eine ärztliche Eingangsuntersuchung (Ganzkörperstatus) durchgeführt. Pathologische Befunde der körperlich-neurologischen Untersuchung führten zum Ausschluß von der Studienteilnahme. Die Screening-Befunde sämtlicher Studienteilnehmer wiesen keine besonderen Auffälligkeiten auf und lagen meist im Normbereich – eine Ausnahme bildete lediglich die Sehleistung: 61,8% der Probanden benötigten eine Sehhilfe.

Die gleichzeitige Einnahme von Psychopharmaka oder von Arzneimitteln, welche das Schlafverhalten oder Reaktionsvermögen beeinflussen - in den vier Wochen vor Beginn oder während der Teilnahme an der Studie - waren laut Studienplan untersagt und führten zum Ausschluß des Probanden. Alle anderen Begleittherapien mussten seit mindestens vier Wochen bestehen und durften während der Studienteilnahme möglichst nicht verändert werden (steady state). Keiner der Probanden wies bei der Aufnahmeuntersuchung eine unerlaubte Begleiterkrankung auf oder nahm eine unzulässige Begleitmedikation ein. Im Studienteil 1 wurde eine Probandin aufgrund der Einnahme eines tricyclischen Antidepressivums von der weiteren Studienteilnahme ausgeschlossen. Für Studienteil 2 mussten fünf Probanden wegen akuter Virusinfektionen und gleichzeitiger Einnahme unerlaubter Begleitmedikation (Protokollverletzung) aus dem ATP-Kollektiv ausgeschlossen werden.

Führte das diagnostische Interview (DIPS) bei der Voruntersuchung zu einem pathologischen Befund, führte dies zur Nichtberücksichtigung des Probanden. Dies war lediglich einmal der Fall. Bei 102 eingeschlossenen Probanden des ITT-Kollektives zeigten sich bei 101 Probanden keine psychiatrischen Auffälligkeiten.

Das Drogenscreening mittels Drogenschnelltest TRIAGE™ (immunologischer Test, mit dem die acht häufigsten Drogen und missbräuchlich verwendeten Medikamente bzw. deren Hauptmetaboliten im Urin nachgewiesen werden können) wurde vor Aufnahme des Probanden und vor der Abschlußuntersuchung durchgeführt. Probanden mit positivem Befund wurden nicht in die Studie eingebunden bzw. die abschließenden Tests wurden nicht mehr durchgeführt. Bis auf drei Ausnahmen war das Ergebnis des Drogenscreenings – zu beiden Untersuchungszeitpunkten – bei sämtlichen Studienteilnehmern negativ. Zwei Probanden wurden aufgrund des positiven Testbefundes bei der Voruntersuchung nicht in die Studie aufgenommen. Bei einem Probanden erbrachte das bei der Aufnahme durchgeführte Drogenscreening – nach vorheriger Einnahme eines codeinhaltigen Hustensaftes – ein positives Testergebnis. Nach Absetzen des Hustensaftes und erneutem Drogenscreening mit unauffälligem Befund wurde dieser Proband dann vierzehn Tage später in die Studie eingeschlossen.

Da alkoholisierte Probanden keine verwertbaren Studienergebnisse generieren, wurden beim Screening nur solche Probanden eingeschlossen, deren Mess-

ergebnisse keine Alkoholisierung ergab. Das Ergebnis der AAK-Doppelmessung (Siemens ALCOMAT Infrarot-Analysengerät) durfte >0.00 mg/l nicht überschreiten. Zu den Folgeuntersuchungen wurde bei sämtlichen eingeschlossenen Probanden – vor Durchführungen der psychometrischen Verfahren – eine mögliche Alkoholisierung überprüft. Wurde dabei eine Alkoholisierung festgestellt, schied der Proband unverzüglich aus der Studie aus; die psychometrischen Tests wurden nicht mehr durchgeführt.

Mittels des ATV wurden beim Screening Einstellung und Umgang mit Alkohol evaluiert. Der speziell für alkoholabhängige Kraftfahrer konzipierte Test wird seit langem erfolgreich in der verkehrspsychologischen Begutachtung eingesetzt und ist auch im Rahmen von Arzneimittelstudien verwendbar. Auffällige Probanden wurden bei der Voruntersuchung von der Studienteilnahme ausgeschlossen.

AAK-Doppelmessungen und ATV-Testergebnisse erbrachten bei keinem Studienteilnehmer Hinweise auf eine Alkoholisierung. Kein Proband musste von der Studienteilnahme bzw. den Folgeuntersuchungen ausgeschlossen werden.

## **7.5 Analyisierte Populationen**

Insgesamt lagen 111 Dokumentationen zur Auswertung vor. Neun Probanden wurden lediglich voruntersucht (gescreent). Aufgrund von Verstößen bezüglich der Ein- oder Ausschlusskriterien wurden diese Probanden von der Studienteilnahme ausgeschlossen.

Insgesamt 102 Probanden wurden in die Studie aufgenommen (ITT-Kollektiv). Keiner dieser Probanden wies bei der Aufnahme eine unerlaubte Begleiterkrankung oder –medikation auf. Sämtliche Einschlußkriterien wurden erfüllt und keines der Ausschlusskriterien wurde verletzt.

Insgesamt vier Probanden brachen die Studie vorzeitig ab (Studienteil 1: 3 Probanden/Studienteil 2: 1 Proband). Da nicht für alle Kontrollzeitpunkte eines Studienteils Messwerte vorlagen, wurden die Dokumentationen dieser vier Probanden für die jeweiligen ATP-Analysen nicht berücksichtigt.

### *Kollektiv Studienteil 1*

Für die ATP-Auswertung des Studienteils 1 (N=99) konnten 33 Probanden der Gruppe "Baldrian-Extrakt", 34 Probanden der Placebo- und 32 Probanden der Flunitrazepam-Gruppe herangezogen werden. Insgesamt drei Probanden (ein Proband unter Baldrian-Extrakt und 2 Probanden unter Flunitrazepam) wurden von der ATP-Analyse ausgeschlossen.

### *Kollektiv Studienteil 2*

Im ATP-Kollektiv des Studienteils 2 (N=91) wurden die Daten von 47 Probanden der Baldrian-Extrakt-Gruppe und 44 Probanden der Placebo-Gruppe berück-

sichtigt. Die Daten von 11 Probanden blieben aufgrund von Protokollverstößen (10 Probanden) bzw. vorzeitigem Behandlungsende (1 Proband) bei der ATP-Auswertung unberücksichtigt.

<b>Kollektiv</b>	Baldrian	Placebo	Flunitrazepam + Baldrian	Flunitrazepam + Placebo
<b>Gesamt (N=102)</b>	34 33,33%	34 33,33%	17 16,67%	17 16,67%
<b>Studienteil 1 (N=99)</b>	Baldrian	Placebo	Flunitrazepam	
	33 33,33%	34 34,34%	32 32,32%	
<b>Studienteil 2 (N=91)</b>	Baldrian	Placebo	Flunitrazepam	
	47 51,65%	44 48,35%	0 0%	

Tab. 7: Probandenkollektive

#### *Vergleichbarkeit der Untersuchungsgruppen in der Ausgangslage*

Die Homogenität des Gesamtkollektivs (N=102) sowie der ATP-Kollektive (ST1:N=99/ST2:N=91) wurde mit dem Mann-Whitney-U-Test überprüft. Die berechneten p-Werte zeigen keine signifikanten Unterschiede zwischen den Behandlungsgruppen.

Lediglich im Studienteil 1 bestanden beim medizinischen Sicherheitsparameter „Puls“ sowohl im ITT- als auch im ATP-Kollektiv Unterschiede in der Ausgangslage zwischen den Teilstichproben Flunitrazepam/Baldrian-Extrakt (ITT\_p=0.0309/ATP\_p=0.0341) und Flunitrazepam/Placebo (ITT\_p=0.0439/ATP\_p=0.0784). Die Baselinedifferenzen hinsichtlich der Herzfrequenz sind jedoch für die Aussagekraft der psychometrischen Ergebnisse ohne Relevanz.

Zur Baseline war beim psychometrischen Leistungstest „DT“ in der Hauptzielvariablen (Median der Reaktionszeit) ein Trend im Vergleich der Behandlungsgruppen Flunitrazepam vs. Baldrian-Extrakt festzustellen (p=0.0754+). Nach Rückfrage bei der Herstellerfirma des Wiener Testsystems (Dr. Schuhfried GmbH in Mödling, Österreich), stellt dieser Trend die Vergleichbarkeit der Medikationsgruppen keinesfalls in Frage. Der gefundene Trend ist psychologisch irrelevant und hatte keinen Einfluss auf die Untersuchungsergebnisse.

Die Homogenitätsprüfung mittels Mann-Whitney-U-Test belegt für das Gesamtkollektiv (N=102) bzgl. Alter und Geschlechtsverteilung keine signifikanten Unterschiede zwischen den Untersuchungsgruppen.

Um die Verteilung der Geschlechter auf die Medikationsgruppen auf Homogenität zu prüfen, wurde der Chi-Quadrat-Test verwendet. Sowohl für das ITT-Kollektiv als auch für das ATP-Kollektiv wurden keine beachtenswerten Unterschiede objektiviert.

<b>Gesamtkollektiv (N=102)</b>	<b>MANN-WHITNEY-U-Test (p-Werte für den Vergleich der Medikationsgruppen)</b>			
	Studienteil 1			Studienteil 2
	Baldrian vs. Placebo	Baldrian vs. Flunitrazepam	Placebo vs. Flunitrazepam	Baldrian vs. Placebo
Alter (Jahre)	0.5348	0.8538	0.3798	0.3656
Größe (cm)	0.5512	0.8203	0.7353	0.4428
Gewicht (kg)	0.2722	0.4077	0.8540	0.1930
Syst. RR (mmHg)	0.7036	0.2694	0.1600	0.6156
Diast. RR (mmHg)	0.3791	0.5843	0.2303	0.2846
Puls (Schläge/min)	0.6984	0.0309	0.0439	0.7959
<b>Auswertekollektive (ST1:N=99/ST2:N=91)</b>				
Alter (Jahre)	0.6330	0.6550	0.3256	0.2440
Größe (cm)	0.4778	0.9948	0.4518	0.5351
Gewicht (kg)	0.2511	0.5593	0.6533	0.2818
Syst. RR (mmHg)	0.6290	0.2644	0.1234	0.2094
Diast. RR (mmHg)	0.3423	0.4653	0.1604	0.1836
Puls (Schläge/min)	0.5973	0.0341	0.0784	0.6124
DT	0.8360	0.0754	0.1638	0.7357
<b>CHI-QUADRAT-Test (p-Werte)</b>				
<b>Geschlecht Gesamtkollektiv (N=102)</b>	Studienteil 1			Studienteil 2
	0.3660			0.2289
<b>Geschlecht Auswertekollektive (ST1:N=99/ST2:N=91)</b>	0.3186			0.4366

Tab. 8: Homogenitätsprüfung

## 7.6 Soziodemographische Charakteristika

Der Anteil 30-39jähriger Studienteilnehmer beträgt im ITT-Kollektiv 52,0%, der Anteil 40-49jähriger Probanden 24,5% und der Anteil älterer Teilnehmer (50-60 Jahre) umfasst 23,6%. Die Dekaden sind mit einer ausreichenden Gruppenstärke besetzt. Jüngere Probanden stellen zwar die Hälfte der Teilnehmer, doch ist diese Altersgruppe nicht überrepräsentiert.

Altersgruppe	Baldrian	Placebo	Flunitrazepam + Baldrian	Flunitrazepam + Placebo	Gesamt
30-39 Jahre	18 52,9%	16 47,1%	11 64,7%	8 47,1%	53 52,0%
40-49 Jahre	8 23,5%	7 20,6%	4 23,5%	6 35,3%	25 24,5%
50-59 Jahre	8 23,5%	9 26,5%	2 11,8%	3 17,6%	22 21,6%
60 Jahre	0 0%	2 5,9%	0 0%	0 0%	2 2,0%
Gesamt	34	34	17	17	102

Tab. 9: Medikation und Altersgruppen (Gesamtkollektiv -N=102)

Bezogen auf das ITT-Kollektiv (N=102) betrug der Anteil weiblicher Probanden 57,8%, der Anteil männlicher Probanden lag bei 42,2% - kein Geschlecht war deutlich überproportional vertreten, ein geschlechtsspezifischer Bias der Ergebnisse damit eher unwahrscheinlich.

Altersgruppe	Baldrian	Placebo	Flunitrazepam + Baldrian	Flunitrazepam + Placebo	Gesamt
männlich	16 47,1%	11 32,4%	9 52,9%	7 41,2%	43 42,2%
weiblich	18 52,9%	23 67,6%	8 47,1%	10 58,8%	59 57,8%
Gesamt	34	34	17	17	102

Tab. 10: Medikation und Geschlecht (Gesamtkollektiv - N=102)

## **8 ERGEBNISSE ZU DEN PSYCHOMETRISCHEN VERFAHREN UND SCHLUSSFOLGERUNGEN**

Aus Gründen der besseren Übersichtlichkeit wird im Rahmen der Ergebnisdarstellung auf eine Präsentation sämtlicher Resultate bzw. generierter Tabellen und Aufstellungen verzichtet.

Der interessierte Leser wird auf den separaten Materialienband verwiesen. Dem Materialienband sind zu beiden Studienteilen folgende Informationen und ausführliche Dokumentationen zu entnehmen:

- Angaben zur Studienpopulation
- KFPT- Ergebnisse  
(Tabellen und p-Werte der Cluster und Einzelfarben sowie Korrelationen zwischen Valenzwert im ED und Wahlhäufigkeit im KFPT)
- ED-Resultate  
(Tabellen und p-Werte nach Geschlecht, Altersklassen und Medikationsgruppen)
- DT-Ergebnisse incl. p-Werte.

## 8.1 Primärer Endpunkt

### 8.1.1 Kölner Farbpräferenztest (KFPT)

#### Baseline

In der Ausgangslage des Gesamtkollektivs wurden für die schönen/hässlichen Pyramiden folgende Farbwahlhäufigkeiten der Cluster/Einzelfarben evaluiert. Bei der Farbwahl für die schönen Pyramiden wurde das Cluster 2 bevorzugt (M=4,37), gefolgt vom roten Farbton 1b (M=3,95). Bei der Wahl der hässlichen Pyramiden wurde der Farbton 10c (schwarz) eindeutig präferiert (M=9,13) – die restlichen Cluster/Einzelfarben folgten mit deutlichem Messwertabständen, angeführt von Cluster 3 (M=4,28). Bei Betrachtung der Ergebnisse ist die mitunter große Streuung beachtenswert. So ist z.B. die Standardabweichung bei den Farbtönen 1b und 5b größer als das arithmetische Mittel dieser Farbtöne. Die Spannweite der Extrema war gleichfalls zwischen den Clustern bzw. Farbtönen sehr unterschiedlich. Minimum und Maximum lagen bei Cluster 2 am engsten zusammen und bei Farbton 10c am weitesten auseinander. Aufgrund dieser großen Varianz der Messwerte sind Generalisierungen zum Farbwahlverhalten nur sehr eingeschränkt möglich.

Cluster/Einzelfarbtone	M	S	Min	Max	N
Cluster 1 (1a/4a/6a/10b)					
Schöne Pyramiden	1,95	1,97	0	9	96
Hässliche Pyramiden	3,45	2,37	0	10	96
Cluster 2 (6b/8b/9b/8c/10a)					
Schöne Pyramiden	4,37	1,57	0	8	96
Hässliche Pyramiden	2,08	1,39	0	6	96
Cluster 3 (1c/5c)					
Schöne Pyramiden	2,75	2,74	0	11	96
Hässliche Pyramiden	4,28	2,83	0	11	96
Farbton 1b					
Schöne Pyramiden	3,95	4,24	0	17	96
Hässliche Pyramiden	1,63	2,06	0	9	96
Farbton 5b					
Schöne Pyramiden	3,00	3,92	0	20	96
Hässliche Pyramiden	1,51	2,61	0	17	96
Farbton 10c					
Schöne Pyramiden	2,91	4,14	0	16	96
Hässliche Pyramiden	9,13	7,37	0	28	96

Tab. 11: Häufigkeit der Cluster/Einzelfarbtöne in schönen und hässlichen Pyramiden (Baseline)

## Studienteil 1

### *Cluster 1*

Schöne Pyramiden:

Die Baseline-Messwerte der drei Medikationsgruppen waren hinsichtlich der Parameter Median, Streuung und Extrema vergleichbar. Bei Retest 1 zeigten sich in der Gegenüberstellung der drei Behandlungsgruppen gleichfalls keine statistisch zu beachtenden Gruppenunterschiede.

Auch innerhalb der Medikationsgruppen ergaben sich keine statistisch signifikanten Messwertveränderungen von der Baseline zur Folgetestung.

Hässliche Pyramiden:

In der Ausgangslage zeigte sich ein signifikanter Unterschied zwischen den Behandlungsgruppen Baldrian-Extrakt und Flunitrazepam ( $p=0.024^*$ ), der bei Retest 1 abgeschwächt bestehen blieb ( $p=0.084^+$ ). Obwohl unter Baldrian-Extrakt bei der Folgetestung eine signifikante Zunahme ( $p=0.019^*$ ) der Wahlhäufigkeit von Cluster 1 gemessen wurde, bestanden die Gruppenunterschiede fort.

In beiden Testdurchführungen wurde das Cluster 1 in der Flunitrazepamgruppe häufiger gewählt (Baseline  $M=4.13$  / Retest 1  $M= 4.16$ ) als in der Baldrian-Extrakt-Gruppe (Baseline:  $M= 2.87$  / Retest 1:  $M= 3.48$ ). Die anderen Gruppenvergleiche erbrachten sowohl in der Baseline als auch bei Retest 1 statistisch signifikante und klinisch relevante Unterschiede.

Innerhalb der Medikationsgruppen Placebo und Flunitrazepam wurden bei Retest 1 keine beachtenswerten Veränderungen gegenüber der Ausgangslage gemessen. Lediglich unter Baldrian-Extrakt wurde eine signifikante Zunahme ( $p=0.019^*$ ) der Wahlhäufigkeit ( $M=0.62$ ) von Cluster 1 evaluiert.

<b>Stichprobe</b>	<b>Zeitpunkt</b>	<b>M</b>	<b>S</b>	<b>Min</b>	<b>Max</b>	<b>N</b>
Gesamtkollektiv	Baseline	3,45	2,37	0	10	96
	Retest 1	3,72	2,22	0	10	96
	Abnahme/Zunahme	,27	1,82	-5	5	96
Baldrian	Baseline	2,87	2,03	0	8	32
	Retest 1	3,48	2,04	0	8	32
	Abnahme/Zunahme	,62	1,68	-5	4	32
Placebo	Baseline	3,38	2,62	0	10	33
	Retest 1	3,53	2,27	0	8	33
	Abnahme/Zunahme	,15	1,87	-5	5	33
Flunitrazepam	Baseline	4,13	2,32	1	10	31
	Retest 1	4,16	2,36	1	10	31
	Abnahme/Zunahme	,03	1,90	-4	5	31

Tab. 12: Häufigkeit des Clusters 1 (1a/4a/6a/10b) in häßlichen Pyramiden (Studienteil 1)

## Cluster 2

Schöne Pyramiden:

Hinsichtlich der Häufigkeit des Clusters 2 ergaben sich in der Ausgangslage keine beachtenswerten Unterschiede zwischen den drei Behandlungsgruppen.

Bei Retest 1 führte die signifikant häufigere Wahlhäufigkeit von Flunitrazepam ( $M=0,72$  /  $p=0.016^*$ ) zu statistisch relevanten Unterschieden im Vergleich zu Baldrian-Extrakt ( $p=0.023^*$ ) und Placebo ( $p=0.040^*$ ).

Zwischen Baldrian-Extrakt und Placebo bestanden aufgrund der Messwertveränderungen in Relation zur Baseline bei Retest 1 weiterhin keine bedeutenden Unterschiede.

Stichprobe	Zeitpunkt	M	S	Min	Max	N
Gesamtkollektiv	Baseline	4,37	1,57	0	8	96
	Retest 1	4,50	1,67	0	9	96
	Abnahme/Zunahme	,14	1,63	-4	5	96
Baldrian	Baseline	4,61	1,34	2	7	32
	Retest 1	4,45	1,66	0	8	32
	Abnahme/Zunahme	-,16	1,33	-4	2	32
Placebo	Baseline	4,43	1,93	0	8	33
	Retest 1	4,30	1,80	1	8	33
	Abnahme/Zunahme	-,13	1,76	-4	3	33
Flunitrazepam	Baseline	4,05	1,33	1	6	31
	Retest 1	4,77	1,56	2	9	31
	Abnahme/Zunahme	,72	1,65	-4	5	31

Tab. 13: Häufigkeit des Clusters 2 (6b/8b/9b/8c/10a) in schönen Pyramiden (Studienteil 1)

Hässliche Pyramiden:

Der Gruppenvergleich zur Baseline erbrachte keine relevanten Unterschiede in der Ausgangslage. Die statistischen Parameter lagen gruppenübergreifend eng zusammen. Aufgrund der geringfügigen Messwertveränderungen bei Retest 1 wurde auch bei der Folgetestung keine Gruppenunterschiede deutlich.

Innerhalb der drei Medikationen zeigten sich bei Retest 1 im Vergleich zur Baseline ebenfalls keine ausgeprägten Veränderungen.

## Cluster 3

Schöne Pyramiden:

Der Gruppenvergleich zeigte in der Ausgangslage keine zu beachtenden Gruppenunterschiede. Zwischen Placebo und Flunitrazepam bestand ein geringer Unterschied im Farbwahlverhalten (Placebo  $M=2,39$  / Flunitrazepam  $M=3,21$ ), der jedoch statistisch zu vernachlässigen ( $p=0.128$ ) und auch klinisch irrelevant war. Bei der Folgetestung nivellierten sich diese geringfügigen Gruppenunterschiede. Insgesamt waren bei Retest 1 keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen zu objektivieren.

Bei Retest 1 wurde Baldrian-Extrakt im Vergleich zu Placebo und Flunitrazepam zwar etwas häufiger gewählt (Trends für die Vergleiche Baldrian-Extrakt vs. Placebo  $p=0.072+$  und Baldrian-Extrakt vs. Flunitrazepam  $p=0.082+$ ), doch waren diese Messwertveränderungen zu gering ( $p=0.135$ ), um von Relevanz für das Farbwahlverhalten zu sein.

Hässliche Pyramiden:

In der Ausgangslage und bei Retest 1 zeigten sich keine aussagekräftigen Unterschiede zwischen den drei Therapiegruppen. Die Messwertveränderungen bei Retest 1 waren insgesamt gering ausgeprägt. Auch innerhalb der Medikationen wurden keine statistisch relevanten Veränderungen deutlich.

#### *Farbton 1b*

Schöne Pyramiden:

Sowohl bei Baseline als auch bei Retest 1 wurden keine statistisch bedeutsamen Gruppenunterschiede evaluiert, wobei bei der Folgetestung lediglich unter Placebo eine ausgeprägte Veränderung gemessen wurde (Zunahme der Wahlhäufigkeit um  $M=1,12$  Punkte), die jedoch keine statistische Signifikanz erreichte. Innerhalb der drei Medikationsgruppen wurden bei Retest 1 keine psychologisch relevanten Veränderungen gegenüber der Baseline festgestellt.

Hässliche Pyramiden:

Die Gruppenvergleiche erbrachten keine nennenswerten Unterschiede in der Ausgangslage. Bei Retest 1 zeigten sich gleichfalls keine signifikanten Unterschiede zwischen den Therapiegruppen. Lediglich statistisch und klinisch irrelevante Unterschiede für die Vergleiche Baldrian-Extrakt vs. Placebo ( $p=0.151$ ) und Baldrian-Extrakt vs. Flunitrazepam ( $p=0.194$ ) wurden evaluiert. Innerhalb der Medikationen zeigte sich bei Retest 1 für Baldrian-Extrakt ein statistischer Trend ( $p=0.063+$ ), der durch die geringe Abnahme der Wahlhäufigkeit von Farbton 1b ( $M=-0.53$ ) bedingt war. Unter Baldrian-Extrakt und Placebo traten bei der Folgetestung keine bedeutsamen Messwertveränderungen gegenüber den Ausgangsbefunden auf.

#### *Farbton 5b*

Schöne Pyramiden:

Im Baseline-Vergleich zwischen Placebo und Flunitrazepam ergab sich ein statistischer Trend ( $p=0.074+$ ), da in der Flunitrazepamgruppe der Farbton 5b in der Ausgangslage häufiger gewählt wurde als in der Placebogruppe (Flunitrazepam  $M=3,42$  / Placebo  $M=2,64$ ). Zwischen Baldrian-Extrakt und Flunitrazepam bestand ein marginaler und statistisch zu vernachlässigender Unterschied ( $p=0.157$ ). Der Baseline-Vergleich zwischen Baldrian-Extrakt und Placebo zeigte keine statistischen Unterschiede.

Bei Retest 1 wurden zwischen den drei Behandlungsgruppen keine aussagekräftigen Gruppenunterschiede evaluiert. Aufgrund einer mäßigen Zunahme der Wahlhäufigkeit unter Placebo ( $M=0,55$ ) und einer gegenläufigen Entwicklung unter Flunitrazepam ( $M=-,55$ ) nivellierten sich bei Retest 1 die in der Baseline evaluierten Gruppenunterschiede beider Behandlungsgruppen. Der marginale Baseline-Unterschied zwischen Baldrian-Extrakt und Flunitrazepam ( $p=0.157$ ) blieb in Folge geringer Messwertvariationen bei Retest 1 nahezu unverändert bestehen ( $p=0.167$ ).

Innerhalb der drei Medikationen zeigten sich bei Retest 1 – keine relevanten Messwertveränderungen im Vergleich zur Baseline. Unter Flunitrazepam wurde der Farbton 5b bei Retest 1 zwar geringfügig häufiger gewählt als zur Baseline, doch ist die Zunahme der Wahlhäufigkeit statistisch irrelevant ( $p= 0.142$ ) und zu vernachlässigen.

Hässliche Pyramiden:

Die Gruppenvergleiche erbrachten weder in der Ausgangslage noch aufgrund der geringen Messwertveränderungen bei Retest 1 statistisch relevante Unterschiede. Auch innerhalb der Medikationen traten keine bedeutsamen Veränderungen von der Baseline zum Retest 1 auf.

#### *Farbton 10c*

Schöne Pyramiden:

Die Gegenüberstellung der Medikationen erbrachte sowohl in der Ausgangslage und in Folge der zu vernachlässigenden Messwertvariationen auch bei Retest 1 keine statistischen Unterschiede. Innerhalb der Medikationsgruppen traten unter Baldrian-Extrakt und Placebo bei der Folgetestung gegenüber der Baseline keine beachtenswerten Veränderungen im Farbwahlverhalten auf. Lediglich unter Flunitrazepam war bei Retest 1 eine Abnahme der Wahlhäufigkeit des Farbtönen 10c zu beobachten ( $M=-1.10$  / Trend  $p=0.061+$ ). Diese Abnahme war jedoch zu schwach ausgeprägt, um klinische Relevanz zu erlangen.

Hässliche Pyramiden:

In der Ausgangslage zeigten sich mäßig ausgeprägte Differenzen im Farbwahlverhalten der Teilstichproben, die jedoch statistisch zu vernachlässigen waren (Baldrian-Extrakt vs. Placebo  $p=0.168$ ; Baldrian-Extrakt vs. Flunitrazepam  $p=0.105$ ). Bei Retest 1 bestanden die geringen Gruppenunterschiede fort (Baldrian-Extrakt vs. Placebo  $p=0.166$ ; Baldrian-Extrakt vs. Flunitrazepam  $p=0.180$ ). Insgesamt zeigten sich bei Retest 1 zwischen den drei Medikationen in Relation zur Baseline keine bedeutsamen Messwertveränderungen. Unter Baldrian-Extrakt war lediglich eine mäßige ( $M=-1,28$ ) und statistisch irrelevante Abnahme ( $p=0.118$ ) der Wahlhäufigkeit bei Retest 1 zu beobachten.

Für die Berechnung der p-Werte wurden innerhalb der Medikationsgruppen für jeden Parameter die Zeitpunkte Baseline und Retest 1 mittels Wilcoxon-Test miteinander verglichen. Die drei Behandlungsgruppen wurden anhand der Parameter zur Baseline mit dem Mann-Whitney-U-Test, zum Retest 1 mit dem Paardifferenzen-U-Test (Differenz zur Baseline) verglichen.

Insgesamt betrachtet zeigten die p-Werte der Cluster bzw. Einzelfarben bei schönen Pyramiden keine bedeutsamen Gruppenunterschiede, auch die Messwertveränderungen innerhalb der Behandlungsgruppen waren meist gering ausgeprägt. Lediglich bei Cluster 2 kam es nach vorheriger Einnahme von Flunitrazepam zu signifikanten Veränderungen bei Retest 1, die auch im Gruppenvergleich zu den beiden an deren Medikationen statistisch abgesichert sind.

Cluster/ Einzelfarbe	Zeitpunkt	Baldrian	Placebo	Flunitra.	Baldrian vs Placebo	Baldrian vs Flunitra.	Placebo vs Flunitra.
Cluster 1	Baseline	-	-	-	n.s.	n.s.	n.s.
	Retest 1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Cluster 2	Baseline	-	-	-	n.s.	n.s.	n.s.
	Retest 1	n.s.	n.s.	0.016*	n.s.	0.023*	0.040*
Cluster 3	Baseline	-	-	-	n.s.	n.s.	0.128
	Retest 1	0.135	n.s.	n.s.	0.072+	0.082+	n.s.
1b	Baseline	-	-	-	n.s.	n.s.	n.s.
	Retest 1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
5b	Baseline	-	-	-	n.s.	0.157	0.074+
	Retest 1	n.s.	n.s.	0.142	n.s.	0.167	n.s.
10c	Baseline	-	-	-	n.s.	n.s.	n.s.
	Retest 1	n.s.	n.s.	0.061+	n.s.	n.s.	n.s.

Tab. 14: p-Werte der KFPT-Cluster/Einzelfarben - schöne Pyramiden (Studienteil 1)

Cluster/ Einzelfarbe	Zeitpunkt	Baldrian	Placebo	Flunitra.	Baldrian vs Placebo	Baldrian vs Flunitra.	Placebo vs Flunitra.
Cluster 1	Baseline	-	-	-	n.s.	0.024*	0.192
	Retest 1	0.019*	n.s.	n.s.	n.s.	0.084+	n.s.
Cluster 2	Baseline	-	-	-	n.s.	n.s.	n.s.
	Retest 1	0.166	n.s.	n.s.	n.s.	n.s.	n.s.
Cluster 3	Baseline	-	-	-	0.153	n.s.	0.120
	Retest 1	n.s.	n.s.	n.s.	0.116	n.s.	n.s.
1b	Baseline	-	-	-	n.s.	n.s.	n.s.
	Retest 1	0.063+	n.s.	n.s.	0.151	0.194	n.s.
5b	Baseline	-	-	-	n.s.	n.s.	n.s.
	Retest 1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
10c	Baseline	-	-	-	0.168	0.105	n.s.
	Retest 1	0.188	n.s.	n.s.	0.166	0.180	n.s.

Tab.15: p-Werte der KFPT-Cluster/Einzelfarben – hässliche Pyramiden - (Studienteil 1)

## Studienteil 2

Aufgrund der statistisch und klinisch zu vernachlässigenden Unterschiede zwischen den beiden Behandlungsgruppen in der Ausgangslage als auch der insgesamt geringen Messwertveränderungen und Medikationsunterschiede bei Retest 2 nach zweiwöchiger Therapie werden die Ergebnisse von schönen und hässliche Pyramiden – mit Ausnahme von Cluster 3 - gemeinsam dargestellt.

### *Cluster 1*

Die Ausgangsbefunde (Baseline) zeigten keine beachtenswerten Unterschiede zwischen Baldrian-Extrakt und Placebo. Dieses Ergebnis wurde bei Retest 2 aufgrund der geringen Messwertveränderungen bestätigt.

Auch innerhalb der Medikationen zeigten sich keine statistisch signifikanten Veränderungen von Baseline zu Retest 2. Lediglich eine mäßige Zunahme der Wahlhäufigkeit des Clusters 1 ( $M=0,28$ ) bei hässlichen Pyramiden unter Baldrian-Extrakt war festzustellen, die jedoch statistisch ( $p=0.114$ ) irrelevant und klinisch ohne Bedeutsamkeit ist.

### *Cluster 2*

In der Baseline wurden keine zu beachtenden Unterschiede zwischen den Behandlungsgruppen evident. Aufgrund der geringen Messwertveränderungen wurden auch bei Retest 2 keine statistischen Gruppenunterschiede deutlich.

Die Gegenüberstellung von Baseline und Retest 2 innerhalb der Medikationen erbrachte gleichfalls keine Hinweise auf bedeutsame Veränderungen im Farbwahlverhalten der Probanden. Lediglich unter Baldrian-Extrakt war bei Retest 2 eine geringfügige Zunahme der Wahlhäufigkeit von Cluster 2 bei schönen Pyramiden festzustellen ( $M=0,23$ ), die jedoch statistisch ( $p=0.141$ ) zu vernachlässigen ist.

### *Cluster 3*

Die Baselinewerte beider Therapiegruppen waren hinsichtlich der statistischen Parameter vergleichbar und auch bei Retest 2 ergaben sich keine statistisch signifikanten Gruppenunterschiede im Farbwahlverhalten sowohl bei schönen als auch hässlichen Pyramiden. Innerhalb der Medikationen wurden jedoch deutlich ausgeprägte Messwertveränderungen bezüglich der Farbwahl bei schönen als auch hässlichen Pyramiden gemessen.

### Schöne Pyramiden:

Bei Retest 2 ergaben sich statistisch hochsignifikante und konträre Messwertveränderungen in beiden Therapiegruppen gegenüber der Ausgangslage.

Unter Baldrian-Extrakt wurde eine deutliche Abnahme der Wahlhäufigkeit des Clusters 3 ( $M=-0,43$ ) objektiviert, die auch statistische Signifikanz erreichte ( $p=0.002^{**}$ ). Unter Placebo wurde hingegen eine stark ausgeprägte Zunahme der Wahlhäufigkeit des Clusters 3 beobachtet ( $M=0,51$ ), die hochsignifikant war ( $p=0.000^{***}$ ).

Stichprobe	Zeitpunkt	M	S	Min	Max	N
Gesamtkollektiv	Baseline	2,71	2,78	0	11	88
	Retest 2	2,73	3,04	0	15	88
	Abnahme/Zunahme	,02	3,00	-11	11	88
Baldrian	Baseline	3,00	2,90	0	11	46
	Retest 2	2,57	3,24	0	15	46
	Abnahme/Zunahme	-,43	3,35	-11	11	46
Placebo	Baseline	2,39	2,63	0	10	42
	Retest 2	2,90	2,85	0	9	42
	Abnahme/Zunahme	,51	2,52	-5	8	42

Tab. 16: Häufigkeit des Clusters 3 (1c/5c) in schönen Pyramiden (Studienteil 2)

### Hässliche Pyramiden:

Auch hier zeigten sich statistisch hochsignifikante und konträre Messwertveränderungen im Vergleich zur Ausgangslage. Nach zweiwöchiger Einnahme des Baldrian-Extraktes wurde eine Zunahme der Wahlhäufigkeit des Clusters 3 ( $M=0,17$  /  $p=0.000^{***}$ ) gemessen; nach Placebogabe wurde eine Abnahme der Wahlhäufigkeit des Clusters 3 ( $M=-0.6$  /  $p=0.000^{***}$ ) evaluiert. Beide Ergebnisse wurden mit statistisch hoher Signifikanz gesichert.

Stichprobe	Zeitpunkt	M	S	Min	Max	N
Gesamtkollektiv	Baseline	4,30	2,82	0	11	88
	Retest 2	4,36	3,05	0	12	88
	Abnahme/Zunahme	,06	2,75	-8	8	88
Baldrian	Baseline	3,90	2,80	0	11	46
	Retest 2	4,08	3,04	0	12	46
	Abnahme/Zunahme	,17	2,14	-4	6	46
Placebo	Baseline	4,73	2,81	0	10	42
	Retest 2	4,67	3,08	0	12	42
	Abnahme/Zunahme	-,06	3,31	-8	8	42

Tab. 17: Häufigkeit des Clusters 3 (1c/5c) in hässlichen Pyramiden (Studienteil 2)

Die obigen Resultate belegen innerhalb der Medikationsgruppen eine gegenläufige Tendenz im Farbwahlverhalten. Unter Baldrian-Extrakt erfolgte bei den schönen Pyramiden eine Abnahme der Wahlhäufigkeit des Clusters 3. Bei den hässlichen Pyramiden wurde hingegen eine Zunahme der Wahlhäufigkeit bei Retest 2 gemessen. Unter Placebo ging die Zunahme der Wahlhäufigkeit von

Cluster 3 bei schönen Pyramiden mit der Abnahme der Wahlhäufigkeit im Retest 2 bei hässlichen Pyramiden einher.

#### *Farbton 1b*

Die Baselinewerte bei schönen bzw. hässlichen Pyramiden waren in beiden Behandlungsgruppen gut vergleichbar – Mittelwert, Streuung und Extrema lagen eng zusammen. In Folge geringer Messwertveränderungen bei Retest 2 bestanden bei zu diesem Untersuchungszeitpunkt gleichfalls keine statistischen Unterschiede. In den Medikationsgruppen konnten in der Gegenüberstellung von Baseline und Retest 2 Befunden keine signifikanten Veränderungen gemessen werden.

#### *Farbton 5b*

Der Vergleich beider Behandlungsgruppen erbrachte in der Ausgangslage bei schönen und hässlichen Pyramiden keine statistischen Unterschiede. Bei Retest 2 wurden ebenfalls keine bedeutsamen Messwertveränderungen objektiviert, die zu Gruppenunterschieden führen.

Unter Placebo wurde bei den schönen Pyramiden bei Retest 2 eine geringe Zunahme ( $M=0,60$  /  $p=0.190$ ) der Wahlhäufigkeit des Farbtons 5b beobachtet. Ebenfalls unter Placebo konnte bei den hässlichen Pyramiden bei Retest 2 eine Abnahme ( $M=-0,69$  /  $p=0.143$ ) der Wahlhäufigkeit des Farbtons 5b gemessen werden.

Da beide Ergebnisse korrespondieren, waren die Messwertveränderungen innerhalb der Placebogruppe von Baseline zu Retest 2 nicht von statistischer Signifikanz. Die Messwertvariationen unter Baldrian-Extrakt waren gleichfalls statistisch zu vernachlässigen.

#### *Farbton 10c*

Die Baselinebefunde beider Medikationen erbrachten keine statistisch zu beachtenden Gruppenunterschiede sowohl bei schönen als auch bei hässlichen Pyramiden. Aufgrund gleichgerichteter und gering ausgeprägter Messwertveränderungen bei schönen Pyramiden, konnten im Retest 2 keine relevanten Gruppenunterschiede evaluiert werden.

Bei der Frabwahl der hässlichen Pyramiden zeigte sich im Retest 2 ein statistischer Trend ( $p=0.071+$ ) in der Gegenüberstellung von Baldrian-Extrakt und Placebo, welcher durch die signifikante Zunahme der Wahlhäufigkeit ( $M=1,93$  /  $p=0.028^*$ ) von Farbton 10c bei Retest 2 unter Placebo bedingt war. Die Messwertveränderungen in der Baldrian-Extrakt-Gruppe von der Baseline zu Retest 1 waren hingegen statistisch unbedeutend.

Stichprobe	Zeitpunkt	M	S	Min	Max	N
Gesamtkollektiv	Baseline	3,13	4,24	0	16	88
	Retest 2	2,90	4,83	0	21	88
	Abnahme/Zunahme	-,23	3,96	-9	17	88
Baldrian	Baseline	3,52	4,59	0	16	46
	Retest 2	3,33	4,94	0	19	46
	Abnahme/Zunahme	-,20	4,09	-9	13	46
Placebo	Baseline	2,69	3,82	0	15	42
	Retest 2	2,43	4,73	0	21	42
	Abnahme/Zunahme	-,26	3,87	-7	17	42

Tab. 18: Häufigkeit des Farbtons 10c in schönen Pyramiden (Studienteil 2)

Stichprobe	Zeitpunkt	M	S	Min	Max	N
Gesamtkollektiv	Baseline	9,09	7,57	0	28	88
	Retest 2	9,88	6,92	0	30	88
	Abnahme/Zunahme	,78	6,38	-18	17	88
Baldrian	Baseline	9,76	8,15	0	28	46
	Retest 2	9,50	7,62	0	30	46
	Abnahme/Zunahme	-,26	6,58	-18	17	46
Placebo	Baseline	8,36	6,89	0	28	42
	Retest 2	10,29	6,13	0	24	42
	Abnahme/Zunahme	1,93	6,01	-15	17	42

Tab 19.: Häufigkeit des Farbtons 10c in hässlichen Pyramiden (Studienteil 2)

Zur Berechnung der p-Werte wurde auch im Studienteil 2 innerhalb der Medikationsgruppen für jeden Parameter die Zeitpunkte Baseline und Retest 2 mittels Wilcoxon-Test miteinander verglichen. Die drei Behandlungsgruppen wurden anhand der Parameter zur Baseline mit dem Mann-Whitney-U-Test, zum Retest 2 nach 14tägiger medikamentöser Behandlung mit dem Paar-Differenzen-U-Test (Differenz zur Baseline) verglichen.

Cluster/ Farbton	Zeitpunkt	Schöne Pyramiden - Medikationen -			Hässliche Pyramiden - Medikationen-		
		Baldrian	Placebo	Baldrian vs Placebo	Baldrian	Placebo	Baldrian vs Placebo
Cluster 1	Baseline	-	-	n.s.	-	-	n.s.
	Retest 2	n.s.	n.s.	n.s.	0.114	n.s.	n.s.
Cluster 2	Baseline	-	-	n.s.	-	-	n.s.
	Retest 2	0.141	n.s.	n.s.	n.s.	n.s.	n.s.
Cluster 3	Baseline	-	-	n.s.	-	-	0.101
	Retest 2	0.002*	0.000*	n.s.	0.000*	0.000*	n.s.
1b	Baseline	-	-	n.s.	-	-	n.s.
	Retest 2	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
5b	Baseline	-	-	n.s.	-	-	n.s.
	Retest 2	n.s.	0.190	n.s.	n.s.	0.143	n.s.
10c	Baseline	-	-	n.s.	-	-	n.s.
	Retest 2	n.s.	n.s.	n.s.	n.s.	0.028*	0.071+

Tab. 20: p-Werte der KFPT-Cluster/Einzelfarben – schöne/hässliche Pyramiden – (Studienteil 2)

## Weitere Ergebnisse

Um die Frage zu klären, ob zwischen individuellen Farbpräferenzen und Farbwahlverhalten ein Zusammenhang besteht, wurden ergänzend zur Auswertung der KFPT-Zielvariablen (Cluster und Einzelfarbtöne) Spearman-Rangkorrelationskoeffizienten zwischen Valenzwert im ED und Wahlhäufigkeit des Clusters bzw. des Einzelfarbtönen im KFPT berechnet.

Für den Rangkorrelationskoeffizient gilt, dass er nicht kleiner als  $-1$  und nicht größer als  $+1$  werden kann. Ein Wert von  $r_s = 0$  bedeutet, dass zwischen den Rangordnungen beider Variablen keine Systematik besteht. Ein Wert von  $r_s = +1$  impliziert eine perfekte Übereinstimmung der beiden Rangordnungen. Entsprechend gilt bei einem Wert von  $r_s = -1$ , dass sich die Rangplätze perfekt diametral gegenüberstehen. Generell legt ein Betrag nahe bei  $1$  einen starken und ein Betrag nahe bei der Null einen schwachen oder überhaupt keinen Zusammenhang nahe. Ist der Korrelationskoeffizient negativ, weist dies auf einen gegenläufigen Zusammenhang hin: je größer der Wert der einen Variablen (Valenzwert) ist, desto kleiner ist der Wert der anderen (Wahlhäufigkeit).

Im Studienteil 1 fanden sich bei schönen Pyramiden positive Zusammenhänge zwischen Farbpräferenzen im ED und Wahlhäufigkeit im KFPT. Deren Ausprägungsgrad war jedoch bei Cluster und Einzelfarbtönen unterschiedlich. Bei sämtlichen Clustern zeigten sich - unabhängig von Untersuchungszeitpunkt – im Gesamtkollektiv deutliche Zusammenhänge zwischen Valenzwert und Wahlhäufigkeit, wobei die Korrelationen bei Cluster 1 am deutlichsten waren. Demgegenüber wurden bei den drei Einzelfarbtönen zu beiden Messzeitpunkten die statistischen Zusammenhänge insgesamt nur mäßig deutlich.

Bei den hässlichen Pyramiden ergaben sich im Studienteil 1 bei Clustern und Einzelfarbtönen zu beiden Testabnahmen gegenläufige Korrelationen zwischen Valenzwert und Wahlhäufigkeit. Cluster und Einzelfarbtöne unterschieden sich hinsichtlich der Ausprägung der statistischen Zusammenhänge. Bei den Clustern zeigten sich zum Teil signifikante (Cluster 1) bis hochsignifikante (Cluster 3) Ergebnisse. Bei den Farbtönen ergab sich nur bei Farbton 10c in der Baseline eine statistische Signifikanz. Im Gegensatz zu den schönen Pyramiden konnte bei den hässlichen Pyramiden sowohl bei Clustern als auch bei Einzelfarbtönen insgesamt kein positiver Zusammenhang zwischen subjektiven Farbpräferenzen und Farbwahlverhalten hergestellt werden.

Cluster bzw. Einzelfarbtone	Schöne Pyramiden		Hässliche Pyramiden	
	Baseline	Retest 1	Baseline	Retest 1
Cluster 1	0.306 **	0.345 **	-0.221 *	-0.247 *
Cluster 2	0.241 *	0.214 *	-0.201 +	-0.247 *
Cluster 3	0.095	0.215 *	-0.358 ***	-0.477 ***
1b	0.102	0.136	-0.086	-0.101
5b	-0.004	0.118	-0.049	-0.092
10c	0.128	0.031	-0.236 *	-0.113

Tab. 21: Spearman-Rangkorrelationskoeffizienten zwischen Valenzwert im Eindrucksdifferential und Wahlhäufigkeit des Clusters/Einzelfarbtönen im Kölner Farbpräferenztest (Studienteil 1, Gesamtkollektiv – N=96)

Cluster bzw. Einzelfarbtone	Schöne Pyramiden		Hässliche Pyramiden	
	Baseline	Retest 2	Baseline	Retest 2
Cluster 1	0.319 **	0.320 **	-0.272 *	-0.263 *
Cluster 2	0.210 *	0.360 **	-0.198 +	-0.234 *
Cluster 3	0.133	0.212 *	-0.403 ***	-0.331 **
1b	0.057	0.134	-0.053	0.061
5b	0.025	0.033	-0.061	-0.080
10c	0.147	0.127	-0.257 *	-0.244 *

Tab. 22: Spearman-Rangkorrelationskoeffizienten zwischen Valenzwert im Eindrucksdifferential und Wahlhäufigkeit des Clusters/Einzelfarbtone im Kölner Farbpräferenztest (Studienteil 2, Gesamtkollektiv – N=88)

Im Studienteil 2 ergaben sich zu beiden Messzeitpunkten bei schönen Pyramiden bei Clustern und Einzelfarbtönen – mit unterschiedlicher statistischer Ausprägung – positive Zusammenhänge zwischen Farbpräferenz und Farbwahlverhalten. Die Korrelationen waren bei den Clustern signifikant, jedoch bei den Einzelfarbtönen vergleichsweise gering ausgeprägt.

Die Ergebnisse zu den hässlichen Pyramiden erbrachten zu beiden Untersuchungszeitpunkten gegenläufige Korrelationen, die bei Clustern zum Teil hochsignifikant (Cluster 3, Baseline) und bei Einzelfarben relativ geringer ausgeprägt waren.

Eine Ausnahme bildete der Farbtone 10c, wo zu beiden Messzeitpunkten eine negativ signifikante Korrelation bestand. Ein Zusammenhang zwischen persönlichen Farbpräferenzen und Farbwahlverhalten konnte bei hässlichen Pyramiden weder bei Clustern noch bei Einzelfarben gezeigt werden.

Die Resultate beider Studienteile korrespondierten miteinander. Die gefundenen - positiven wie negativen Korrelationen - waren bei Clustern weitaus deutlicher ausgeprägt als bei Einzelfarben. Bei der Farbzusammenstellung der schönen Pyramiden wurden positive Zusammenhänge zwischen Farbpräferenz im ED und Farbwahlverhalten im KFPT evaluiert. Bei den hässlichen Pyramiden bestand ein diametraler Zusammenhang. Die Probanden bevorzugten für die hässlichen KFPT-Pyramiden offensichtlich nicht diejenigen Farben, die Ihnen im ED gefallen haben. Eine mögliche Erklärung könnte sein, dass die Farbwahl stark abhängig ist vom Selbstbild der Person und getrennt von ihren Farbpräferenzen erfolgte. Die Ergebnisse für die einzelnen Medikationsgruppen reflektieren die in den Gesamtkollektiven beider Studienteile objektivierten Befunde.

## 8.2 Sekundäre Endpunkte

### 8.2.1 Semantisches Eindrucksdifferential (ED)

#### *Baseline*

In der Ausgangslage (Baseline) dominierte innerhalb eines Clusters bzw. einer Einzelfarbe folgende Dimension:

Cluster 1 (1a/4a/6a/10b)	=	Valenz (M=4,6933)
Cluster 2 (6b/8b/8c/9b/10a)	=	Valenz (M=5,6922)
Cluster 3 (1c/5c)	=	Potenz (M=5,3280)
Farbton 1b	=	Potenz (M=6,1224)
Farbton 5b	=	Erregung (M=5,6401)
Farbton 10c	=	Potenz (M=6,5166)

Cluster/Farbton	ED-Dimension	M	S	Min	Max	N
Cluster 1 (1a/4a/6a/10b)	Erregung	3,2700	,7187	1,38	5,34	99
	Valenz	4,6933	,9166	3,04	6,96	99
	Potenz	3,0170	,6434	1,38	4,88	99
Cluster 2 (6b/8b/9b/8c/10a)	Erregung	4,1558	,7121	,63	5,86	99
	Valenz	5,6922	,7399	3,80	7,37	99
	Potenz	4,6330	,7127	2,67	6,50	99
Cluster 3 (1c/5c)	Erregung	4,4082	1,2279	1,00	,09	99
	Valenz	3,2509	1,2580	1,00	5,92	99
	Potenz	5,3280	1,1221	2,08	8,00	99
Farbton 1b	Erregung	6,0623	1,3750	1,50	8,00	99
	Valenz	5,9525	1,3758	2,50	8,00	99
	Potenz	6,1224	1,1078	2,33	8,00	99
Farbton 5b	Erregung	5,6401	1,5392	,00	8,00	99
	Valenz	4,4019	1,7553	,00	8,00	99
	Potenz	5,3248	1,3822	,00	7,67	99
Farbton 10c	Erregung	3,6140	,9285	2,33	6,50	99
	Valenz	4,7590	1,4581	1,83	6,83	99
	Potenz	6,5166	,7143	3,83	8,00	99

Tab. 23: Cluster/Farbton und ED Dimensionswerte  
(Baseline/Gesamtkollektiv – N=102)

### Geschlecht und Alter

Für Cluster 2 sowie den Einzelfarben 1b und 10c wurden keine beachtenswerten Geschlechtsunterschiede hinsichtlich der E/V/P-Dimensionen in der Ausgangslage deutlich. Bei Farbton 1b war die Potenzwirkung unabhängig vom Geschlecht vordergründige Dimension, doch ist zu betonen, dass die beiden anderen Dimensionen auf hohem Ausprägungsniveau vergleichbare Messwerte erzielten. Bei den Cluster 1 und 3 sowie dem Farbton 5b zeigten sich hingegen geschlechtsspezifische Unterschiede: Cluster 1 wurde von Männern (M=3,4923) im Vergleich zu Frauen (M=3,0993) als signifikant ( $p=0.007^*$ ) erregender erlebt. Cluster 3 erlebten Männer (M=3,5527) im Vergleich zu Frauen (M=3,0191) als signifikant ( $p=0.041^*$ ) angenehmer. Farbton 5b wurde von Frauen (M=5,8395) im Vergleich zu Männern (M=5,3805) als signifikant ( $p=0.036^*$ ) erregender erlebt. Bei Cluster 2, Cluster 3 sowie dem Farbton 1b waren keine altersspezifischen Unterschiede im Farberleben erkennbar. Das Cluster 1 wurde von Älteren (50-60jährige / M=3,6993) signifikant erregender erlebt ( $p=0.012^*$ ) als von Jüngeren (30-39jährige / M=3,0869). Bei Farbton 1b dominierte bei Probanden unter 40 Jahre die Potenzwirkung; ab 40 Jahre aufwärts wurde der Farbton sowohl als potent als auch erregend wahrgenommen - wobei die Unterschiede zwischen beiden Dimensionen marginal und statistisch irrelevant sind. Unabhängig von Alter und Geschlecht lagen alle drei Dimensionen auf hohem Ausprägungsniveau eng beisammen und bestätigen den Farbton „rot“ als „Superfarbe“, die dimensionsübergreifend als erregend, angenehm und kraftvoll erlebt wird. Der Farbton 5b wurde von Jüngeren (30-39jährige / M=6,0417) hochsignifikant erregender erlebt ( $p=0.003^{**}$ ) als von Älteren (50-60jährige / M=4,9108). Mit zunehmendem Alter als weniger angenehm erlebt wurde der Farbton 10c (30-39/50-60  $p=0.002^{**}$  // 40-49-50-60  $p=0.037^*$ ).

Cluster / Farbton	Dimension	Geschlecht	Alter (30-39/40-49)	Alter (30-39/50-60)	Alter 40-49/50-60)
Cluster 1	Erregung	0.007 **	0.104	0.012 *	n.s.
	Valenz	0.115	n.s.	n.s.	n.s.
	Potenz	0.056 +	0.132	0.066 +	n.s.
Cluster 2	Erregung	n.s.	n.s.	n.s.	0.125
	Valenz	n.s.	0.138	n.s.	0.072 +
	Potenz	n.s.	n.s.	n.s.	n.s.
Cluster 3	Erregung	0.054 +	n.s.	n.s.	n.s.
	Valenz	0.041 *	n.s.	n.s.	n.s.
	Potenz	0.051 +	n.s.	n.s.	n.s.
1b	Erregung	n.s.	n.s.	n.s.	n.s.
	Valenz	n.s.	0.132	n.s.	0.072+
	Potenz	n.s.	n.s.	n.s.	n.s.
5b	Erregung	0.036 *	0.050+	0.003 **	n.s.
	Valenz	n.s.	n.s.	n.s.	n.s.
	Potenz	0.138	0.149	0.007 **	n.s.
10c	Erregung	0.140	n.s.	n.s.	n.s.
	Valenz	0.162	n.s.	0.002 **	0.037 *
	Potenz	n.s.	n.s.	0.182	n.s.

Tab. 24: ED p-Werte für den Vergleich der Geschlechter und Altersklassen (Baseline)

## Studienteil 1

### *Cluster 1 und Cluster 2*

Nach einmaliger Einnahme der Medikation blieb die Valenzwirkung bei Retest 1 in beiden Clustern bei sämtlichen Medikationsgruppen weiterhin die dominante Dimension im Farberleben der Probanden. Insgesamt waren die E/V/P-Messwertveränderungen (Baseline-Retest 1) in beiden Clustern minimal.

### *Cluster 3*

Die Potenzwirkung war – in Relation zur Ausgangslage im Gesamtkollektiv – auch bei Retest 1 die dominante Dimension im Farberleben der Probanden aller drei Medikationsgruppen.

Im Vergleich Baldrian-Extrakt vs. Placebo erlebten die Probanden nach Baldrian-einnahme das Cluster 3 als potenter ( $p=0.020^*$ ). Die anderen Gruppenvergleiche zeigten statistisch unauffällige Ergebnisse. Insgesamt waren die E/V/P-Messwertveränderungen (Baseline-Retest 1) geringfügig ausgeprägt.

### *Farbton 1b*

Zur Baseline als auch bei Retest 1 blieb die Potenzwirkung unter Baldrian-Extrakt und Placebo die dominante Dimension. Unter Flunitrazepam war Valenz vordergründig ( $M=6,1094$ ) wenn auch nur in geringem Abstand zur Potenzwirkung ( $M=6,0941$ ). Insgesamt waren die E/V/P-Messwertveränderungen (Baseline-Retest 1) marginal.

### *Farbton 5b*

„Erregung“ bleibt – in Relation zur Ausgangslage im Gesamtkollektiv – auch bei Retest 1 die dominante Dimension im Farberleben der Probanden aller drei Medikationsgruppen. Die geringen E/V/P-Messwertveränderungen (Baseline-Retest 1) sind statistisch zu vernachlässigen und ohne Relevanz für das Farberleben.

Innerhalb der Medikationsgruppen zeigen sich Unterschiede hinsichtlich des Ausprägungsgrades der Potenzwirkung: Unter Flunitrazepam wird der Farbton 5b als am kraftvollsten erlebt ( $M=5,6406$ ), während der Farbton 5b unter Placebo ( $M=5,4662$ ) und Baldrian-Extrakt ( $M=4,8730$ ) relativ weniger potent erlebt wird. Der Unterschied im Farberleben zwischen Flunitrazepam und Baldrian-Extrakt ist statistisch signifikant ( $p=0.015^*$ ), während sich gegenüber Placebo nur ein Trend ( $p=0.071+$ ) zeigte.

Medikation	Dimension	M	S	Min	Max	N
Baldrian	Erregung	5,1618	1,8535	,00	7,67	33
	Valenz	4,8930	1,9004	,00	8,00	33
	Potenz	4,8730	1,4873	,00	7,50	33
Placebo	Erregung	5,8638	1,3408	3,00	8,00	34
	Valenz	4,1859	1,5065	1,17	6,50	34
	Potenz	5,4662	1,3531	2,50	7,67	34
Flunitrazepam	Erregung	5,8956	1,2900	2,83	7,67	32
	Valenz	4,1250	1,7915	1,17	7,67	32
	Potenz	5,6406	1,2102	2,33	7,67	32

Tab. 25: Dimensionswerte des Farbtons 5b nach Medikation (Studienteil 1)

### Farbton 10c

In Relation zur Ausgangslage im Gesamtkollektiv war auch bei Retest 1 die Potenzwirkung dominante Dimension im Farberleben der Probanden aller drei Medikationsgruppen. Insgesamt waren die E/V/P-Messwertveränderungen (Baseline-Retest 1) geringfügig.

Medikation	Dimension	M	S	Min	Max	N
Baldrian	Erregung	3,6203	,9595	2,33	6,50	33
	Valenz	4,4497	1,4707	2,00	6,50	33
	Potenz	6,3579	,6831	4,67	7,83	33
Placebo	Erregung	3,6965	,9302	2,33	6,50	34
	Valenz	4,7253	1,4448	2,17	6,83	34
	Potenz	6,6224	,6590	5,50	8,00	34
Flunitrazepam	Erregung	3,5200	,9150	2,33	6,17	32
	Valenz	5,1138	1,4259	1,83	6,83	32
	Potenz	6,5678	,7918	3,83	7,83	32

Tab. 26: Dimensionswerte des Farbtons 10c nach Medikation (Studienteil 1)

Im Gruppenvergleich Baldrian-Extrakt vs. Flunitrazepam zeigten sich in den Dimensionen Valenz ( $p=0.053+$ ) und Potenz ( $p=0.091+$ ) statistische Trends. Rückschlüsse auf ein unterschiedliches Farberleben unter Einfluss der beiden Verum-Medikationen waren daraus jedoch nicht abzuleiten.

<b>Cluster bzw. Farbton</b>	<b>Dimension</b>	<b>Baldrian vs. Placebo</b>	<b>Baldrian vs. Flunitrazepam</b>	<b>Placebo vs. Flunitrazepam</b>
Cluster 1	Erregung	n.s.	n.s.	n.s.
	Valenz	n.s.	0.061 +	n.s.
	Potenz	n.s.	n.s.	n.s.
Cluster 2	Erregung	n.s.	n.s.	n.s.
	Valenz	n.s.	n.s.	n.s.
	Potenz	0.196	n.s.	0.052 +
Cluster 3	Erregung	n.s.	n.s.	n.s.
	Valenz	n.s.	n.s.	n.s.
	Potenz	0.020 *	n.s.	n.s.
1b	Erregung	n.s.	n.s.	n.s.
	Valenz	n.s.	n.s.	n.s.
	Potenz	n.s.	n.s.	n.s.
5b	Erregung	0.171	0.127	n.s.
	Valenz	0.070+	0.071+	n.s.
	Potenz	0.130	0.015 *	n.s.
10c	Erregung	n.s.	n.s.	n.s.
	Valenz	n.s.	0.053+	n.s.
	Potenz	0.130	0.091+	n.s.

Tab. 27: ED p-Werte für den Vergleich der Medikationsgruppen (Studienteil 1)

## Studienteil 2

### *Cluster 1*

Nach 14tägiger Einnahme der beiden Medikationen wurden nur geringfügige Messwertveränderungen der Dimensionen gegenüber der Ausgangslage objektiviert. Die Valenzwirkung blieb bei Retest 2 – unabhängig von der Medikationsgruppe - dominierende Dimension im Farberleben. Eine Beeinflussung des Farberlebens durch die zweiwöchige Einnahme der Medikationen war nicht feststellbar. Gruppenunterschiede zwischen Baldrian-Extrakt und Placebo hinsichtlich der Dimensionsausprägungen waren statistisch nicht signifikant und auch klinisch zu vernachlässigen.

### *Cluster 2*

Im Vergleich zur Baseline zeigten sich nach Therapie kaum Messwertveränderungen in den drei Dimensionen. Bei Retest 2 bestand die in der Ausgangslage evaluierte Dominanz der Valenzwirkung in beiden Behandlungsgruppen fort. Auch im Gruppenvergleich wurden keine relevanten Messwertunterschiede deutlich. Eine Beeinflussung des Farberlebens unter Medikationseinnahme war nicht objektivierbar.

### *Cluster 3*

Insgesamt wurden bei Retest 2 nur geringe Veränderungen in den Dimensionsausprägungen gegenüber der Baseline 2 gemessen. Die potente Wirkung blieb auch unter Therapie weiterhin vordergründige Dimension im Farberleben beider Medikationsgruppen.

Bezüglich der Potenzwirkung war ein Unterschied zwischen Placebo und Baldrian-Extrakt erkennbar. Nach Einnahme des Baldrian-Extraktes wurde das Cluster 3 vergleichsweise kraftvoller erlebt als unter Placebo, doch lieferte dieser statistische Trend ( $p=0.068+$ ) lediglich einen Hinweis auf ein evtl. differierendes Farberleben in dieser Dimension. In den beiden anderen Dimensionen (Erregung und Valenz) zeigten sich keine statistisch fassbaren Medikationsunterschiede, welche den Trend bei der Potenzwirkung bestätigen. Insgesamt war eine Beeinflussung des Farberlebens unter Medikation nicht nachweisbar.

Medikation	Dimnsion	M	S	Min	Max	N
Baldrian	Erregung	4,2891	1,2911	1,00	7,09	47
	Valenz	3,4150	1,2609	1,00	5,92	47
	Potenz	5,5139	1,1199	2,59	8,00	47
Placebo	Erregung	4,4147	1,1169	1,92	7,00	44
	Valenz	3,1722	1,2717	1,17	5,75	44
	Potenz	5,0376	1,1316	2,08	7,34	44

Tab. 28: Dimensionswerte des Clusters 3 (1c/5c) nach Medikation (Studienteil 2)

### *Farbton 1b*

Nach Therapie zeigten sich bei Retest 2 nur marginale Messwertveränderungen in den Behandlungsgruppen gegenüber der Ausgangslage im Gesamtkollektiv. Die Dimensionsausprägungen lagen weiterhin auf hohem Niveau eng zusammen. Unter Placebo bestand die leichte Präferenzierung der potenten Farbwirkung fort, während unter Baldrian-Extrakt der Farbton 1b primär als erregend erlebt wurde. Die Differenzen zu den Mittelwerten der Dimensionen Potenz und Valenz waren jedoch minimal. Insgesamt erzielte der Farbton 1b auch nach 14tägiger Behandlung auf allen drei Dimensionen eine nahezu gleichförmige intensive Wirkung und bestätigt die Baseline-Befunde („rot“ als Superfarbe). Medikationsbedingte Veränderungen des Farberlebens waren nicht festzustellen.

### *Farbton 5b*

Gemessen wurden lediglich minimale Variationen innerhalb der Dimensionsausprägungen im Vergleich zwischen Ausgangslage und Retest 2. Die erregende Wirkung blieb auch nach Therapie und unabhängig von der Medikation vordergründige Erlebnisqualität des Farbtons 5b. Die geringen Messwertunterschiede zwischen den beiden Behandlungsgruppen waren statistisch nicht signifikant. Eine Beeinflussung des Farberlebens durch Baldrian-Extrakt ist eher unwahrscheinlich.

### *Farbton 10c*

Die Gegenüberstellung von Baseline und Retest 2 erbrachte nur geringe Messwertveränderungen in den drei Dimensionen. Die Potenzwirkung blieb unabhängig von der Medikation weiterhin dominant, wobei die Wirkung des Farbtons 10c unter Baldrian-Extrakt etwas weniger kraftvoll war als unter Placebo ( $p=0.074+$ ). Medikationsbedingte Unterschiede im Farberleben lassen sich aus diesem Befund jedoch nicht ableiten.

<b>Medikation</b>	<b>Dimnsion</b>	<b>M</b>	<b>S</b>	<b>Min</b>	<b>Max</b>	<b>N</b>
Baldrian	Erregung	3,6091	,9859	2,33	6,50	47
	Valenz	4,6345	1,4460	2,00	6,83	47
	Potenz	6,3966	,7890	3,83	7,83	47
Placebo	Erregung	3,6286	,9237	2,33	6,50	44
	Valenz	4,8709	1,5460	1,83	6,83	44
	Potenz	6,6666	,6295	5,50	8,00	44

Tab. 29: Dimensionswerte des Farbtons 10c nach Medikation (Studienteil 2)

<b>Cluster bzw. Farbton</b>	<b>Dimension</b>	<b>Baldrian vs. Placebo</b>
Cluster 1	Erregung	n.s.
	Valenz	n.s.
	Potenz	n.s.
Cluster 2	Erregung	n.s.
	Valenz	n.s.
	Potenz	n.s.
Cluster 3	Erregung	n.s.
	Valenz	n.s.
	Potenz	0.068 +
1b	Erregung	n.s.
	Valenz	n.s.
	Potenz	n.s.
5b	Erregung	n.s.
	Valenz	n.s.
	Potenz	n.s.
10c	Erregung	n.s.
	Valenz	n.s.
	Potenz	0.074+

Tab. 30: ED p-Werte für den Vergleich der Medikationsgruppen (Studienteil 2)

## 8.2.2 Determinationstest (DT)

### Studienteil 1

#### *Reaktionszeit (Median) im Modus "Adaptiv" (Hauptzielvariable des DT)*

Die Ausgangswerte des „Medians der Reaktionszeit“ betragen in der Baldrian-Extrakt-Gruppe 807.15 ms, in der Placebo-Gruppe 800.56 ms und in der Flunitrazepam-Gruppe 776.38 ms. Die Baldrian-Extrakt-Gruppe weist die relativ schlechteste, die Flunitrazepam-Gruppe die vergleichsweise beste Reaktionszeit aus. Die Range von 30.77 ms zwischen den Baselineresultaten der Gruppen Flunitrazepam und Baldrian-Extrakt bewegt sich an der Obergrenze des noch zulässigen Toleranzbereiches für einen Gruppenvergleich. Mit diesem Befund korrespondierend wurde ein statistischer Trend ( $p=0.0754+$ ) im Vergleich der Behandlungsgruppen Flunitrazepam vs. Baldrian-Extrakt objektiviert. Dieser Trend ist jedoch psychologisch irrelevant und schränkt die Vergleichbarkeit der beiden Medikationsgruppen keinesfalls ein. Eine Gegenüberstellung der Medikationsgruppen ist möglich, da die Baselinewerte der Parameter Median, Standardabweichung und Extrema gut vergleichbar sind. Die Vergleiche Baldrian-Extrakt vs. Placebo ( $p=0.8360$ ) und Placebo vs. Flunitrazepam ( $p=0.1638$ ) belegen keine statistisch signifikanten Gruppenunterschiede in der Ausgangslage.

Bei der Gegenüberstellung von Leertest und Retestdaten ist zu bedenken, dass beim DT durch Übungs- und Lerneffekte eine Leistungssteigerung im Retest zu erwarten ist. Da nur ein zeitlicher Abstand von einem Tag zwischen beiden Testabnahmen liegt, ist von einem deutlichen Übungsfortschritt auszugehen. Bei negativer Beeinflussung der kognitiven Leistungsfähigkeit durch die Verumpräparate ist dieser Leistungszuwachs allerdings nur in der Placebogruppe zu erwarten.

Wenn man Leistungsminderung als Absinken unter das Niveau der Ausgangslage definiert, sind in allen drei Teilstichproben bei Retest 1 im Mittel keine Leistungsdefizite messbar. Das Ausmaß der Leistungssteigerung ist jedoch in den Behandlungsgruppen unterschiedlich ausgeprägt. Die Differenz zwischen den Gruppen Baldrian-Extrakt und Flunitrazepam beträgt im Mittel 36.40 ms und belegt einen psychologisch relevanten Unterschied hinsichtlich der Leistungssteigerung beim Retest. Der verminderte Leistungszuwachs der Flunitrazepam-Gruppe verweist auf einen deutlich geschmäleren Übungsgewinn, auch wenn der Median der Reaktionszeit sich von den Testergebnissen der Vergleichsgruppen nur gering unterscheidet. Der Umfang des Lerneffektes der Gruppen Baldrian-Extrakt und Placebo ist vergleichbar. Die Differenz zwischen den Gruppen beträgt bei Retest 1 durchschnittlich 10.06 ms. Beide Gruppen profitieren gleichermaßen vom Übungseffekt.

<b>Medikation</b>	<b>M</b>	<b>S</b>	<b>Median</b>	<b>Min</b>	<b>Max</b>	<b>N</b>
<b>Baldrian</b>						
Leertest	807,15	81,79	791,00	630,00	1006,00	33
Retest 1	745,12	72,79	739,00	605,00	941,00	33
Abs.- Veränderung	62,03	26,49	60,00	8,00	121,00	33
%-Veränderung	7,60	3,05	7,59	1,09	14,77	33
<b>Placebo</b>						
Leertest	800,56	86,68	808,00	625,00	1003,00	34
Retest 1	748,59	82,95	734,50	600,00	957,00	34
Abs.- Veränderung	51,97	26,80	49,00	10,00	110,00	34
%-Veränderung	6,45	3,21	6,01	1,59	13,88	34
<b>Flunitrazepam</b>						
Leertest	776,38	73,69	759,50	665,00	988,00	32
Retest 1	750,75	74,62	735,50	612,00	992,00	32
Abs.- Veränderung	25,63	36,77	34,50	-80,00	82,00	32
%-Veränderung	3,22	4,77	4,32	-11,82	10,72	32

Tab. 31: DT / Median der Reaktionszeit in ms (Studienteil 1)

Eine detaillierte Betrachtung der "Veränderungen der mittleren Gesamtdauer" zeigt, dass bei 25% der Probanden der Flunitrazepam-Gruppe eine Leistungseinbuße vorliegt, die bei alleiniger Betrachtung des Medians für das gesamte Subkollektiv (Flunitrazepam-Gruppe) nicht evident wird: Bei 15.6% der Probanden verschlechterte sich der Median der Reaktionszeit um bis zu 25 ms; bei zwei Probanden (6.3%) zwischen 25-50 ms und bei einem Probanden (3.1%) um 80 ms. Damit sind bei einem Viertel der Probanden der Flunitrazepam-Gruppe Leistungsdefizite nachweisbar.

Der Median der gesamten Teilstichprobe "Flunitrazepam" belegt hingegen eine - hinter den beiden Vergleichsgruppen zurückbleibende - Leistungsverbesserung. Der Median der Reaktionszeit geht um 25.63 ms zurück, da sich insgesamt 75% der Probanden in ihrer Leistung steigern, wobei sich der Großteil (40.6%) im Bereich zwischen 25-50 ms verbessert. Der für Flunitrazepam typische "Hangover-Effekt" kommt nicht bei allen Probanden der Flunitrazepam-Gruppe zum Tragen. Lediglich 25% der Studienteilnehmer dieser Gruppe zeigen gegenüber der Ausgangslage eine reduzierte Testleistung.

Sämtliche Probanden unter Placebo und Baldrian-Extrakt weisen hingegen bei Restest 1 eine Leistungssteigerung auf (Abnahme der Reaktionszeit). Der Übungseffekt wird nicht durch einen Medikamenteneffekt überlagert, wobei der Median der Reaktionszeit in der Baldrian-Extrakt-Gruppe sogar noch deutlicher abnimmt als in der Placebo-Gruppe. Der Großteil der Probanden verbessert sich in beiden Medikationsgruppen (Placebo 76,5%; Baldrian-Extrakt 84.8%) zwischen 25-100 ms.

Gruppierung	Baldrian		Placebo		Flunitrazepam		Gesamt	
> -100 bis -75	0	0%	0	0%	1	3,1%	1	1,0%
> - 50 bis -25	0	0%	0	0%	2	6,3%	2	2,0%
> - 25 bis 0	0	0%	0	0%	5	15,6%	5	5,1%
> 0 bis 25	3	9,1%	6	17,6%	3	9,4%	12	12,1%
> 25 bis 50	11	33,3%	11	32,4%	13	40,6%	35	35,4%
> 50 bis 75	8	24,2%	9	26,5%	6	18,8%	23	23,2%
> 75 bis 100	9	27,3%	6	17,6%	2	6,3%	17	17,2%
> 100 bis 125	2	6,1%	2	5,9%	0	0%	4	4,0%
Gesamt	33	100%	34	100%	32	100%	99	100%

Tab. 32: DT / Veränderung der mittleren Gesamtdauer, ms/25er Einteilung (Studienteil 1)

Der mit dem Mann-Whitney-U-Test berechnete p-Wert für den Vergleich Placebo vs. Baldrian-Extrakt ( $p=0.1184$ ) belegt keinen statistischen Unterschied zwischen den Gruppen. Das 95%-Konfidenzintervall liegt bei  $[-2.9508, 23.0702]$ .

Der Unterschied Flunitrazepam vs. Baldrian-Extrakt ist auf dem 0.1%-Niveau hochsignifikant ( $p=0.0001$ ; Konfidenzintervall  $[-52.3756, -20.4350]$ ), der Unterschied Flunitrazepam vs. Placebo ist auf dem 5%-Niveau signifikant ( $p=0.0115$ ; Konfidenzintervall  $[-42.2987, -103925]$ ).

Die gesicherten Signifikanzen spiegeln die psychologisch relevanten Gruppenunterschiede bezüglich des Übungseffektes wider.

### *Richtige Reaktionen*

Diese Variable misst die Leistungsfähigkeit des Probanden, bei Reaktionsketten (längerandauernde Folgen von einfachen Reaktionsaufgaben) unter erheblicher Belastung rasch und adäquat zu reagieren.

Der Baseline-Mittelwert richtiger Reaktionen beträgt für die Baldrian-Extrakt-Gruppe 241.24 und für die Placebo-Gruppe 239.97. Die Flunitrazepam-Gruppe erzielt im Durchschnitt 247.53 richtige Reaktionen. Bezogen auf die Gesamtnorm ist die Ausgangslage der drei Teilstichproben als leicht überdurchschnittlich einzustufen: Baldrian-Extrakt-Gruppe PR 60; Placebo-Gruppe PR 60; Flunitrazepam-Gruppe PR 65. Die Positionsbestimmung in der Normstichprobe belegt eine vergleichbare Ausgangslage der Medikationsgruppen bezüglich dieser Testvariablen. Die berechneten p-Werte (s. Tab. 42) lassen keine statistisch signifikanten Gruppenunterschiede in der Ausgangslage erkennen.

Die durchschnittliche Anzahl der richtigen Reaktionen nimmt von der Aufnahmetestung zu Retest 1 in allen drei Medikationsgruppen zu, wobei der Leistungszuwachs der Flunitrazepam-Gruppe hinter den beiden Vergleichsgruppen zurückbleibt. Die durchschnittliche Anzahl richtiger Reaktionen liegt beim Retest 1 in allen drei Medikationsgruppen auf vergleichbarem Niveau. Die Range beträgt beim Leertest 7.56 und beim Retest 5.6 richtige Reaktionen.

Der Leistungsanstieg in der Flunitrazepam-Gruppe ist nur halb so groß wie in den beiden anderen Medikationsgruppen. Die mittlere Anzahl richtiger Reak-

tionen bleibt in der Flunitrazepam-Gruppe signifikant (5%-Niveau) hinter dem Ergebnis der Baldrian-Extrakt-Gruppe zurück ( $p=0.0166$ ). Die Testergebnisse der Probanden in den Gruppen Baldrian-Extrakt und Placebo stimmen bezüglich Ausgangslage und Leistungszuwachs fast überein ( $p=0.3729$ ). Eine Beeinträchtigung unter Baldrian-Extrakt ist nicht feststellbar. Auch bei den mittleren prozentualen Veränderungen zeigt sich bei der Gegenüberstellung von Flunitrazepam und Baldrian-Extrakt auf dem 5%-Niveau ein statistisch signifikanter Unterschied (0.0209). Die beiden anderen Vergleiche belegen keine statistisch relevanten Gruppenunterschiede. Die Leistungsfähigkeit der Probanden (bei Reaktionsketten unter erheblicher Belastung rasch und adäquat zu reagieren) ist bei den Medikationsgruppen Baldrian-Extrakt und Placebo nach einmaliger vorabendlicher Einnahme beim Retest vergleichbar.

Medikation	M	S	Median	Min	Max	N
<b>Baldrian</b>						
Leertest	241,24	36,17	239,00	166,00	332,00	33
Retest 1	264,85	35,82	266,00	142,00	325,00	33
Abs.- Veränderung	-23,61	18,83	-24,00	-58,00	36,00	33
%-Veränderung	-10,25	9,19	-9,39	-26,85	20,22	33
<b>Placebo</b>						
Leertest	239,97	39,23	234,00	164,00	330,00	34
Retest 1	263,76	25,11	262,00	192,00	337,00	34
Abs.- Veränderung	-23,79	20,44	-19,00	-85,00	9,00	34
%-Veränderung	-10,83	10,34	-9,25	-39,63	3,31	34
<b>Flunitrazepam</b>						
Leertest	247,53	31,95	245,50	174,00	301,00	32
Retest 1	259,25	32,14	256,50	187,00	333,00	32
Abs.- Veränderung	-11,72	20,65	-18,50	-38,00	42,00	32
%-Veränderung	-5,13	8,16	-7,08	-18,18	14,63	32

Tab. 33: DT / Richtige Reaktionen (Studienteil 1)

### *Falsche Reaktionen*

Falsche Reaktionen zeigen die Tendenz zu Verwechslungen an. Zu falschen Reaktionen kommt es dann, wenn es dem Probanden nicht mehr gelingt, die adäquate Reaktion vor der Einwirkung konkurrierender irrelevanter Reaktionen abzuschirmen. Damit steht diese Variable in enger Beziehung zur Aufmerksamkeitsfunktion. Eine hohe Zahl falscher Reaktionen ist als Indikator für eine Beeinträchtigung der Aufmerksamkeit zu betrachten. Bezogen auf die Gesamtnorm ist die Ausgangslage der drei Teilstichproben leicht unterdurchschnittlich. Der mittleren Anzahl falscher Reaktionen sind folgende Baseline Prozentränge zuzuordnen: Baldrian-Extrakt-Gruppe (M=12.52) PR 40, Placebo-Gruppe (M=12.68) PR 40, Flunitrazepam-Gruppe (M=15.25) PR 30.

Die Positionsbestimmung in der Normstichprobe belegt eine noch vergleichbare Ausgangslage der Medikationsgruppen. Die schlechtere Ausgangslage der

Flunitrazepam-Gruppe ist zu vernachlässigen. Die berechneten p-Werte lassen keine statistisch signifikanten Gruppenunterschiede in der Ausgangslage erkennen (s. Tab. 42). Die mittlere Anzahl falscher Reaktionen nimmt in allen drei Medikationsgruppen von Leertest zu Retest 1 geringfügig zu. Dieses Ergebnis steht im Zusammenhang mit der Steigerung in der Variablen "Anzahl der bearbeiteten Reize". Die mittlere Gesamtzahl der bearbeiteten Reize nimmt in allen drei Gruppen bei der ersten Folgetestung zu, was mit einem Anstieg des Fehleranteils verbunden ist. Das Ausmaß des Fehleranstiegs ist in den Gruppen vergleichbar, wobei die Zunahme falscher Reaktionen in der Baldrian-Extrakt-Gruppe absolut gesehen geringer ausfällt als in den Vergleichsgruppen. Mit durchschnittlich 13.15 falschen Reaktionen (Retest 1) liegt das Leistungsniveau der Baldrian-Extrakt-Gruppe oberhalb der Placebo-Gruppe (M=14.06) und der Flunitrazepam-Gruppe (M=16.28). Falsche Reaktionen und die Tendenz zu Verwechslungen nehmen im Vergleich zum Leertest in allen drei Medikationsgruppen beim ersten Retest zu. Der Anstieg falscher Reaktionen weist isoliert betrachtet auf eine mögliche geringe Beeinträchtigung der Aufmerksamkeit hin. Die Ergebnisse der anderen Testvariablen stützen dies jedoch nicht. Auch ist die Zunahme der DT-Fehlreaktionen im Retest ein häufig gemessenes Testresultat in anderen Studien. Eine Erklärung könnte sein, dass die Probanden beim Retest unbefangener sind und weniger Angst haben, Fehler zu machen. Die Zunahme des Fehleranteils ist in allen drei Medikationsgruppen verhältnismäßig gering und verteilt sich gleichmäßig auf die Gruppen, so dass keine psychologisch relevante Beeinträchtigung der Aufmerksamkeitsfunktion festzustellen ist. Statistisch signifikante Gruppenunterschiede sind nicht objektivierbar (Placebo vs. Baldrian-Extrakt  $p=0.9199$ ; Placebo vs. Flunitrazepam  $p=0.7086$ ; Flunitrazepam vs. Baldrian-Extrakt  $p=0.7721$ ).

Medikation	M	S	Median	Min	Max	N
<b>Baldrian</b>						
Leertest	12,52	11,69	9,00	,00	60,00	33
Retest 1	13,15	10,85	11,00	,00	45,00	33
Abs.-	-,64	7,52	,00	-18,00	17,00	33
Veränderung	-17,93	79,71	,00	-250,00	100,00	33
%-Veränderung						
<b>Placebo</b>						
Leertest	12,68	8,13	11,50	2,00	37,00	34
Retest 1	14,06	10,56	10,50	1,00	36,00	34
Abs.-	-1,38	6,79	-1,00	-17,00	9,00	34
Veränderung	-15,93	65,57	-3,13	-242,86	80,00	34
%-Veränderung						
<b>Flunitrazepam</b>						
Leertest	15,25	21,43	8,00	1,00	125,00	32
Retest 1	16,28	22,69	10,00	1,00	131,00	32
Abs.-	-1,03	6,21	-,50	-22,00	20,00	32
Veränderung	-16,62	63,91	-2,40	-314,29	90,91	32
%-Veränderung						

Tab. 34: DT / Falsche Reaktionen (Studienteil 1)

### *Ausgelassene Reaktionen*

Diese Variable beschreibt, ob Reaktionen unter Zeitdruck abgebrochen werden. Sie liefert damit einen Hinweis auf Resignationstendenzen des Probanden.

Bezogen auf die Gesamtnorm ist die Ausgangslage der drei Teilstichproben als unterdurchschnittlich einzustufen: Baldrian-Extrakt-Gruppe PR 35, Placebo-Gruppe PR 35, Flunitrazepam-Gruppe PR 25. Die Positionsbestimmung in der Normstichprobe belegt eine noch vergleichbare Ausgangslage der Medikationsgruppen - die Range beträgt 10 Prozentränge. Die schlechtere Ausgangslage der Flunitrazepam-Gruppe ist zu vernachlässigen. Die berechneten p-Werte lassen keine statistisch signifikanten Gruppenunterschiede in der Ausgangslage erkennen (s. Tab. 42).

Hinsichtlich der ausgelassenen Reaktionen ist bei allen drei Teilstichproben im Retest 1 im Mittel eine geringe Leistungssteigerung messbar. Die Standardabweichungen in den Teilstichproben sind im Verhältnis zu den Mittelwerten relativ groß (ca. die Hälfte des Gruppenmittelwertes). Die Leistungsfähigkeit innerhalb der Kollektive ist sehr unterschiedlich, was auch die Extrema verdeutlichen. Die Standardabweichungen der Medikationsgruppen bewegen sich aber in vergleichbarer Größenordnung. Die Mediane (absolute Veränderungen) der einzelnen Behandlungsgruppen belegen (1,00/0,50/1,00), das sich jeweils die Hälfte der Probanden verbessert bzw. verschlechtert hat. Die in den Mittelwerten bei allen drei Medikationsgruppen erkennbare Abnahme ausgelassener Reaktionen zeigt sich im Median nur bei zwei Teilstichproben (Placebo, Baldrian-Extrakt). Ein geringer Anstieg des Median ist in der Flunitrazepam-Gruppe feststellbar. Der Median relativiert den Mittelwert und verdeutlicht, daß sich die Testleistung in der Flunitrazepam-Gruppe bei Retest 1 leicht verschlechtert.

Die berechneten p-Werte für die durchschnittlichen absoluten Veränderungen von Leertest zu Retest 1 belegen keine statistisch signifikanten Gruppenunterschiede (Baldrian-Extrakt vs. Placebo  $p=0.8112$ ; Baldrian-Extrakt vs. Flunitrazepam  $p=0.1000$ ; Flunitrazepam vs. Placebo  $p=0.6663$ ). Die Ergebnisse der Flunitrazepam-Gruppe sind mit denen der beiden anderen Gruppen vergleichbar.

Zusammenfassend ist in den Medikationsgruppen Baldrian-Extrakt und Placebo eine Abnahme ausgelassener Reaktionen zu beobachten. Resignationstendenzen sind bei Besuch 2 nicht festzustellen und es werden weniger Reaktionen abgebrochen. Diese Aussage ist für die Flunitrazepam-Gruppe nur bedingt gültig, da sich diese Tendenz nur im Mittelwert und nicht im Median widerspiegelt.

Medikation	M	S	Median	Min	Max	N
<b>Baldrian</b>						
Leertest	12,88	7,44	13,00	2,00	33,00	33
Retest 1	11,88	6,55	11,00	2,00	25,00	33
Abs.- Veränderung	1,00	6,18	1,00	-13,00	12,00	33
%-Veränderung	-19,04	84,31	9,09	-250,00	84,62	33
<b>Placebo</b>						
Leertest	12,56	5,56	12,00	2,00	24,00	34
Retest 1	11,71	5,97	10,00	3,00	25,00	34
Abs.- Veränderung	,85	5,96	,50	-16,00	15,00	34
%-Veränderung	-5,91	68,45	3,33	-320,00	71,43	34
<b>Flunitrazepam</b>						
Leertest	14,19	6,94	12,00	2,00	31,00	32
Retest 1	13,00	5,15	12,50	2,00	24,00	32
Abs.- Veränderung	1,19	5,58	1,00	-10,00	17,00	32
%-Veränderung	-10,35	76,80	13,39	-350,00	77,78	32

Tab. 35: DT / Ausgelassene Reaktionen (Studienteil 1)

### *Bearbeitete Reize*

Diese Testvariable korrespondiert mit der Variable „Reaktionszeit“, da beide in einem unmittelbaren Zusammenhang stehen. Die Baselinewerte der drei Medikationsgruppen liegen eng beieinander und sind somit gut vergleichbar, wobei die Ausgangslage der Flunitrazepam-Gruppe leicht oberhalb des Niveaus der Vergleichsgruppen einzuordnen ist. Standardabweichungen und Extrema sind in der Ausgangslage vergleichbar. Die berechneten p-Werte lassen keine statistisch signifikanten Gruppenunterschiede erkennen (s. Tab. 42).

In allen drei Medikationsgruppen wurde eine deutliche Leistungssteigerung in der Variablen "Anzahl bearbeiteter Reize" von Leertest zu Retest 1 gemessen. Der Leistungszuwachs (mittlere absolute Veränderung) der Gruppen Baldrian-Extrakt (M=-22.64) und Placebo (M=-22.26) ist fast identisch. Die Steigerung des Leistungsumfangs der Flunitrazepam-Gruppe (M=-10.69) bleibt deutlich hinter den beiden anderen Medikationsgruppen zurück. Die berechneten p-Werte belegen folglich statistisch signifikante Gruppenunterschiede. Hinsichtlich der mittleren absoluten Veränderungen von der Aufnahme zur Folgetestung ergeben sich für den Vergleich Placebo vs. Flunitrazepam  $p=0.0013$  und für den Vergleich Placebo vs. Baldrian-Extrakt  $p=0.0208$ . Die Unterschiede sind auf dem 1%-Niveau bzw. auf dem 5%-Niveau signifikant. Die Gegenüberstellung Baldrian-Extrakt vs. Placebo zeigt keinen statistisch relevanten Unterschied ( $p=0.3401$ ).

Vom Übungseffekt profitieren die Gruppen Baldrian-Extrakt und Placebo in gleichem Umfang, während der Leistungszuwachs in der Flunitrazepam-Gruppe vergleichsweise nur halb so groß ausfällt und auch statistisch signifikant hinter den Ergebnissen der Vergleichsgruppen zurückbleibt. Die Ergebnisse dieser Variablen stützen und ergänzen die Ergebnisse zur DT-Hauptzielvariablen.

<b>Medikation</b>	<b>M</b>	<b>S</b>	<b>Median</b>	<b>Min</b>	<b>Max</b>	<b>N</b>
<b>Baldrian</b>						
Leertest	257,48	31,68	251,00	193,00	334,00	33
Retest 1	280,12	31,50	277,00	180,00	340,00	33
Abs.- Veränderung	-22,64	12,90	-24,00	-42,00	17,00	33
%-Veränderung	-9,01	5,66	-9,16	-17,60	8,63	33
<b>Placebo</b>						
Leertest	257,65	34,22	253,50	193,00	340,00	34
Retest 1	279,91	32,95	281,50	218,00	351,00	34
Abs.- Veränderung	-22,26	14,78	-21,00	-64,00	-1,00	34
%-Veränderung	-9,02	6,55	-8,18	-25,39	-,34	34
<b>Flunitrazepam</b>						
Leertest	265,75	28,61	268,00	196,00	311,00	32
Retest 1	276,44	29,07	278,50	202,00	341,00	32
Abs.- Veränderung	-10,69	16,26	-15,50	-30,00	36,00	32
%-Veränderung	-4,22	5,97	-5,53	-13,45	11,61	32

Tab. 36: DT / Bearbeitete Reize (Studienteil 1)

## Studienteil 2

### *Reaktionszeit (Median) im Modus "Adaptiv" (Hauptzielvariable des DT)*

Hinsichtlich relevanter Parameter (Standardabweichung, Extrema, Median) sind die beiden Kollektive gut vergleichbar. Der berechnete p-Wert ( $p=0.7357$ ) belegt die gute Vergleichbarkeit in der Baseline. Etwaige marginale Messwertunterschiede sind psychologisch zu vernachlässigen. Gegenüber der Ausgangslage (Leertest) ist in beiden Medikationsgruppen bei der Abschlussuntersuchung (Retest 2) eine Leistungssteigerung messbar. Die Range zwischen den beiden Medikationsgruppen bezüglich des Hauptzielkriteriums beträgt 4.34 ms und bleibt somit deutlich unterhalb der zulässigen Toleranzgrenze.

Die Ergebnisse zeigen, dass sich die Leistungsfähigkeit beider Gruppen bei Retest 2 in vergleichbarem Umfang verbessert hat. Die berechneten p-Werte belegen keine statistisch signifikanten Gruppenunterschiede. Der p-Wert für die durchschnittlichen absoluten Veränderungen von Leertest zu Retest 2 beträgt  $p=0.4481$  (das 95%-Konfidenzintervall liegt bei  $[-10.2712, 18.9540]$ ) und der p-Wert für die mittleren prozentualen Veränderungen beträgt  $p=0.4317$ . Unter Baldrian-Extrakt ist die Reaktionszeit (Median) nachweislich um maximal 10 ms schlechter als unter Placebo (Untergrenze des Konfidenzintervalls), bei einer gemeinsamen Standardabweichung um 35 ms. Mit den vorliegenden Daten ist die Äquivalenz zwischen Baldrian-Extrakt und Placebo nachgewiesen. Die Auswirkungen einer 14tägigen Baldrian-Therapie auf das Reaktionsvermögen unterscheiden sich offensichtlich nicht von denen eines Placebos.

Medikation	M	S	Median	Min	Max	N
<b>Baldrian</b>						
Leertest	793,26	77,96	791,00	630,00	1006,00	47
Retest 2	724,60	63,28	720,00	574,00	909,00	47
Abs.- Veränderung	68,66	36,77	76,00	-7,00	135,00	47
%-Veränderung	8,47	4,36	9,42	-1,07	15,43	47
<b>Placebo</b>						
Leertest	799,00	86,95	801,50	625,00	1003,00	44
Retest 2	734,68	77,80	724,50	598,00	929,00	44
Abs.- Veränderung	64,32	33,35	59,00	-4,00	136,00	44
%-Veränderung	7,93	3,90	7,37	-,63	15,42	44

Tab. 37: DT / Median der Reaktionszeit in ms (Studienteil 2)

### *Richtige Reaktionen*

Die Positionsbestimmung in der Normstichprobe zeigt eine vergleichbare Ausgangslage der Medikationsgruppen: Baldrian-Extrakt-Gruppe PR 65, Placebo-Gruppe PR 60. Auch bezüglich der Parameter Standardabweichung, Median und Extrema bestehen keine relevanten Gruppenunterschiede. Der berechnete p-

Wert ( $p=0.4187$ ) belegt die gute Vergleichbarkeit der Medikationsgruppen in der Ausgangslage.

Die durchschnittliche Anzahl der richtigen Reaktionen nimmt von der Aufnahme-  
testung zu Retest 2 in beiden Medikationsgruppen zu. Standardabweichungen  
und Extrema zeigen keine Gruppenunterschiede. Die Range beträgt beim Leer-  
test 5.21 und beim zweiten Retest 6.2 richtige Reaktionen. Die Testergebnisse  
der Probanden in den Gruppen Baldrian-Extrakt und Placebo stimmen bezüglich  
Ausgangslage und Leistungszuwachs fast überein. Eine Beeinträchtigung ist  
unter Baldrian-Extrakt nicht feststellbar. Die Leistungsfähigkeit der Probanden  
hat sich in den Gruppen in vergleichbarem Umfang gesteigert. Die mittels Mann-  
Whitney-U-Test vorgenommenen Gruppenvergleiche ergaben hinsichtlich der  
mittleren absoluten Veränderungen ( $p=0.6056$ ) sowie der durchschnittlichen  
prozentualen Veränderungen ( $p=0.6855$ ) richtiger Reaktionen keine signifikanten  
Gruppenunterschiede.

Medikation	M	S	Median	Min	Max	N
<b>Baldrian</b>						
Leertest	244,57	33,71	244,00	166,00	332,00	47
Retest 2	278,09	30,43	276,00	189,00	346,00	47
Abs.- Veränderung	-33,51	21,73	-35,00	-72,00	24,00	47
%-Veränderung	-14,52	9,79	-14,40	-33,80	7,23	47
<b>Placebo</b>						
Leertest	239,36	37,86	238,00	164,00	330,00	44
Retest 2	271,89	32,22	272,50	198,00	349,00	44
Abs.- Veränderung	-32,52	18,75	-32,00	-64,00	2,00	44
%-Veränderung	-14,60	9,70	-13,17	-39,02	,69	44

Tab. 38: DT / Richtige Reaktionen (Studienteil 2)

### *Falsche Reaktionen*

Die Positionsbestimmung in der Normstichprobe ergibt eine vergleichbare Aus-  
gangslage der Medikationsgruppen bezüglich dieser Testvariablen: Baldrian-  
Extrakt-Gruppe PR 32, Placebo-Gruppe PR 35. Bei den Parametern Standard-  
abweichung und Extrema zeigen sich Gruppenunterschiede. Leistungsunter-  
schiede innerhalb der Baldrian-Extrakt-Gruppe führen zu einer relativ großen  
Streuung. Die Extrema verdeutlichen die große Spannweite der gemessenen  
Fehlreaktionen in der Baldrian-Extrakt-Gruppe und die vergleichsweise einheit-  
lichere Leistungsfähigkeit des Probandenkollektivs der Placebogruppe. Diese  
Gruppenunterschiede schränken jedoch die grundsätzliche Vergleichbarkeit  
nicht ein. Die sich bei Retest 2 ergebenden Tendenzen in Richtung Zunahme  
(Baldrian-Extrakt) bzw. Abnahme der Fehlreaktionen (Placebo) sind zu gering  
ausgeprägt, um eine unterschiedliche Entwicklung der Gruppen ableiten zu  
können. Insgesamt betrachtet ist der Anteil von Fehlreaktionen und die Tendenz  
zu Verwechslungen im Gruppenvergleich kaum verändert.

Die Range zwischen den Gruppen beträgt bei Retest 2 im Mittel 1.48 Reaktion. Die in der Ausgangslage unterschiedliche Standardabweichung findet sich auch beim Retest 2 wieder. Die Extrema zeigen, dass die Leistungsunterschiede innerhalb des Baldrian-Extrakt-Kollektivs fast unverändert fortbestehen. Psychologisch relevante Gruppenunterschiede sind nicht festzustellen. Die berechneten p-Werte für die mittleren absoluten Veränderungen ( $p=0.5066$ ) und der durchschnittlichen prozentualen Veränderungen ( $p=0.68805$ ) zeigen keine signifikanten Gruppenunterschiede.

Medikation	M	S	Median	Min	Max	N
<b>Baldrian</b>						
Leertest	14,17	19,64	9,00	,00	125,00	47
Retest 2	14,28	20,03	9,00	1,00	128,00	47
Abs.- Veränderung	-,11	6,58	,00	-15,00	17,00	47
%-Veränderung	-10,27	64,61	,00	-275,00	87,50	47
<b>Placebo</b>						
Leertest	13,16	7,86	12,00	2,00	37,00	44
Retest 2	12,80	7,92	11,00	,00	27,00	44
Abs.- Veränderung	,36	7,50	1,50	-18,00	16,00	44
%-Veränderung	-8,14	73,88	18,75	-257,14	100,00	44

Tab. 39: DT / Falsche Reaktionen (Studienteil 2)

### *Ausgelassene Reaktionen*

Bezogen auf die Gesamtnorm ist die Ausgangslage der Gruppen als unterdurchschnittlich einzustufen: Baldrian-Extrakt-Gruppe PR 30, Placebo-Gruppe PR 30. Bezüglich Mittelwert, Standardabweichung, Median und Extrema differieren die Messwerte beider Gruppen nur geringfügig. Etwaige Unterschiede sind psychologisch zu vernachlässigen. Die Gruppen sind gut vergleichbar ( $p=0.9810$ ).

Hinsichtlich der ausgelassenen Reaktionen ist eine geringe Leistungssteigerung im Retest 2 messbar. In beiden Gruppen ist eine unbedeutende Abnahme ausgelassener Reaktionen zu beobachten. Resignationstendenzen sind nicht festzustellen und es werden weniger Reaktionen abgebrochen. Die berechneten p-Werte belegen keine statistisch signifikanten Gruppenunterschiede. Der auf die durchschnittlichen absoluten Veränderungen bezogene p-Wert beträgt  $p=0.5479$ . Auch hinsichtlich der durchschnittlichen prozentualen Veränderungen zeigt sich bei dieser Variable kein statistisch auffälliger Gruppenunterschied ( $p=0.5434$ ).

Medikation	M	S	Median	Min	Max	N
<b>Baldrian</b>						
Leertest	13,53	7,31	13,00	2,00	33,00	47
Retest 2	10,68	6,22	9,00	2,00	37,00	47
Abs.- Veränderung	2,85 -,16	7,52 80,27	2,00 21,43	-25,00 -300,00	18,00 75,00	47 47
%-Veränderung						
<b>Placebo</b>						
Leertest	13,20	5,84	12,00	2,00	26,00	44
Retest 2	11,20	5,07	10,00	2,00	23,00	44
Abs.- Veränderung	2,00 5,24	5,18 45,71	2,00 18,33	-7,00 -100,00	14,00 77,78	44 44
%-Veränderung						

Tab. 40: DT / Ausgelassene Reaktionen (Studienteil 2)

### *Bearbeitete Reize*

Die Baselinewerte der Medikationsgruppen liegen eng beieinander und sind somit gut vergleichbar. Unterschiede im Median und den Extrema sind zu vernachlässigen. Statistisch zeigt sich kein relevanter Gruppenunterschied in der Ausgangslage ( $p=0.4481$ ).

In beiden Gruppen wurde beim Retest 2 eine deutliche Leistungssteigerung gemessen. Ausgangslage und Leistungszuwachs beider Gruppen sind fast identisch. Die Leistungsstärke der Gruppen hat sich auf hohem Niveau konsolidiert. Psychologisch relevante Gruppenunterschiede sind nicht festzustellen. Auch statistisch zeigen sich keine signifikanten Gruppenunterschiede.

Da die Bearbeitungsdauer der Reize so geregelt ist, dass sie der durchschnittlichen Bearbeitungszeit des Probanden entspricht, ist die subjektive Belastung für den Probanden gleichgehalten. Eine Aufschlüsselung der richtigen Reaktionen in "zeitgerecht Richtige" und "verspätet Richtige" ist nicht mehr sinnvoll, da das Verhältnis dieser beiden Variablen zueinander im Wesentlichen durch den adaptiven Algorithmus bestimmt wird. Eine detaillierte Betrachtung dieser beiden Variablen erübrigt sich deshalb.

Die berechneten p-Werte für die mittleren absoluten Veränderungen ( $p=0.5646$ ) sowie für die durchschnittlichen prozentualen Veränderungen ( $p=0.6972$ ) ergeben keine statistisch abzusichernden Hinweise auf relevante Unterschiede zwischen den beiden Medikationsgruppen oder auf bedeutsame Messwertveränderungen innerhalb der Medikationsgruppen im Studienverlauf.

<b>Medikation</b>	<b>M</b>	<b>S</b>	<b>Median</b>	<b>Min</b>	<b>Max</b>	<b>N</b>
<b>Baldrian</b>						
Leertest	258,74	40,72	254,00	185,00	393,00	47
Retest 2	292,36	35,11	288,00	211,00	383,00	47
Abs.- Veränderung	-33,62	21,86	-36,00	-87,00	21,00	47
%-Veränderung	-13,88	9,15	-13,64	-34,25	6,25	47
<b>Placebo</b>						
Leertest	252,52	39,18	245,50	180,00	357,00	44
Retest 2	284,68	34,32	284,00	202,00	372,00	44
Abs.- Veränderung	-32,16	20,54	-32,00	-68,00	9,00	44
%-Veränderung	-13,61	9,19	-12,83	-30,36	2,61	44

Tab. 41: DT / Bearbeitete Reize (Studienteil 2)

Determinationsstest (DT)	MANN-WHITNEY-U-Test p-Werte			
	Studienteil 1			Studienteil 2
	Baldrian vs. Placebo	Baldrian vs. Flunitrazepam	Placebo vs. Flunitrazepam	Baldrian vs. Placebo
<b>Median der Reaktionszeit (Hauptzielvariable)</b>				
Leertest		0.0754	0.1638	0.7357
Retest 1 bzw. 2	0.8360	0.9425	0.9846	0.6825
abs. Veränderung	0.9850	0.0001	0.0115	0.4481
% - Veränderung	0.1184 0.1229	0.0001	0.0111	0.4317
<b>Bearbeitete Reize</b>				
Leertest	0.9950	0.2101	0.2040	0.5434
Retest 1 bzw. 2	0.7442	0.5459	0.7193	0.3822
abs. Veränderung	0.3401	0.0013	0.0208	0.4844
% - Veränderung	0.3501	0.0014	0.0195	0.6057
<b>Richtige Reaktionen</b>				
Leertest	0.8508	0.4704	0.3077	0.4178
Retest 1 bzw. 2	0.5940	0.2454	0.6907	0.2835
abs. Veränderung	0.3729	0.0166	0.1249	0.6056
% - Veränderung	0.5184	0.0209	0.1205	0.6855
<b>Falsche Reaktionen</b>				
Leertest	0.5136	0.7375	0.7577	0.2117
Retest 1 bzw. 2	0.7968	0.7031	0.9129	0.2975
abs. Veränderung	0.9199	0.7721	0.7086	0.5066
% - Veränderung	0.8222	0.5493	0.8120	0.6805
<b>Ausgelassene Reaktionen</b>				
Leertest	0.9599	0.4618	0.4557	0.9810
Retest 1 bzw. 2	0.8850	0.2778	0.2213	0.3450
abs. Veränderung	0.8112	1.0000	0.6663	0.5479
% - Veränderung	0.7971	0.8183	0.8373	0.5434

Tab. 42: Determinationsstest (DT) - p-Werte -  
(Leertest und Retest 1 für Studienteil 1 / Leertest und Retest 2 für Studienteil 2)

### 8.3 Schlussfolgerungen aus den Ergebnissen psychometrischer Verfahren

#### Kölner Farbpräferenztest (KFPT)

##### *Fazit Studienteil 1*

Die Baselinebefunde des Gesamtkollektivs waren durch eine große Range der Messwerte geprägt; Streuung und Extrema lagen zum Teil über dem arithmetischen Mittel (Cluster 1, 1b, 5b, 10c), was auf große interindividuelle Varianzen im Farbwahlverhalten hinweist, welche die Generalisierbarkeit der erhobenen Ergebnisse (Mittelwerte) einschränkt. Trotzdem bleibt als Befund festzuhalten, dass bei den schönen Pyramiden das Cluster 2 und der Farbton 1b (rot) bevorzugt wurden, während bei den hässlichen Pyramiden der Farbton 10c (schwarz) eindeutig präferiert wurde.

In der Placebo-Kontrollgruppe wurden bei der Folgetestung keine psychologisch relevanten Messwertveränderungen gegenüber der Ausgangslage objektiviert. Auch in den beiden Verumgruppen wurden nur wenige ausgeprägte Veränderungen der Farbpräferenzen beim Retest evaluiert. Unter Flunitrazepam zeigte sich bei der Wahl der schönen Pyramiden im Cluster 2 eine bei der Folgetestung signifikante Veränderung ( $p=0.016^*$ ) und bei Farbton 10c ein Trend ( $p=0.061+$ ). Unter Baldrian-Extrakt wurde bei der Wahl der hässlichen Pyramiden im Cluster 1 signifikante Veränderungen bei Retest 1 ( $p=0.019^*$ ) und bei Farbton 1b ein Trend ( $p=0.063+$ ) gemessen.

Vor dem Hintergrund dieser Ergebnisse erscheint die mögliche Beeinflussung des Farbwahlverhaltens durch die beiden aktiven pharmakologischen Substanzen unwahrscheinlich. Die im DT objektivierete Beeinträchtigung der kognitiven Leistungsfähigkeit durch das Benzodiazepin korrespondiert nicht mit dem KFPT-Befund. Farbwahlverhalten und emotionale Befindlichkeit wurden durch Flunitrazepam offensichtlich nicht verändert.

Im Gruppenvergleich Placebo vs. Baldrian-Extrakt ergaben sich insgesamt keine bedeutsamen Messwertunterschiede. Die Gegenüberstellung von Flunitrazepam und Baldrian-Extrakt zeigte nur bei den schönen Pyramiden für das Cluster 2 bei Retest 1 ( $p=0.023^*$ ) und bei den hässlichen Pyramiden für das Cluster 1 in der Ausgangslage ( $p=0.024^*$ ) signifikante Unterschiede. Für Flunitrazepam und Placebo ergab sich bei den schönen Pyramiden ein signifikanter Unterschied bei Retest 1 im Cluster 2 ( $p=0.40^*$ ) und für den Farbton 5b ein Trend ( $p=0.074+$ ). Insgesamt wurden nach einmaliger abendlicher Einnahme der Medikation bei der Folgetestung am darauf folgenden Morgen weder innerhalb der drei Behandlungsgruppen noch im Gruppenvergleich gravierende Veränderungen des Farbwahlverhaltens deutlich.

### *Fazit Studienteil 2*

Medikationsunterschiede (Baldrian-Extrakt/Placebo) hinsichtlich des Farbwahlverhaltens ließen sich nicht belegen. Ausgeprägte Veränderungen der Farbpräferenzen nach 2wöchiger Therapie waren nur bei Cluster 3 - bei schönen und hässlichen Pyramiden – in beiden Medikationsgruppen festzustellen. Diese Befunde wurden nicht durch die Ergebnisse anderer Cluster bzw. Einzelfarben bestätigt. Insgesamt kam es innerhalb der Medikationsgruppen nur zu geringen Messwertveränderungen.

In der Gegenüberstellung von Baldrian-Extrakt und Placebo zeigten sich nach 14tägiger Therapie keine relevanten Unterschiede im Farbwahlverhalten. Ein medikamentöser Einfluss durch das Baldrianpräparat auf die emotionale Befindlichkeit ist auszuschließen.

### **Semantisches Eindrucksdifferential (ED)**

#### *Fazit Studienteil 1*

Gegenüber der Ausgangslage im Gesamtkollektiv (Baseline) veränderte sich nach einmaliger Einnahme der Medikationen (Retest 1) das Farberleben – unabhängig von der Behandlungsgruppe - hinsichtlich der E/V/P-Dimensionen nur unwesentlich.

Im Vergleich der drei Medikationsgruppen ergaben sich nur bei Cluster 3 für den Vergleich Baldrian-Extrakt vs. Placebo in der Dimension „Potenz“ ( $p=0.020^*$ ) und beim Farbton 5c für den Vergleich Baldrian-Extrakt vs. Flunitrazepam ( $p=0.015^*$ ) signifikante Gruppenunterschiede. Diese beiden Befunde alleine sind jedoch kein Indiz für ein generell verändertes Farberleben. Insgesamt zeigten die Gruppenvergleiche keine statistisch signifikanten und klinisch relevanten Veränderungen in der Farbwahrnehmung der Probanden.

#### *Fazit Studienteil 2*

In Relation zu den Ausgangsbefunden im Gesamtkollektiv veränderte sich nach 14tägiger Einnahme der Medikationen das Farberleben hinsichtlich der E/V/P-Dimensionen nur unwesentlich. Klinisch relevante Unterschiede zwischen den Behandlungsgruppen waren nicht objektivierbar. Eine Beeinträchtigung des Farberlebens unter Baldriantherapie ist damit unwahrscheinlich.

## Determinationstests (DT)

### *Studienteil 1*

In drei Variablen des Determinationstests – Reaktionszeit (Median), Anzahl der richtigen Reaktionen, Anzahl der bearbeiteten Reize - war der Leistungszuwachs der Flunitrazepam-Gruppe signifikant geringer als in den beiden Vergleichsgruppen. Dabei waren die unter Flunitrazepam aufgetretenen Einschränkungen der Leistungsfähigkeit psychologisch relevant. Im DT zeigte sich bei keiner Variablen ein statistisch auffälliger Unterschied zwischen Placebo und Baldrian-Extrakt. Nach einmaliger vorabendlicher Einnahme waren beim Retest am Morgen des Folgetages Reaktionsgüte und Leistungsfähigkeit der Probanden beider Medikationsgruppen in der Hauptzielvariablen (Median der Reaktionszeit) gleichwertig.

Die einmalige Einnahme eines Baldrian-Extraktes führte zu keiner Beeinträchtigung des Reaktionsvermögens, der Aufmerksamkeit oder der Koordinationsleistung. Ein vigilanzmindernder Hangover-Effekt wurde nicht objektiviert. Hinsichtlich des erzielten Übungsfortschrittes waren Baldrian-Extrakt und Placebo-Kontrollgruppe in ihrer Leistungsfähigkeit vergleichbar.

Für das Phytopharmakon ließen sich keine sedativen Residualwirkungen belegen, wohl aber für die chemisch definierte Substanz. Unter Flunitrazepam war das Reaktionsvermögen der Probanden in der DT-Hauptzielvariablen beeinträchtigt. Auch bei der niedrigen Dosierung von 1 mg/die Flunitrazepam konnte ein Hangover-Effekt bei 25% der Probanden nachgewiesen werden, auch war der Übungseffekt signifikant geringer als in den Vergleichsgruppen.

Verkehrsrelevante kognitive Leistungsminderungen am Folgetag sind auch nach nur einmaliger vorabendlicher Flunitrazepam-Einnahme nicht auszuschließen.

Die DT-Ergebnisse für Placebo und Baldrian-Extrakt korrespondieren mit den farbpsychologischen Resultaten von KFPT und ED. In allen Verfahren zeigten sich in beiden Behandlungsgruppen keine Hinweise auf kognitive Beeinträchtigungen bzw. emotionale Veränderungen. Die im DT objektivierete Beeinträchtigung des Reaktionsvermögens unter Flunitrazepam wurde durch die farbpsychologischen Ergebnisse nicht gestützt. Im ED und KFPT zeigten sich keine bedeutsamen Veränderungen des Farberlebens oder der Farbpräferenzen.

### *Studienteil 2*

Die Studienergebnisse belegten keine Vigilanzminderung nach einer 14tägigen Baldriantherapie. Im DT wurden in keiner Variablen signifikante Unterschiede zu Placebo objektiviert. Das Konfidenzintervall schließt für die variable

Reaktionszeit (Median) eine relevante Beeinträchtigung unter Baldrian-Extrakt mit 97.5%iger Sicherheit aus. Die Äquivalenz zwischen Baldrian-Extrakt und Placebo konnte bestätigt werden. Das Ausmaß der gemessenen Leistungssteigerungen im Abschlusstest nach zweiwöchiger Behandlung war in beiden Untersuchungsgruppen vergleichbar.

DT-Ergebnisse und farbpsychologische Resultate korrespondieren miteinander. Sowohl kognitive Leistungsfähigkeit als auch emotionale Befindlichkeit werden durch das Baldrian-Präparat offensichtlich nicht beeinträchtigt.

## 9 ERGEBNISSE ZUR ARZNEIMITTELSICHERHEIT UND SCHLUSSFOLGERUNGEN

### 9.1 Unerwünschte Ereignisse (Adverse Events)

#### Studienteil I

Von der Mehrheit der Probanden (N=60) wurden nach einmaliger Einnahme der Studienmedikation keine Unverträglichkeiten berichtet. Baldrian-Extrakt (N=11) und Placebo (N=12) waren gleich gut verträglich, lediglich unter Flunitrazepam (N=19) traten relativ mehr unerwünschte Ereignisse auf.

Anzahl der UE	Baldrian	Placebo	Flunitrazepam	Total
0	23	22	15	60
1	7	9	8	24
2	2	3	3	8
3			5	5
4	1		2	3
5				0
6	1			1
7			1	1
<b>Total</b>	<b>34</b>	<b>34</b>	<b>34</b>	<b>102</b>

Tab. 43: Anzahl der UE pro Proband bei Besuch 2 (Studienteil 1)

Bei Visite 2 wurden von elf Probanden (32%) der Baldrian-Extrakt-Gruppe, zwölf (35%) der Placebo-Gruppe und 19 (56%) der Flunitrazepam-Gruppe mindestens ein UE angegeben. Mit Ausnahme der drei häufigsten UE Müdigkeit bzw. starke Müdigkeit, Kopfschmerzen und Schwindel wurde kein UE von mehr als drei Probanden einer Gruppe genannt. Müdigkeit oder starke Müdigkeit wurden unter Flunitrazepam am häufigsten angegeben (N=10, 29.4%), gefolgt von Baldrian-Extrakt (N=4, 11.8%) und Placebo (N=1, 2.9%). Auch Kopfschmerzen und Schwindel waren unter Flunitrazepam mit fünf bzw. vier Nennungen relativ häufiger als unter Baldrian-Extrakt (je eine Nennung) und Placebo (zwei bzw. null Nennungen). Bei den einzelnen UE sind die Zahlen jedoch zu klein, um daraus Gruppenunterschiede abzuleiten. Auffällig ist jedoch die insgesamt gegenüber beiden Referenzgruppen höhere Zahl von Probanden mit UE und Zahl der Nennungen unter Flunitrazepam.

<b>Unerwünschtes Ereignis (UE)</b>	<b>Baldrian</b>	<b>Placebo</b>	<b>Flunitrazepam</b>	<b>Total</b>
Müdigkeit	4	1	8	13
Kopfschmerzen	1	2	5	8
Schwindel	1		4	5
Dösigkeit	2		3	5
Verlangsamung	1		3	4
Schlafstörungen	3	1		4
Schwäche (allgemein)	2		1	3
Starke Müdigkeit			2	2
Konzentrationsstörungen			2	2
Übelkeit	1	1		2
Exzessive Träume			2	2
Einschlafschwierigkeiten		2		2
Geschmacksstörung	2			2
Juckreiz	1			1
Verschwommenes Sehen			1	1
Merkschwäche			1	1
Kopfdruckgefühl		1		1
Geschmacksstörung (unspezifisch)			1	1
Kolik (abdominal)			1	1
Denkverlangsamung			1	1
Taube Zunge	1			1
Völlegefühl			1	1
Kreislaufstörungen			1	1
Taubheitsgefühl (lokal)		1		1
Koordinationsstörungen			1	1
Magenbrennen			1	1
Rülpfen	1			1
Magendruck		1		1
Wirre Träume		1		1
Sturz		1		1
Gangunsicherheit			1	1
Gleichgültigkeitsgefühl	1			1
Atemwegsinfektion (obere)		1		1

<b>Unerwünschtes Ereignis (UE)</b>	<b>Baldrian</b>	<b>Placebo</b>	<b>Flunitrazepam</b>	<b>Total</b>
Rückenschmerzen		1		1
Verkürzte Schlafzeit			1	1
Unruhiger Schlaf		1		1
Kribbeln in den Gliedern		1		1
Sodbrennen			1	1
<b>Total</b>	<b>11</b>	<b>12</b>	<b>19</b>	<b>42</b>

Tab. 44: Unerwünschte Ereignisse – Studienteil 1 (Mehrfachnennungen möglich)

Schwerwiegende UE traten nicht auf. Der Schweregrad der meisten UE wurde als "gering" oder "mäßig" angegeben.

In einem möglichen oder wahrscheinlichen Zusammenhang mit der Medikation gesehen wurden unter Baldrian-Extrakt 1x Geschmacksstörung, taube Zunge und Müdigkeit (möglich), 1x Dösigkeit (wahrscheinlich), unter Placebo je 1x Müdigkeit, Kopfschmerzen, Schlafschwierigkeiten, Übelkeit, Taubheitsgefühl, Kribbeln, Kopfdruckgefühl und wirre Träume (möglich), unter Flunitrazepam je 1x Müdigkeit, Dösigkeit, Hypokinesie und Konzentrationsstörung (wahrscheinlich) sowie weitere 33 UE (möglich).

### Auswaschphase

Bei der Untersuchung am Tag 9 wurden UE erfasst, die in der medikationsfreien Zeit aufgetreten waren. Insgesamt wurden bei 24 Probanden (23.5%) UE beobachtet, wobei am häufigsten Kopfschmerzen genannt wurden. Ein kausaler Zusammenhang mit der am Tag 1 erfolgten einmaligen Medikationseinnahme war nicht gegeben.

### Studienteil 2

Insgesamt wurden im zweiten Studienteil 80 UE berichtet die sich gleichmäßig auf beide Behandlungsgruppen verteilen – 41 UE traten unter Baldrian-Extrakt und 39 UE unter Placebo auf (siehe Tab.46). Von den nach 14tägiger Behandlung (Besuch 4) genannten UE waren Kopfschmerzen (66x), allgemeiner Schwäche (17x), Übelkeit sowie Dösigkeit und Schwindel (je 16x), Atemwegsinfektion (12x) und "Müdigkeit" (7x) am häufigsten. Bei keiner der Beschwerden gab es einen Hinweis auf einen relevanten Unterschied zwischen den Medikationsgruppen. Schlafstörungen wurden nur in Einzelfällen berichtet.

Die als "sedativa-typischen" Begriffe zusammengefassten Nennungen erscheinen im Studienteil 2 relativ seltener als im Teil 1. Dösigkeit, Müdigkeit, Gang- oder Bewegungsunsicherheit und Konzentrationsstörung wurden insgesamt 22mal unter Baldrian-Extrakt und zehnmal unter Placebo genannt. Schwerwiegende UE traten nicht auf. Der Schweregrad der meisten UE wurde als "gering" oder "mäßig" eingestuft. Bis auf wenige Ausnahmen waren die dokumentierten Beschwerden bis zum Abschlussbesuch abgeklungen. Die o.g. UE wurden weder unter Baldrian-Extrakt noch unter Placebo in einem wahrscheinlichen Zusammenhang mit der Prüfmedikation gesehen. In einem möglichen Zusammenhang stehend wurden bewertet unter Baldrian-Extrakt 15 (14.1%) Nennungen [Dösigkeit (2x), Müdigkeit (2x), je 1x Kopfschmerzen, Geschmacksstörung, Mundtrockenheit, Zunge taub, Übelkeit, Schwindel, Kopfdruck, Schlafzeit verkürzt, Schwäche allgemein, Gangunsicherheit, Konzentrationsstörung], unter Placebo 10 (8.0%) Nennungen [Dösigkeit (2x), je 1x Kopfschmerzen, Mundtrockenheit, Schwindel, Kribbeln, Müdigkeit, Quaddeln, Appetitzunahme und Juckreiz]. Zu einem Therapieabbruch wegen UE kam es nur bei einer Probandin (Nr. 59, Placebo). Ausschlaggebend für den Abbruch der Studie war das Auftreten von Quaddeln an Hals und Beinen sowie unerträglicher Juckreiz.

Anzahl der UE	Baldrian	Placebo	Total
0	10	12	22
1	13	11	24
2	13	6	19
3	5	6	11
4	4	8	12
5	2	3	5
6	2	1	3
7	2	2	4
8		1	1
9		1	1
<b>Total</b>	<b>51</b>	<b>51</b>	<b>102</b>

Tab. 45: Anzahl der UE pro Proband bei Besuch 4 (Studienteil 2)

<b>Unerwünschtes Ereignis (UE)</b>	<b>Baldrian</b>	<b>Placebo</b>	<b>Total</b>
Kopfschmerzen	30	36	66
Schwäche (allgemein)	8	9	17
Übelkeit	8	8	16
Dösigkeit	10	6	16
Schwindel	8	8	16
Atemwegsinfektion (obere)	5	7	12
Müdigkeit	5	2	7
Gangunsicherheit	4	2	6
Rhinitis		4	4
Dymenorrhoe	1	2	3
Durchfall	2	1	3
Mundtrockenheit	1	1	2
Juckreiz	1	1	2
Kopfdruckgefühl	2		2
Ekzem		2	2
Virusinfekt	1	1	2
Bewegungsunsicherheit	2		2
Kribbeln in den Gliedern		2	2
Larynxschmerzen		2	2
Heuschnupfen		2	2
Geschmackstörung	1	1	2
GGT Anstieg		1	1
GPT Anstieg	1		1
Halsentzündung	1	1	2
Zahnschmerzen		1	1
Beinkrämpfe		1	1
Husten		1	1
Ausschlag (Erythematoses)		1	1
Tachykardie	1		1
Schlafstörungen		1	1
Quaddeln		1	1
Nervosität	1		1
Periphere Ischämie		1	1
Leukopenie	1		1

<b>Unerwünschtes Ereignis (UE)</b>	<b>Baldrian</b>	<b>Placebo</b>	<b>Total</b>
Herzrhythmusstörung		1	1
Gelenkschmerz		1	1
Atemnot		1	1
Vermehrtes Schwitzen		1	1
Saures Aufstoßen		1	1
Kreislaufstörungen	1		1
Schlafschwierigkeiten		1	1
Erbrechen		1	1
Zystitis	1		1
Kolik (abdominal)		1	1
Ischialgie		1	1
Konzentrationsstörungen	1		1
Blutplättchen (vermehrt)		1	1
Verkürzte Schlafzeit	1		1
Hämatom		1	1
Appetitzunahme		1	1
Herpes Labialis		1	1
Taube Zunge	1		1
Verdauungsbeschwerden	1		1
Alpträume	1		1
Entzündungsprozess (lokal)	1		1
Hornhautentzündung		1	1
Geschmacksstörung (unspezifisch)	1		1
Sprachstörung		1	1
Furunkulose	1		1
Ekzem (allergisch)	1		1
Verstopfung		1	1
Depression		1	1
Abgeschlagenheit	1		1
<b>Total</b>	<b>41</b>	<b>39</b>	<b>80</b>

Tab. 46: UE-Nennungen – Studienteil 2 (Mehrfachnennungen möglich)

## 9.2 Klinisches Labor

Die Labordaten lagen vollständig vor. Abweichungen vom Normbereich wurden durch + bzw. - markiert (zugrunde gelegt wurde der Normbereich des in die Durchführung involvierten Labors). Relevante Abweichungen von der Norm waren als Ausschlusskriterium definiert. Folglich wiesen keine an der Studie teilnehmenden Probanden bei der Aufnahme deutlich pathologische Werte auf.

Bei der ATP-Auswertung unberücksichtigt blieben Probanden, die bei der Aufnahme und Abschlussuntersuchung klinisch relevante Normabweichungen zeigten. Dies war definiert als Abweichungen > 10% bei Hb, Ery, Leuko und Thrombozyten. Relevante Abweichungen für Transaminasen und Kreatinin waren vorab nicht definiert. Sie wurden ex post mit über 100% bei den Transaminasen und Kreatinin > 2 mg/dl festgelegt.

Kein Proband musste aufgrund klinisch relevanter Normabweichungen von der ATP-Auswertung ausgeschlossen werden. Bei den meisten Probanden waren die Normabweichungen, soweit überhaupt aufgetreten, tolerabel. Vermutlich waren die beobachteten Normabweichungen zufällig, es gab keine Anhaltspunkte für diesbezügliche unerwünschte Arzneimittelwirkungen. Bei der Abschlussuntersuchung waren in den Laborbefunden nur selten Auffälligkeiten zu erkennen; es wurden nur wenige unerwünschte Ereignisse bzw. Nebenwirkungen dokumentiert.

## 9.3 Schlussfolgerungen zur Arzneimittelsicherheit

Sowohl das Phytopharmakon als auch das Benzodiazepin erwiesen sich als sichere und insgesamt gesehen gut verträgliche und tolerable Medikamente, wobei die Quote unerwünschter Ereignisse im ersten Studienteil höher war als im zweiten, was durch die Mitführung des synthetischen Arzneimittels begründet ist (Suggestiveffekt bei den Probanden).

Ein möglicher kausaler Zusammenhang mit berichteten Nebenwirkungen wurde nur für einen Probanden unter Baldrian-Extrakt objektiviert (Schwindelgefühl). Die sehr niedrige Häufigkeit unerwünschter Ereignissen unter Baldrian-Extrakt deckt sich mit den bisherigen Erfahrungsberichten (Hüther, 1989; Maisenbacher, 1992). Im Gegensatz hierzu wurden bei vier Probanden unter Flunitrazepam für Benzodiazepine typische Nebenwirkungen beobachtet (Schwindelgefühl, Müdigkeit, Hypokinesie und Konzentrationsschwäche). Die gefundene Nebenwirkungsrate für das Benzodiazepinpräparat entspricht den Erwartungen und steht im Einklang mit den Angaben der vorliegenden Literatur (Müller, 1990).

## 10 DISKUSSION

Das gewählte Studiendesign der randomisierten, placebokontrollierten Doppelblindstudie entsprach hinsichtlich Planung, Durchführung und Auswertung den aktuellen wissenschaftlichen Anforderungen und ist „state of the art“.

Die für pharmakopsychologische Studien gültigen ethischen und rechtlichen Anforderungen des deutschen Arzneimittelgesetzes (AMG) und die Regeln der „Good Clinical Practice“ (GCP) wurden erfüllt. Methodischen Forderungen an farbpsychologische Studien, wie z.B. Spezifizierung der Farben und mehrdimensionale Erfassung der Gefühlserlebnisse, wurde gleichfalls entsprochen.

Sämtliche Untersuchungen wurden gemäß den Planungsvorgaben durchgeführt, fehlende Angaben bzw. Messwerte waren aufgrund der sorgfältigen Dokumentation und des engmaschigen Monitorings selten. Protokollverstöße wurden nur in wenigen Ausnahmefällen dokumentiert und waren ohne Relevanz für die Aussagekraft der Studie. Die hohe Qualität bei Planung, Durchführung und Datenerhebung tragen maßgeblich dazu bei, die wissenschaftliche Aussagekraft der evaluierten Ergebnisse zu erhöhen.

Die Objektivierung möglicher medikamentöser Effekte auf Kognition und Emotion erfordert komplexe und hoch sensitive Messinstrumente. Die eingesetzten psychometrischen Verfahren repräsentieren entweder etablierte Testprozeduren für die Evaluierung der mentalen Leistungsfähigkeit (DT) oder stellen wie der KFPT die Weiterentwicklung eines bekannten diagnostischen Verfahrens zur Erfassung der affektiven Persönlichkeitsstruktur dar (Brickenkamp, 1986; Heiss & Halder, 1975; Quatember & Maly, 1980; Stephan, 1980; Ruß & Fischer, 1989).

Da sich die Farbpalette des computergestützten KFPT von seinem „klassischen“ Vorläufer FPT unterscheidet und sich Farbverbände (Cluster) als stabilste und aussagekräftigste Merkmale erwiesen haben (Heiss, Halder & Höger, 1975) mussten Vorgaben zur standardisierten Durchführung, zur Festlegung der Zielvariablen sowie zur Auswertung und Interpretation der Ergebnisse für den KFPT neu entwickelt werden. Angaben zu psychometrischen Gütekriterien und Einsatzmöglichkeiten des Verfahrens lagen bis dato nicht vor.

Vor diesem Hintergrund erfolgte im Rahmen der Studiendurchführung zunächst eine farbpsychologische Exploration der computergenerierten Farbpalette des KFPT mittels der Baselinewerte des semantischen Eindrucksdifferentials (ED). Die anschließende Clusteranalyse führte zur Einteilung der KFPT-Farbpalette in drei Cluster und drei Einzelfarbtöne, die sich wie folgt charakterisieren: Die Cluster 1 und 2 sowie der Farbton 5b (gelb) zeigen eine vorrangige Valenzwirkung; Cluster 3 und insbesondere der Farbton 10c (schwarz) werden als primär kraftvoll eingeschätzt. Der Farbton 1b (rot) nimmt eine Sonderstellung ein, da er fast gleichermaßen als erregend, angenehm und kraftvoll erlebt wird („Superfarbe rot“).

Die drei explorierten Cluster bzw. Einzelfarbtöne bilden das Farberleben mit unterschiedlichen Akzenten ab – so fokussieren die Einzelfarbtöne auf die Dimensionen Potenz (1b/10c) und Erregung (5b), die Cluster auf die Dimensionen Valenz (Cluster 1+2) und Potenz (Cluster 3). Dabei ergänzen sich Cluster und Einzelfarben und spiegeln die Dreidimensionalität des EVP-Modells wider. Hervorhebung oder Verzicht auf eines der Cluster bzw. einen speziellen Farbton erscheinen wenig sinnvoll – auch ginge die Affinität zum Farbensatz des Farbpyramidentest (FPT) verloren.

Unter Berücksichtigung der farbpsychologischen Besonderheiten des achromatischen Farbtons 10c (schwarz) sowie der beiden Farbtöne mit maximaler Sättigung und mittlerer Helligkeit 1b (rot) und 5b (gelb) wurden die drei Cluster bzw. verbleibenden Einzelfarbtöne als KFPT-Zielvariablen für die statistische Auswertung festgelegt. Im Rahmen des Studienverlaufs wurden Messwertveränderungen der ex-post definierten KFPT-Zielvariablen unter medikamentösem Einfluss durch ein Benzodiazepin und ein Phytopharmakon sowie eines Scheinmedikamentes (Placebo) objektiviert.

Aussagekraft und Interpretation der KFPT-Befunde basieren auf den Ergebnissen der Clusteranalyse und dem speziell entwickelten Auswertekonzept mit drei Clustern und drei Einzelfarbtönen. Die in der Ausgangslage vorherrschenden Dimensionsausprägungen charakterisieren die Cluster/ Einzelfarbtöne auch nach Medikationseinnahme (Studienteile 1 und 2) – relevante Veränderungen des Farberlebens lassen sich nicht feststellen und die Akzentuierung der Dimensionen bleibt bei Clustern/Einzelfarben im gesamten Studienverlauf bei allen drei Untersuchungszeitpunkten konstant.

Gemäß dem hypothesenbildenden Pilotcharakter der Studie wurden a priori keine Hypothesen zum Farbwahlverhalten formuliert, die ex-post confirmatorisch geprüft werden sollten. Sämtliche statistische Analysen wurden deskriptiv durchgeführt, ggf. wurden zusätzlich p-Werte berechnet. Abschließend wurden die Befunde zur Farbwahrnehmung mit den DT-Ergebnissen zur kognitiven Leistungsfähigkeit der „Valerian-Treatment Study“ (Kuhlmann et al. 1999) verglichen.

Eine erste Beurteilung der Pharmakosensitivität des KFPT erfolgte durch den Vergleich mit den DT-Ergebnissen zur kognitiven Leistungsfähigkeit der parallel durchgeführten „Valerian-Treatment-Studie“ (Kuhlmann et al. 1999). Korrespondierend fanden sich in beiden Studien weder Hinweise auf eine Veränderungen der Farbwahrnehmung bzw. des Farbwahlverhaltens, noch konnte eine Beeinträchtigung der kognitiven Leistungsfähigkeit durch die Einnahme eines Baldrianpräparates festgestellt werden.

Während sich unter Baldrian-Extrakt keine Leistungsminderungen zeigten, wurde unter Flunitrazepam in der „Valerian-Treatment“-Studie bei drei Variablen im DT ein signifikant zurückbleibender Übungsfortschritt objektiviert. So konnte z.B. in der Hauptzielvariablen des DT (Median der Reaktionszeit) bei 25% der

Probanden der für Flunitrazepam typische Hangover-Effekt (Dreßing, 1992; Gerhard et al., 1996) mit einer Vigilanzminderung objektiviert werden. In der farbpsychologischen Kontrollstudie wurde dieses Ergebnis jedoch nicht bestätigt – Veränderungen der Farbwahrnehmung und des Farbwahlverhaltens durch das Benzodiazepin wurden nicht beobachtet.

Die Resultate zur Pharmakovigilanz belegen insgesamt eine gute Verträglichkeit in Verbindung mit einer geringen Nebenwirkungsrate. Die Anzahl der Nebenwirkungen war erwartungsgemäß im ersten Studienteil höher als im zweiten Studienteil, was durch die Mitführung des synthetischen Arzneimittels begründet ist (Suggestiveffekt beim Benzodiazepin).

Diese Befunde stehen im Einklang mit bis der bis dato vorliegenden Literatur. Bereits 1962 untersuchte Bühring die Wirkung von Baldrian auf die Reaktionsgeschwindigkeit von Kur-Patienten. Er stellte bei Dauermedikation keine signifikante Änderung – insbesondere keine Zunahme – der Reaktionsgeschwindigkeit fest und empfahl deshalb Baldrian als Tagessedativum. Stephan (1982) konnte zeigen, dass die Einnahme von Baldrian zu keiner Beeinflussung der Verkehrstüchtigkeit führt. Es wurde im Gegenteil in einer Reihe bedeutsamer Leistungsvariablen eine Steigerung der Leistungsfähigkeit beobachtet. Eine negative Auswirkung von Baldrian auf die Verkehrstüchtigkeit konnte daher weitgehend ausgeschlossen werden. Allerdings ergaben sich bei höheren Dosierungen medikamentöse Nebenwirkungen, die zu einer Beeinträchtigung des körperlichen Wohlbefindens führten. Eine aktuellere Untersuchung von Albrecht et al. (1995) belegt gleichfalls, dass eine Baldrian-Melisse Kombination kraftfahrwesentliche Leistungen nicht mindert.

Eine Gegenüberstellung der KFPT-Befunde mit den Ergebnissen anderer farbpsychologischer Untersuchungen ist nicht möglich, da der KFPT in der vorliegenden Studie erstmalig eingesetzt wurde und somit keine weiteren Studienergebnisse vorliegen. Ein Vergleich mit historischen FPT-Daten aus pharmakopsychologischen Studien (Fahrenberg & Prystav, 1966; Lienert, 1961) ist aufgrund der Unterschiede hinsichtlich Testmaterial, Farbpalette, Zielvariablen und Auswertung unzulässig. Eine Beurteilung der Pharmakosensitivität des KFPT ist deshalb auch nur eingeschränkt – auf Basis der aktuell vorliegenden Ergebnisse - möglich. Erst in weiteren pharmakopsychologischen Studien kann letztlich die Eignung des KFPT als therapiesensitives Verfahren für den Einsatz in klinischen Arzneimittelstudien geklärt werden.

Eine abschließende Beurteilung hinsichtlich der psychometrischen Gütekriterien des KFPT ist momentan gleichfalls nicht möglich. Der KFPT ist ein Verfahren, welches speziell zur Erfassung von Affekten konstruiert wurde. Bei Folgetestungen sollten demnach auch affektive Veränderungen erfasst werden, was die Eignung des KFPT als Verlaufsparemeter impliziert. Folglich sollten Veränderungen im Farbwahlverhalten umso größer sein, je stärker die Veränderungen der Affektivität bei einem Probanden angenommen werden können. Bereits Houben (1962) konnte nachweisen, dass eine signifikante Beziehung zwischen Veränderungen des FPT und den Affekten des Probanden

besteht. Spezielle Test-Retest-Reliabilitäten zur vorliegenden Studie wurden nicht berechnet. Die geringen Messwertveränderungen in der Placebo-Kontrollgruppe verdeutlichen jedoch Konstanz und geringe Varianz der KFPT Messergebnissen bei allen drei Testabnahmen – was für die Reliabilität des Verfahrens spricht.

Hinsichtlich der Beurteilung von Gültigkeit und Zuverlässigkeit liefert die vorliegende Arbeit nur einen ersten Beitrag auf dem langen Weg der Validierung - der KFPT befindet sich noch im Stadium der Erprobung. Die Validierung eines projektiven Verfahrens wie dem KFPT ist ein jahrelanger Prozess und erfordert eine Vielzahl weiterer Einzeluntersuchungen. Nur so kann die ausreichende Stabilität der Farbvariablen als notwendige Voraussetzung der Validität empirisch gesichert werden.

Wie jeder andere Persönlichkeitstest erlaubt auch der KFPT nur Aussagen, die mit einer relativen Unsicherheit behaftet sind. Als Einzeltest sind z.B. die diagnostischen Möglichkeiten des KFPT beschränkt. Nur im Rahmen eines umfassenden psychodiagnostischen Settings erlangen seine Befunde klinische Bedeutsamkeit. Die Befunde der einzelnen Verfahren einer Testbatterie ergänzen und korrigieren sich; erst dadurch erzeugen sie eine übergreifende Validität. Vor diesem Hintergrund ist die Gegenüberstellung der KFPT-Befunde mit den DT-Ergebnissen von besonderem Interesse, erhält man doch Aufschluss über Korrespondenz oder Divergenz.

Weiterhin ist zu beachten, dass die Validierung von Software sich stark von der Validierung von Hardware bzw. Testmaterial unterscheidet. Bei Hardwareprodukten wird bei Auslieferung einer Fehlerfreiheit angenommen und der Alterungsprozess ist von entscheidender Bedeutung für die Zuverlässigkeit. Bei Softwareprodukten gibt es dagegen keine vergleichbaren Alterungsprozesse, jedoch wird aufgrund der Komplexität eine Restfehlermenge vermutet, denn die Fehlerfreiheit einer komplexen Software ist praktisch nicht beweisbar (Asam, Drenkhard, & Maier, 1986).

Um die Restfehlermenge möglichst klein zu halten, müssen die essentiellen funktionalen Bereiche einer Software möglichst fehlerfrei sein. Bei einem computergestützten Testverfahren für die Steuerung von psychologischen Reaktionszeitexperimenten sind dies z.B. Funktionen, welche die korrekte Zeit- und Reaktionserfassung betreffen. Bei farbpsychologischer Software müssen die Bedingungen einer spezifizierter Farbcodierung sowie der standardisierten Generierung der einzelnen Farbtöne gewährleistet sein – beides ist beim KFPT erfüllt.

Die Benutzung validierter Software reduziert in jedem Fall die Fehleranfälligkeit der Versuchssteuerung. Eine Standardisierung der Testverfahren wird durch einheitliche Softwarepakete zwar gefördert, die Software allein ist aber nicht entscheidend. Bei computergestützten Testverfahren spielt zusätzlich die Standardisierung der Hardware (z.B. Bildschirm, Maus) eine entscheidende Rolle, so sind z.B. Helligkeit, Kontrast und Farbsättigung des verwendeten

Bildschirms vorzuschreiben. Hier sind beim KFPT noch weitere Präzisierungen und Vorgaben notwendig, da die Verwendung handelsüblicher Monitore kaum zu kontrollierende Varianzen (z.B. verwendete Bildschirmmaße) ermöglicht, welche die beabsichtigte Standardisierung unterläuft. Sind Studienvergleiche angestrebt, sollte deshalb auf vergleichbare bzw. möglichst identische Hardwarekomponenten geachtet werden. Dieser absolute Vergleich ist aber bei den meisten pharmakopsychologischen Untersuchungen nicht erforderlich, weil der Pharmakoeffekt in Relation zu einer Kontrollgruppe geschätzt wird. Lediglich Faktoren, die theoretisch mit den Effekten interagieren können, wie z.B. starke Helligkeitsunterschiede, sollten auch in diesen Fällen möglichst konstant gehalten werden (Behringer, 1992).

Um die Einsatz- und Anwendungsmöglichkeiten des KFPT auf eine breite Basis zu stellen, sind weitere farbpsychologische Untersuchungen notwendig. So sind z.B. individuelle Farbpräferenzen ohne Berücksichtigung des soziokulturellen Hintergrundes nicht generalisierbar. Choungourian (1968) konnte kulturelle Differenzen und geschlechtsspezifische Unterschiede in der Bevorzugung bestimmter Farben aufzeigen. Basierend auf seinen Befunden folgerte Choungourian, dass man „keine universelle Regel für Farbpräferenzen aufstellen kann“. Blau und rot - welche häufig als universelle Farbtöne gelten - wurden in seiner Untersuchung von US-Amerikanern anders beurteilt als von Orientalen. Nur der grüne Farbton zeigte eine kulturübergreifend konsistent hohe Farbpräferenz im untersuchten Kollektiv. Für die Konstruktion farbpsychologischer Tests empfiehlt Choungourian daher die Entwicklung kulturspezifischer Normwerte für Frauen und Männer – für den KFPT steht dieser kultur- und geschlechtsspezifische Normierungsprozess noch aus.

Die meisten Theorien zu Geschlechtsunterschieden in der Emotionalität betonen sowohl biologische als auch soziokulturelle Faktoren. Ein weit verbreitetes Vorurteil in westlichen Kulturen unterstellt Frauen mehr Emotionalität als Männern (Fischer & Manstead, 2000) und weiterhin schreibt man Frauen stärkere Gefühlsreaktionen auf unangenehme Ereignisse bzw. Situationen zu, insbesondere wenn diese bedrohlich oder traumatisch sind (Allen & Haccoun, 1976; Kring & Gordon, 1998). Zusätzlich deuten epidemiologische Studienergebnisse auf ein höheres Risiko bei Frauen für affektive Störungen hin, wie z.B. Angst- und Depressionserkrankung (Nolen-Hoeksma, 1987; Sachs-Ericsson & Ciarlo, 2000).

Den Befund, dass Frauen mehr Angst in bedrohlichen Situationen entwickeln erklären Gordon & Riger (1991) dadurch, dass Frauen sich weniger gut gegen körperliche Angriffe wehren können als Männer. Dieser schlichte biologische Ansatz wirkt jedoch wenig überzeugend und eine psychologische Erklärung fehlt gänzlich.

Eine Vielzahl wissenschaftlicher Untersuchungen belegen eindeutig, dass Männer und Frauen sich hinsichtlich Mimik und Expressivität des Gesichtsausdrucks stark unterscheiden (Kring & Gordon, 1998; O’Gorman, 1983; Schwartz, Brown & Ahern, 1980). Auch gibt es Hinweise für geschlechts-

spezifische Unterschiede in der Emotionalität nach dem Betrachten von Bildern (Lang, Greenwald, Bradley & Hamm, 1993), insbesondere erotische Bilder werden von Männern als stärker emotional erregend erlebt als von Frauen (Murnen & Stockton, 1997).

Die Ergebnisse der vorliegenden Untersuchung sind im Hinblick auf mögliche Geschlechtsunterschiede im Farberleben inkonsistent. Bei Cluster 1 und 3 sowie bei Farbton 5b wurden im semantischen Eindrucksdifferential geschlechtsspezifische Unterschiede deutlich. Cluster 1 wurde von Männern als erregender erlebt und Cluster 3 als angenehmer eingestuft. Demgegenüber beurteilten Frauen den Farbton 5b als erregender. Für Cluster 2 sowie den Einzelfarben 1b und 10c wurden keine bedeutsamen Geschlechtsunterschiede in den E/V/P-Dimensionen objektiviert.

Die Annahme von Cheskin (1948), dass Männer generell den Farbton „blau“ und Frauen „rot“ bevorzugen, ließ sich auf Basis des vorliegenden Datenmaterials nicht bestätigen. Bei Farbton 1b (rot) war die Potenzwirkung unabhängig vom Geschlecht die dominante Dimension, doch erzielten auch die beiden anderen Dimensionen annähernd hohe Dimensionswerte bei Frauen und Männern. Sowohl die erhobenen Daten als auch die Ergebnisse anderer Untersuchungen (Bradley, Codispoti, Sabatinelli & Lang, 2001) belegen, dass trotz offensichtlich vorhandener spezifischer Geschlechtsunterschiede die affektiven Reaktionen von Frauen und Männer auf Bildreize mit angenehmen oder bedrohlichen Situationen insgesamt durchaus vergleichbar sind, weshalb diese Annahme auch auf das Farberleben zutreffen sollte.

Wie erleben Menschen Farben und welche Gefühle lösen sie aus? Der psychologischen Farbforschung mangelte es bisher an aktuellen und methodisch sauberen Studienergebnissen. Mit der vorliegenden Arbeit ist ein erster Schritt getan hier Abhilfe zu schaffen, doch sind weitere intensive Forschungsanstrengungen erforderlich, um der Farbpsychologie in Zukunft wieder mehr wissenschaftliches Gewicht zu verleihen – und somit einen qualifizierten Beitrag zur Erforschung von Emotionen und Affekten zu leisten.

Bei Emotionen handelt es sich um personeninterne, psychische Phänomene. Infolgedessen ist eine direkte Diagnostik emotionaler Befindlichkeiten bzw. emotionaler Störungen auf Selbstauskünfte des Klienten angewiesen. Farben gelten als besonders gut geeignete Indikatoren des emotionalen Geschehens. Allen Farbtests liegt das Konstrukt zugrunde, „dass Farben in einer besonderen, hervorgehobenen und markanten Weise Beziehungen zur Affektivität zeigen“ und „Farben der nächste und offenbar adäquate Ausdruck und Repräsentant von Affekten sind“ (Heiss, Halder & Höger 1975). Der KFPT ist folglich ein prinzipiell geeignetes Messinstrument zur Erfassung von Gefühlszuständen. Experimentelle und empirische Farbforschung ist zugleich auch immer experimentelle und empirische Emotionsforschung. Vor diesem Hintergrund sollen die vorliegenden Ergebnisse in einen größeren Zusammenhang gestellt und bewertet werden.

Darüber was eine Emotion ist, worin ihre Funktion besteht und in welchem Verhältnis sie zur Kognition steht, wird in der psychologischen Forschung seit Jahrzehnten kontrovers diskutiert (Bischof, 1989; Dörner, 1989; Schneider, 1989). So bedauert bereits Fantino (1973), dass Emotionen deshalb nicht mit der wissenschaftlichen Breite und Tiefe untersucht wurden, wie andere psychologische Bereiche. Den Grund sieht Fantino im Mangel an Übereinstimmung darüber, wie eine Emotion zu definieren ist. Auch Kleinginna und Kleinginna (1981) betonen, dass in der Varianz von Definitionen eines der Hauptprobleme der Emotionsforschung besteht. In ihrer Übersichtsarbeit führen sie insgesamt 92 Definitionen auf, die sie in 11 Kategorien ordnen – eine umfassende und allgemein akzeptierte Definition lässt sich jedoch nicht extrahieren.

Der Emotionsbegriff wird in der psychologischen Forschung sehr unterschiedlich verwendet. Typisch sind Mehr-Komponenten Definitionen, die Emotionen mehrere oder alle der folgenden Bestandteile zuschreiben (Kleinginna & Kleinginna, 1981):

1. Kognitive Situationseigenschaften (z.B. Einschätzungen einer Situation als bedrohlich oder unbedrohlich und unkontrollierbar).
2. Subjektive Gefühlszustände (z.B. Anspannung und innere Unruhe)
3. Physiologische Veränderungen ( z.B. erhöhte Herzfrequenz, Blutdruckanstieg, Schweißsekretion)
4. Motivationale Tendenzen (z.B. Wünsche, aus der Situation zu fliehen)
5. Ausdrucksverhalten (z.B. ängstlicher Gesichtsausdruck)
6. Instrumentelles Verhalten (z.B. Flucht)

Bei diesen Mehr-Komponenten-Definitionen bleibt jedoch im konkreten Einzelfall völlig unklar, welche und wie viele dieser – empirisch meist nur mäßig korrelierten – Komponenten vorhanden sein müssen, damit eine „Emotion“ vorliegt. Zudem ergeben sich Abgrenzungsprobleme zu verwandten Phänomenen und psychologischen Konstrukten (wie z.B. Motivation).

Durch das Fehlen einer allgemein verbindlichen Definition besteht folglich kein Konsens darüber, was die primären Dimensionen emotionalen Erlebens sind. Weitgehend Einigkeit besteht nur darin, dass Emotionen Reaktionsmuster auf diskrete, auslösende Ereignisse (interne oder externe) sind, die sich wenigstens in drei Komponenten manifestieren, nämlich: a) in den verbalen Äußerungen über subjektive Erlebnisweisen, b) dem motorisch-expressiven Verhalten (Mimik, Gestik, Vokalisation) und c) den vom autonomen Nervensystem gesteuerten vegetativen Veränderungen (Hamm & Vaitl, 1993).

Auch Pekrun (1990) fasst den Emotionsbegriff enger und fokussiert wie Hamm und Vaitl (1993) auf drei Komponenten subjektiven Erlebens: a) affektive Komponente (Gefühlserleben im engeren Sinne), b) körperperzeptive Komponente (Wahrnehmungen von physiologischen Veränderungen und Ausdrucks-

motorik), und c) kognitive Komponente (z.B. Sich-Sorgen-Machen bei Angst). Üblicherweise werden dabei unterschiedliche „Primäremotionen“ differenziert, denen jeweils ein spezifisches Ausdrucksverhalten zugeordnet ist (insbesondere Freude, Traurigkeit, Angst, Ärger, Ekel, Scham). Andere Emotionen werden dann als Mischungen oder als kognitive Differenzierungen von Primäremotionen aufgefasst (z.B. Stolz als Freude, deren kognitive Komponente eine internale Kausalattribution eines positiven Ereignisses beinhaltet.)

In der vorliegenden Literatur wird immer wieder betont, dass alle drei Reaktionskomponenten sowie ihre Bedingungen zur Beschreibung des emotionalen Geschehens herangezogen werden müssen (Ekman, 1984; Lang, 1985; Öhmann, 1987; Scherer, 1990). Problematisch ist jedoch, dass die drei Reaktionssysteme kaum oder gar nicht miteinander kovariieren (Eysenck, 1975; Öhmann, 1987; Lang, 1979). Eine der Gründe für dieses „Kovariationsproblem“ (Fahrenberg, 1982), ist sicherlich, dass alle drei Reaktionskomponenten neben der gemeinsamen Varianzquelle – der ausgelösten Emotion – ein hohes Maß an spezifischer Variation besitzen. Ist man also an einer hohen konvergenten und diskriminierenden Validität interessiert – welche für die Interrelationen der drei Reaktionskomponenten von großer Bedeutung ist – sollte eine möglichst große Kontrolle all jener Varianzquellen angestrebt werden, die nichts mit der ausgelösten Emotion zu tun haben. Dies bezieht sich aber nicht allein auf die experimentelle Kontrolle der abhängigen Variablen (z.B. Definition adäquater Zeitfenster), sondern eben auch auf die genaue Definition und Abstufung der unabhängigen Variablen, also der Verfahren, die zur Provokation von Emotionen eingesetzt werden. Hier sieht man sich aber ein schier unübersehbares Vielfalt von Methoden gegenüber, was eine kontrollierte und systematische Variation der Bedingungen – neben der Wiederholbarkeit das zweite wichtige Kriterium experimenteller Forschung (McGuigan, 1968) – nur schwer möglich macht. Standardisierte Verfahren, also Verfahren mit bekannten Auswirkungen auf alle drei Reaktionskomponenten fehlen jedoch weitgehend. Vor diesem Hintergrund ist der standardisierte Farbensatz des KFPT bzw. die standardisierte Generierung visueller Reize zur Auslösung von emotionaler Reaktionen und der Beurteilung von Farbpräferenzen von besonderem Interesse.

Die Beschreibung von subjektiven Erlebnisswelten, wie z.B. der Farbwahrnehmung mittels semantischen Eindrucksdifferentials basiert auf den klassischen sprachpsychologischen Untersuchungen von Osgood, Suci und Tannenbaum (1957). Der gesamte semantische Raum wird durch drei zentrale Dimensionen beschrieben: Erregung (beruhigend – erregend), Valenz (angenehm – unangenehm), Potenz (stark – schwach). Auch die mit dem semantischen Differential vorgenommenen Beurteilungen von Emotionen ließen sich mit faktorenanalytischen Methoden auf drei Basisdimensionen zurückführen, welche Ertel (1965) als fundamentale Dimensionen affektiven Erlebens interpretierte. Diese dimensional organisierten emotionalen Erlebens wurden ebenfalls durch Studien von Mehrabian und Russel (1974) bestätigt. Sie gaben Probanden 64 Situationsbeschreibungen vor, die eine Palette verschiedenster emotionaler Reaktionen auslösen sollten. Zur Erfassung der emotio-

nenalen Befindlichkeit entwickelten die Autoren ein eigenes Polaritätsprofil, bestehend aus 18 Skalen mit bipolaren Adjektivpaaren. Eine faktorenanalytische Auswertung dieser Befindlichkeitseinstufungen ergab wiederum eine Dreifaktorenlösung, wobei der Faktor „Valenz“ (angenehm – unangenehm) 27%, der Faktor Erregung (ruhig – erregt) 23% und der Faktor Dominanz (dominant – submissiv) 14% der gesamten Varianz aufklärte.

Diese Resultate belegen, dass das semantische Eindrucksdifferential offensichtlich ein breites Spektrum der affektiven Befindlichkeit zuverlässig und zeitökonomisch erfasst. Natürlich sollte in diesem Zusammenhang bedacht werden, dass die dreidimensionale faktorenanalytische Struktur methodenimmanent sein kann, also nur für das verwendete Polaritätsprofil spezifisch ist (Hamm & Vaitl, 1993) – Verallgemeinerungen sind also nur unter Vorbehalt möglich. Dennoch legen die vorliegenden Befunde von Russel und Mehrabian (1974) nahe, dass zumindest ein großer Teil der emotionalen Befindlichkeiten, welche über verbale Äußerungen anhand verschiedenster Skalen erfasst werden, sich im Wesentlichen durch drei zentrale Basisdimensionen beschreiben lassen.

Der im Rahmen der vorliegenden Arbeit gewählte Ansatz, die farbpsychologische Exploration des KFPT durch das ED vorzunehmen, wird durch die Resultate der oben angeführten Autoren bestätigt – die Übertragung von ED-Befunden auf die Erfassung emotionalen Geschehens und Farbwahrnehmung erscheint mithin methodisch abgesichert. Das semantische Eindrucksdifferential beschreibt auch unter Berücksichtigung aktuellerer Forschungsergebnisse subjektive Erlebniswelten adäquat.

Der generelle Anspruch der Dreidimensionalität des semantischen Eindrucksdifferentials wurde später jedoch - wiederum von Russel (1979) - relativiert. Seine neueren Befunde belegen eine große Varianz der Dominanzdimension, welche diese als wenig geeignet zur aussagekräftigen Beschreibung des emotionalen Geschehens erscheinen lassen. Russel beurteilt deshalb die Relevanz der Dimension „Dominanz“ eher kritisch, was ihn zukünftig von nur zwei Dimensionen emotionalen Geschehens ausgehen lässt (Russel, 1980). Die Schwierigkeit der konsistenten Identifizierung einer dritten Basisdimension – neben Valenz und Erregung - zur Verteilung der Reize im affektiven Raum betonen auch Storm und Storm (1987).

Ein aktuellerer Ansatz zur Beschreibung von Emotionen im semantischen Raum ist das Modell von Watson und Tellegen (1985). Die Autoren postulieren ein zweidimensionales System und differenzieren eine positive vs. eine negative Affektdimension, wobei die orthogonale Anordnung der Valenz- und Erregungsdimension um 45 Grad rotiert. Nach diesem Modell ist negative Affektivität immer mit einer Mobilisierung energetischer Ressourcen – also mit einem erhöhten Erregungsniveau – verbunden. Gefühle wie Trauer, Depression oder Hilflosigkeit sind demnach eher durch eine geringe Ausprägung positiver Affektivität bei einem mittleren Erregungsniveau gekennzeichnet. Folgt man dem Modell von Watson und Tellegen (1985) lösen neutrale Reize weder positive noch negative Affekte aus und sind von daher auch kaum mit stärkeren Erregungseindrücken

assoziiert. Weniger stark ausgeprägte Gefühle können mit diesem systematischen Ansatz wohlmöglich nicht adäquat beschrieben werden (z.B. leiser Kummer, stille Freude), da es wenig sinnvoll ist, Lust-Aspekte und Erregungs-Aspekte zusammenzufassen. Ob ein solcher Interpretationsansatz auch für die Interkorrelationen derjenigen affektiven Urteile zutreffen, welche mit dem semantischen Differential vorgenommen werden, ist bis dato nicht explizit untersucht worden. Die einfache Übertragung des interpretatorischen Ansatzes erscheint jedoch zweifelhaft, da Aufmerksamkeit, Dominanz oder Potenz nicht erfasst und beschrieben werden.

Demgegenüber belegt eine empirische Arbeit von Feist (2000) mittels Faktorenanalyse die Dreidimensionalität des emotionalen Raumes. Feist betont, dass die Dimension „attention-rejection“ bzw. „Spannung-Lösung“ auf drei Manifestationsebenen von Emotionen – der des subjektiven Erlebens, der Wahrnehmung des Ausdrucksverhaltens anderer und der physiologischen Ebene – validiert ist. Die Items des verwendeten Fragebogens wurden jedoch speziell für das emotionale Erleben von Fernsehtalkshows konzipiert. Für eine breitere Anwendung der Skala zur Generierung allgemeiner Aussagen bzw. der Einsatz im Rahmen farbpsychologischer Untersuchungen bedarf es einer Modifikation und neuerlichen Validierung.

Auch die Untersuchungen von Valdez und Mehrabian (1994) zu Effekten von Farben auf Emotionen basierten auf dem dreidimensionalen PAD-Modell (Pleasure-Arousal-Dominance). Die Farbmuster waren aus dem Farbsystem von Munsell (1915) abgeleitet und kategorisieren sich in Wellenlänge (hue), Sättigung (saturation) und Helligkeit (brightness). Die Ergebnisse stützen die Annahme einer Dreidimensionalität des semantischen Raumes. Befunde von zwei Studien der Autoren belegen konsistent, dass dunkle Farben (einschließlich dunkelgrau und schwarz) als weniger angenehm, mehr erregend und deutlich mehr dominant erlebt werden als hellere Farben (einschließlich hellgrau und weiß). Dunklere Farben induzieren eher Gefühle wie Angst (-P), Feindschaft (+A) oder Aggression (+D). Diese Befunde korrespondieren mit den Resultaten von Frank und Gilovich (1988) – die Farbe schwarz wird auch hier eher assoziiert mit Aggression und aggressivem Verhalten.

Ein leicht einzusetzendes non-verbales Verfahren zur direkten Erfassung des subjektiven Erlebens auf den drei Dimensionen Valenz (Pleasure), Erregung (Arousal) und Dominanz (Dominance) mittels standardisierter Bildreize ist das von Lang (1980) konstruierte Self-Assessment Manikin (SAM). Beurteilungen hinsichtlich der Dimensionen „Pleasure“ und „Arousal“ beim Betrachten der Bildvorlagen korrelieren hoch mit Ratings des semantischen Eindrucksdifferentials (Bradley & Lang, 1994). Unterschiede zum ED bestehen in der Dimension „Dominance“, die wohlmöglich durch die präzisere Erfassung emotionaler Reaktionen bzw. deren Dimensionalität durch das SAM begründet sind.

Das Verfahren liegt in zwei Versionen vor. In der ursprünglichen „paper-and-pencil“ Version stufen die Probanden ihren Gefühlseindruck mit einem Kreuz auf

oder zwischen den entsprechenden Reizvorlagen (Figuren) ein, woraus sich eine neunstufige Skala ergibt. Die Faktorenanalyse der resultierenden 18 Ratings generiert Scores in den drei PAD-Dimensionen.

Daneben existiert mittlerweile auch eine von Cook, Atkinson und Lang (1987) entwickelte Computerversion des SAM. Hier werden die drei Dimensionen dynamisch auf einem Display dargestellt und die Probanden nehmen ihre Beurteilungen mit einem Joystick vor. SAM liefert dabei eine 20-stufige Skala, was eine differenziertere Abstufung der jeweiligen Dimension ermöglicht.

SAM wurde seither erfolgreich in einer Vielzahl von Untersuchungen zur Erfassung emotionaler Reaktionen eingesetzt, z.B. auf Abbildungen (Miller, Levin, Kozak, Cook, McLean & Lang, 1987), auf unterschiedliche Geräusche und Klänge (Bradley, 1994) oder schmerzhaft stimuli (McNeil & Brunetti, 1992) sowie bei Angstpatienten (Cook, Melamed, Cuthbert, McNeil & Lang, 1988).

In Ergänzung zum semantischen Eindrucksdifferential erscheint insbesondere die Computerversion des SAM als ein geeignetes Instrument zur Evaluierung der Dimensionsausprägungen des KFPT-Farbensatzes. Dabei ist von wissenschaftlichem Interesse, ob und inwieweit ED-Ergebnisse und SAM-Resultate korrelieren. Falls die ED-Befunde zur farbpsychologischen Exploration des KFPT ihre Bestätigung durch SAM-Resultate erhalten, wäre dies ein wichtiger Fortschritt zur Entwicklung einer standardisierten und normativen Serie von Farbreizen zur Erfassung emotionaler Reaktionen. Gleichgerichtete ED+SAM-Ergebnisse würden belegen, dass mit dem KFPT-Farbensatz zuverlässig Emotionen ausgelöst werden können, welche sich konsistent und unabhängig von der Erhebungsmethode, erfassen lassen. Auch zur Beantwortung der Frage, welche Cluster am zukünftig besten dazu geeignet sind, um Aussagen zur Emotionalität zu generieren, würde die Exploration des KFPT-Farbensatzes mittels SAM einen wertvollen Beitrag leisten. Auf Basis der aktuell noch geringen Datenlage sind solche Aussagen nicht möglich und auch wenig sinnvoll.

Neben dem KFPT ist die Zahl geeigneter Skalen und Testverfahren zur Erfassung von Gefühlen stark limitiert. Schon allein deshalb stellen die kürzlich von Behr und Becker (2004) vorgestellten SEE (Skalen zum Erleben von Emotionen) eine Bereicherung des testpsychologischen Instrumentariums dar und sollten hier kurz vorgestellt werden - auch wenn die SEE keinen direkten Bezug zur Farbpsychologie bzw. zur Exploration des Farberlebens aufweisen.

Die Autoren wollen mit den SEE Psychotherapeuten jeglicher erlebnisorientierter Psychotherapierichtung ein Verfahren in die Hand geben, das ihnen zeigt, wie der Klient Gefühle wahrnimmt, wie er sie bewertet und mit ihnen umgeht.

Interventionen und Ziele können dadurch individualisiert bestimmt und Veränderungen im Rahmen der Qualitätskontrolle objektiviert werden. Innerhalb der Psychotherapieforschung können die Konstrukte der personenzentrierten Persönlichkeitstheorie gezielt erfasst und überprüft werden, so dass sich auch Weiterentwicklungen der Persönlichkeitstheorie auf empirische Daten stützen können. Bedingt durch die gute praktische Anwendbarkeit der SEE eröffnet sich

mit diesem Verfahren auch die Gelegenheit, eine breitere Personengruppe in Psychotherapie, Forschung und Organisationsentwicklung mit den zentralen Konstrukten des personenzentrierten Ansatz bekannt zu machen (Schlippe-Weinberger, 2004).

Ziel von Behr und Becker war es, mit Hilfe eines Fragebogens die zentralen Konstrukte der klientenzentrierten Persönlichkeitstheorie und des Konzepts der Emotionalen Intelligenz zu erfassen. Nach dem Konzept der klassischen Testtheorie entwickelten sie insgesamt 7 Skalen, die im Einzelnen folgende inhaltliche Bereiche erfassen: Akzeptanz eigener Emotionen, Erleben von Emotionsüberflutung, Erleben von Emotionsmangel, körperbezogene Symbolisierung von Emotionen, imaginative Symbolisierung von Emotionen, Erleben von Emotionsregulierungen und Erleben von Selbstkontrolle. Es zeigte sich eine hohe Reliabilität über längere Zeiträume und eine gute Validität. Bei den SEE handelt es sich um einen „paper-and-pencil“ Test. Insgesamt 42 Items werden nach einer 5-stufigen Antwortkategorie von „stimmt gar nicht“ bis „stimmt völlig“ beurteilt. Die Skalen können sowohl von einzelnen Personen als auch von Gruppen ausgefüllt werden. Die Auswertung erfolgt manuell unter Zuhilfenahme eines Auswertungsbogens, einer Auswertungsschablone und des Manuals.

Was bei den SEE bereits vorliegt, gehört sicherlich auch zu den wichtigsten Aufgaben für die methodische Weiterentwicklung des KFPT - die Erstellung eines speziellen Testmanuals. Dieses Manual sollte Informationen über das zu messende Konstrukt, Zusammensetzung der Testform, Durchführungshinweise für den Versuchsleiter sowie Auswertungs- und Interpretationsbeispiele enthalten. Darüber hinaus sollten eine Eichstichprobe und relevante Teilpopulationen (z.B. nach Alter und Geschlecht) festgelegt und Normwerttabellen, Angaben zur Reliabilität und Validität für das eigentliche Testverfahren sowie für die Eich- und Kontrollstichprobe gemacht werden. Dies ist jedoch Aufgabe und Inhalt zukünftiger Untersuchungen.

## 11 Literaturverzeichnis

- Albrecht, M., Berger, W., Laux, P., Schmidt, U. & Martin, C. (1995). Psychopharmaka und Verkehrssicherheit: Der Einfluss von Euvegal-Dragees forte auf die Fahrtüchtigkeit und Kombinationswirkungen mit Alkohol. *Zeitschrift für Allgemeinmedizin*, 71, 1215-1228.
- Allen, J.G. & Haccoun, D.M. (1976). Sex differences in emotionality: A multidimensional approach. *Human Relations*, 29, 711-722.
- Ammon, H.P.T. (1982). Unerwünschte Arzneimittelinteraktionen und Verkehrstauglichkeit. In K. Luff & A. Schrey (Hrsg.). *Arzneimittel und Verkehrssicherheit* (124-136). München: Universitätsdruckerei und Verlag Dr. C. Wolf & Sohn.
- Asam, R., Drenkhard, N. & Maier, H.-H. (1986). *Qualitätsprüfung von Softwareprodukten*. Siemens AG: Berlin und München.
- Behr, M. & Becker, M. (2004). *SEE – Skalen zum Erleben von Emotionen*. Göttingen: Hogrefe-Verlag für Psychologie.
- Behringer, J. (1992). Computergestützte kognitive Testverfahren. In J. Oldigs-Kerber & J.P. Leonard (Hrsg.). *Pharmakopsychologie – Experimentelle und Klinische Aspekte* (134-149). Jena-Stuttgart: Gustav Fischer Verlag.
- Biehl, F. (1972). *Der ATV als Hilfsmittel in der Fahreignungsdiagnostik*. Arbeiten aus dem Verkehrspsychologischen Institut. Wien: Kleine Fachbuchreihe (Band 11).
- Bischof, N. (1989). Emotionale Verwirrungen – oder: Von den Schwierigkeiten im Umgang mit der Biologie. *Psychologische Rundschau*, 40, 188-205.
- Borbély, A. (1986). Benzodiazepinhypnotika. In H. Hippius, R. Engel & G. Laakmann (Hrsg.). *Benzodiazepine* (268-285). Berlin-Heidelberg: Springer-Verlag.
- Bradley, M.M. (1994). Emotional memory: A dimensional analysis. In S. van Goozen, N.E. van de Poll & J.A. Sergeant (Hrsg.). *The emotions: Essays on emotion theory* (97-134). Hillsdale, NJ: Erlbaum.
- Bradley, M.M. & LANG, P.J. (1994). Measuring Emotion: The Self-Assessment-Manikin And The Semantic Differential. *Journal of Behav. Ther. & Exp. Psychiatry*, Vol. 25, No. 1, 49-59.
- Bradley, M.M., Codispoti, M., Sabatinelli, D. & Lang, P.J. (2001). Emotion and Motivation II: Sex Differences in Picture Processing. *Emotion*, Vol. 1, No. 3, 300-319.
- Brengelmann, J. C. (1957). Farbwahl und Variabilität im Farbpapientest. *Arch. Psychiat. und Z. ges. Neurol., Bd. 196*, 297-306.
- Brickenkamp, R. (1975). *Handbuch apparativer Verfahren in der Psychologie*. Verlag für Psychologie Dr. C.J. Hogrefe: Göttingen.
- Brickenkamp, R. (1986). *Handbuch psychologischer und pädagogischer Tests*. Göttingen: Verlag für Psychologie Dr. C.J. Hogrefe.
- Bühning, M. (1962). Zur Wirkung eines Baldrian-Hopfenpräparates auf die Reaktionsgeschwindigkeit bei Kurpatienten. *Kassenarzt*, 16, 18-19.
- Cheskin, L. (1948). *Colors*. New York: Liveright.
- Cook, E. W., Atkinson, L. & Lang, P. J. (1987). Stimulus control and data acquisition for IBM PCs and compatibles. *Psychophysiology*, 24, 726.
- Cook, E. W., Melamed, B. G., Cuthbert, B. N., McNeil, D. W. & Lang, P. J. (1988). Emotional imagery and the differential diagnosis of anxiety. *Journal of Consulting and Clinical Psychology*, 56, 734-740.

- Choungourian, A. (1968). Color Preferences and Cultural Variation. *Perceptual and motor Skills*, 26, 1203-1202.
- Diester, J. & Stapf, B. (1969). *Untersuchung über die Beziehung zwischen den physikalischen Eigenschaften von Oberflächenfarben und ihrer emotionalen Eindrucksqualität*. Diplomarbeit, Freiburg: Albert-Ludwigs Universität.
- Dörner, D. (1989). Emotion, Kognition und Begriffsverwirrungen: Zwei Anmerkungen zur Köhler-Vorlesung von Norbert Bischof. *Psychologische Rundschau*, 40, 206-209.
- Dreßing, H. (1992). Baldrian-Melisse-Kombinationen versus Benzodiazepin: Bei Schlafstörungen gleichwertig?. *Therapiewoche*, 42, 726-736.
- Ekman, P. (1984). Expression and the nature of emotion. In K.R. Scherer & P. Ekman (Hrsg.). *Approaches to emotions* (319-343). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ertel, S. (1964). Die emotionale Natur des "semantischen " Raumes. *Psychologische Forschung*, 28, 1-32.
- Ertel, S. (1965). Standardisierung eines Eindrucksdifferentials. *Zeitschrift für experimentelle und angewandte Psychologie*, 12, 22-58.
- ERTEL, S. (1965). Weitere Untersuchungen zur Standardisierung eines Eindrucksdifferentials. *Zeitschrift für experimentelle und angewandte Psychologie*, 12, 177-208.
- Eysenck, H.J. (1941). A critical and experimental study of color preferences. *Amer. J. Psychology*, Vol. 54, 385-394.
- Eysenck, H.J. (1975). The measurements of emotion: Psychological parameters and methods. In L. Levi (Hrsg). *Emotions: Their parameters and measurements* (51-96). Amsterdam: Elsevier.
- Fahrenberg, J. & Prystav, G. (1966). Psychophysiologische Untersuchung eines Tranquilizers nach kovarianzanalytischem Plan zur Kontrolle von Ausgangswerten und Persönlichkeitsdimensionen. *Arzneimittelforschung*, 6, 16 Jg., 754-759.
- Fahrenberg, J. (1982). *Probleme der Mehrebenen-Beschreibung und Prozeß-Forschung*. Freiburg: Forschungsbereiche des Psychologischen Instituts der Albert-Ludwigs-Universität Freiburg im Breisgau.
- Fantino, E. (1973). Emotion. In J.A. Nevin (Hrsg). *The study of behavior: Learning, motivation, emotion, and instinct*. Glenview, Illinois: Scott, Foresman.
- Feist, A. (2000). *Emotionale Wirkungen von Fernsehtalkshows*. Aachen: Shaker Verlag.
- Ferrara, S.D. (1994). Drugs-of-Abuse testing in urine: statistical approach and experimental comparison of immunochemical and chromatographic techniques. *Journal of Analytical Toxicology*, 18, 278-291.
- Fischer, A.H. & Manstead, A.S.R. (2000). The relation between gender and emotions in different cultures. In A.H. Fischer (Hrsg). *Gender and emotion: Social psychology perspectives* (71-94). New York: Cambridge University Press.
- Frank, O.M.G. & Gilovich, T. (1988). The dark side of self- and social perception: Black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology*, 54, 74-85.
- Gerhard, U., Linnenbrink, N., Georghiadou, C. & Hobi, V. (1996). Vigilanzmindernde Effekte zweier pflanzlicher Schlafmittel. *Schweizerische Rundschau für Medizin*, 85, Nr. 15, 473-481.
- Gordon, M.T. & Riger, S. (1991). *The female fear: The social costs of rape*. Chicago: University of Illinois Press.
- Guilford, J.P. & Smith, P.C. (1959). A system of color preferences. *American Journal of Psychology*, 72, 487-502.
- Hamm, A. & Vaitl, D. (1993). Emotionsinduktion durch visuelle Reize: Validierung einer Stimulationsmethode auf drei Reaktionsebenen. *Psychologische Rundschau*, 44, 143-161.
- Hartung, J. (1982). *Lehr- und Handbuch der angewandten Statistik*. München-Wien: R. Oldenbourg Verlag.
- Heiss, R., Honsberg J. & Karl, H. (1955). Vorläufige Mitteilungen über die Verwendung von „hässlichen Pyramiden“ im Farbpyramiden-Test. *Zeitschrift für diagnostische Psychologie*, 3, 106-124.

- Heiss, R., Fahrenberg, J. & Seuss, R. (1971). Eine Studie zur Prüfung der psychodynamischen Wirkung von Dihydroergotoxin. *Arzneimittelforschung*, 6, 21. Jg., 797-800.
- Heiss, R., Halder, P. & Höger, D. (1975). *Der Farbpolymeridentest*. Bern: Verlag Hans Huber, 2. Auflage.
- Hemmelgarn, B., Suissa, S., Huang, A., Boivin, J.-F. & Pinard, G. (1997). Benzodiazepine Use and the Risk of Motor Vehicle Crash in the Elderly. *JAMA Abstracts*, July 2m.
- Herberg, K.-W. (1994). Psychotrope Phytopharmaka im Test: Alternative zu synthetischen Psychopharmaka?. *Therapiewoche*, 44, 704-713.
- Hiltmann, H. (1953). Zur Syndromatik der Verlaufsformel im Farbpolymeridentest. In H. Hiltmann (Hrsg.). *Verlaufsanalyse in der psychologischen Diagnostik* (53-58). Bern-Stuttgart.
- Hindmarch, I. (1987). Three antidepressants (amitryptiline, dothiepin, fluoxetine) with and without alcohol, compared with placebo on tests of psychomotor ability related to car driving. *Human Psychopharmacology*, 2, 177-183.
- Hofstätter, P.R. & Lübbert, H. (1958). Die Eindrucksqualitäten von Farben. *Zeitschrift für diagnostische Psychologische und Persönlichkeits Forschung.*, 6, 211-227.
- Höger, D. (1968). *Grundlagen des Farbwahlverhaltens bei Farbgestaltungsverfahren*. Habil.-Schrift, Freiburg.
- Höger, D. (1980). *Zur Psychologie des Farberlebens älterer Menschen als Grundlage der farblichen Gestaltung ihrer Wohnwelt*. Bielefeld.
- Houben, A.M.J. (1962). Die Reliabilität des Farbpolymeridentests. *Diagnostica*, 8, 110-124.
- Hüther, K.J. (1989). Phytotherapie von psychovegetativen Dysregulationen und Schlafstörungen. Eine Alternative ohne Vigilanzinbuße und Abhängigkeitspotential. *Ngm* 5, 150-155.
- Janke, W. (1988). Pharmaka als Forschungswerkzeuge in der Leistungsforschung. Zu Heinrich Dükers Bedeutung in der Pharmakopsychologie. *Archiv für Psychologie*, 140, 223-245.
- Karl, H. (1953). Die Diagnostik der Antriebsstruktur im Farbpolymeridentest. *Zeitschrift für experimentelle und angewandte Psychologie*, 1, 524-567.
- Kautz, S. (1994). Baldrian - mehr als nur ein Hausmittel. *Therapiewoche*, 44, 804-807.
- Kring, A.M. & Gordon, A.H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology*, 74, 686-703.
- Kleinginna, P.R. & Kleinginna, A.M. (1981). A Categorized List of Emotional Definitions, with Suggestions for a Consensual Definition. *Motivation and Emotion*, Vol. 5, No. 4, 345-379.
- Kraepelin, E. (1892). *Über die Beeinflussung einfacher psychischer Vorgänge durch einige Arzneimittel*. Jena: Gustav Fischer Verlag.
- Kraepelin, E. (1927). *Psychiatrie*. Leipzig: Barth-Verlag.
- Krieglstein, J.K. & Grusla, D. (1988). Zentral dämpfende Inhaltsstoffe im Baldrian - Valepotriate, Valerensäure, Valeranon und ätherisches Öl sind jedoch unwirksam. *Deutsche Apotheker Zeitschrift*, 128, 2041-2046.
- Kring, A.M. & Gordon A.H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology*, 74, 686-703.
- Kuhlmann, J., Berger, W., Podzuweit, H. & Schmidt, U. (1999). The Influence of Valerian Treatment on „Reaction Time, Alertness an Concentration“ in Volunteers. *Pharmacopsychiatry*, 32, 1-7.
- Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 475-512.
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson & T. A. Williams (Hrsg.). *Technology in mental health care delivery systems* (119-137). Norwood, NJ: Ablex.

- Lang, P. J. (1985). The cognitive psychophysiology of emotion: Fear and anxiety. In A. H. Tuma & J. D. Maser (Hrsg.). *Anxiety and anxiety disorders* (131-170). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lang, P. J., Greenwald, M. K., Bradley, M. M. & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioural reactions. *Psychophysiology*, *30*, 261-273.
- Leathwood, P.D. (1982). Aqueous Extracts of Valerian Roots improves Sleep Quality in Man. *Pharmacology, Biochemecical & Behavior*, *17*, 65-71.
- Leathwood, P.D. & Chauffard, F. (1982). Quantifying the Effects of Mild Sedatives. *Journal of J. Psychiatric Research*, *17*, 116-122.
- Lienert, G.A. (1961). Die Farbwahl im Farbpyramidentest unter Lysergsäurediätylamid (LSD). *Zeitschrift für angewandte und experimentelle Psychologie*, *8*, 110-121.
- Lienert, G.A. (1969). *Testaufbau und Testanalyse*. Weinheim.
- Luria, R. E. (1975). The validity and reliability of the visual analogue mood scale. *Journal of Psychiatric Research*, *12*, 51-57.
- Maisenbacher, J. (1992). Baldrian und Melisse – milde Psychopharmaka. *Therapiewoche*, *42*, 2140-2144.
- Margraf, J., Scheiner, S. & Ehlers, A. (1994). *DIPS - Diagnostisches Interview bei psychischen Störungen*. Berlin-Heidelberg: Springer Verlag, 2. Auflage.
- McGuigan, F. J. (1968). *Experimental Psychology*. Englewood Cliffs.
- McNeil, D.W. & Brunetti, D.G. (1992). Pain and fear: a bioinformatical perspective on repositivity to imagery. *Behavioral Research Therapy*, *30*, 513-520.
- Mehrabian, A. & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge, MA: MIT Press.
- Michel, L. (1959). Eine Modifikation des Farbpyramidentests: Der 14-Farbensatz. *Diagnostica*, *5*, 66-86.
- Michel, L. (1962). Die Reliabilität des Farbpyramidentests. *Psychologische Rundschau*, *13*, 20.
- Mierke, K. (1959). Fahrtauglichkeitsprüfung mit dem Wiener Determinationsgerät. *Zeitschrift für experimentelle Psychologie, Bd V*.
- Miller, G. A., Levin, D. N., Kozak, M. J., Cook, E.W., McLean, A. & Lang, P.J. (1987). Individual differences in emotional imagery. *Cognition and Emotion*, *1*, 367-390.
- Moser, L. Hüther, K.J., Koch-Weser, J. & Lundt, P.V. (1978). Effects of Terfenadine and Diphenhydramine Alone or in Combination with Diazepam or Alcohol on Psychomotor Performance and Subjective Feelings. *European Journal of Clinical Pharmacology*, *14*, 417-423.
- Müller, J. & Romeyke, A. (1971). Über die psychodynamische Wirkung eines Thymoleptikums im Langzeitversuch. Freiburg: Diplom-Arbeit.
- Müller, W.E. (1990). Benzodiazepine – Eine nicht unproblematische Arzneigruppe. Abhängigkeit und Suchtproblematik, mögliche Alternativen. *Therapiewoche*, *40*, 3635-3641.
- Munsell, A.H. (1915). *Atlas of the Munsell Color System*. Boston.
- Murnen, S.K. & Stockton, M. (1997). Gender and self-reported sexual arousal in response to sexual stimulation: A meta-analytic review. *Sex Roles*, *37*, 135-153.
- Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin*, *101*, 259-282.
- Öhmann, A. (1987). The psychophysiology of emotion: An evolutionary-cognitive perspective. In P. K. Ackles, J. R. Jennings & M. G. H. Coles (Hrsg.). *Advances in psychophysiology*, Vol. 2, 79-127. Greenwich: JAI Press

- O’Gorman, J.G. (1983). Individual differences in the orienting response. In D. Siddle (Hrsg.). *Orienting and habituation: Perspectives in human research*. New York: Wiley.
- O’Reilly, P.O. & Blewett, D. (1959). Color Analysis of the Pyramid Test. *Disorders of Nervous System*, 20, 211-213.
- Osgood, C. E., Suci, G. J. & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois.
- Oyama, T., Tanaka, Y. & Chiba, Y. (1962). Affective dimensions of colors: A cross-cultural study. *Jap. Psychology Research*, Vol. 4, 78-91.
- Pekrun, R. (1990). Emotion: Klassifikation und Diagnostik. In U. Baumann & M. Perrez (Hrsg.). *Klinische Psychologie*. Bern: Verlag Hans Huber.
- Quatember, R. & MALY, J. (1980). Neuropsychologische Untersuchungsmethoden altersspezifischer Leistungsparameter. *Wiener Medizinische Wochenzeitschrift*, 130, 688-692.
- Rainio, K. & Matikainen, R. (1954). Eine Faktorenanalyse im Farbpyramidentest. *Zeitschrift für diagnostische Psychologie*, 2, 292-308.
- Raskin, A. (1982). Cognitive and performance deficits in depression. *Psychopharmacological Bulletin* 18, 4, 196-202.
- Russel, J. A. (1979). Affective space is bipolar. *Journal of Personality and Social Psychology*, 37, 345-356.
- Russel, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Ruß, M. & Fischer, P.A. (1989). Reaktionszeit und Aufgabenkomplexität: Der Komplexitätseffekt als ein neuropsychologischer Indikator für den Schweregrad der zerebralen Beeinträchtigung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 10, Heft 3, 145-153.
- Sachs-Ericsson, N. & Ciarlo, J. A. (2000). Gender, social roles and mental health: An epidemiological perspective. *Sex Roles*, 43, 605-628.
- Schaie, K.W. (1962). Die Validität des Farbpyramiden-Tests. *Diagnostica*, 8, 141-181.
- Scherer, K. R. (1990). Theorien und aktuelle Probleme der Emotionspsychologie. In K. R. Scherer (Hrsg.). *Psychologie der Emotion: Enzyklopädie der Psychologie. Bd. 3*, 345-422. Göttingen: Hogrefe.
- Schneider, K. (1989). Zur Lage der Emotionsforschung oder der Kampf gegen Strohpuppen. *Psychologische Rundschau*, 40, 216-218.
- Schwartz, G. E., Brown, S. L. & Ahern, G. L. (1980). Facial muscle patterning and subjective experience during affective imagery: Sex differences. *Psychophysiology*, 17, 75-82.
- Siegfried, K. & O’Conolly, M. (1986). Cognitive and psychomotor effects of different antidepressants in the treatment of old age depression. *International Clinical Psychopharmacology*, 1, 231-243.
- Spiegel, R. (1991). Psychopharmakotherapie. In M. Perez & U. Baumann (Hrsg.). *Lehrbuch Klinische Psychologie*. Bern: Verlag Hans Huber.
- Stephan, E. (1980). Beeinflussung der Verkehrstüchtigkeit durch die Einnahme eines pflanzlichen Sedativums. *Therapie-woche*, 30, 3662-3678.
- Schlippe-Weinberger, S. (2004). SEE – Skalen zum Erleben von Emotionen. *Gesprächspsychotherapie und Personenzentrierte Beratung*, 4, 297-298.
- Schuhfried, G. (1996). *Wiener Testsystem. Computergestützte psychologische Diagnostik*. Katalog 11/95 und Manuale 1995/1996. Wien.
- Storm, C. & Storm, T. (1987). A taxonomic study of the vocabulary of emotions. *Journal of Personality and Social Psychology*, 53, 805-816.
- Valdez, P. & Mehrabian, A. (1994). Effects of color on emotion. *Journal of Experimental Psychology*, 123, 394-409.

- Vorbach, E.U., Görtelmeyer, R. & Brüning, J. (1996). Therapie von Insomnien: Wirksamkeit und Verträglichkeit eines Baldrianpräparates. *Psychopharmakotherapie*, 3.Jg, Heft 3, 109-115.
- Watson, D. & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98, 219-235.
- Wegener, R. & Leonard, J.P. (1992). Anwendungsgebiete und Nutzen der Pharmakopsychologie. In J. Oldigs-Kerber & J.P. Leonard (Hrsg.). *Pharmakopsychologie – Experimentell und Klinische Aspekte* (471-490). Jena-Stuttgart: Gustav Fischer Verlag.
- Weingarter, H. (1982). Psychobiological determinants of memory failure. In L.R. Squire & N. Buttes (Hrsg.). *Neuropsychology of Memory*. New York: Guilford.
- Whitfield, T.W. & Wiltshire, T.J. (1990). Color Psychology: a critical review. *Genetic, Social and General Psychology Monographs, Volume 116* (4), 385-411.
- Wittling, A. (1976). E. Kraepelin. In: W. Arnold, H.J. Eysenck & R. Meili (Hrsg.). *Lexikon der Psychologie* (355-357), Freiburg: Herder.