

4 Ergebnisse

4.1 Demographie

An der Studie nahmen 13 Studenten und 17 Ärzte (9 Ärzte „Normal“-Gruppe, 8 Ärzte „Cross-Over“-Gruppe) teil. Die demographische Verteilung der in den Fragebögen erfassten Parameter sowie Angaben zum Nachtschlaf zwischen BD8 und FD8 sind der Tabelle 2 zu entnehmen:

Tabelle 2: Demographische Übersicht, Die Werte sind ausgedrückt in MW \pm SD bzw. Median (halber Interpezentilenabstand), keine signifikanten Unterschiede zwischen den Gruppen mit Ausnahme folgender Werte:

	Kontroll- gruppe (Studenten)	Testgruppe (Ärzte)						
		alle	Cross- Over- Analyse	Subgruppe Erfahrung		Subgruppe Schlaf		
				Ärzte Normal	Ärzte Cross- Over	Lernende	Experten	
n	13	17	9	8	9	8	7	10
Geschlecht	4m : 9w	12m : 5w	5m : 4w	7m : 1w	5m : 4w	7m : 1w	5m : 2w	7m : 3w
Alter	25.2 \pm 3.6 ¹	31.6 \pm 4.2 ¹	31.3 \pm 5.5	31.9 \pm 2.4	29.4 \pm 3.5 ²	34.0 \pm 3.6 ²	30.4 \pm 2.8	32.4 \pm 4.9
Erfahrung i.d. laparoskopischen Chirurgie	keine	1.0 (1.5)	1.0 (1.75)	1.5 (1.25)	1.0 (0) ³	4.0 (1.375) ³	2.0 (1.5)	1.0 (1.625)
Berufsjahre i.d. Chirurgie	keine	3.4 \pm 3.9	4 \pm 5.2 ⁴	2.8 \pm 1.6 ⁴	1.2 \pm 1.3 ⁵	5.9 \pm 4.4 ⁵	2.3 \pm 1.7 ⁶	4.2 \pm 4.8 ⁶
Erfahrung mit Videospiele	2.0 (0.75)	2.0 (0.75)	2.0 (0.75)	2.0 (0.375)	2.0 (0.75)	2.0 (0.75)	2.0 (1.0)	2.0 (0.625)
Schlafstunden (zwischen BD16/FD8)	6.5 \pm 0.9 ⁷	2.9 \pm 1.4 ⁷	3.1 \pm 1.7 ⁸	2.7 \pm 1.0 ⁸	3.0 \pm 1.6 ⁹	2.8 \pm 1.3 ⁹	1.6 \pm 1.1 ¹⁰	3.8 \pm 0.7 ¹⁰

¹ **Alter:** Studenten gegen Ärzte, $p \leq 0.001$; ² **Alter:** Lernende gegen Experten, $p = 0.059$

³ **Berufsjahre in der Chirurgie:** Ärzte Normal gegen Ärzte Cross-Over, $p = 0.815$

⁴ **Berufsjahre in der Chirurgie:** Lernende gegen Experten, $p \leq 0.001$

⁵ **Berufsjahre in der Chirurgie:** <3h Schlaf gegen \geq 3h Schlaf, $p = 0.601$

⁶ **Schlafstunden zwischen BD16/FD8:** Studenten gegen Ärzte, $p \leq 0.001$

⁷ **Schlafstunden zwischen BD16/FD8:** Ärzte Normal gegen Ärzte Cross-Over, $p = 0.423$

⁸ **Schlafstunden zwischen BD16/FD8:** Lernende gegen Experten, $p = 0.606$

⁹ **Schlafstunden zwischen BD16/FD8:** Die Subgruppe Schlaf war aufgeteilt in <3h (< 3 Schlafstunden) und \geq 3h (\geq 3 Schlafstunden), $p \leq 0.001$

In der Gruppe der Studenten waren die Frauen in der Überzahl, bei den Ärzten waren es die Männer.

Das Alter, die Berufsjahre in der Chirurgie sowie der Nachtschlaf sind als $MW \pm SD$ angegeben. Für die Erfahrung in der Laparoskopie und die Erfahrung mit Videospielen sind der Median und in Klammern der halbe Interperzentilenabstand angegeben.

Die Studenten haben im Mittel 6,54 Stunden geschlafen mit geringer Standardabweichung. Sie hatten demnach keinen Schlafentzug.

Die Ärzte haben in der Nacht des Bereitschaftsdienstes (zwischen BD8 und FD8) durchschnittlich 2,9 Stunden geschlafen, wobei eine Unterteilung in < 3 h sowie ≥ 3 h bis < 6 h Schlaf möglich war (7:10 Probanden).

Auch die Ärzte der Subgruppe ≥ 3 h Schlaf wiesen mit im Mittel 3,9 Stunden Schlaf einen deutlichen Schlafentzug auf.

Bei allen Gruppen zeigte sich eine geringe Erfahrung mit Videospielen, die unter den Probanden etwa gleich verteilt war.

Bis auf drei Ärzte hatte kein Proband Erfahrung mit einem OP-Simulator.

4.2 Trainingsphase

Die Diagramme der Einzelparameter der Trainingsphase ließen keine „Ausreißer“ erkennen. Dies ist hier am Beispiel des Parameters Kamera_Zeit dargestellt, in dem die Leistungen aller Probanden einzeln graphisch dargestellt sind (siehe Abb. 11). Dieses Diagramm zeigt einen homogenen Lernzuwachs sowie eine geringe Varianz in den Trainingsleistungen.

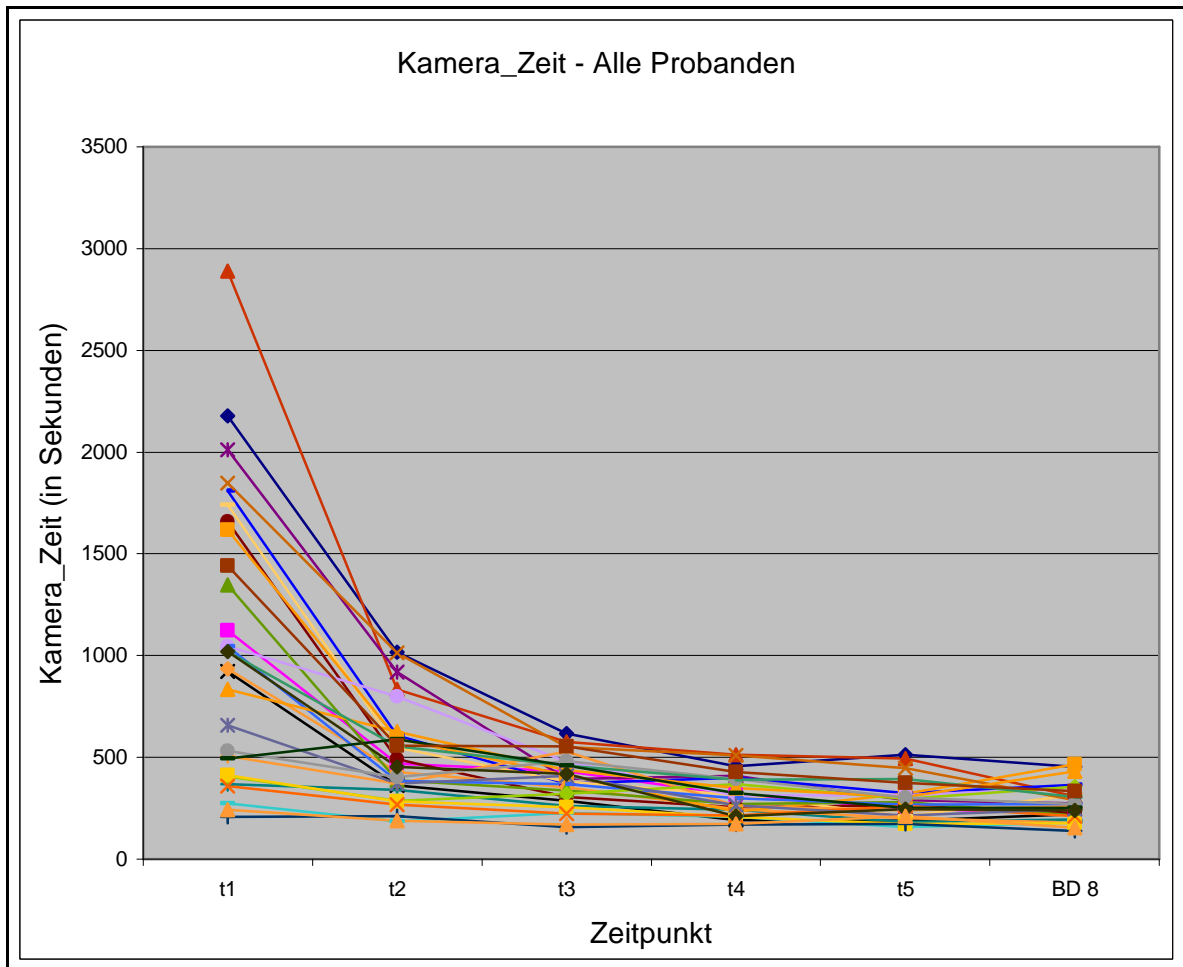


Abb. 11: Parameter Kamera_Zeit für alle Probanden einzeln, keine „Ausreißer“

Der Vergleich Studenten gegen Ärzte zeigte deutliche Lernkurven am OP-Simulator. Dies wird am Beispiel des Parameters Kamera_Zeit (als Maß für die Bearbeitungsgeschwindigkeit) deutlich (siehe Abb. 12).

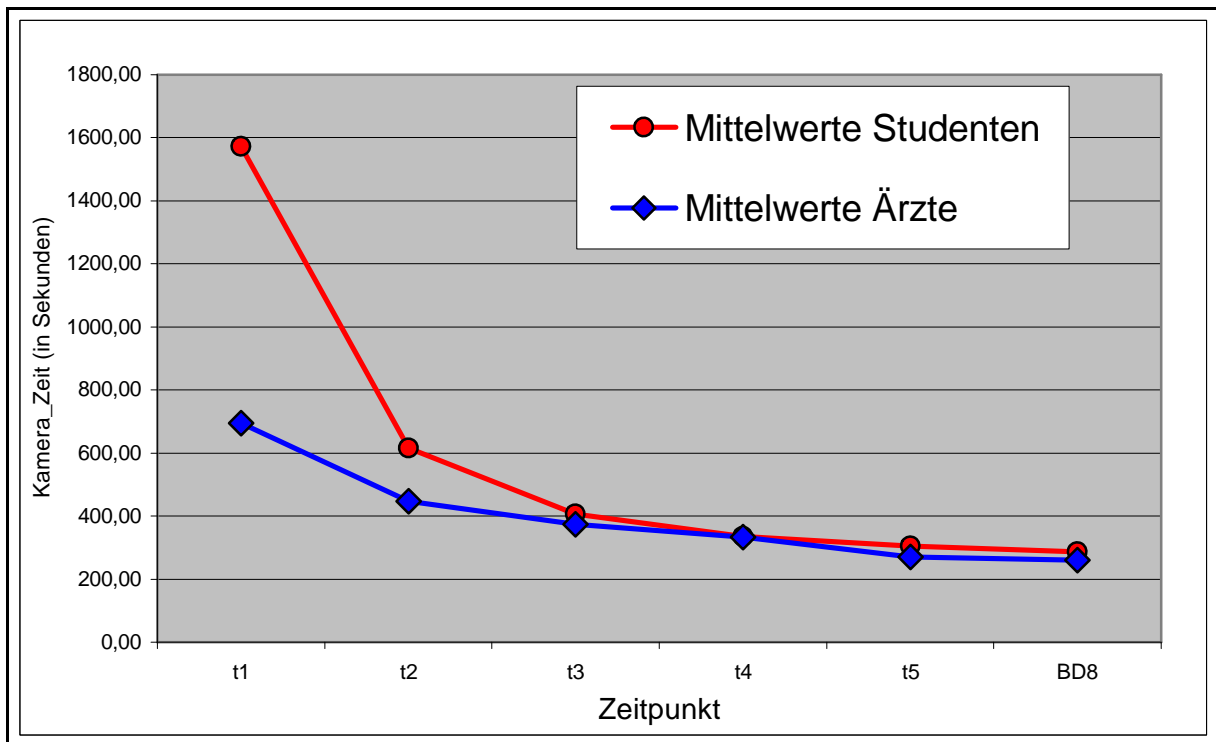


Abb. 12: Einzelparameter Kamera_Zeit, Studenten gegen Ärzte (Normal und Cross-Over), deutliche Lernkurve, ab Tag 4 Übergang in ein Plateau

Die Ausgangswerte der Studenten waren sehr hoch (im Mittel etwa 1600 Sekunden zum Durchführen der Übung, also langsame Geschwindigkeit). Jedoch zeigte sich rasch eine starke Verbesserung und ein Übergang in eine Plateauphase ab dem vierten Tag als Ausdruck eines deutlichen Lerneffektes.

Die Ärzte waren zu Beginn wesentlich schneller, zeigten aber auch eine Lernkurve, jedoch mit deutlich geringeren Ausgangswerten.

Als weiteres Beispiel ist der Parameter Kamera_Weg dargestellt (siehe Abb. 13). Dieser zeigte einen sehr ähnlichen Verlauf.

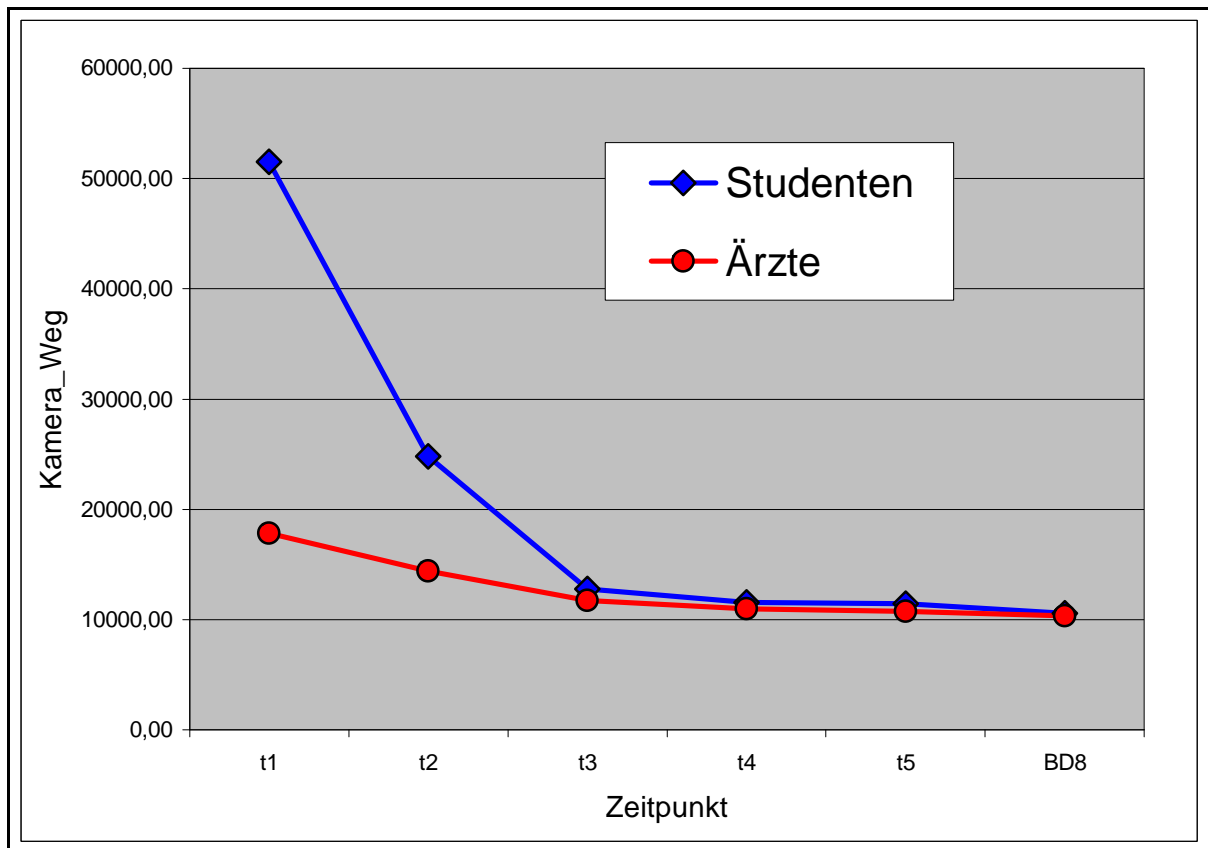


Abb. 13: Einzelparameter Kamera_Weg, Studenten gegen Ärzte (Normal und X-Over), deutliche Lernkurve

Bis zum Tag 3 glichen sich alle Gruppen in einer Plateauphase einander an. Ab dem Tag 4 ließen sich keine signifikanten Unterschiede mehr zwischen den beiden Gruppen erkennen.

Auch die Gesamtscores der Trainingsphase zeigten für den Gesamtscore 1 (Weg + Zeit), dass die Ärzte an den ersten beiden Tagen signifikant besser waren (siehe Abb. 14).

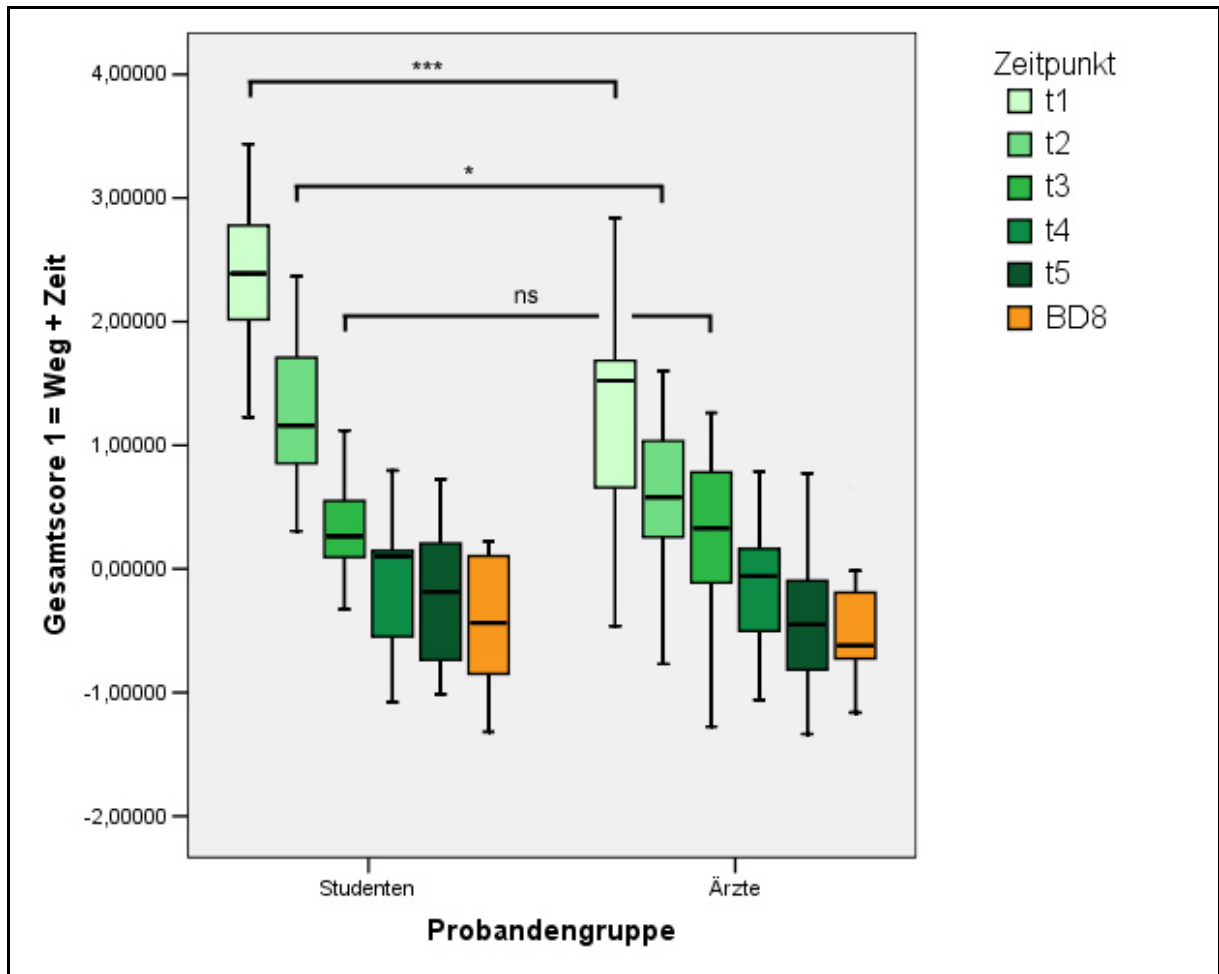


Abb. 14: Boxplot des Gesamtscores 1 = Weg + Zeit, Studenten gegen Ärzte, (je niedriger der Wert, desto besser die Leistung); Ärzte sind signifikant besser an den Trainingstagen 1 und 2; t1: $p \leq 0,001$; t2: $p = 0,025$

Der Gesamtscore 2 (Fehler) hingegen zeigte an allen fünf Trainingstagen keine signifikanten Unterschiede (siehe Abb. 15).

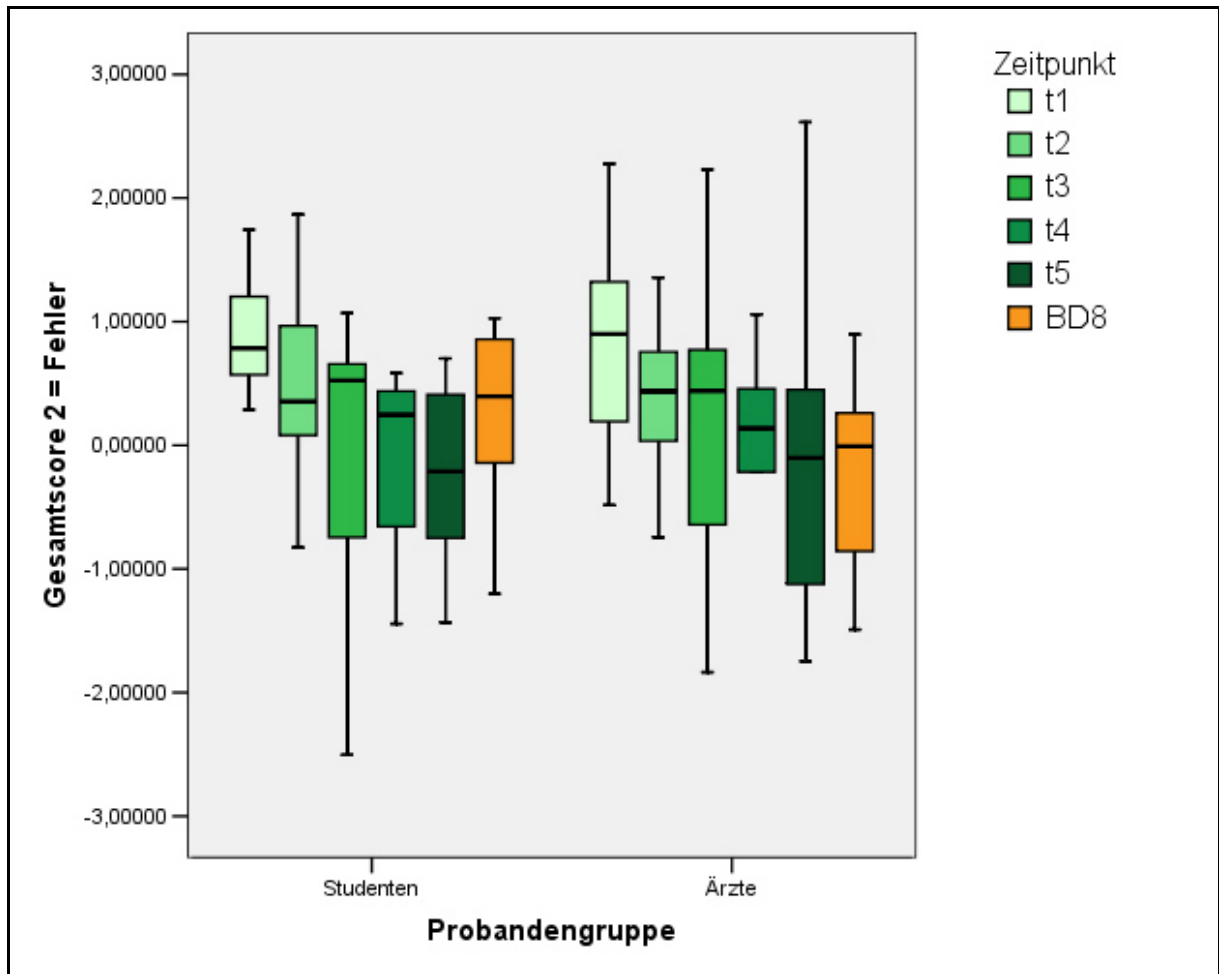


Abb. 15: Boxplots des Gesamtscores 2 = Fehler, Studenten gegen Ärzte, (je niedriger der Wert, desto besser die Leistung); keine signifikanten Unterschiede

Weiterhin ließ sich eine Abnahme der Steigung der Lernkurven mit zunehmendem Erfahrungsgrad erkennen. D.h., die Probanden mit mehr Erfahrung absolvierten das Training besser/schneller. Dies ist anhand des Gesamtscores 1 verdeutlicht (siehe Abb. 16).

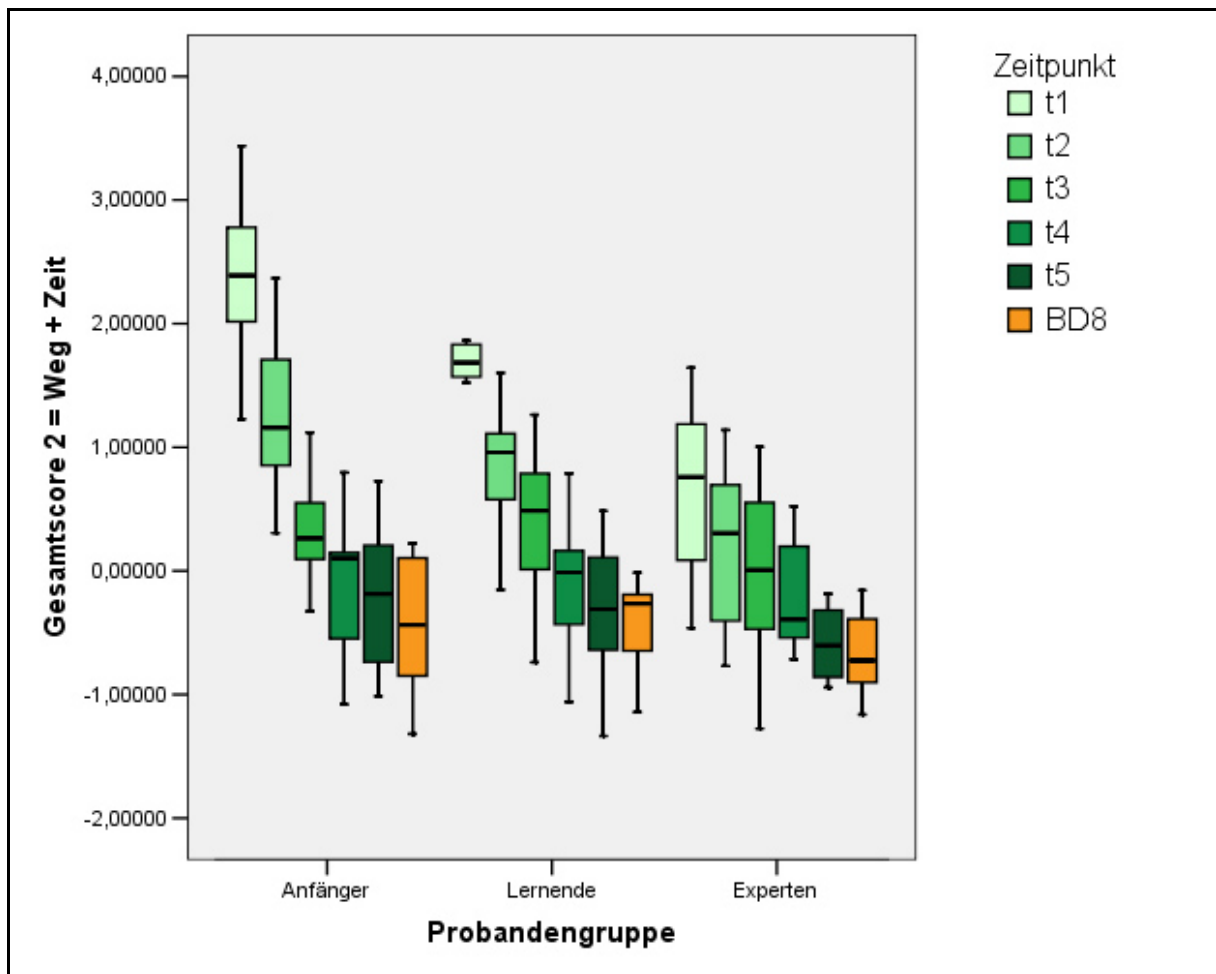


Abb. 16: Boxplot Gesamtscore 1 = Weg + Zeit, Subgruppen nach Erfahrungsgrad, (je niedriger der Wert, desto besser die Leistung); je größer der Erfahrungsgrad, desto geringer der Lernzuwachs

4.3 Bereitschaftsdienst

Die Bereitschaftsdienstphase war die Testphase der Studie, die eventuelle Leistungsdefizite nach Schlafentzug untersuchen sollte. Die Testzeitpunkte BD8 und BD16 erfolgten zu Beginn und nach 8 Stunden Dienst ohne Schlafentzug. Der Zeitpunkt FD8 fand am Folgetag statt nach Schlafentzug. Am Tag darauf, dem Recovery Day (RD8) erfolgte eine erneute Testung, wieder ohne Schlafentzug.

FD8 als entscheidender Testzeitpunkt nach Schlafentzug ist in den folgenden Abbildungen rot dargestellt.

4.3.1 Gruppenanalyse: Studenten und Ärzte

Betrachtet man die Ergebnisse der Simulatorübungen für die Studenten und Ärzte, so findet man auch hier eine Tendenz zur Verbesserung zu jedem weiteren Testzeitpunkt, die zum Teil signifikant zum Ausgangszeitpunkt BD8 ist. In den Abbildungen 17 und 18 sind die Gesamtscores 1 (Weg+Zeit) und 2 (Fehler) dargestellt.

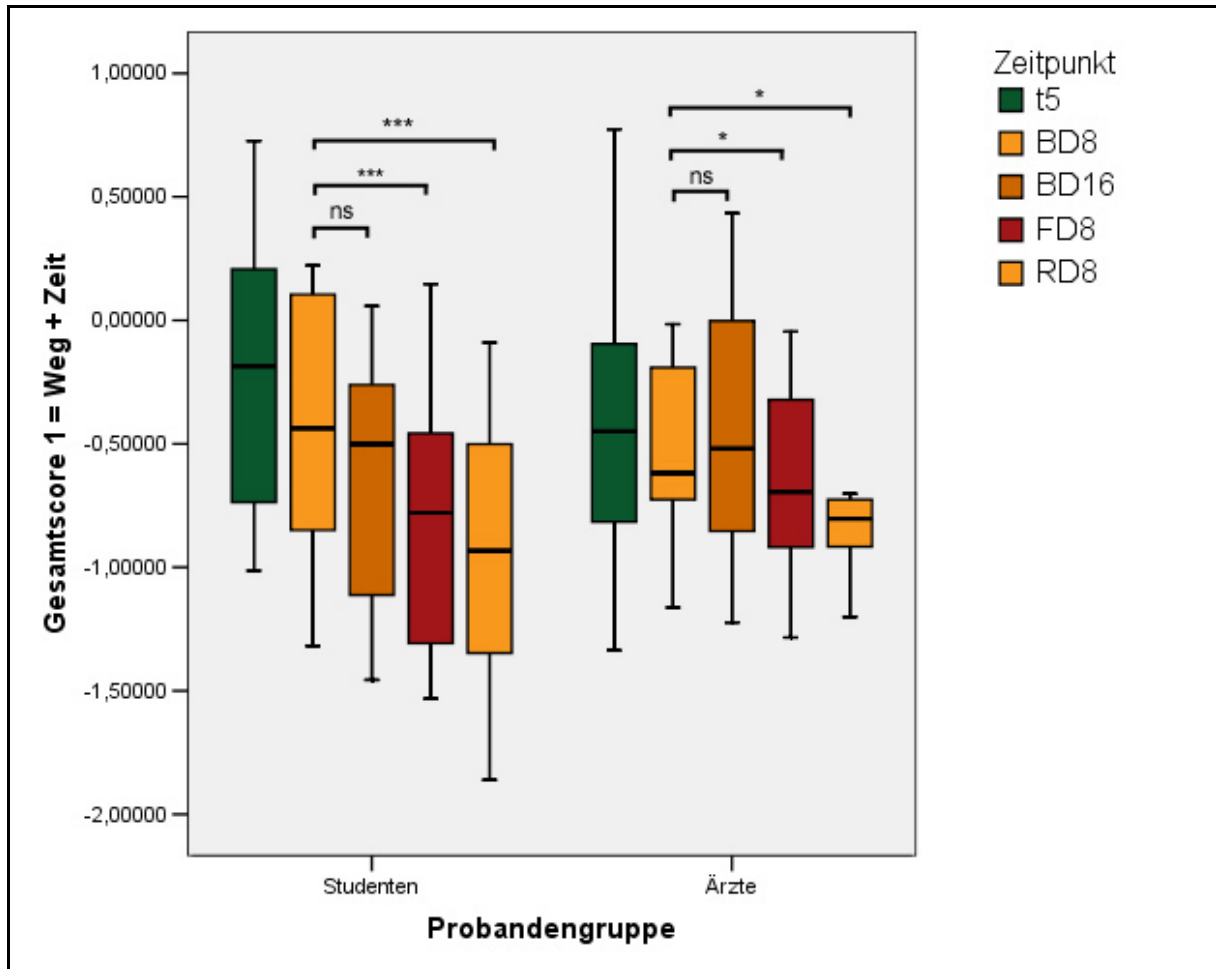


Abb. 17: Boxplot des Gesamtscores 1 = Weg + Zeit, Bereitschaftsdienstphase, Vergleich zwischen BD8 und den Folgetestzeitpunkten, Studenten und Ärzte, (je niedriger der Wert, desto besser die Leistung); keine signifikante Verschlechterung der Folgezeitpunkte im Vergleich zum Ausgangszeitpunkt BD8, Tendenz zur Verbesserung:

Studenten: BD8 zu BD16 ns, BD8 zu FD8 p = 0,001, BD8 zu RD8 p = 0,001
 Ärzte: BD8 zu BD16 ns, BD8 zu FD8 p = 0,028, BD8 zu RD8 p = 0,046

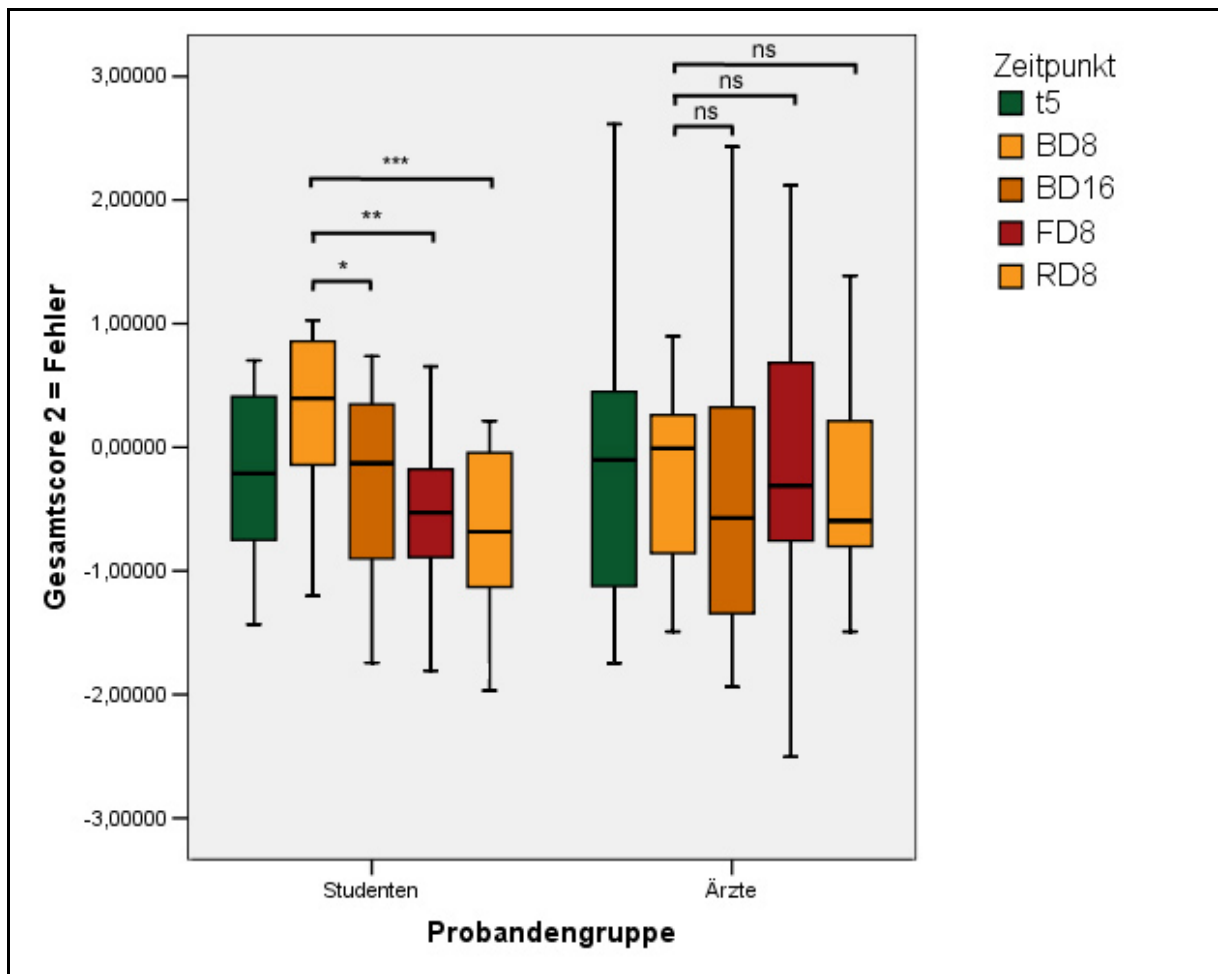


Abb. 18: Boxplot des Gesamtscores 2 = Fehler, Bereitschaftsdienstphase, Vergleich zwischen BD8 und den Folgetestzeitpunkten, Studenten und Ärzte, (je niedriger der Wert, desto besser die Leistung); keine signifikante Verschlechterung der Folgezeitpunkte im Vergleich zum Ausgangszeitpunkt BD8, Tendenz zur Verbesserung:

Studenten: BD8 zu BD16 $p = 0,016$, BD8 zu FD8 $p = 0,002$, BD8 zu RD8 $p = 0,001$
 Ärzte: BD8 zu BD16 ns, BD8 zu FD8 ns, BD8 zu RD8 ns

Im Vergleich innerhalb der Gruppen verbesserten sich die Studenten während der Testphase in beiden Gesamtscores der Simulatorübungen. Fast alle Folgezeitpunkte waren im Vergleich zum Ausgangszeitpunkt BD8 signifikant besser.

Auch die Ärzte verbesserten sich weiter in allen Tests, mit Ausnahme des Gesamtscores 2, also in der Anzahl der Fehler. Eine zu BD8 signifikante Verschlechterung ließ sich jedoch zu keinem Zeitpunkt und in keinem Parameter oder Gesamtscore finden (auch nicht bei den Ärzten zu FD8, dem Zeitpunkt nach Schlafentzug).

Der Vergleich Studenten gegen Ärzte zeigte auch in der Bereitschaftsdienstphase keine signifikanten Unterschiede. Gleiches galt für beide psychometrische Tests.

In den Abbildungen 19 und 20 sind die Ergebnisse der Konzentrationstests für die Bereitschaftsdienstphase dargestellt.

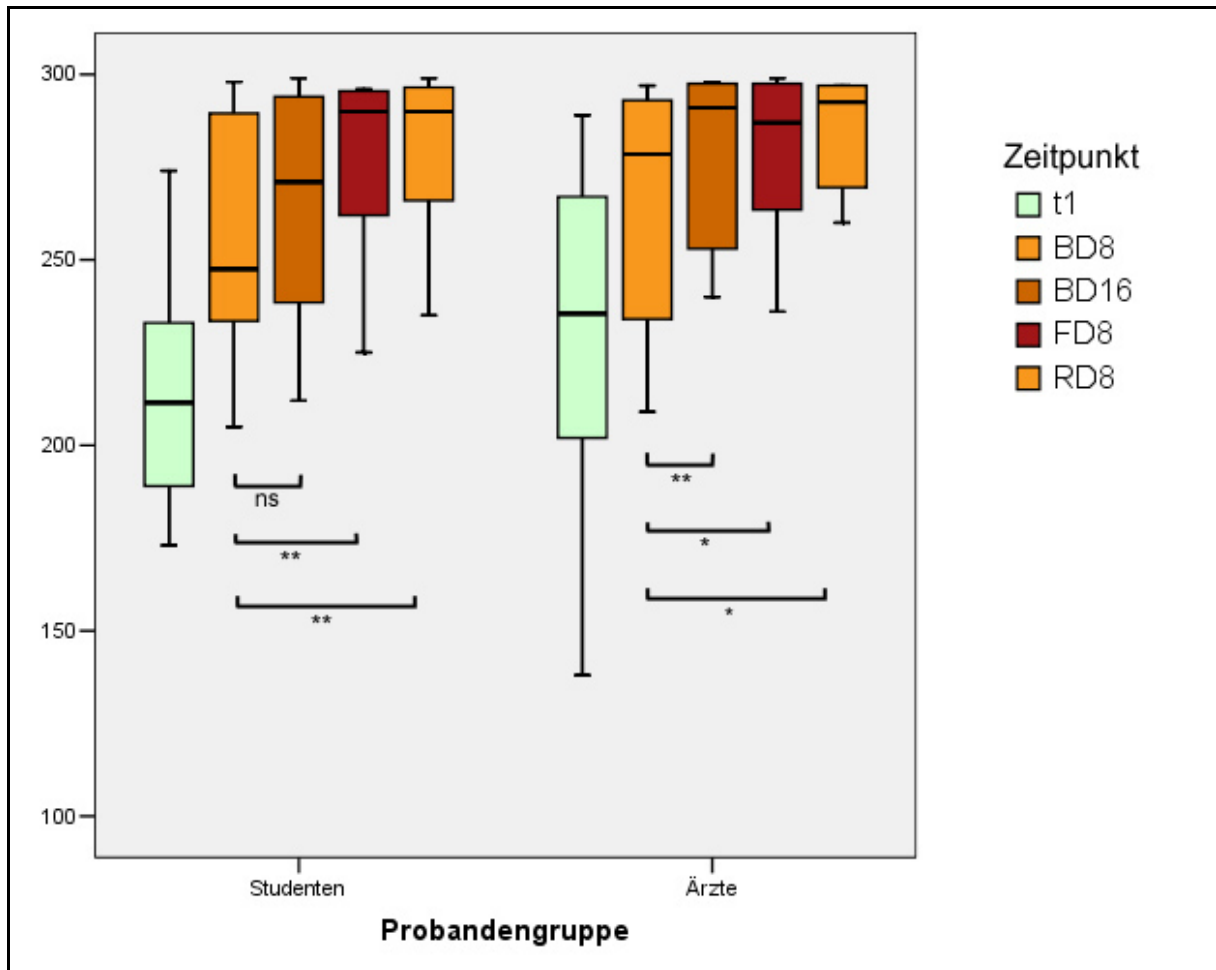


Abb. 19: Boxplot d2-KL, Bereitschaftsdienstphase, Vergleich zwischen BD8 und den Folgetestzeitpunkten, Studenten und Ärzte, (je höher der Wert, desto besser die Leistung); bei beiden Gruppen zeigte sich eine signifikante Verbesserung fast aller Folgezeitpunkte im Vergleich zum Ausgangszeitpunkt BD8.

Studenten: BD8 zu BD16 ns, BD8 zu FD8 $p = 0,005$, BD8 zu RD8 $p = 0,002$

Ärzte: BD8 zu BD16 $p = 0,002$, BD8 zu FD8 $p = 0,023$, BD8 zu RD8 $p = 0,013$

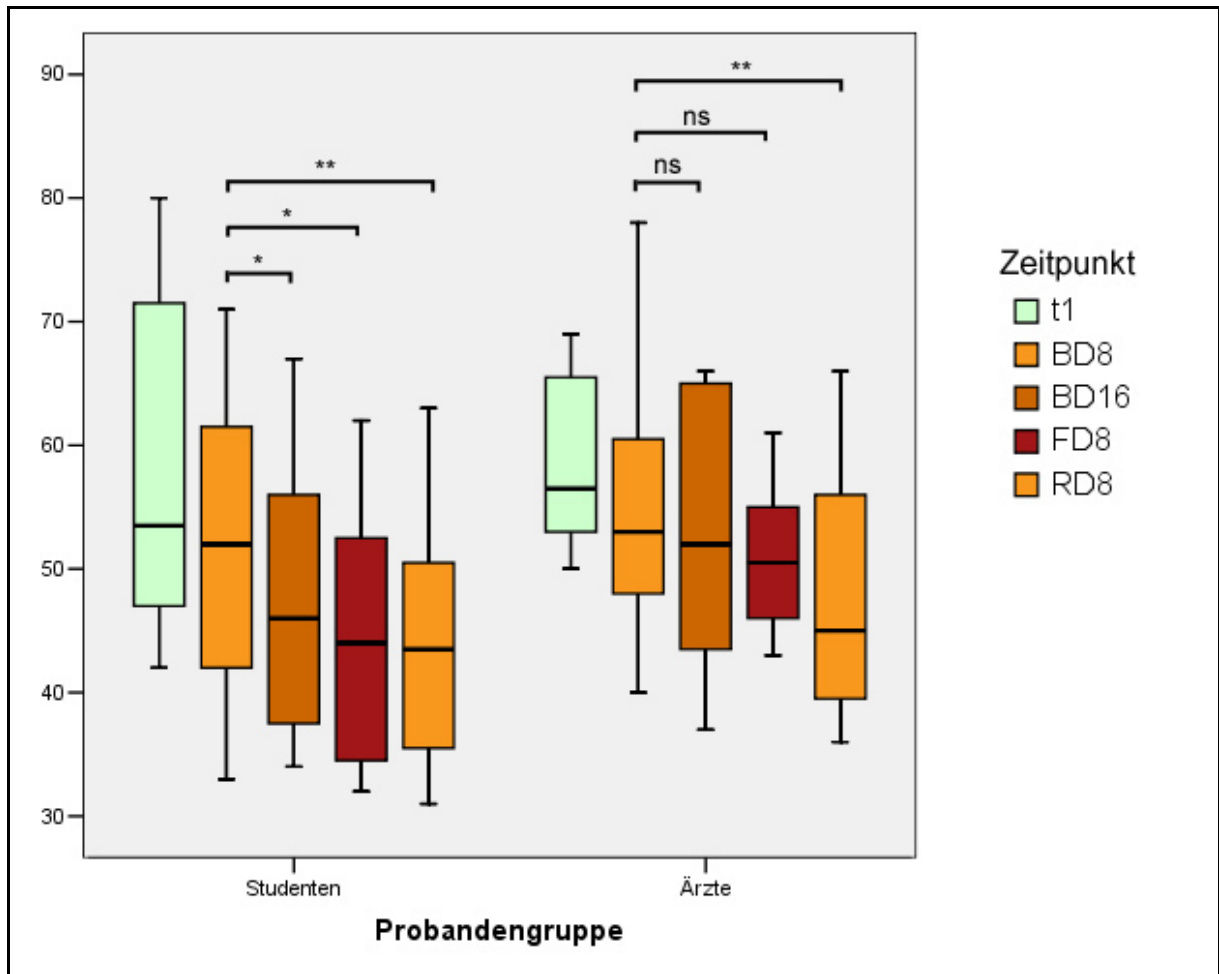


Abb. 20: Boxplot ZVT-Zeit, Bereitschaftsdienstphase, Vergleich zwischen BD8 und den Folgetestzeitpunkten, Studenten und Ärzte, (je niedriger der Wert, desto besser die Leistung); bei beiden Gruppen zeigte sich eine signifikante Verbesserung fast aller Folgezeitpunkte im Vergleich zum Ausgangszeitpunkt BD8.

Studenten: BD8 zu BD16 $p = 0,005$, BD8 zu FD8 $p = 0,012$, BD8 zu RD8 $p = 0,009$
 Ärzte: BD8 zu BD16 ns, BD8 zu FD8 ns, BD8 zu RD8 $p = 0,009$

In beiden psychometrischen Tests lässt sich eine signifikante Verbesserung fast aller Folgezeitpunkte zum Ausgangszeitpunkt BD8 feststellen.

4.3.2 Subgruppenanalyse Erfahrung

Mit der Aufteilung der Probanden in die Subgruppe Erfahrung wurde untersucht, ob es Leistungsunterschiede abhängig vom Erfahrungsgrad gibt. Es zeigte sich im Vergleich der Subgruppen untereinander, dass in der Bereitschaftsdienstphase keine signifikanten Leistungsunterschiede mehr bestanden. Lediglich bei Betrachtung des

Gesamtscores 2, also der Fehler, fällt auf, dass die Experten zu BD8 signifikant weniger Fehler machten, als die Anfänger und die Lernenden (siehe Abb. 21).

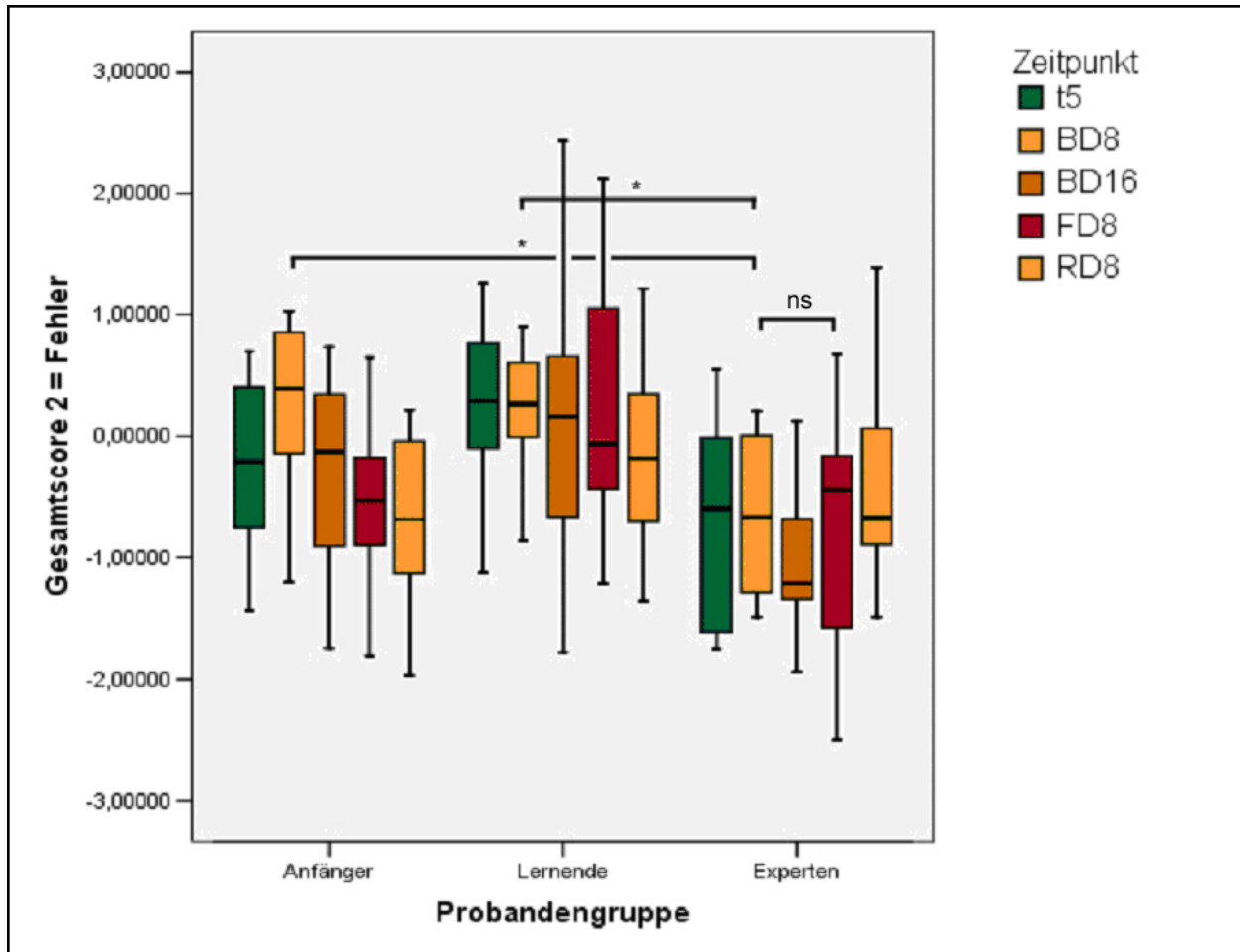


Abb. 21: Boxplot Gesamtscore 2 = Fehler, Anfänger gegen Lernende gegen Experten, (je niedriger der Wert, desto besser die Leistung); signifikant weniger Fehler bei den Experten im Vergleich zu den Anfängern und Lernenden zu BD8:

Lernende gegen Experten zu BD8 $p = 0,021$; Anfänger gegen Experten zu BD8 $p = 0,014$

Innerhalb der Expertengruppe: BD8 zu BD16: ns

Im Vergleich der Testzeitpunkte innerhalb der Gruppen zeigten sich signifikante Veränderungen in allen drei Gruppen lediglich im d2-Test (siehe Abb. 22).

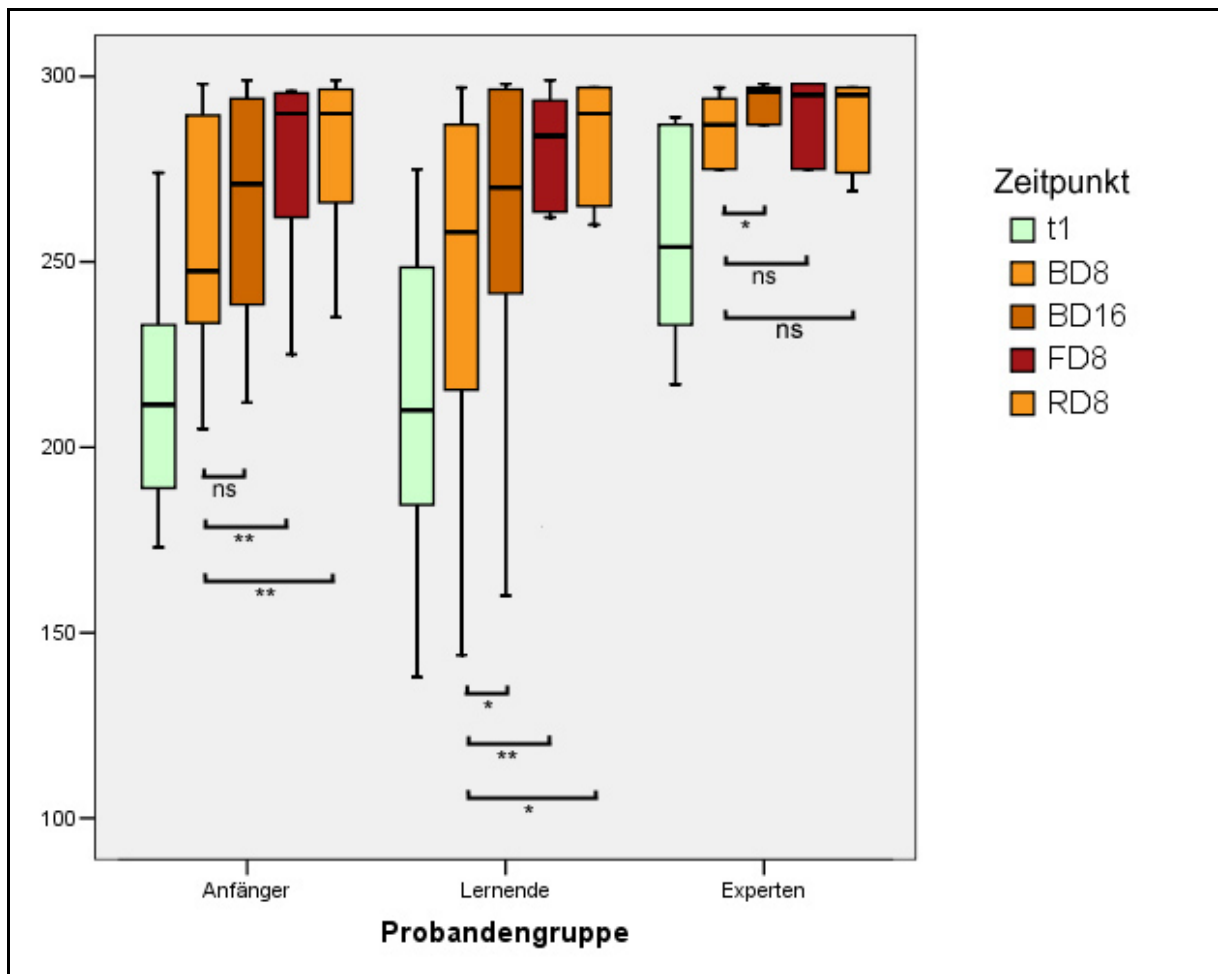


Abb. 22: Boxplot D2-Test, Vergleich der Testzeitpunkte innerhalb der Gruppen Anfänger, Lernende, Experten: (je höher der Wert, desto besser die Leistung); signifikante Verbesserung aller Gruppen auch in der Bereitschaftsdienstphase:

Anfänger: BD8 zu BD16 ns, BD8 zu FD8 $p = 0,005$, BD8 zu RD8 $0,002$
 Lernende: BD8 zu BD16 $p = 0,015$, BD8 zu FD8 $p = 0,008$, BD8 zu RD8 $p = 0,028$
 Experten: BD8 zu BD16 $p = 0,043$, BD8 zu FD8 ns, BD8 zu RD8 ns,
 keine Verschlechterungen

Die Verbesserung der Anfänger wurde bereits im Vergleich Studenten gegen Ärzte erwähnt.

Die Lernenden verbesserten sich an allen drei Folgezeitpunkten signifikant zu BD8.

Die Experten verbesserten sich lediglich an BD16 signifikant zu BD8.

Eine Verschlechterung der Folgetestzeitpunkte in Vergleich zu BD8 konnte in allen drei Gruppen nicht festgestellt werden (auch nicht bei den Lernenden und Experten zu FD8, dem Zeitpunkt nach Schlafentzug).

Im Zahlenverbindungstest zeigten sich weder signifikante Verschlechterungen noch Verbesserungen (siehe Abb. 23)

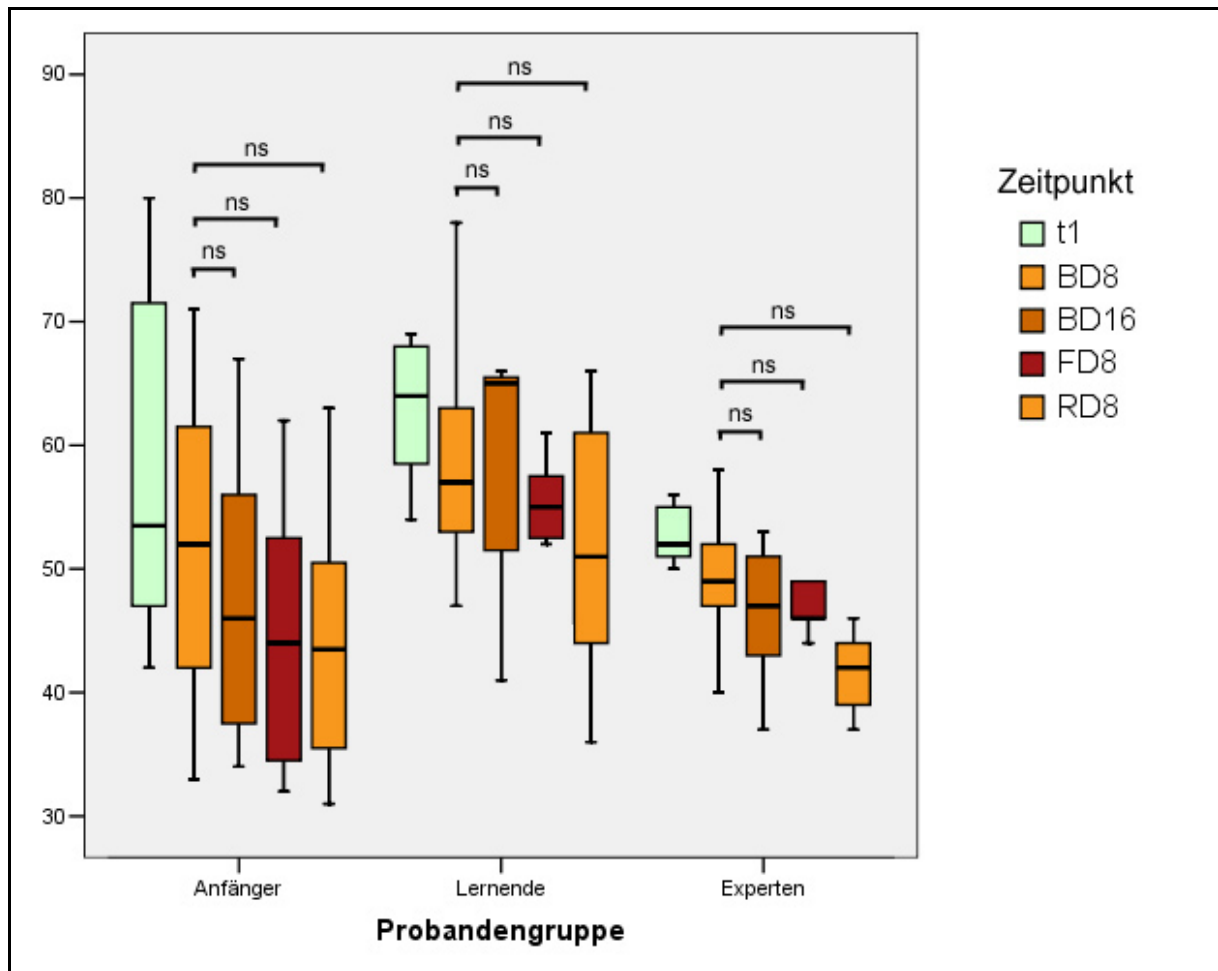


Abb. 23: Boxplot ZVT-Zeit, Vergleich der Testzeitpunkte innerhalb der Gruppen Anfänger, Lernende, Experten: (je niedriger der Wert, desto besser die Leistung); keine signifikante Verbesserung/Verschlechterung aller Gruppen in der Bereitschaftsdienstphase

4.3.3 Subgruppenanalyse Schlaf

Die Subgruppenanalyse Schlaf teilte die Probanden nach den Schlafstunden zwischen BD8 (dem Beginn des Bereitschaftsdienstes) und FD8 (dem Folgemorgen) ein.

- 1) Studenten: Normalschlaf, $\geq 6h$
- 2) Ärzte: $< 3h$
- 3) Ärzte: $\geq 3h$

Damit sollten die Leistungsunterschiede in Abhängigkeit vom Schlafentzug genauer untersucht werden.

In der Gruppe der Studenten zeigten sich keine Verschlechterungen der Folgezeitpunkte (siehe Abb. 24). Alle Zeitpunkte waren im Vergleich zu BD8 signifikant besser.

Keine Verschlechterungen zeigte ebenfalls die Gruppe der Ärzte mit ≥ 3 h und < 6 h Schlaf. Gesamtscore 1 sowie ZVT-Zeit waren zum Zeitpunkt RD8 sogar signifikant besser im Vergleich zu BD8 (siehe Abb. 25).

Die Gruppe mit dem größten Schlafentzug (< 3 h) zeigte in der Bereitschaftsdienstphase weder eine signifikante Verbesserung noch eine Verschlechterung, auch nicht zu FD8 (nach Schlafentzug) und zwar in allen Übungen und Tests. Lediglich in BD16 zu BD8 im d2-Test zeigte sich eine signifikante Verbesserung (siehe Abb. 26).

Zwischen BD8 und FD8 gab es also in allen Gruppen keine Verschlechterungen, sondern lediglich Verbesserungen, die zum Teil signifikant waren.

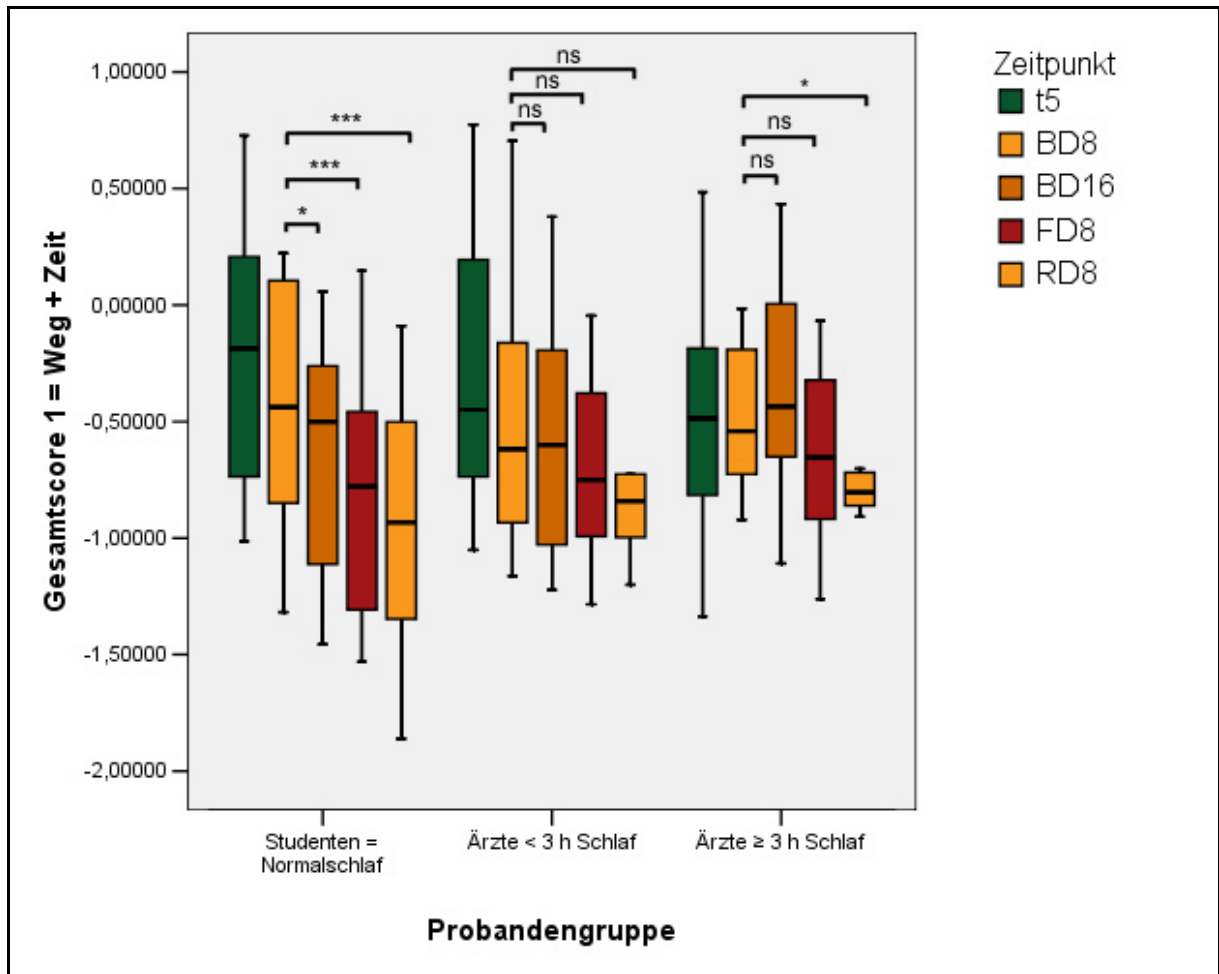


Abb. 24: Boxplot Gesamtscore 1 = Weg + Zeit, Normalschlaf gegen < 3h Schlaf gegen ≥ 3h Schlaf, (je niedriger der Wert, desto besser die Leistung); zum Teil signifikante Verbesserungen, keine signifikante Verschlechterung nach Schlafentzug;

Normalschlaf: BD8 zu BD16 $p = 0,04$, BD8 zu FD8 $p = 0,001$, BD8 zu RD8 $p = 0,001$
 < 3h Schlaf: BD8 zu BD16 ns, BD8 zu FD8 ns, BD8 zu RD8 ns
 ≥ 3h Schlaf: BD8 zu BD16 ns, BD8 zu FD8 ns, BD8 zu RD8 $p = 0,018$,

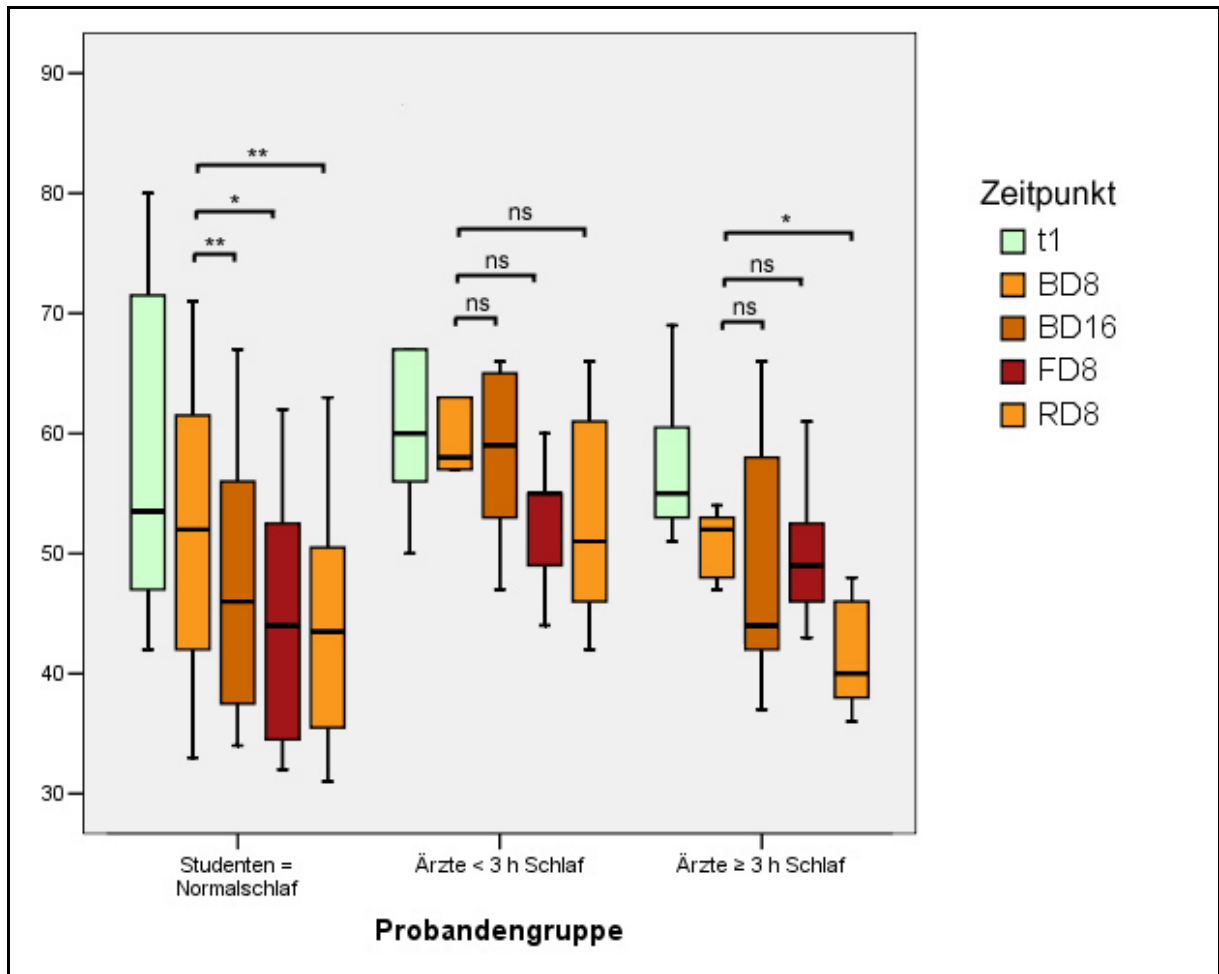


Abb. 25: Boxplot ZVT-Zeit, Normalschlaf gegen < 3h Schlaf gegen ≥ 3h Schlaf; (je niedriger der Wert, desto besser die Leistung); zum Teil signifikante Verbesserungen, keine signifikante Verschlechterung nach Schlafentzug;

Normalschlaf: BD8 zu BD16 $p = 0,005$, BD8 zu FD8 $p = 0,012$, BD8 zu RD8 $p = 0,009$
 < 3h Schlaf: BD8 zu BD16 $p = ns$, BD8 zu FD8 ns , BD8 zu RD8 ns
 ≥ 3h Schlaf: BD8 zu BD16 ns , BD8 zu FD8 ns , BD8 zu RD8 $p = 0,028$

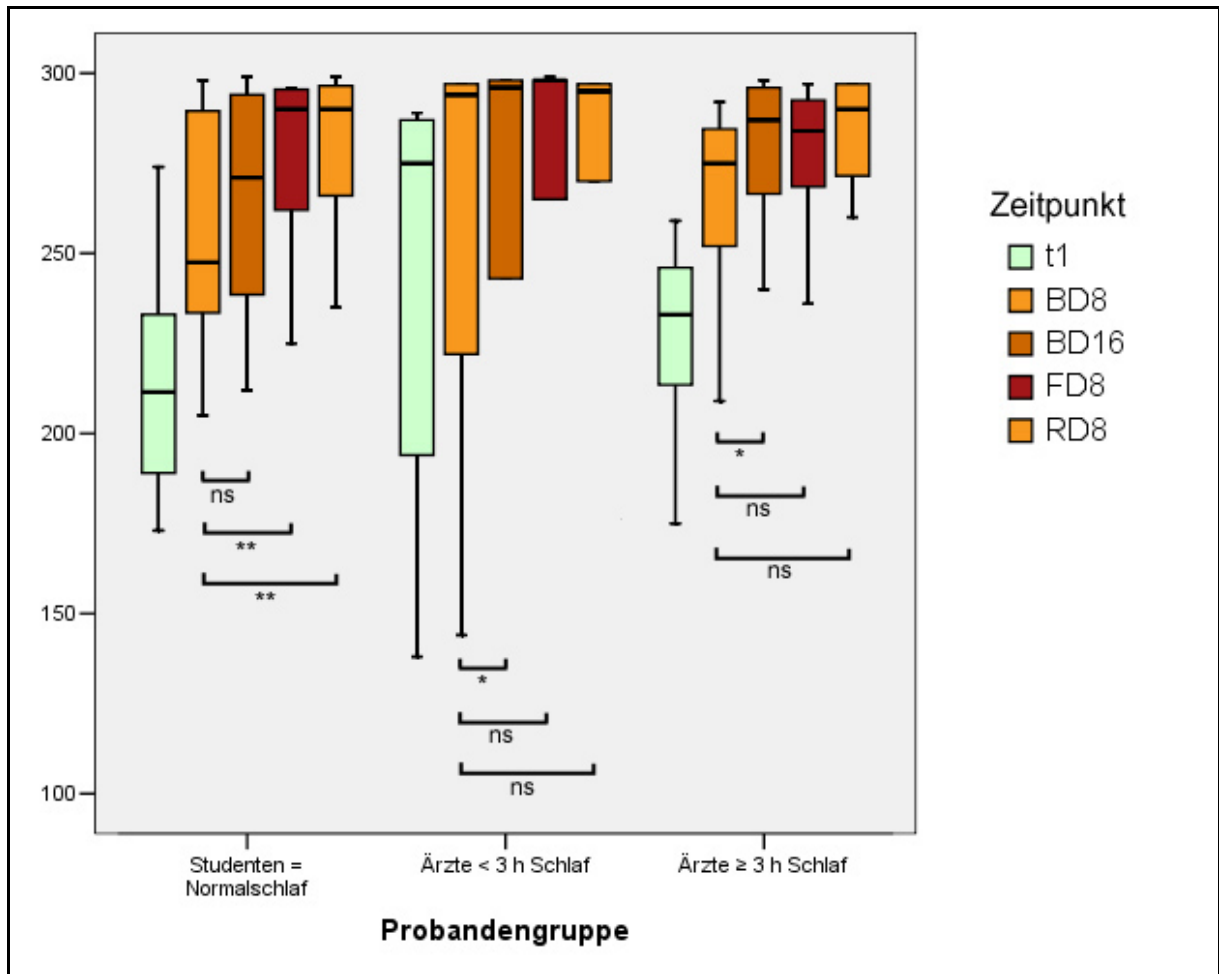


Abb. 26: Boxplot d2-KL, Normalschlaf gegen < 3h Schlaf gegen ≥ 3h Schlaf; (je höher der Wert, desto besser die Leistung); zum Teil signifikante Verbesserungen, keine signifikante Verschlechterung nach Schlafentzug;

Normalschlaf: BD8 zu BD16 ns, BD8 zu FD8 $p = 0,005$, BD8 zu RD8 $p = 0,002$
 < 3h Schlaf: BD8 zu BD16 $p = 0,027$, BD8 zu FD8 ns, BD8 zu RD8 ns
 ≥ 3h Schlaf: BD8 zu BD16 $p = 0,017$, BD8 zu FD8 ns, BD8 zu RD8 ns

4.3.4 Testdauer

Die Dauer der einzelnen Tests wurde sowohl vom Simulator als auch bei den psychometrischen Tests vom Testleiter genau erfasst. Somit konnte die Gesamttestdauer zu jedem Testzeitpunkt berechnet werden (siehe Tabelle 3).

Tabelle 3: Durchschnittliche Testdauer zu den Testzeitpunkten, Simulatorübungen + psychometrische Tests

	Testdauer in Sekunden Studenten	Testdauer in Sekunden Ärzte
BD8	1697,5	1610,4
BD16	1623,2	1594,6
FD8	1605,6	1579,6
RD8	1574,6	1578,8
Mittelwert	1625,2 = 27,1 min.	1578,8 = 26,3 min.

4.4 Cross-Over-Analyse

Beim Vergleich der Gruppe Ärzte „Normal“ mit der „Cross-Over“-Gruppe zeigten sich keine signifikanten Unterschiede, sowohl in der Trainingsphase als auch in der Testphase (siehe Abb. 27). Dies bedeutet, dass bei den Testzeitpunkten BD8 und BD16 also vor dem Schlafentzug keine weiteren signifikanten Lerneffekte mehr vorhanden waren.

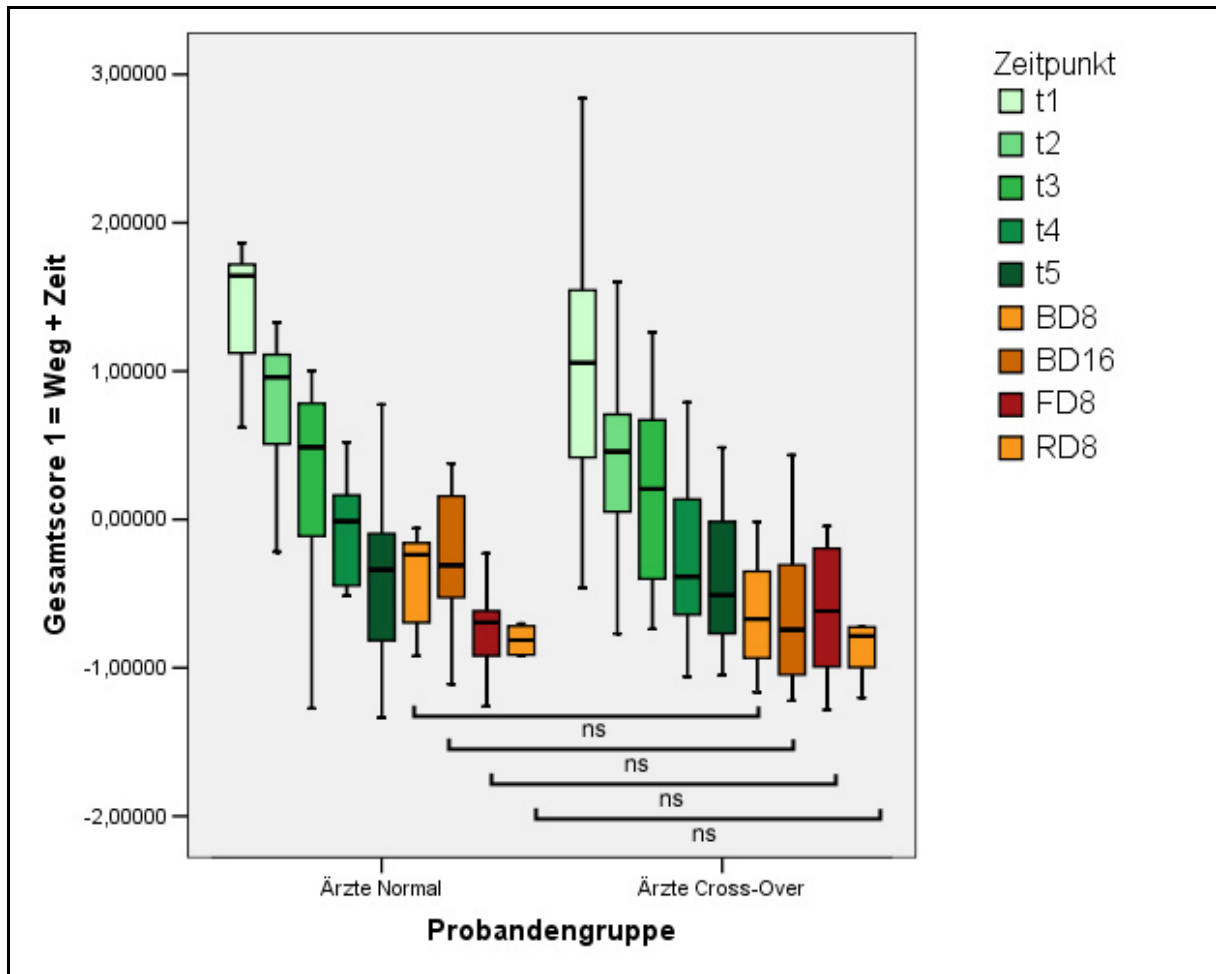


Abb. 27: Vergleich Ärzte Normal gegen Ärzte Cross-Over, (je niedriger der Wert, desto besser die Leistung); es zeigten sich keine Signifikanten Unterschiede: $p < 0,05$ = signifikant, p -Werte des Vergleichs Ärzte Normal gegen Ärzte Cross-Over an BD8 = 0,178; BD16 = 0,141, FD8 = 0,7, RD8 = 0,568

4.5 Auswertung der Fragebögen

Die meisten Probanden hatten in den letzten fünf Stunden vor den Testzeitpunkten etwas gegessen. Zum Zeitpunkt FD8 hatten diverse Ärzte jedoch eine Nahrungskarenz von mehr als 5 bis zu 10 Stunden.

Die Studenten fühlten sich stets gut bis etwas leistungseingeschränkt. Bei den Ärzten findet man lediglich zu FD8 einen subjektiven Leistungsknick (siehe Abb. 28). 10 Ärzte fühlten sich etwas müde, die restlichen sieben Ärzte fühlten sich müde und leistungseingeschränkt.

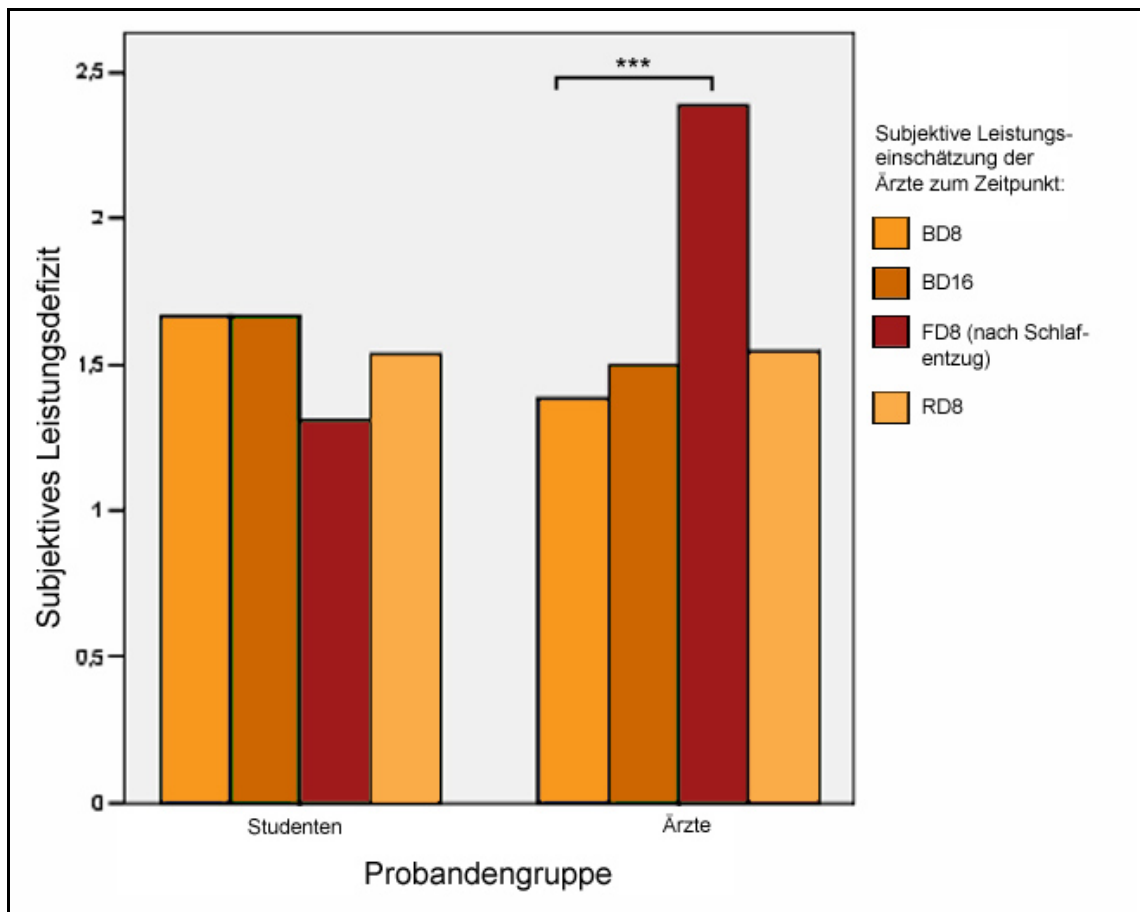


Abb. 28: Säulendiagramm der Subjektiven Leistungseinschätzung von Ärzten und Studenten, signifikanter subjektiver Leistungsknick der Ärzte nach Schlafentzug an FD8

Sowohl die Ärzte als auch die Studenten waren in der Regel guter Stimmung.

Die Studenten fühlten sich zu allen Testzeitpunkten gesund. Gleiches galt auch für die Ärzte. Nur zu FD8 fühlte sich eine Person krank.

Die Frage, ob sie jetzt genauso gut operieren könnten, wie zu BD8 beantworteten vier Ärzte mit ja, sieben beurteilten sich als in einer Operation leistungseingeschränkt, wobei sie sich jedoch in der Lage fühlten, diese Einschränkung zu kompensieren. Die restlichen sechs Ärzte empfanden ebenfalls eine leichte Leistungseinschränkung, von der sie nicht glaubten, diese anderweitig kompensieren zu können.

Etwa die Hälfte der Ärzte gab an, in einem Patientengespräch zu FD8 genauso geduldig sein zu können, wie zu BD8. Die andere Hälfte beurteilte sich als ungeduldiger und gereizter.