

Aus der
- Klinik und Poliklinik für Kinderchirurgie -
der Universitätsmedizin der Johannes Gutenberg-Universität Mainz

„Der Einfluss von Alkohol auf die operative Performance am MIMIC dV-Trainer®“

Inauguraldissertation
zur Erlangung des Doktorgrades der Medizin
der Universitätsmedizin
der Johannes Gutenberg-Universität Mainz

Vorgelegt von
Corinna Schnorpfeil, geb. Schmitz
aus Koblenz

Mainz, 2023

Tag der Promotion:

28. Juni 2023

Eidesstattliche Erklärung

„Ich versichere, dass ich die Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen oder anderen Quellen (auch Internet) entnommen sind, habe ich als solche eindeutig kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht und noch nicht zur Anerkennung oder Bewertung vorgelegt worden. Mir ist bekannt, dass Verstöße gegen diese Anforderungen zur Bewertung der Arbeit mit der Note „nicht ausreichend“ führen sowie die Nichterteilung des angestrebten Leistungsnachweises zur Folge haben.“

Münstermaifeld, den 27.06.2023



Inhaltsverzeichnis

Abkürzungsverzeichnis	IV
Abbildungsverzeichnis	V
Tabellenverzeichnis	VII
1 Einleitung	1
1.1 Allgemeines	1
1.2 Problemstellung.....	2
1.3 Zielsetzung und Aufbau der Studie	2
2 Literaturdiskussion	3
2.1 Definition des Begriffs „Simulation“ und Einordnung des MIMIC dV-Trainers®	3
2.2 Physiologische Effekte von Alkohol und Auswirkungen von Alkoholkonsum	5
2.3 Alkoholeinfluss bei Ärzten und Chirurgen.....	7
3 Material und Methoden	13
3.1 Aufbau MIMIC dV-Trainer®.....	13
3.2 Übungen und Auswertungsparameter.....	15
3.2.1 <i>Allgemeines</i>	15
3.2.2 <i>Erläuterung des Bewertungssystems</i>	15
3.2.3 <i>Übungen</i>	16
3.2.3.1 Ring Tower	16
3.2.3.2 Suture Sponge	18
3.2.3.3 Stacking Challenge	19
3.2.3.4 Vitruvian Operation.....	21
3.3 Probandenkollektiv und Probandenakquise	22
3.4 Intervention Alkohol	23
3.5 Studienbeschreibung.....	23
3.5.1 <i>Vorstudie zur Lernkurve</i>	23
3.5.2 <i>Studiendurchführung</i>	24
3.6 <i>Studienfinanzierung</i>	26
3.7 Statistische Testverfahren	26
4 Ergebnisse	28
4.1 Primäre Messparameter	28

4.1.1	<i>Ring Tower</i>	28
4.1.1.1	Time to complete.....	29
4.1.1.2	Wire collisions	31
4.1.1.3	Wire contact duration	33
4.1.1.4	Tower Knockoff	35
4.1.1.5	Economy of motion.....	37
4.1.1.6	Instrument collisions.....	39
4.1.1.7	Instruments out of view	41
4.1.1.8	Drops.....	43
4.1.2	<i>Suture Sponge</i>	45
4.1.2.1	Time to complete.....	46
4.1.2.2	Economy of motion.....	48
4.1.2.3	Instrument collisions.....	50
4.1.2.4	Excessive instrument force	52
4.1.2.5	Instruments out of view	54
4.1.2.6	Master workspace range	56
4.1.2.7	Drops.....	58
4.1.2.8	Missed Targets.....	59
4.1.3	<i>Stacking Challenge</i>	61
4.1.4	<i>Vitruvian Operation</i>	63
4.2	Tabellarische Zusammenfassung der Ergebnisse	65
4.3	Sekundäre Messparameter	66
4.3.1	<i>Blutalkoholkonzentration</i>	66
4.3.2	<i>Trinkgewohnheiten</i>	66
5	Diskussion	68
5.1	Zusammenfassung und Interpretation der Ergebnisse	68
5.2	Vergleich mit bestehender Literatur.....	69
5.3	Praktische Implikationen und Bedeutung in der Chirurgie	74
5.4	Limitierung und Ausblick.....	75
6	Zusammenfassung	77
7	Literaturverzeichnis	78

8	Anhang	81
8.1	Tabelle zu den Leistungsmetriken.....	81
8.2	Bewertungsschlüssel der einzelnen Leistungsmetriken.....	83
8.3	Bestimmung der benötigten Alkoholmenge.....	92
8.4	Flowchart zum geplanten Studienablauf	93

Abkürzungsverzeichnis

BZgA	Bundeszentrale für gesundheitliche Aufklärung
bzw.	beziehungsweise
dV-Trainer	da Vinci -Trainer®
et al.	“et alii”: und andere
BAK	Blutalkoholkonzentration
Op-Saal	Operationssaal
VR-Simulator	Virtual Reality Simulator
RT	Ring Tower
SS	Suture Sponge
ST	Stacking Challenge
VO	Vitruvian Operation

Abbildungsverzeichnis

Abbildung 1. Widmark-Formel zur Berechnung der Alkoholkonzentration. Übernommen aus Graw (2015).	6
Abbildung 2. MIMIC dV-Trainer®. Entnommen aus Mimic Technologies (2018).	13
Abbildung 3. „Foodswitch Panel“ des MIMIC dV-Trainer. Entnommen aus Mimic Technologies (2018).	14
Abbildung 4. Beispiel des Bewertungsschlüssels einer Metrik.	16
Abbildung 5. Auszug aus Ring Tower Bild 1. Entnommen aus Mimic Technologies (2016).	17
Abbildung 6. Auszug aus Ring Tower Bild 2. Entnommen aus Mimic Technologies (2016).	17
Abbildung 7. Metriken der Auswertung zu Ring Tower. Eigene Darstellung.	18
Abbildung 8. Auszug aus Suture Sponge. Entnommen aus Mimic Technologies (2016).	19
Abbildung 9. Metriken der Auswertung zu Suture Sponge.	19
Abbildung 10. Auszug aus Stacking Challenge. Entnommen aus Mimic Technologies (2016).	20
Abbildung 11. Auswertung Stacking Challenge.	20
Abbildung 12. Auszug aus Vitruvian Operation. Entnommen aus Mimic Technologies (2016).	21
Abbildung 13. Auswertung Vitruvian Operation.	22
Abbildung 14. Streudiagramm zu „Time to complete“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention	30
Abbildung 15. Streudiagramm zu „Wire collisions“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	32
Abbildung 16. Streudiagramm zu „Wire contact duration“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	34
Abbildung 17. Streudiagramm zu „Tower knockoff“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	36
Abbildung 18. Streudiagramm zu „Economy of motion“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	38
Abbildung 19. Streudiagramm zu „Instrument collisions“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	40
Abbildung 20. Streudiagramm zu „Instruments out of view“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	42
Abbildung 21. Streudiagramm zu „Drops“ der Übung Ring Tower mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	44
Abbildung 22. Streudiagramm zu „Time to complete“ der Übung Suture Sponge mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	47
Abbildung 23. Streudiagramm zu „Economy of motion“ der Übung Suture Sponge mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	49
Abbildung 24. Streudiagramm zu „Instrument collisions“ der Übung Suture Sponge mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	51
Abbildung 25. Streudiagramm zu „Excessive instrument force“ der Übung Suture Sponge mit (<u>_1</u>) und ohne (<u>_0</u>) Intervention.	53

Abbildung 26. Streudiagramm zu „Instruments out of view“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.....	55
Abbildung 27. Streudiagramm zu „Master workspace range“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.....	57
Abbildung 28. Streudiagramm zu „Drops“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.....	58
Abbildung 29. Streudiagramm zu „Missed targets“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.....	60
Abbildung 30. Streudiagramm zu der Übung Stacking Challenge mit (_1) und ohne (_0) Intervention.....	62
Abbildung 31. Streudiagramm zu der Übung Vitruvian Operation mit (_1) und ohne (_0) Intervention.....	64
Abbildung 32. Häufigkeitsverteilung zur Trinkgewohnheit.....	67

Tabellenverzeichnis

<i>Tabelle 1: Metriken und dazugehörige Einheiten</i>	28
<i>Tabelle 2: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken der Übung Ring Tower</i>	29
<i>Tabelle 3: Häufigkeiten der Differenzen von „Time to complete“ der Übung Ring Tower</i>	30
<i>Tabelle 4: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Time to complete“ der Übung Ring Tower</i>	31
<i>Tabelle 5: Häufigkeiten der Differenzen von „Wire collisions“ der Übung Ring Tower</i>	32
<i>Tabelle 6: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Wire collisions“ der Übung Ring Tower</i>	33
<i>Tabelle 7: Häufigkeiten der Differenzen von „Wire contact duration“ der Übung Ring Tower</i>	34
<i>Tabelle 8: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Wire contact duration“ der Übung Ring Tower</i>	35
<i>Tabelle 9: Häufigkeiten der Differenzen von „Tower knockoff“ der Übung Ring Tower</i>	36
<i>Tabelle 10: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Tower knockoff“ der Übung Ring Tower</i>	37
<i>Tabelle 11: Häufigkeiten der Differenzen von „Economy of motion“ der Übung Ring Tower</i>	38
<i>Tabelle 12: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Economy of motion“ der Übung Ring Tower</i>	39
<i>Tabelle 13: Häufigkeiten der Differenzen von „Instrument collisions“ der Übung Ring Tower</i>	40
<i>Tabelle 14: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instrument collisions“ der Übung Ring Tower</i>	41
<i>Tabelle 15: Häufigkeiten der Differenzen von „Instruments out of view“ der Übung Ring Tower</i>	42
<i>Tabelle 16: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instruments out of view“ der Übung Ring Tower</i>	43
<i>Tabelle 17: Häufigkeiten der Differenzen von „Drops“ der Übung Ring Tower</i>	44
<i>Tabelle 18: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Drops“ der Übung Ring Tower</i>	45
<i>Tabelle 19: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken der Übung Suture Sponge</i>	46
<i>Tabelle 20: Häufigkeiten der Differenzen von „Time to complete“ der Übung Suture Sponge</i>	47
<i>Tabelle 21: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Time to complete“ der Übung Suture Sponge</i>	48
<i>Tabelle 22: Häufigkeiten der Differenzen von „Economy of motion“ der Übung Suture Sponge</i>	49
<i>Tabelle 23: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Economy of motion“ der Übung Suture Sponge</i>	50
<i>Tabelle 24: Häufigkeiten der Differenzen von „Instrument collisions“ der Übung Suture Sponge</i>	51

<i>Tabelle 25: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instrument collisions“ der Übung Suture Sponge</i>	<i>52</i>
<i>Tabelle 26: Häufigkeiten der Differenzen von „Excessive instrument force“ der Übung Suture Sponge</i>	<i>53</i>
<i>Tabelle 27: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Excessive instrument force“ der Übung Suture Sponge</i>	<i>54</i>
<i>Tabelle 28: Häufigkeiten der Differenzen von „Instruments out of view“ der Übung Suture Sponge</i>	<i>56</i>
<i>Tabelle 29: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instruments out of view“ der Übung Suture Sponge</i>	<i>56</i>
<i>Tabelle 30: Häufigkeiten der Differenzen von „Master workspace range“ der Übung Suture Sponge.....</i>	<i>57</i>
<i>Tabelle 31: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Master workspace range“ der Übung Suture Sponge</i>	<i>58</i>
<i>Tabelle 32: Häufigkeiten der Differenzen von „Drops“ der Übung Suture Sponge</i>	<i>59</i>
<i>Tabelle 33: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Drops“ der Übung Suture Sponge</i>	<i>59</i>
<i>Tabelle 34: Häufigkeiten der Differenzen von „Missed targets“ der Übung Suture Sponge</i>	<i>60</i>
<i>Tabelle 35: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Missed targets“ der Übung Suture Sponge</i>	<i>61</i>
<i>Tabelle 36: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der Übung Stacking Challenge.....</i>	<i>61</i>
<i>Tabelle 37: Häufigkeiten der Differenzen der Übung Stacking Challenge.....</i>	<i>62</i>
<i>Tabelle 38: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert der Übung Stacking Challenge</i>	<i>63</i>
<i>Tabelle 39: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der Übung Vitruvian Operation</i>	<i>63</i>
<i>Tabelle 40: Häufigkeiten der Differenzen der Übung Vitruvian Operation</i>	<i>64</i>
<i>Tabelle 41: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert der Übung Vitruvian Operation</i>	<i>65</i>
<i>Tabelle 42: Darstellung der Metriken mit signifikanter sowie ohne signifikante Verschlechterung</i>	<i>65</i>
<i>Tabelle 43: Blutalkoholkonzentration aller Probanden.....</i>	<i>66</i>
<i>Tabelle 44: Deskriptive Statistik zu Blutalkoholkonzentration</i>	<i>66</i>
<i>Tabelle 45: Auflistung zu den Trinkgewohnheiten der Probanden</i>	<i>67</i>

1 Einleitung

1.1 Allgemeines

In der modernen Chirurgie werden zunehmend Operationen in Kooperation mit RobotikSystemen durchgeführt. Bei den Robotik-Systemen unterscheidet man zwischen aktiven, semi-aktiven und Master-Slave-Systemen. Aktive Systeme können selbstständig unter Aufsicht operieren. Die semi-aktiven Systeme beruhen auf einer Interaktion von Operateur¹ und Roboter. Beim Master-Slave-System wird der Roboter vollständig durch den Operateur gesteuert. Eine Variante dieses Master-Slave-Robotersystems stellt das Da Vinci System dar, welches heutzutage in der roboterunterstützten Chirurgie mit etwa 80 verwendeten Systemen in Deutschland die größte Bedeutung hat. Der Operateur bedient die Steuerkonsole, die sich im Operationssaal, jedoch nicht im sterilen Bereich befindet. Die drei Arbeitsarme und ein Kameraarm werden über sogenannte Trokare in den Bauchraum des Patienten eingebracht; diese Instrumente werden vom Operateur über die Steuerkonsole bedient. Ein Vorteil des Da Vinci Systems liegt in der guten Visualisierung des Operations-Gebietes mit einer gleichzeitigen Vergrößerung in Dreidimensionalität. Außerdem verfügen die Arbeitsarme über sieben Freiheitsgrade und sind somit beweglicher als die menschliche Hand. Durch einen Tremorfilter wird ein potentiell Zittern des Operateurs ausgeglichen (Feussner & Wilhelm, 2016). Zu den Nachteilen gehört, dass diese Robotertechniken sehr teuer sind (Arden, 2018) und dass die Bedienung einer gründlichen Einarbeitung und Übung bedarf (Bendel, 2018).

Diese Übung kann zum Beispiel an einem Roboterchirurgie-Simulator, wie dem MIMIC dV-Trainer[®] erlangt werden. Der MIMIC dV-Trainer[®] ist der Konsole des Da Vinci-Roboters in Hard- und Software nachempfunden und bietet dadurch ein originalgetreues Medium, um kostengünstig ein zugängliches und effektives Training außerhalb des Operationssaals zu ermöglichen. Der MIMIC dV-Trainer[®] ist der einzige eigenständige Robotersimulator, der in vielen publizierten Studien unabhängig voneinander validiert (Geuenich, 2011) wurde (Mimic Technologies, Inc, 2018). Die chirurgische Leistung ist nicht zuletzt an Operationssimulatoren häufig Gegenstand vieler Forschungsarbeiten. Auch die Wirkung gewisser Einflüsse wie zum Beispiel Alkoholkonsum auf die operative Performance im Allgemeinen wurden bereits in der Literatur publiziert (siehe 2.3). Spezielle Untersuchungen am MIMIC dV-Trainer[®] unter Alkoholeinfluss liegen jedoch bisher nicht in bekannter und publizierter Version vor.

¹ Zur Erleichterung der Lesbarkeit wird in dieser Arbeit insgesamt nicht in geschlechtsspezifischen Personenbezeichnungen unterschieden. Die gewählte männliche Form schließt eine weibliche Form gleichberechtigt ein.

Alkohol ist ein breit akzeptiertes Genussmittel, das in unserer Gesellschaft eine nicht unerhebliche Rolle spielt (Graw, 2015). Der jährliche durchschnittliche Konsum von reinem Alkohol lag in Deutschland zum Beispiel im Jahr 2016 bei 17 Litern pro Einwohner. Verglichen mit dem weltweiten Durchschnitt ist dieser Pro-Kopf-Verbrauch etwa dreimal so viel (Wengel, 2020). Jedoch sind die Beeinträchtigungen, die bei einem Alkoholkonsum entstehen können, nicht zu unterschätzen. Dazu zählen unter anderem (je nach Höhe der Blutalkoholkonzentration) die Abnahme der Konzentrationsfähigkeit, Störungen der Bewegungskoordination, Beeinträchtigung der Sehleistung und Schwächung des Urteilsvermögens (BZgA, 2020). Abgesehen von diesen kognitiven Beeinträchtigungen kann ein regelmäßiger Alkoholkonsum durchaus auch zu nachhaltigen körperlichen Funktionseinbußen führen (Wengel, 2020).

1.2 Problemstellung

Eine Studie zeigt, dass viele Ärzte beruflichen Stress mit dem Konsum von Alkohol oder Medikamenten kompensieren (Geuenich, 2011). Zusätzlich zu der Stresskomponente haben Personen, deren tatsächliche Arbeitszeit oberhalb der Standardempfehlungen liegt, wozu auch häufig Ärzte beziehungsweise Chirurgen gehören, ein höheres Risiko, ihren Alkoholkonsum zu erhöhen, als Personen, die diese Arbeitszeitüberschreitung nicht regelmäßig vollziehen (Virtanen, 2015). Demzufolge ist der Anteil an potenziell missbrauchsgefährdeten Personen in der Ärzteschaft höher als in der Normalbevölkerung (Geuenich, 2011). Die Dunkelziffer von Ärzten, die im Dienst alkoholisiert sind, ist unklar. Daher ist die Gefahr von einem möglicherweise alkoholisierten Arzt behandelt beziehungsweise operiert zu werden, im täglichen Leben auch nicht ganz unwahrscheinlich.

1.3 Zielsetzung und Aufbau der Studie

Unsere Hypothese war es, dass Alkoholgenuss die operative Performance eines Chirurgen im Allgemeinen und bei der Roboterchirurgie verschlechtert. Um diese Hypothese bestätigen zu können, führten wir eine prospektive Interventionsstudie am MIMIC dV-Trainer® als Simulationstrainer für spätere reelle Operationen mit dem Da Vinci® Roboter durch. Die wissenschaftliche Fragestellung, die wir im Näheren für den Beleg unserer Hypothese beantworten möchten ist, ob akuter Alkoholkonsum einen Einfluss auf die chirurgische Performance am MIMIC dV-Trainer® hat.

2 Literaturdiskussion

2.1 Definition des Begriffs „Simulation“ und Einordnung des MIMIC dV-Trainers®

Simulationen sind grundlegend definiert als „eine Technik, um reale Erfahrungen durch geführte Erfahrungen zu ersetzen oder zu verstärken, die oft in die Natur eingetaucht sind und wesentliche Aspekte der realen Welt auf vollständig interaktive Weise hervorrufen oder replizieren“ (Agha & Fowler, 2015, S. 350). In diesem Kontext gibt die Arbeit von MacCraith et al. (2019) einen Überblick über Kosten, Leistung und Herausforderung des Roboter-Simulations-training. Die Simulation stellt eine sichere und kostengünstige Trainingsmethode dar und hat sich aus ebendiesen Gründen in den letzten Jahren auf den chirurgischen Bereich ausgeweitet. In der robotergestützten Chirurgie stehen 6 VR-Simulatoren zur Verfügung. Diese sollen zunehmend weltweit in die Ausbildung zur roboterassoziierten Chirurgie, vor allem in der Urologie, aufgenommen werden. Hierdurch könne unter anderem die Lernkurve bei chirurgischen Eingriffen verkürzt werden, ohne die Patientensicherheit zu gefährden. Laut MacCraith et al. (2019) werden chirurgische Fehler in 53 Prozent der Fälle auf Unerfahrenheit zurückgeführt und „Unerfahrenheit führe zu einer größeren Anzahl von Fehlern als Müdigkeit, schlechte Kommunikation und übermäßige Arbeitsbelastung“ (MacCraith, Forde, & Davis, 2019, S. 371). Mittels Simulation werde eine künstliche Umgebung geschaffen, um die Reaktion eines Individuums zu bewerten und zu schulen. Randomisierte Studien ergaben nach dem Training am Simulator signifikant bessere Ergebnisse der Operationsleistung (MacCraith, Forde, & Davis, 2019).

Kenney et al. (2009) publizierten ihre Studienergebnisse in Bezug auf die Oberflächen-, Inhalts- und Konstruktvalidität des MIMIC dV-Trainer®. Bei diesem prospektiven Studienmodell haben Anfänger und erfahrene Chirurgen verschiedene Übungen am MIMIC dV-Trainer® bearbeitet. Durch den Vergleich der Leistungen der beiden Gruppen Anfänger und erfahrene Chirurgen wurde die Konstruktvalidität bestimmt. Anhand einer Umfrage innerhalb der beiden Gruppen wurde die Gültigkeit für die Oberfläche und den Inhalt des MIMIC dV-Trainer® festgelegt. Alle Anfänger stufen den MIMIC dV-Trainer® als realistische Übungsvariante ein und alle erfahrenen Chirurgen beurteilten den Simulator als nützlich für die Schulung von Ärzten und Chirurgen. Die Befragung der erfahrenen Gruppe ergab jedoch, dass einige der Module nicht akzeptabel sind und mittels Befragung von Fachchirurgen konnten die Mängel dieser Module detektiert werden. Der Vergleich zwischen Anfängern und erfahrenen Chirurgen als Testung der Konstruktvalidität ergab, dass die erfahrenen Chirurgen alle Module deutlich schneller absolvierten. Außerdem wurden die Anfänger in gewissen Übungen ebenfalls in der Instrumentenbewegung, der Anzahl der Instrumentenkollision und dem Bewegungsradius

übertroffen. Schlussfolgernd ergab sich eine hohe Validität von Inhalt, Oberfläche und Konstruktion des MIMIC dV-Trainer® und könnte somit als nützlicher Simulator für das Training und die Bewertung in der Roboterchirurgie dienen. Einige Module müssen jedoch inhaltlich angepasst und verbessert werden (Kenney, Wszolek, Gould, Libertino, & Moinzadeh, 2009).

Seymour et al. (2002) bewerteten den Erwerb von chirurgischen Fähigkeiten durch Simulation mithilfe eines Vergleiches von chirurgischen Assistenten mit traditionellem Training und mit einem laparoskopischen Simulatortraining. Die Probanden führten eine Cholezystektomie durch, welche auf Video aufgezeichnet wurde. Die Simulatorgruppe seziierte die Gallenblase zu 29 Prozent schneller, machte fünfmal weniger Fehler und neunmal häufiger Fortschritte (Seymour, et al., 2002).

Agha et al. (2015) publizierten ein Review über die wachsende Relevanz der Simulation, auch als Schlüsselinstrument für die Ausbildung von Chirurgen. Die Simulation wird hier als umfassende und realistische Möglichkeit beschrieben, technische Fähigkeiten zu erlernen und bietet eine standardisierte und sichere Methode zur Ausbildung von Chirurgen. Simulation ermöglicht es, Fehler zu machen, aus solchen Situationen zu lernen und zu reflektieren, ohne die Risiken, die mit einer Operation an echten Patienten verbunden sind. Weiterhin können mithilfe von Simulatoren komplizierte Operationen geplant werden und perioperative Risiken bewertet werden. Für eine handwerkliche Tätigkeit wie die Chirurgie, welche in hohem Maß feinmotorische Fähigkeiten, wie Geschicklichkeit und Koordination des Operateurs bedarf, ist die Simulation eine perfekte Möglichkeit, eben diese Fähigkeiten zu erlernen und zu optimieren (Agha & Fowler, 2015).

Über den MIMIC dV-Trainer® gibt es bereits einige Studien. Unter anderem auch darüber, dass ein „Roboter-Warm up“ durch den MIMIC dV-Trainer® vor einer Operation die Leistung verbessert und Fehler minimiert. Lendvay et al. (2013) haben in dieser randomisierten Studie 51 Assistenzärzte aus den Bereichen Allgemeinchirurgie, Gynäkologie und Urologie einen validierten Lehrplan für Roboterchirurgie an einem VR-Robotersimulator und am DaVinci-Chirurgieroboter absolvieren lassen. Dabei wurden die Chirurgen randomisiert zu zwei Gruppen zugeordnet, welche entweder eine fünfminütige Simulator-Aufwärmphase oder eine zehnminütige Lesephase in einem Freizeitbuch durchführten. Es war hier eine signifikante Leistungsverbesserung und Fehlerreduktionsrate bei den Chirurgen nach der Simulator-Aufwärmphase zu verzeichnen. Außerdem reduzierte das Aufwärmen am VR-Simulator Fehler bei einer komplexeren Aufgabe wie zum Beispiel dem Nähen (Lendvay, et al., 2013).

2.2 Physiologische Effekte von Alkohol und Auswirkungen von Alkoholkonsum

Der Begriff Alkohol steht im Kontext von Genussmitteln übergeordnet für die chemische Verbindung Ethanol. Darüber hinaus gibt es jedoch noch eine Vielzahl weiterer unterschiedlicher Alkohole, die für den menschlichen Genuss nicht geeignet sind. Chemisch gesehen ist es ein Nervengift, welches zu komplexen Schädigungen führen kann. Infolgedessen kann es zu Störungen des sinnesphysiologischen Systems und einer Alteration der Psyche kommen (Mußhoff, Madea, & Heß, 2014). Dabei werden unter anderem Störungen des Sehvermögens, im Sinne von einer herabgesetzten Geschwindigkeit des Tiefenschärfesehvorgangs was zu Fehleinschätzungen der Geschwindigkeiten und Entfernungen bewegter Objekte führt. Eine Einschränkung der koordinierten Blick- und Führungsbewegungen der Augen verstärkt diese Fehleinschätzung. Auch Störungen des Gleichgewichts und ein Verlust der Muskelkoordination mit konsekutivem Nachlassen der Feinmotorik werden als sinnesphysiologische Folgen beschrieben (Mußhoff, Madea, & Heß, 2014). Ethanolbedingte geistig-seelische Mängel können bereits bei niedrigen Promillewerten (0,2-0,5) zu einer veränderten Verhaltensweise führen, die beispielsweise eine erhebliche Gefährdung für die Teilnahme am Straßenverkehr darstellt. Alkohol führt zu einem Abbau der Persönlichkeit, wodurch es zu einer Enthemmung und Kritikminderung kommt, was die Risikobereitschaft bei erhöhtem Selbstwert- und Machtgefühl steigert. Weitere Schädigungen der Psyche bedingen vor allem ein Nachlassen der Aufmerksamkeit, der Auffassungsgabe, der Umsicht und der Besonnenheit. Das Nachlassen der Aufmerksamkeit und der Auffassungsgabe beeinflussen maßgeblich die Qualität der Konzentration und Reaktionsfähigkeit (Mußhoff, Madea, & Heß, 2014). „Sowohl das Vermögen, die Aufmerksamkeit auf ein Ereignis bzw. einen Gegenstand gerichtet zu halten (Tenazität), als auch die Fähigkeit, die Aufmerksamkeit neuen Ereignissen bzw. Gegenständen zuzuwenden (Vigilanz), ist unter Ethanoleinfluss herabgesetzt“ (Mußhoff, Madea, & Heß, 2014, S. 721). Nicht nur die Reaktionszeit in ihrer Dauer ist heraufgesetzt, sondern auch die Genauigkeit wird herabgesetzt. Es kommt zu unpräzisen Handlungen bis hin zu Fehlreaktionen. Eine zusätzlich durch den Einfluss von Ethanol bedingte Ermüdung wirkt sich ebenfalls nachteilig auf die Leistung aus. Die psychischen Alterationen präsentieren sich bei gleicher Ethanolkonzentration im Blut in der Resorptionsphase erheblich stärker ausgeprägt als in der Eliminationsphase (Mußhoff, Madea, & Heß, 2014).

Schon ab 0,3 Promille Ethanolkonzentration im Blut kommt es häufig zu einer verminderten Sehleistung und die Aufmerksamkeit und das Reaktionsvermögen lassen nach. Weiterhin steigt die Risikobereitschaft. Bei 0,5 Promille sind häufig Einschränkungen der visuellen und der auditiven Sinneswahrnehmung zu verzeichnen, was unter anderem die Einschätzung von

Entfernungen beeinträchtigen kann. Ab 0,8 Promille kommt es gehäuft zu massiven Konzentrationseinbußen und zu einem Tunnelblick. Die Reaktionszeit verlängert sich deutlich und Gleichgewichtsstörungen kommen gehäuft vor. Enthemmung und Überschätzung treten bei diesem Alkoholisierungsgrad häufig auf. Ab einer Blutalkoholkonzentration von 1 Promille kommt es zu einem Rauschstadium mit Verwirrheitszuständen und weiteren Einschränkungen der kognitiven und sinnesphysiologischen Funktionen. Ein Alkoholisierungsgrad von 2 bis 3 Promille wird auch Betäubungsstadium genannt, in dem man kaum noch koordinierte Bewegungen ausführen kann und welches zu rezidivierendem Erbrechen führt. Ab 3 Promille spricht man von einem Lähmungsstadium, in dem es zu Bewusst- und Reflexlosigkeit, Gedächtnisverlust und einer flachen Atmung kommen kann (BZgA, 2020).

Hauptsächlich resorbiert wird konsumierter Alkohol über den Dünndarm und wird anschließend durch den Pfortaderkreislauf im menschlichen Körper verteilt. Hier dienen die wässrigen Kompartimente als Verteilungsraum, welcher bei Männern circa 70 Prozent der Körpermasse und bei Frauen etwa 60 Prozent der Körpermasse ausmachen. Erst kurz nach dem Trinkende wird bei üblichen Trinkverläufen der höchste Blutalkoholwert erreicht. Je nach Alkoholgehalt, Trinkvolumen und -zeit sowie der Magenfüllung kann eine Menge von bis etwa 30 Prozent unter anderem durch First-pass-Effekte und regionale Stoffwechselforgänge nicht im Blutwert nachgewiesen werden. Dies wird als Resorptions- oder Alkoholdefizit bezeichnet und muss bei der Berechnung der Alkoholkonzentration im Körper mitbedacht werden. Bereits zum Aufnahmezeitpunkt des Alkohols beginnt auch dessen Elimination durch die Alkoholdehydrogenase der Leber. Intraindividuell liegt die Eliminationsrate durch die Dehydrogenase etwa konstant bei 0,1-0,25 Prozent in der Stunde (Graw, 2015). Bei chronischem Alkoholkonsum sind höhere Abbauraten bis circa 0,36 Prozent zu verzeichnen, da hier von einer zusätzlichen Enzyminduktion des mikrosomalen Ethanol-oxidierenden Systems auszugehen ist, welches bei Gelegenheitskonsumenten nur wenig zur Elimination beiträgt. Die Berechnung („Widmark-Formel“) der Blutalkoholkonzentration setzt sich somit als Quotient aus der Alkoholmenge in Gramm zum Produkt aus Körpergewicht und dem oben erklärten Verteilungsfaktor r (je nach prozentual wässrigen Körperanteilen bei Frauen 0,6 und bei Männern 0,7):

$$C_0 = \frac{\text{Alkoholmenge (in g)}}{\text{Körpergewicht (in kg) * Reduktionsfaktor } r}$$

Abbildung 1. Widmark-Formel zur Berechnung der Alkoholkonzentration. Übernommen aus Graw (2015).

Daraus ergibt sich eine theoretische Blutalkoholkonzentration (C0-Wert), wobei noch die während des Trinkvorganges eliminierte Alkoholmenge abgezogen werden muss (Graw, 2015). Weiss et al. (2013) veröffentlichten eine Zusammenfassung einiger Studien über Alkohol induzierte kognitive Dysfunktionen. Akuter Alkoholkonsum führt demgemäß zu nachweisbaren Veränderungen der Hirnleistung und neben reversiblen Störungen der motorischen Regulation und des Gleichgewichts, auch zu Beeinträchtigungen von Gedächtnis und der Aufmerksamkeit. Durch chronischen Alkoholkonsum kann es zu erheblichen neuropsychologischen Defiziten kommen. In einer zerebralen Bildgebung von Alkoholabhängigen sind bei 50 Prozent eine Atrophie mit dem klinischen Korrelat von Einbußen der Kognition zu verzeichnen. Durch eine Interaktion des Ethanols mit der neuronalen Lipidmembran kommt es zu einer Störung der Membranfunktion. Weiterhin sind bei langanhaltendem Konsum nachweislich neurochemische Prozesse gestört und es kommt infolgedessen zu einer Down-Regulation des GABA-ergen Systems und einer Dopaminunterfunktion. Diese Störungen können im Zusammenhang mit parkinsonähnlichen Symptomen stehen. Außerdem konnte durch neurophysiologische Studien nachgewiesen werden, dass vor allem der präfrontale Kortex und seine neuronalen Verbindungen anfällig für toxische Effekte des Alkohols sind. Dabei kann es vor allem zu Einschränkungen bei der Problemlösefähigkeit, dem Entscheidungstreffen, Planen und Organisieren kommen. Weiterhin zeigen sich der inferiore mediale präfrontale Cortex und der orbitofrontale Cortex besonders sensibel gegenüber alkoholtoxischen Effekten. Bei Läsionen in diesen Gebieten zeigte sich, dass die Betroffenen besonders risikohafte Entscheidungen ohne Berücksichtigung langfristiger Konsequenzen treffen (Weiss, Singewald, Ruepp, & Marksteiner, 2013).

2.3 Alkoholeinfluss bei Ärzten und Chirurgen

Es gibt keine definierten Kriterien für akzeptable Blutalkoholspiegel während der Arbeit als Chirurg und keine gesetzliche Regulierung des Trinkverhaltens von Ärzten (Kirby, Kapoor, Das-Purkayastha, & Harries, 2012). Dabei ist die Einstellung zum Alkoholkonsum während eines Bereitschaftsdienstes teilweise unterschiedlich - von der absoluten Abstinenz bis hin zu Werten um die gesetzliche Grenze für Autofahrer. In der Luftfahrt gibt es strenge Richtlinien; Piloten wird geraten vor einem Flug mindestens 24 Stunden auf Alkohol zu verzichten (Kirby, Kapoor, Das-Purkayastha, & Harries, 2012). Die beruflichen Aktivitäten und Anforderungen haben viele Gemeinsamkeiten zu denen eines Chirurgen. Dazu gehören unter anderem komplexe Verfahren bewältigen und manuelle Geschicklichkeit in unter Druck stehender Atmosphäre beweisen, in der Fehler zum Verlust von Menschenleben führen können. Die folgenden

Veröffentlichungen haben sich mit den Auswirkungen von Alkohol bei Ärzten und Chirurgen befasst und bieten einige Erkenntnisse.

In einer Studie von Geuenich et al. (2011) publiziert im Deutschen Ärzteblatt 2011 wurde über eine Online-Befragung berichtet, die aufzeigte, dass viele Ärzte beruflichen Stress mit dem Konsum von Alkohol oder Medikamenten kompensieren. Dabei wurden 1287 Ärzte und Ärztinnen anonymisiert über ihre Belastungen und deren Kompensation befragt. Insgesamt gaben hier rund 30 Prozent der Befragten an, berufsbedingten Stress teilweise, überwiegend, häufig oder gar immer durch Alkohol oder anderen Medikamenten abzubauen. Bei Ärzten wurde im Vergleich zur Normalpopulation ein signifikant erhöhter Dauerstress durch lange Arbeitstage festgestellt. Durch einen positiven Zusammenhang von Dauerstress und Substanzkonsum besitzen Ärzte somit eine potenziell höhere Gefährdung von stressbedingtem Alkoholkonsum (Geuenich, 2011).

Kocher et al. (2006) untersuchten mit ihrer Studie die Auswirkungen von Alkohol auf die chirurgische Geschicklichkeit unmittelbar nach und am Morgen nach dem Alkoholkonsum. Im Näheren wurden fünf Chirurgen auf einem laparoskopischen chirurgischen Simulator (MIST-VR) geschult, bis sie damit vertraut waren. Anschließend wurden drei Situationen untersucht: Eine Ausgangssituation, in der die Probanden ausgeruht und ausgeschlafen ohne vorherigen Alkoholkonsum in den Versuch starteten; eine zweite Ausgangssituation, wo die Probanden nur wenig Schlaf hatten und eine dritte Ausgangssituation, in der die Probanden Alkohol am Abend konsumierten und ebenfalls nur wenig Schlaf hatten. In jeder dieser Situationen wurden die Testpersonen während des Tages, um Mitternacht und am folgenden Morgen untersucht. Bei der simulierten Aufgabe mussten zwei Objekte, die sich auf einem anderen Objekt befanden diathermisiert werden, während das Hauptobjekt auf engstem Raum gehalten wurde. Ziel war es, das Hauptobjekt dabei nicht durch Diathermie zu beschädigen. Bewertet wurde die Bearbeitungszeit, die Anzahl von Fehlern, die Diathermiezeit und die Diathermiezeit am Hauptobjekt (Schadenszeit). Weiterhin wurden der Atemalkoholwert und die Schlafdauer notiert. Die Ergebnisse zeigen, dass der Alkoholeinfluss den größten nachteiligen Einfluss auf die chirurgische Leistung zu haben scheint. Die Kombination aus Alkohol und Schlafentzug wirkte sich schädlich auf einige psychomotorische Parameter wie der Leistungsgenauigkeit, der Reaktionszeit und der Wachsamkeit aus. Weiterhin waren am Morgen nach dem Konsum und nach einer kurzen Schlafdauer eine längere Gesamtdauer und längere Diathermiezeiten zu verzeichnen. Jedoch hat sich die Anzahl der Fehler und die Verletzungszeit (Diathermiezeit am Hauptobjekt) am Morgen danach im Gegensatz zum alkoholisierten Zustand verringert, was

auf eine höhere Wachsamkeit zurückzuführen war (Kocher, Warwick, Al-Ghnam, & Patel, 2006).

Kirby et al. (2012) publizierten ihre Ergebnisse zu einer Studie über die Wirkung von unterschiedlichen Alkoholmengen auf die chirurgischen Fähigkeiten. Diese behandelt im Speziellen die Testung der Geschicklichkeit, der kognitiven Fähigkeiten und soziale Interaktionen, die zur Gewährleistung der Patientensicherheit während einer HNO-Operation beitragen. Die Alkoholmenge wurde hier in konsumierten Gläsern von Wein statt der Blutalkoholkonzentration gemessen. Diese Maßeinheit wurde gewählt, um die Studie und deren Ergebnisse für Chirurgen zugänglicher und auf die eigene Person besser projizierbar zu machen. Vier gesunde Teilnehmer mit unterschiedlicher chirurgischer Erfahrung wurden getestet. Mindestens 24 Stunden vor dem Experiment sollte auf Alkohol verzichtet werden und mindestens vier Stunden vor Beginn der Testung sollte nichts gegessen werden. Die Probanden sollten schließlich anhand eines Kehlkopf-Trainingsmodells mit pathologischen Kehlkopfbefunden (validiertes Trainingsinstrument) eine Mikrolaryngoskopie durchführen. Dabei wurde diese Fähigkeit insgesamt zu vier Zeitpunkten überprüft: Nüchtern und 45 Minuten nach einem, drei und sechs Gläsern Wein. Während der Testung war ein Gutachter anwesend und der operative Vorgang wurde auf Video aufgezeichnet, was im Anschluss von zwei HNO-Chirurgen beurteilt wurde. Die Auswertung zeigte, dass ein Glas Wein 45 Minuten vor der simulierten Operation keine nachteiligen Auswirkungen auf die chirurgische Fähigkeit hatten. Nach drei oder mehr Gläsern Wein wurden jedoch alle Teilnehmer als unsicher für die Operation eingestuft. Auch die Kommunikation mit der Assistenz und dem anwesenden Personal wurde ab drei oder mehr Gläsern negativ beeinflusst. Frühere Untersuchungen belegten schon, dass Alkohol die grundlegenden Mechanismen der Verhaltenskontrolle selektiv beeinträchtigen und zu Aggressionen und sozial abweichenden Verhaltensweisen führen kann. Da Kommunikation und Teamarbeit eine wichtige Rolle bei chirurgischen Eingriffen, vor allem in Bezug auf die Patientensicherheit, spielen, stellt diese Wirkung des Alkoholkonsums ein erhebliches Sicherheitsrisiko dar (Kirby, Kapoor, Das-Purkayastha, & Harries, 2012).

Bei der Studie, die Dorafshar et al. (2002) publizierten wurden 28 Probanden zufällig einer Alkohol- oder einer Placebogruppe zugeordnet. Die Testpersonen wurden hierbei eine Stunde nach dem Konsum von Alkohol, sowie zehn Stunden nach dem Konsum und einer Nacht Schlaf an einem VR-Simulator getestet. Es erfolgte mit einer Atemalkoholüberprüfung nur eine subjektive Schätzung der Blutalkoholkonzentration von 85 mg/100 ml und eine subjektive Bewertung mittels einer visuellen Analogskala. Es wurde die Anzahl der Fehler, der Zeitaufwand, die Bewegungsökonomie und der übermäßige Gebrauch von Diathermie überprüft. In allen

vier Kategorien war eine signifikant alkoholbedingte Beeinträchtigung der simulierten laparoskopischen Operation 60-90 Minuten nach Trinkbeginn zu verzeichnen. Die Testung zehn Stunden nach Trinkbeginn ergab jedoch kein signifikantes Ergebnis. Eine Verblindung der Studie wurde aufgrund der allgemein in der Bevölkerung bekannten physiologischen und verhaltensbezogenen Effekte von Alkohol als ungeeignet betrachtet (Dorafshar, O'Boyle, & McCloy, 2002).

Die bislang größte prospektive Studie mit erfahrenen laparoskopischen Chirurgen von van Dyken et al. (2013) überprüfte die Geschicklichkeit von Chirurgen nach einem Abend mit mäßigem Alkoholkonsum. Es wurde die Hypothese aufgestellt, dass sich die Feinmotorik am Morgen nach einer Nacht mit mäßigem Alkohol ändern würde, was sich in einer Abnahme der Geschwindigkeit und Genauigkeit niederschlägt. Dazu wurden 27 Chirurgen als Versuchspersonen rekrutiert. Es wurden die Instrumentenhandhabung, das Schnittmuster und das intrakorporale Nähen als Grundlagen der laparoskopischen Chirurgie überprüft. Testmedium stellte ein Forschungscurriculum, das vom American College of Surgeons und der Society of American Gastrointestinal and Endoscopic Surgeons gesponsert wird, eines laparoskopischen Simulators dar. Die Versuchspersonen waren an diesem Curriculum alle zertifiziert und hatten Erfahrungen mit laparoskopischen Operationen. Dabei wurden die drei oben genannten Fähigkeiten abgefragt und zeitlich festgelegt. Weiterhin wurde in Bezug auf die Genauigkeit bewertet, ob außerhalb eines Grenzbereiches geschnitten wurde und ob beim intrakorporalen Nähen die Nadel neben dem schwarzen Punkt eingeführt wurde. In beiden Fällen wurde dies als ungenau bewertet. Weiterhin wurde der beim Nähen gebildete Knoten nach Abschluss der Aufgabe auf seine Haltbarkeit getestet. Bei Insuffizienz wurde dies als ungenau bewertet. Es erfolgte zum einen eine Basismessung der Fähigkeiten ohne Alkoholeinfluss. Zum anderen wurde das Curriculum von den Probanden am Morgen nach einem Abend mit Alkoholkonsum absolviert. Dabei wurde nicht festgelegt, wieviel Alkohol die Probanden zu trinken hatten, dies blieb den Versuchspersonen freigestellt. Die geschätzte Blutalkoholkonzentration wurde schließlich 20 Minuten nach Beendigung des Trinkens und am Morgen nach acht bis zehn Stunden Schlaf kurz vor der Testung am Simulator mittels eines Atemalkoholtesters bestimmt. Die geschätzte Blutalkoholkonzentration lag in der Nacht des Konsums im Mittel bei 1 Promille und am nächsten Morgen bei 0 Promille. Es wurden in der Zeiterfüllung als auch in der Schnittführung keine signifikanten Unterschiede festgestellt. Beim intrakorporalen Nähen jedoch ergaben sich signifikante Unterschiede; am Morgen nach dem Alkoholkonsum waren die Probanden weniger genau. Korrelationen nach Spearman ergaben sich zwischen der BAK und den Zeiten zum Abschließen der Übungen zur Schnittführung, zur Instrumentenhandhabung

und zum intrakorporalen Nähen nicht. Mittels eines Fragebogens wurden zusätzlich Schlafdauer und Qualität bewertet. Die Auswertung der Bögen zeigte, dass der Schlaf nach dem Alkoholkonsum vergleichsweise von kürzerer Dauer und schlechterer Qualität zeugte, was ergänzend zu dem „Katereffekt“ des Alkohols zu weniger Erholung führen kann. Hier muss beachtet werden, dass die schlechte Schlafqualität und kurze Schlafdauer einen Confounder der Studie darstellt (Dorafshar, O'Boyle, & McCloy, 2002).

Die Studie, die von Gallagher et al. (2011) veröffentlicht wurde, befasste sich mit der Fragestellung, wie sich übermäßiger Alkoholkonsum auf die laparoskopisch chirurgische Leistung am nächsten Tag auswirkt. Dazu wurden insgesamt zwei Studien mithilfe des „Minimal Invasive Surgical Trainers“, einem Laparoskopie-Simulator, durchgeführt. Der MIST-VR, der bereits von Wilson et al. (1997) beschrieben wurde bietet sechs Aufgaben mit schrittweise steigender Komplexität und steigenden Schwierigkeitsstufen (Wilson, Middlebrook, Sutton, & Stone, 1997). Dabei basiert jede Aufgabe auf einer wichtigen chirurgischen Technik, welche bei einer laparoskopischen Gallenblasenentfernung verwendet werden. Bewertet wurden diese Aufgaben wurden mithilfe von Leistungskennzahlen. Eine randomisiert kontrollierte Studie mit 16 Studenten der Naturwissenschaften, die zuvor noch nie einen laparoskopischen Simulator bedient haben, bildete Studie 1. Hier wurden unter zufälliger Zuteilung zwei Gruppen gebildet; eine Interventionsgruppe, die Alkohol konsumierte und eine Kontrollgruppe. Bei Studie 2 wurden acht Chirurgen getestet, die in der Laparoskopie und an dem Virtual Reality Simulator Erfahrung haben. Im Vorfeld absolvierten alle Probanden eine an deren Ausbildungsstand angepasste Übungseinheit, um eine mögliche Verzerrung durch Lernen zu torpedieren. Weiterhin führte jeder Proband einen Durchgang der Übungen ohne Alkoholeinfluss durch. Am Abend vor dem Versuchstag nahmen die Interventionsgruppe aus Studie 1 und alle Versuchsteilnehmer der Studie 2 Alkohol zu sich, bis sich die Probanden subjektiv betrunken fühlten. Es wurde keine zu konsumierende Alkoholmenge vorgegeben, um somit die Realität am ehesten widerspiegeln zu können. Alle Teilnehmer schliefen vor 24 Uhr und ohne Unterbrechung. Am nächsten Morgen wurden die Übungen um 9 Uhr, um 13 Uhr und um 16 Uhr getestet. Die zu bewertenden Parameter bestanden aus der Zeit zum Beenden der Aufgaben in Sekunden, der mittleren Fehlerzahl (ein Fehler ist hierbei als eine vom Ziel abweichende Bewegung definiert) und der Effizienz der Diathermieanwendung (die mittlere Verbrennungszeit in Relation zur optimalen Verbrennungszeit). Die Ergebnisse zeigten, dass die Ausgangsleistungen der beiden Gruppen von Studie 1 keine statistisch signifikanten Unterschiede aufzeigten. Bei der Auswertung des Testtages zeigten sich große Unterschiede zwischen den beiden Gruppen; die Interventionsgruppe schnitt bei allen drei Leistungsparametern schlechter ab. Jedoch erwies sich hier nur der Unterschied zum Messzeitpunkt um 9 Uhr als statistisch

signifikant, was in dem Fall auf eine große Leistungsvariabilität innerhalb der Interventionsgruppe zurückgeführt wurde. Anhand der Messwerte wurde eine zweifaktorielle Varianzanalyse durchgeführt. Einerseits wurden hier Unterschiede der Leistungsvariablen zwischen den beiden Gruppen und andererseits Unterschiede der drei Messzeitpunkte zur Ausgangsleistung untersucht. Die Analyse zeigte, dass sich beim Vergleich der beiden Gruppen keine Signifikanz bei der Metrik Zeit ergab. Jedoch war bei den Variablen Bewegungsverhalten und Fehlerzahl ein signifikanter Unterschied der Gruppen zu verzeichnen. Der Vergleich der Messzeitpunkte zeigte hochsignifikante Unterschiede bei der Bearbeitungszeit, der Diathermienutzung und der Fehlerzahl. Bei der zweiten Studie, wo es nur eine Kohorte gab, wurde die Ausgangsleistung mit den Leistungen zu den verschiedenen Messzeitpunkten verglichen. Die Experten­gruppe bearbeitete die Aufgabe um 9 Uhr morgens nach dem Abend Alkoholeinnahme schneller als zum Ausgangszeitpunkt. Um 13 Uhr war die Leistung signifikant schlechter als um 9 Uhr und 16 Uhr ergab sich bei der benötigten Zeit kein Unterschied zum Ausgangswert. In Bezug auf die Diathermienutzung ergab sich zu den verschiedenen Testzeiten eine signifikante Verschlechterung. Bei den Fehlerwerten war ebenfalls eine Verschlechterung nachweisbar. Jedoch zeigte sich nur der Unterschied um 13 Uhr statistisch signifikant zum Ausgangswert (Gallagher, et al., 2011).

Zusammenfassend zeigt sich anhand oben aufgeführter Studien, dass Alkohol einen Einfluss auf manuelle sowie kognitive Fähigkeiten hat. Daher macht eine Evaluation des Einflusses von Alkohol im Sinne einer objektivierbaren Messung der Leistungen durch den MIMIC dV-Trainer® Sinn.

3 Material und Methoden

3.1 Aufbau MIMIC dV-Trainer®

Der MIMIC dV-Trainer® ist ein Trainingssimulator, der dem Da Vinci-Robotersystem nachempfunden wurde und auf das operative Arbeiten an diversen Robotersystemen vorbereiten kann (Mimic Technologies, 2018). Abbildung 2 zeigt den dV-Trainer® mit seinen einzelnen Bestandteilen, welche anschließend stichpunktartig beschrieben werden.



Abbildung 2. MIMIC dV-Trainer®. Entnommen aus Mimic Technologies (2018).

Die Abbildung 2 zeigt unter anderem die „Master Grips“ und „Tension Cables“. Über die beiden Mastergriffe kann der Benutzer mit dem MIMIC dV-Trainer® interagieren. Die „Tension cables“ sind nylonüberzogene Stahlkabel, die mit den „Master Grips“ verbunden sind und unter Spannung stehen. Somit können diese während der Simulation Bewegungen erkennen und spannungsbasierte Informationen rückmelden. Untenstehende Abbildung 3 zeigt die „Master Grips“ und die damit verbundenen „Tension cables“ (Mimic Technologies, 2018). Das Zusammenspiel der „Master Grips“ und der „Tension Cables“ ermöglicht es, die virtuellen Greifer sowie die Kameraoptik in den einzelnen Übungen zu bedienen und zu bewegen. Die „Docking Station“ mit LED ist auf dem mittleren unteren Bildrand der Abbildung 2 zu sehen. Durch diese Andockstelle werden Bewegungserkennungen kalibriert. Diese Kalibrierung muss vor jeder

Benutzung durch den Anwender durchgeführt werden. Das dV-Trainer® Adjustable Stereoscope“, zeigt hochwertige Stereobilder an und erstellt während der Simulation eine Da Vinci®-ähnliche Trainingsumgebung (Mimic Technologies, 2018).



Abbildung 3. „Foodswitch Panel“ des MIMIC dV-Trainer. Entnommen aus Mimic Technologies (2018).

Auf Abbildung 3 ist das „Foodswitch Panel“ mit seinen einzelnen Komponenten zu sehen. Durch das Drücken des Kupplungspedals („Clutch Pedal“) werden die Instrumente in der Virtual-Reality-Umgebung unabhängig von den tatsächlichen Händen „eingefroren“. Auf diese Weise können die „Master Grips“ im physischen Raum neu positioniert werden, ohne die Instrumente in der virtuellen Umgebung zu bewegen. Das Kupplungspedal wird verwendet, um eine angenehme Handhaltung einzustellen und um damit zu verhindern, dass die Master Grips mit dem Boden des MIMIC dV-Trainer® kollidieren. Weiterhin kann der Arbeitsbereich damit effizient genutzt werden. Um die Instrumente „einzufrieren“, lässt man das Kupplungspedal gedrückt, bewegt die Hände in die gewünschte Position und lässt anschließend das Kupplungspedal wieder los. Das Kamerapedal wird verwendet, um die Ansicht relativ zur virtuellen Umgebung zu verschieben. Dafür lässt man das Kamerapedal gedrückt, bewegt die „Master Grips“ in die gewünschte Position (nach hinten, nach vorne, nach oben, nach unten, nach rechts oder nach links) und lässt es wieder los. Wenn das Kamerapedal gedrückt wird, befindet sich der dV-Trainer® im Kameramodus und ein Kamerasymbol, welches die Ausrichtung festhält, wird im Stereoskop angezeigt (Mimic Technologies, 2018).

Weiterhin wird aus Abbildung 3 ersichtlich, dass dem Nutzer Energiepedale zur Verfügung stehen. Durch Drücken des primären (blauen) Energiepedals (s. Abbildung 3) wird bipolare Energie zum Durchtrennen von Gewebe und Blutstillen aktiviert. Um die Energie in einer Übung zu nutzen, drückt und hält man das entsprechende Pedal. Um sie zu beenden, lässt

man das Pedal wieder los. Die monopolaren Instrumente in den Simulationsübungen können entweder eine Koagulationsfunktion („mono coag energy“) oder eine Schneidefunktion („mono cut function“) ausführen. Die Koagulationsfunktion wird vom primären (blauen) Pedal und die Schneidefunktion vom sekundären (gelben) Pedal aktiviert. Um monopolare Energie in einer Übung zu nutzen, drückt und hält man das entsprechende Mono-Energie-Pedal. Zum Beenden der Energie lässt man das Pedal wieder los (Mimic Technologies, 2018).

3.2 Übungen und Auswertungsparameter

3.2.1 Allgemeines

Die Software des MIMIC dV-Trainer® bietet verschiedene Übungen, wobei der Benutzer angewiesen wird, eine Aufgabe zu erledigen. Um diese Aufgaben erledigen zu können, werden die „Master Grips“ auf virtuelle Greifer projiziert. Somit kann man die Grips nutzen, um die Greifwerkzeuge innerhalb der Übung zu steuern (Mimic Technologies, 2018). Am Ende jeder Übung wird dem Benutzer eine Bewertung seiner Leistung mittels einer Vielzahl von Leistungsmetriken und einem Gesamtscore präsentiert. Im Anhang 8.1 werden die einzelnen Leistungsmetriken erläutert.

3.2.2 Erläuterung des Bewertungssystems

Das auf Kompetenz basierende Bewertungssystem verwendet „Proficiency Baselines“ für jede Metrik, welche aus Mittelwerten und Standardabweichungen berechnet werden, die aus Daten von über 100 erfahrenen Chirurgen abgeleitet wurden, sowie einem Kompetenzmultiplikator (Mimic Technologies, 2018).

$$\text{Proficiency Baseline} = \mu + 1 * \sigma$$

- μ = Mittlerer Rohwert des erfahrenen Chirurgen für diese bestimmte Metrik und Übung
- 1 = Leistungsmultiplikator
- σ = Rohwert der Standardabweichung des erfahrenen Chirurgen für diese bestimmte Metrik und Übung

Neben der Proficiency Baseline wurde mithilfe des Durchschnittswertes innerhalb der Stichprobe mathematisch das beste zu erreichende Ziel („Best“) und das schlechteste zu erreichende Ziel („Worst“) für diese Metrik ermittelt. Durch diese drei Werte entsteht eine Matrix, in die man den Rohwert („Value“) des Anwenders einordnen und mit einer absoluten Punktzahl

(Score) bewerten kann (Mimic Technologies, 2018). Abbildung 5 zeigt ein Beispiel dieses Bewertungsschlüssels.

Value: 7.95 cm Score: 22.87

Baseline Values

	Raw Value	Score Value
Best	2.74	40
Proficiency	8.82	20
Worst	14.9	0

Abbildung 4. Beispiel des Bewertungsschlüssels einer Metrik.

Somit wurden für jede Übung die Leistungsmetriken und deren Matrices durch die „Proficiency baseline“, dem besten und dem schlechtesten Wert und deren Scores definiert (Mimic Technologies, 2018). Die verschiedenen Bewertungsschlüssel der einzelnen Leistungsmetriken sind dem Anhang 8.2, A-D, zu entnehmen.

3.2.3 Übungen

3.2.3.1 Ring Tower

Bei der Übung Ring Tower sollen abwechselnd mit dem jeweils rechten und linken Greifer die blinkenden Ringe über einen Draht geführt werden. Dabei soll der Draht möglichst nicht berührt werden. Abbildung 6 veranschaulicht diesen Vorgang. Anschließend wird der Ring abwechselnd auf den an der rechten und linken Seite befestigten „Drahtturm“ aufgesetzt. Die Übung ist beendet, wenn der Ring richtig auf dem letzten Turmsockel sitzt, was in Abbildung 7 zu sehen ist. Die Übung Ring Tower wurde ausgewählt, da hierbei die Hand-Augen-Koordination sowie ein präzises und kontrolliertes Arbeiten gefordert sind. Dazu fließt in die Bewertungskriterien unter anderem mit ein (siehe Abbildung 8), wie ökonomisch und effizient die Probanden die Übung absolvieren konnten („Economy of motion“) und wieviel Präzisionsfehler („Wire Contact Duration“, „Wire Collisions“, „Tower knockoff“) dabei erfolgten (Mimic Technologies, 2016).



Abbildung 5. Auszug aus Ring Tower Bild 1. Entnommen aus Mimic Technologies (2016).

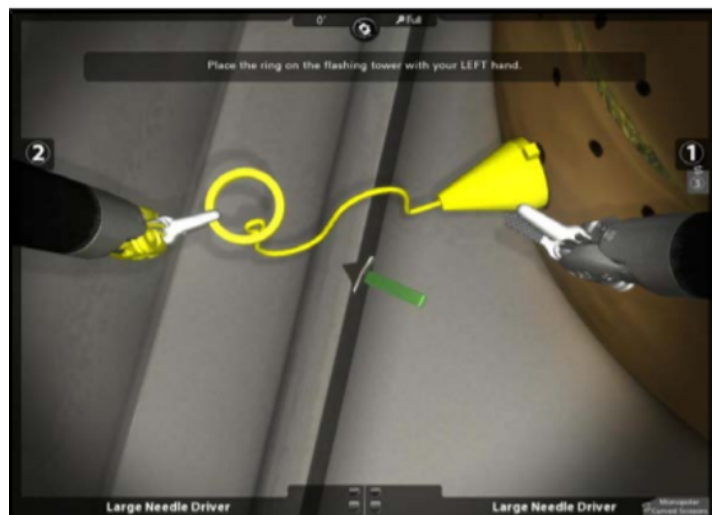


Abbildung 6. Auszug aus Ring Tower Bild 2. Entnommen aus Mimic Technologies (2016).

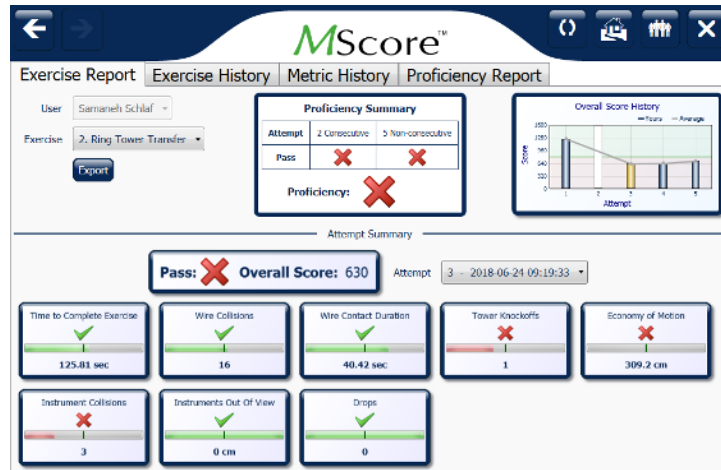


Abbildung 7. Metriken der Auswertung zu Ring Tower. Eigene Darstellung.

3.2.3.2 Suture Sponge

Aufgabe bei dieser Übung ist es, eine Nadel durch vorgegebene Zielscheiben in einem Schwamm ein- und auszuführen, wobei die Positionen der Zielstrukturen variieren. Die Nadel muss in die gelb markierten Zielringe innerhalb des Schwamms eingeführt, mithilfe von Drehbewegungen durch den Schwamm gefädelt und durch ein weiteres gelbes Ziel wieder ausgeführt werden. Die Zielstrukturen werden grün, wenn man die Nadel erfolgreich ein- und ausgeführt hat, was exemplarisch in Abbildung 9 dargestellt wird. Wenn das Ziel rot wird, hat man es verfehlt und muss die Nadel nochmals neu ein- oder ausführen. Wenn die Nadel beide Ziele erfolgreich durchstoßen hat und aus dem Schwamm gezogen wurde, verschwinden die Ziele und ein neues Zielpaar wird angezeigt. Ob für die Aktion der rechte oder der linke Greifer genutzt werden soll, wird signalisiert, indem der betreffende Greifer gelb aufleuchtet. Die Übung ist abgeschlossen, wenn man die Nadel durch insgesamt acht Zielsätze gefädelt hat. Anhand der Übung Suture Sponge kann die Fingerfertigkeit und Geschicklichkeit der Versuchsperson überprüft werden. Auch eine sichere Instrumentenführung ist bei dieser Übung gefragt und wird unter anderem anhand der Metriken „Economy of motion“, „Instrument collisions“ und „Instruments out of view“ bewertet. Da insgesamt 8 Zielpaare bearbeitet werden müssen, wird ein hohes Maß an Konzentrationsfähigkeit gefordert welche durch Metriken wie „Missed targets“ und „Drops“ objektiviert werden kann. Die Gesamtheit aller überprüften Metriken dieser Übung sind in Abbildung 10 dargestellt (Mimic Technologies, 2016).

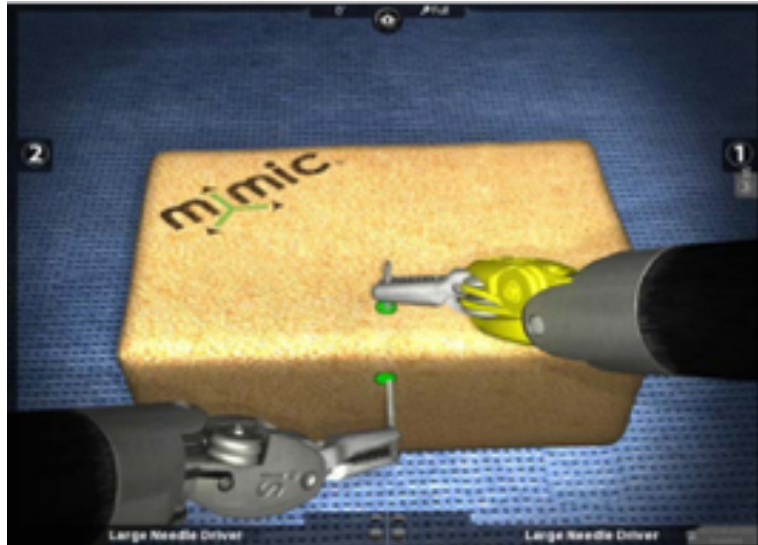


Abbildung 8. Auszug aus Suture Sponge. Entnommen aus Mimic Technologies (2016).

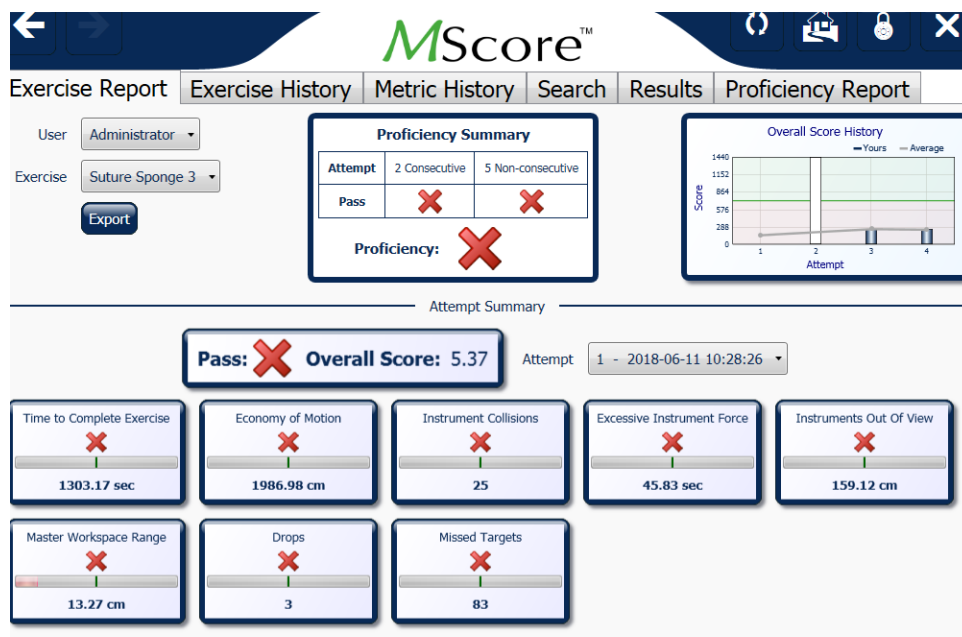


Abbildung 9. Metriken der Auswertung zu Suture Sponge.

3.2.3.3 Stacking Challenge

Durch das Stapeln von Kisten und Dominosteinen soll innerhalb von zwei Minuten ein höchstmöglicher Turm gebaut werden. Die Punktzahl wird durch die erreichte Turmhöhe nach Ablauf der Zeit bedingt. Je höher der Turm ist, desto mehr Punkte können erreicht werden. Daher führt das Stapeln von Dominosteinen „Ende an Ende“ zu einer höheren Punktzahl als das Stapeln von Dominosteinen „flach“ übereinander. In Abbildung 11 ist ein Beispiel eines solchen

Turmes dargestellt. Berührt man einen Teil des Turmes, wenn die Zeit abgelaufen ist, werden keine Punkte vergeben. Auf diese Weise darf der Benutzer keinen instabilen Turm abstützen. Bei der Übung Stacking Challenge geht es vor allem darum, mehrere präzise Objektmanipulationen unter einer Zeitbeschränkung durchzuführen. Weiterhin sind Fingerfertigkeit und eine gute Kameraführung gefragt. Anders als die obenstehend erläuterten Übungen wird diese Aufgabe unter einer zeitlichen Beschränkung absolviert, was im Vergleich zur chirurgischen Tätigkeit das Arbeiten unter einer Drucksituation simuliert. Bei dieser Übung wird nur die Gesamtpunktzahl ausgewertet. Siehe dazu auch Abbildung 12 (Mimic Technologies, 2016).



Abbildung 10. Auszug aus Stacking Challenge. Entnommen aus Mimic Technologies (2016).

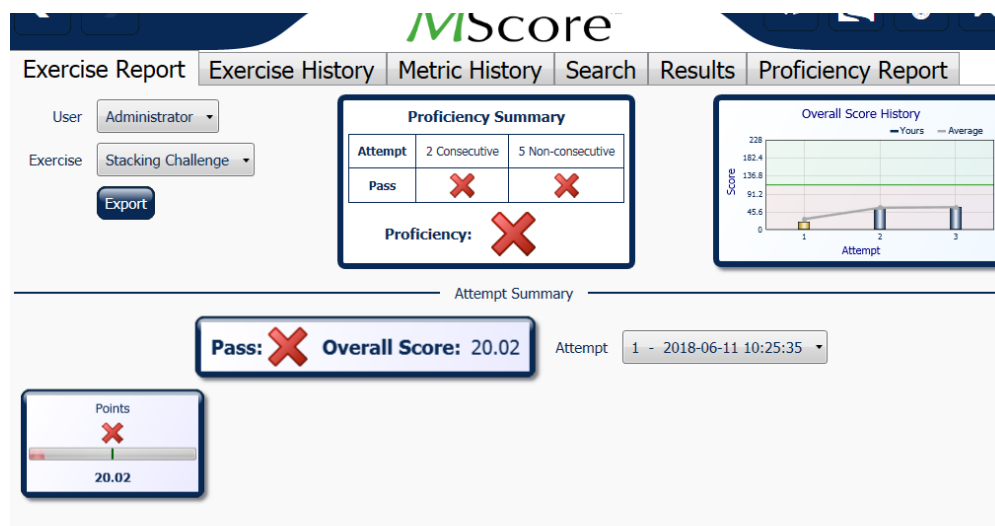


Abbildung 11. Auswertung Stacking Challenge.

3.2.3.4 Vitruvian Operation

Ziel dieser Aufgabe ist es, die Gegenstände, welche von den Dominosteinen umgeben sind, innerhalb von zwei Minuten einzusammeln und in eine Schale zu legen, ohne die Dominosteine dabei umzustößen. Für jedes Objekt, das erfolgreich in die Schüssel gelegt wurde, erhält man Punkte. Der Punktwert jedes Objekts basiert auf der Anzahl der das Objekt umgebenden Dominosteine. Für jeden Domino, den man umwirft, verliert man einen Punkt. Die Übung ist beendet, wenn alle Objekte in der Schüssel platziert wurden oder wenn die Zeit abgelaufen ist. Abbildung 13 zeigt einen beispielhaften Auszug des Spiels. Bei Vitruvian wird ebenfalls wie bei den anderen Übungen eine präzise Fingerfertigkeit und Geschicklichkeit getestet. Jedoch wird dies hierbei unter einer zeitlichen Einschränkung und einem relativ großen Arbeitsbereich untersucht. Dadurch wird ähnlich zu der Übung Stacking Challenge eine Druck-situation simuliert, unter der die Probanden Leistung erbringen sollen. Am Ende der Übung wird die Gesamtpunktzahl als Bewertungsparameter verwendet; siehe dazu Abbildung 14 (Mimic Technologies, 2016).

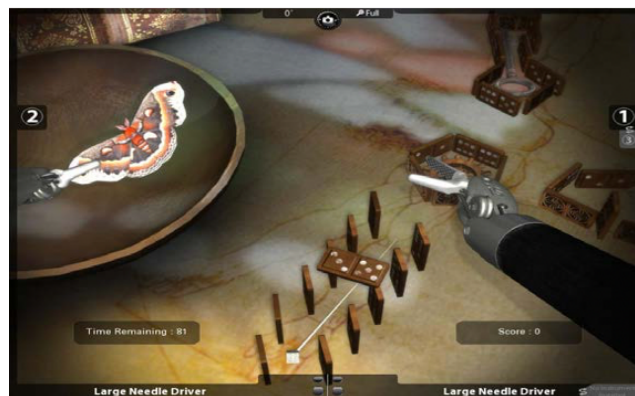


Abbildung 12. Auszug aus Vitruvian Operation. Entnommen aus Mimic Technologies (2016).

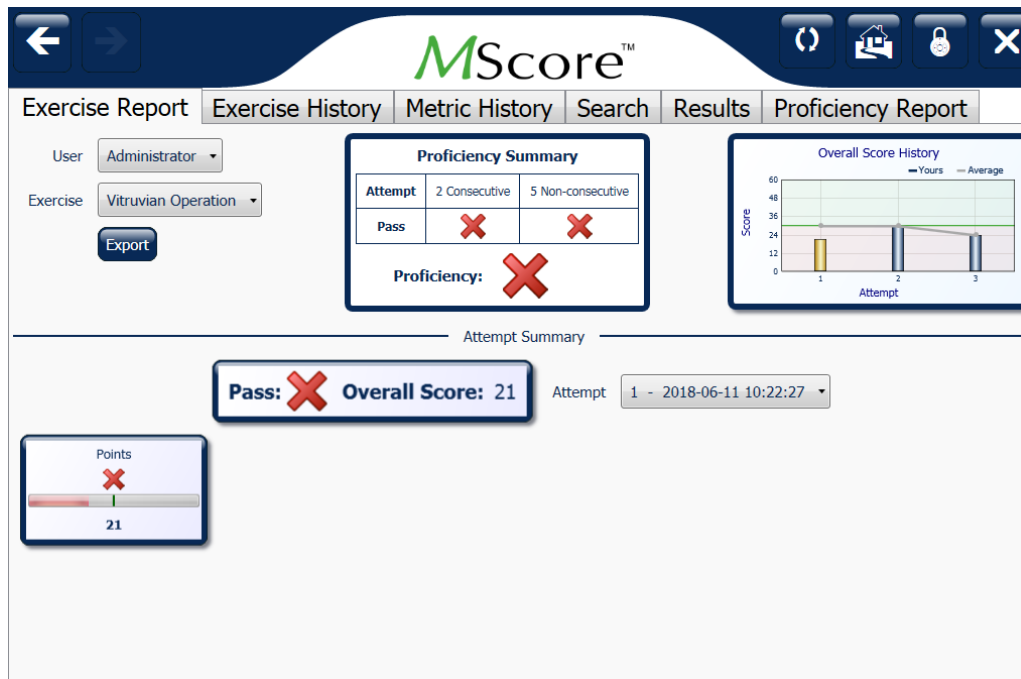


Abbildung 13. Auswertung Vitruvian Operation.

Die obenstehenden vier Übungen wurden ausgewählt, da sie in ihrer Gesamtheit viele technische und kognitive Anforderungen überprüfen, die auch ein Chirurg weitestgehend beherrschen sollte. Darunter zählen eine gute Hand-Augen-Koordination, präzises Arbeiten, eine gute Auffassungsgabe und Konzentrationsfähigkeit, eine sichere Instrumenten- und Kameraführung sowie das Arbeiten unter einer Drucksituation.

3.3 Probandenkollektiv und Probandenakquise

Das Probandenklientel der klinischen Studie beschränkt sich auf Ärzte in der Weiterbildung oder Fachärzte, die in einem chirurgischen Gebiet tätig sind. Akquiriert wurden die Probanden durch eine Vorstellung der geplanten Studie innerhalb der verschiedenen chirurgischen Abteilungen in der Universitätsmedizin Mainz. Zudem wurden Handouts mit Informationen zur Studie und Kontaktdaten ausgehändigt. Ergänzend zu der Studienvorstellung erhielten die chirurgischen Mitarbeiter per E-Mail eine Anfrage zur Studienteilnahme und dem Informationsschreiben. Die Teilnahme erfolgt auf freiwilliger Basis. Insgesamt wurden 20 Probanden als Teilnehmer in die Studie eingeschlossen werden.

Einschlusskriterien zur Studienteilnahme waren die Volljährigkeit, eine angegebener und in der Studieneinwilligung schriftlich bestätigter gesunder Allgemeinzustand ohne Alkoholabhängigkeit, keine regelmäßige Medikamenten- oder Drogeneinnahme und der Ausschluss einer

Schwangerschaft für weibliche Probanden, welches auch in der Studieneinwilligung schriftlich bestätigt werden muss. Gleichfalls musste schon einmal Alkohol genossen und vertragen worden sein.

3.4 Intervention Alkohol

Als zu testendes Alkohollevel wurde für die Studie ein Blutalkoholspiegel von 0,8 Promille (1 Promille = 1 Gramm Alkohol/ pro 1 Kilogramm Blut) definiert. In Deutschland ist die Grenze für die Verkehrsteilnahme auf 0,5 Promille festgesetzt. Ab 0,8 Promille wird ein deutliches Nachlassen der Reaktionsfähigkeit und der Konzentration beschrieben (BZgA, 2020). Dieser Spiegel kann bei einem Durchschnitts-Arzt (Zielproband, männlich, 80kg) durch die Ingestion von 3 Flaschen Bier erreicht werden. Erst ab einem Blutalkoholspiegel von 2 Promille besteht laut der Literatur eine akute Schockgefahr und damit eine erhöhte Gefahr für einen Probanden bei ansonsten vollständig vorliegenden Einschlusskriterien (BZgA, 2020). Dieser Blutalkoholspiegel wird in unserer geplanten Studie nicht annähernd erreicht. Durch eine Zusammenarbeit mit der Untersuchungsstelle für Blutalkohol der Uniklinik Mainz und dessen Leiter konnten weitere Kriterien für eine möglichst valide Durchführung der Studie mit der Intervention Alkohol festgelegt werden, welche unter 3.6.4. näher beschrieben werden. Im Vorfeld an die Testung wurden die Probanden unter anderem darüber aufgeklärt, dass sie nach der Intervention kein Kraftfahrzeug führen dürfen oder anderweitig am Straßenverkehr teilnehmen dürfen, was sie durch den Alkoholgenuss beeinträchtigen oder gefährden könnte.

3.5 Studienbeschreibung

Bei dieser Studie handelt es sich um eine monozentrische und prospektive experimentelle Studie. Die Teilnehmerzahl wurde durch einen statistischen Mitarbeiter des Institutes für medizinische Biometrie und Epidemiologie der Universitätsmedizin Mainz auf insgesamt 20 Probanden festgelegt. Diese Anzahl konnte durch die Ergebnisse der ersten zehn Versuchspersonen ermittelt werden. Jeder Proband selbst absolvierte einen Kontroll- und einen Versuchsdurchgang, was eine direkte Vergleichsmöglichkeit der Ergebnisse jeweils mit und ohne Intervention ermöglichte. Die Studie wurde nicht verblindet, da die Anflutungssymptome des Alkohols nicht unterbunden werden konnten. Im Anhang 8.4 ist ein Überblick zum geplanten Studienablauf zu sehen. Diese Übersicht wurde vor Beginn der Testphase erstellt, um als Orientierungshilfe zu dienen und einen einheitlichen Ablauf an jedem Versuchstag zu gewährleisten.

3.5.1 Vorstudie zur Lernkurve

Zwei weitere Kolleginnen führten am MIMIC dV-Trainer® ebenfalls Interventionsstudien durch und untersuchten, wie sich Schlafentzug und Stressoren auf die Performance am MIMIC dV-

Trainer® auswirken. Im Vorfeld an die insgesamt drei geplanten Interventionsstudien wurde eine Vorstudie zur Lernkurve durchgeführt. Dabei sollte herausgefunden werden, wie oft die Übungen absolviert werden müssen, bis keine signifikante Verbesserung der „Scores“ mehr zu sehen ist und somit ein Lerneffektfehler des Experiments ausgeschlossen werden kann. Für diese Vorstudie wurden 15 Studenten rekrutiert, die drei verschiedene Übungen am MIMIC dV-Trainer® jeweils insgesamt zwölfmal üben sollten. Die Gesamtauswertung zeigte, dass bei allen drei Übungen ab dem fünften Durchgang ein Plateau der Leistung zu erkennen war. Somit wurde für die Interventionsstudien festgelegt, dass die Probanden zum Ausschluss einer möglichen Ergebnisverfälschung durch einen Lerneffekt, im Vorfeld an die Testdurchführungen das geplante Curriculum viermal üben sollten.

3.5.2 Studiendurchführung

Die Probandenakquise anhand der Einschlusskriterien fand über Aushänge, soziale Netzwerke, Vorstellung der Studie in den einzelnen chirurgischen Abteilungen und durch persönliche Einladungen per E-Mail statt. Interessierte Probanden bekamen anschließend Informationen zu der Studie anhand eines Informationsblattes. Bei einer Entscheidung zur Teilnahme an der Studie wurden die Probanden nochmals näher hinsichtlich der Vorbereitung, der Durchführung und der Auswertung des Experimentes aufgeklärt. Nach einer ausreichenden Bedenkzeit für die Studienteilnehmer und einer Möglichkeit, noch offene Fragen klären zu können, mussten die Probanden vor Studienbeginn eine Einwilligungserklärung unterschreiben.

Am Testtag erschienen die Probanden gegen 13 Uhr in den Räumlichkeiten der Rudolf-Frey-Lernklinik der Universitätsmedizin Mainz, wo sich der MIMIC dV-Trainer® befindet. Die Testpersonen sollten darauf geachtet haben, dass die letzte Mahlzeit nicht zu fettreich gewesen ist, um die enterale Aufnahme des Alkohols nicht zu sehr zu beeinflussen (Mußhoff, Madea, & Heß, 2014). Zunächst wurde bei jedem Teilnehmer ein Drogenscreeningtest mittels einer Urinprobe durchgeführt. Dieser wurde von der Firma „Immunanalysis“ zur Verfügung gestellt und beinhaltete eine Testung auf Morphin, Amphetamin, THC und Kokain. Desweiteren wurde bei weiblichen Testpersonen vor Beginn der Studie ein Schwangerschaftstest, ebenfalls mittels eines Teststreifen für Urinproben, durchgeführt. Bei negativen Drogen- und Schwangerschaftstests galten die Einschlusskriterien als erfüllt und die Probanden konnten für die Studie freigegeben werden.

Anhand des Gewichtes des Probanden und weiterer Faktoren wie dem Geschlecht, dem prozentualen Volumenanteil des Alkohols und der Testungszeiten wurde in einer ermittelt, wieviel

Milliliter des alkoholischen Getränkes genossen werden muss, um eine Blutalkoholkonzentration von 0,8 Promille zu erhalten. Mithilfe von Herrn Dr. Kaufmann, dem stellvertretenden Leiter der Blutalkoholstelle der Universitätsklinik Mainz, und basierend auf der „Widmarkschen Formel“, die unter 2.2 näher erläutert wird, konnte eine Excel-Tabelle erstellt werden (siehe Anhang 8.3). Diese diente als Instrument zur Berechnung der benötigten Menge des alkoholischen Getränks in Milliliter. Anschließend sollte der Proband, der in den meisten Fällen noch nicht an dem Simulationsroboter gearbeitet hat, die vier Durchgänge des Curriculums, zum Ausschluss eines möglichen Lerneffektfehlers absolvieren. Das Curriculum beinhaltet die vier Übungen Suture Sponge, Vitruvian Operation, Stacking Challenge und Ring Tower, die unter 3.2. näher beschrieben werden. Im Anschluss daran sollten die Testteilnehmer eine Pause von circa 30 Minuten machen, um die Testung nicht durch Konzentrationsdefizite unbewusst zu beeinflussen.

Nach dieser Pause absolvierten die Probanden die vier Übungen des Curriculums am MIMIC dV-Trainer®. Dieser Durchgang repräsentiert nun den Kontrolldurchgang. Danach erhielten die Teilnehmer die errechnete Menge Alkohol, in den meisten Fällen war dies Wodka mit Orangensaft oder Gin mit Tonic. Die Auswahl des Testgetränkes richtete sich dabei nach einem möglichst hohen Volumenanteil an Alkohol, um mit einer möglichst geringen Menge den gewünschten Blutalkoholspiegel zu erhalten und nach dem Geschmack der Testpersonen. Nachdem der Teilnehmer die vorgegebene Menge an Alkohol innerhalb von 15-30 Minuten zu sich genommen hatte, fand eine weitere Pause statt. Diese Pause diente vor allem dazu, Anflutungssymptome des Alkohols während des Versuchs zu vermeiden. Außerdem sollte sich der Teilnehmer, im Hinblick auf den alkoholphysiologischen Metabolismus, in der sogenannten Eliminationsphase befinden. Diese Phase ist dadurch charakterisiert, dass ein leichtes und kontinuierliches Sinken des Promillewertes stattfindet, was mit mehrmaligem Messen des Atemalkohols verifiziert wurde (Mußhoff, Madea, & Heß, 2014). Die Messungen des Atemalkohols wurden mit einem Atemalkoholtestgerät durchgeführt, welches von der Untersuchungsstelle für Blutalkohol der Universitätsmedizin Mainz zur Verfügung gestellt wurde. Die Daten mit den protokollierten Atemalkoholmessungen wurden pseudonymisiert und digitalisiert gespeichert.

Das Wissen über die alkoholphysiologischen Aspekte zur Durchführung der Studie konnte durch die fachliche Beratung des Leiters der Blutalkoholstelle der Rechtsmedizin der Universitätsmedizin Mainz und durch Literaturrecherche gewonnen werden. Befand sich der Teilnehmer laut Messgerät und zeitlicher Einschätzung in der oben beschriebenen Eliminationsphase, wurde mit dem Versuchsdurchgang gestartet. Der Proband führte nun die vier Übungen des Curriculums unter Alkoholeinfluss, demgemäß mit der Intervention durch. Dies repräsentierte

den Versuchsdurchgang. Im Anschluss an diesen Durchgang wurde dem Probanden eine venöse Blutprobe entnommen und der Atemalkoholwert gemessen. Durch die Untersuchungsstelle für Blutalkohol konnte die tatsächliche Blutalkoholkonzentration in der venösen Probe ermittelt werden. Die Dauer des Prüfvorhabens für einen Einzelnen als Teilnehmer mit Übungsdurchgängen zum Ausschluss des Lerneffektes, Pausen, Getränkegenuss, Alkoholmessungen und Übungen am MIMIC dV-Trainer® betrug zusammen mit der Studienaufklärung und -einwilligung im Durchschnitt drei Stunden. Nach vollständigem Einschluss von 10 Probanden erfolgte eine Zwischenevaluation zur weiterführenden statistischen Beratung und Festlegung der final notwendigen Probandenanzahl von insgesamt 20 Versuchsteilnehmern.

Die Studie wurde in der internationalen Registerdatenbank „Research Registry“ (<https://www.researchregistry.com/>) mit der Nummer „researchregistry3178“ registriert und der Ethikkommission der Landesärztekammer Rheinland-Pfalz zur Begutachtung vorgelegt und unter der Bearbeitungsnummer 837.319.17 (11152) mit zustimmendem Votum belegt.

Die Darstellung der Daten und die Teststatistik dieses Kapitels wurden mithilfe von IBM SPSS und Microsoft Excel angefertigt. Dabei wurden die gewonnenen Rohdaten aus unserer Studie einer deskriptiven als auch explorativen Datenanalyse unterzogen. Die statistische Beratung erfolgte durch das Institut für medizinische Biometrie und Epidemiologie der Universitätsmedizin Mainz. Zum besseren Verständnis der einzelnen Testverfahren und Diagramme werden diese im Folgenden zunächst näher erläutert.

3.6 Studienfinanzierung

Die Studie wurde maßgeblich vom Verein Sterntaler e.V. finanziert. Der Verein Sterntaler e.V. ist ein eingetragener gemeinnütziger Verein, der sich für eine Verbesserung der Rahmenbedingungen von unfallverletzten und operativ zu versorgenden Kindern einsetzt. Dieser Verein vergibt jährlich bis zu zwei Förderstipendien für Forschungsprojekte an der Universitätsmedizin Mainz. Durch die finanzielle Unterstützung mit insgesamt 2.500 € konnten somit insgesamt vier Studien am MIMIC dV-Trainer®, unter anderem die Interventionsstudie mit Alkohol durchgeführt werden.

3.7 Statistische Testverfahren

Zur Darstellung der Messwerte ohne und unter der Intervention innerhalb jeder Metrik wurde ein **Streudiagramm** verwendet. Dadurch sollen potenzielle Beziehungen zwischen den beiden Variablen „ohne Intervention“ und „mit Intervention“ verdeutlicht werden und das Maß der Streubreite kann wieder gespiegelt werden. Die Werte unter der Intervention befinden sich auf der x-Achse und die Werte ohne Intervention befinden sich auf der y-Achse.

Die Anzahl der Probanden, die sich unter der Intervention verbessert oder verschlechtert haben beziehungsweise die darunter keine Leistungsänderung zeigten wurde anhand eines **Vorzeichentests** ermittelt, welcher in tabellarischer Form dargestellt wird. Positive Differenzen heißt dabei, dass diese Probandenanzahl eine schlechtere Leistung unter der Intervention gezeigt hat und negative Differenzen, dass diese Probandenanzahl eine bessere Leistung darunter gezeigt hat. Bei den Übungen Stacking Challenge und Vitruvian Operation verhält sich die Zuteilung umgekehrt, da hier die Punktzahl bewertet wird. Bindung bedeutet, dass die Intervention keinen Leistungsunterschied aufzeigt. Der Vorzeichentest wurde bei jeder Anwendung anhand seiner positiven Differenzen auf eine vorliegende Signifikanz überprüft. Bei einer zu geringen Stichprobengröße von $n < 26$ erfolgte die Testung der Signifikanz mithilfe des z-Wertes. Dieser wird mit dem kritischen Wert einer Standardnormalverteilung verglichen. Bei einem zweiseitigen Signifikanzniveau von ,05 beträgt der z-Wert $\pm 1,96$. Ist der Betrag der Teststatistik höher als der kritische Wert, ist der Test statistisch signifikant (Universität Zürich, 2020).

4 Ergebnisse

4.1 Primäre Messparameter

In Tabelle 1 sind alle im Folgenden analysierten Metriken mit den dazugehörigen Einheiten zusammengefasst.

Tabelle 1: Metriken und dazugehörige Einheiten

Übung	Metrik	Einheit
Ring Tower	RT_Time_to_complete	Sekunden (sec)
	RT_Wire_collisions	Anzahl
	RT_Wire_contact_duration	Sekunden (sec)
	RT_Tower_knockoff	Anzahl
	RT_Economy_of_motion	Zentimeter (cm)
	RT_Instrument_collisions	Anzahl
	RT_Instruments_out_of_view	Zentimeter (cm)
	RT_Drops	Anzahl
Suture Sponge	SS_Time_to_complete	Sekunden (sec)
	SS_Economy_of_motion	Zentimeter (cm)
	SS_Instrument_collisions	Anzahl
	SS_Excessive_instrument_force	Sekunden (sec)
	SS_Instruments_out_of_view	Zentimeter (cm)
	SS_Master_workspace_range	Zentimeter (cm)
	SS_Drops	Anzahl
	SS_Missed_targets	Anzahl
Stacking Challenge	ST_Overall_Score	Punkte
Vitruvian Operation	VO_Overall_Score	Punkte

4.1.1 Ring Tower

Während der Datenerhebungsphase wurde die Software des MIMIC dV-Trainer® auf die Version 3.6 erneuert. Die Übung Ring Tower war nach dem Upgrade nicht mehr vorhanden, weshalb hier nur Ergebnisse von 17 Probanden ausgewertet werden können. In Tabelle 2 sind Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken von Ring Tower zu sehen. Bei insgesamt sieben Metriken zeigen sich die Mittelwerte unter der Intervention (_1) höher als vor der Intervention (_0), was in dem Fall eine schlechtere Leistung bedeutet. Bei einer Metrik („Time to complete“) war zu sehen, dass die Werte unter Alkoholeinfluss im Mittel besser waren, die Probanden haben die Aufgabe schneller absolviert.

Tabelle 2: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken der Übung Ring Tower

Metrik	N	Minimum	Maximum	Mittelwert	Standardabweichung	Variationskoeffizient
RT Time to complete (sec) 0	17	63,92	285,04	151,91	59,705	0,393
RT Time to complete (sec) 1	17	70,57	250,60	146,74	53,589	0,365
RT Wire collisions (Anzahl) 0	17	5,00	20,00	12,06	3,631	0,301
RT Wire collisions (Anzahl) 1	17	3,00	39,00	15,76	7,910	0,501
RT wire contact duration (sec) 0	17	2,26	63,00	17,34	14,486	0,835
RT wire contact duration (sec) 1	17	0,90	110,29	25,16	29,314	1,164
RT Tower knockoff (Anzahl) 0	17	0,00	1,00	0,12	0,332	2,766
RT Tower knockoff (Anzahl) 1	17	0,00	2,00	0,35	0,702	2,005
RT Economy of motion (cm) 0	17	144,24	483,59	239,61	79,954	0,333
RT Economy of motion (cm) 1	17	150,71	519,84	269,35	99,308	0,368
RT Instrument collisions (Anzahl) 0	17	0,00	12,00	2,35	3,278	1,394
RT Instrument collisions (Anzahl) 1	17	0,00	12,00	3,24	3,580	1,104
RT Instruments out of view (cm) 0	17	0,00	33,25	7,40	10,038	1,355
RT Instruments out of view (cm) 1	17	0,02	47,38	9,67	13,042	1,348
RT Drops (Anzahl) 0	17	0,00	2,00	0,65	0,786	1,209
RT Drops (Anzahl) 1	17	0,00	3,00	1,29	1,105	0,856

4.1.1.1 Time to complete

In Abbildung 14 sieht man die Streuung der einzelnen Messwerte von „Time to complete“ der Übung Ring Tower. Die y-Achse zeigt die benötigte Zeit vor der Intervention und die x-Achse zeigt die benötigte Zeit unter der Intervention. Die Skalierung beträgt jeweils 50 Sekunden.

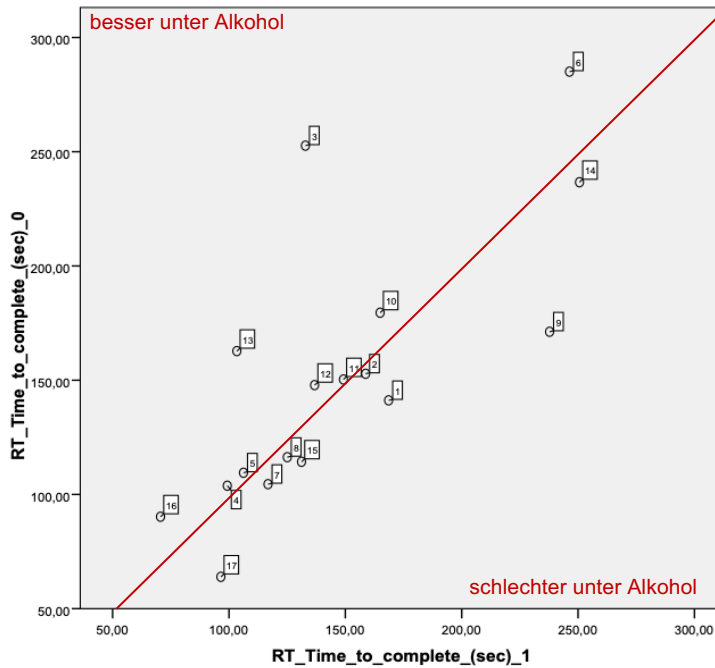


Abbildung 14. Streudiagramm zu „Time to complete“ der Übung Ring Tower mit (_1) und ohne (_0) Intervention

In Tabelle 3 ist zu sehen, dass acht Probanden unter der Intervention in der Metrik „Time to complete“ eine längere Zeit und neun Versuchspersonen eine kürzere Zeit benötigt haben. Der z-Wert als Ausmaß der standardisierten Teststatistik ist in Tabelle 4 dargestellt. Mit einem Wert von $z = ,000$ liegt dieser unter dem kritischen Wert von 1,96. Somit ist das Ergebnis nicht signifikant.

Tabelle 3: Häufigkeiten der Differenzen von „Time to complete“ der Übung Ring Tower

		N
RT_Time_to_complete_(sec)_1 -	Negative Differenzen ^a	9
RT_Time_to_complete_(sec)_0	Positive Differenzen ^b	8
RT_Time_to_complete_(sec)_1 -	Bindungen ^c	0
RT_Time_to_complete_(sec)_0	Gesamt	17

a. $RT_Time_to_complete_1 < RT_Time_to_complete_0$

b. $RT_Time_to_complete_1 > RT_Time_to_complete_0$

c. $RT_Time_to_complete_1 = RT_Time_to_complete_0$

Tabelle 4: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Time to complete“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	8,000 ^a
Standardfehler	2,062
Standardisierte Teststatistik	,000
Asymptotische Sig. (zweiseitiger Test)	1,000
Genauere Sig. (zweiseitiger Test)	1,000

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.1.2 Wire collisions

In der untenstehenden Abbildung 15 ist die Streuung der Messwerte zu „Wire collisions“ der Übung Ring Tower zu sehen. Auf der y-Achse sieht man „Wire collisions“ vor der Intervention und auf der x-Achse „Wire collisions“ unter der Intervention. Die einzelnen Messwerte liegen hier weitestgehend beieinander und weisen keine große Streubreite auf.

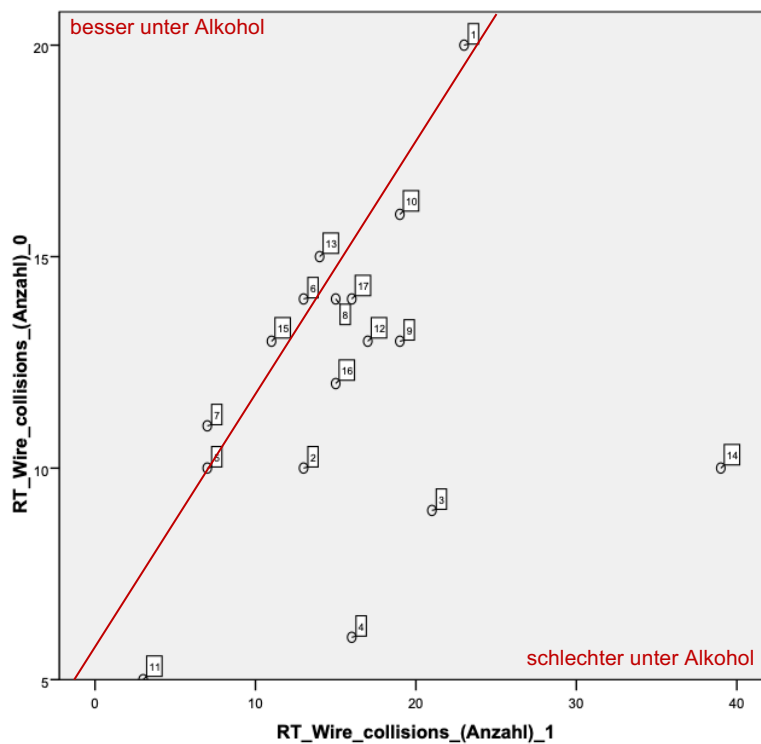


Abbildung 15. Streudiagramm zu „Wire collisions“ der Übung Ring Tower mit (_1) und ohne (_0) Intervention.

Elf Probanden haben laut dem Ergebnis des Vorzeichen-tests, was in Tabelle 5 zu sehen ist, häufigere Kollisionen und sechs Probanden haben seltenere Kollisionen unter der Intervention gezeigt. Eine Signifikanz liegt bei einem Wert von $z = ,970$ nicht vor.

Tabelle 5: Häufigkeiten der Differenzen von „Wire collisions“ der Übung Ring Tower

		N
RT_Wire_collisions_(Anzahl)	Negative Differenzen ^a	6
_1 -	Positive Differenzen ^b	11
RT_Wire_collisions_(Anzahl)	Bindungen ^c	0
_0	Gesamt	17

a. $RT_Wire_collisions_1 < RT_Wire_collisions_0$

b. $RT_Wire_collisions_1 > RT_Wire_collisions_0$

c. $RT_Wire_collisions_1 = RT_Wire_collisions_0$

Tabelle 6: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Wire collisions“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	11,000 ^a
Standardfehler	2,062
Standardisierte Teststatistik	,970
Asymptotische Sig. (zweiseitiger Test)	,332
Genauere Sig. (zweiseitiger Test)	,332

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.1.3 Wire contact duration

Abbildung 16 stellt die Streuung der Leistungen bei der Metrik „Wire contact duration“ dar. Die y-Achse gibt die Ergebnisse ohne Intervention und die x-Achse die Ergebnisse mit Intervention wieder. Ein extremer Ausreißer ist Proband 15, der unter der Intervention eine deutlich geringere Kontaktzeit mit dem Draht hatte als vor der Intervention. Weitere Ausreißer präsentieren die Probanden 2 und 14. Diese haben vor der Intervention eine deutlich geringere Kontaktzeit am Draht gehabt.

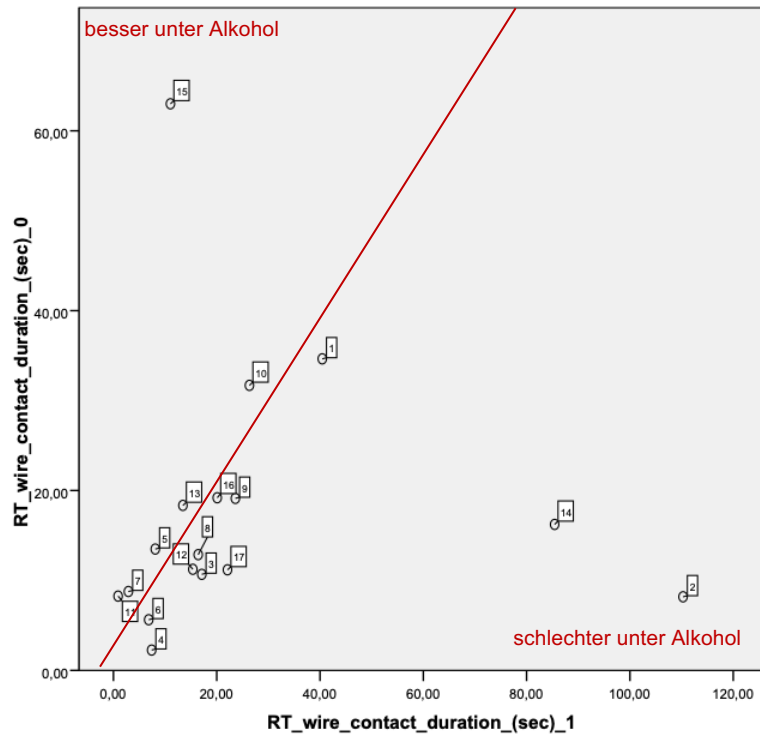


Abbildung 16. Streudiagramm zu „Wire contact duration“ der Übung Ring Tower mit (_1) und ohne (_0) Intervention.

Aus Tabelle 7 ist zu entnehmen, dass elf Probanden unter der Intervention eine längere Kontaktzeit am Draht hatten. Der z-Wert als Ausmaß der standardisierten Teststatistik ist in Tabelle 8 dargestellt. Mit einem Wert von $z = ,970$ liegt dieser unter dem kritischen Wert von 1,96 und erweist keine statistische Signifikanz.

Tabelle 7: Häufigkeiten der Differenzen von „Wire contact duration“ der Übung Ring Tower

		N
RT_wire_contact_duration_(sec)_1 -	Negative Differenzen ^a	6
RT_wire_contact_duration_(sec)_0	Positive Differenzen ^b	11
	Bindungen ^c	0
	Gesamt	17

a. $RT_wire_contact_duration_(\sec)_1 < RT_wire_contact_duration_(\sec)_0$

b. $RT_wire_contact_duration_(\sec)_1 > RT_wire_contact_duration_(\sec)_0$

c. $RT_wire_contact_duration_(\sec)_1 = RT_wire_contact_duration_(\sec)_0$

Tabelle 8: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Wire contact duration“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	11,000 ^a
Standardfehler	2,062
Standardisierte Teststatistik	,970
Asymptotische Sig. (zweiseitiger Test)	,332
Genau Sig. (zweiseitiger Test)	,332

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.1.4 Tower Knockoff

Die Abbildung 17 stellt das Streudiagramm zu der Kategorie Tower Knockoff dar. Die höchsten erbrachten Werte in dieser Kategorie sind ohne den Einfluss von Alkohol ein „Tower knockoff“ und unter dem Einfluss von Alkohol zwei „Tower knockoff“. Einige Probanden haben vor sowie unter der Intervention null „Tower knockoff“ erreicht und somit keine Leistungssteigerung oder -minderung dargeboten.

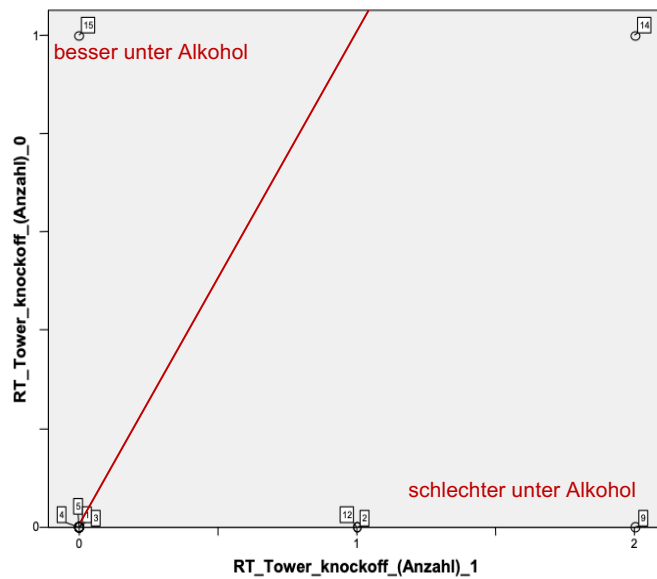


Abbildung 17. Streudiagramm zu „Tower knockoff“ der Übung Ring Tower mit (₁) und ohne (₀) Intervention.

Die Auswertung des Vorzeichenstest (s. Tabelle 9) ergab, dass ein Proband unter der Intervention seltener „Tower knockoff“ zeigte. Vier Probanden zeigten unter der Intervention häufiger „Tower knockoff“ und bei insgesamt zwölf Personen ergab sich durch die Intervention kein Unterschied (= Bindungen) in ihrer Leistung. Die standardisierte Teststatistik (s. Tabelle 10) beträgt $z = ,894$, was keine Signifikanz nachweist.

Tabelle 9: Häufigkeiten der Differenzen von „Tower knockoff“ der Übung Ring Tower

		N
RT_Tower_knockoff_(Anzahl)_1 -	Negative Differenzen ^a	1
RT_Tower_knockoff_(Anzahl)_0	Positive Differenzen ^b	4
	Bindungen ^c	12
	Gesamt	17

a. $RT_Tower_knockoff_ (Anzahl)_1 < RT_Tower_knockoff_ (Anzahl)_0$

b. $RT_Tower_knockoff_ (Anzahl)_1 > RT_Tower_knockoff_ (Anzahl)_0$

c. $RT_Tower_knockoff_ (Anzahl)_1 = RT_Tower_knockoff_ (Anzahl)_0$

Tabelle 10: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Tower knockoff“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	4,000 ^a
Standardfehler	1,118
Standardisierte Teststatistik	,894
Asymptotische Sig. (zweiseitiger Test)	,371
Genauere Sig. (zweiseitiger Test)	,375

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.1.5 Economy of motion

Abbildung 18 zeigt das Streudiagramm zu „Economy of motion“ der Übung Ring Tower. Anhand der Verteilung erkennt man, dass die Streuung der Leistungen vor und unter der Intervention bei der Rubrik „Economy of motion“ nicht sehr ausgeprägt ist. Die Messwerte liegen in ihrer Gesamtheit nah beieinander. Proband 10 kann als Ausreißer definiert werden. Dieser zeigte unter der Intervention eine geringere Zentimeteranzahl bei der Bewegungsökonomie. Die Probanden 9, 12 und 14 zeigten unter der Intervention ein größeres Ausmaß der Bewegung.

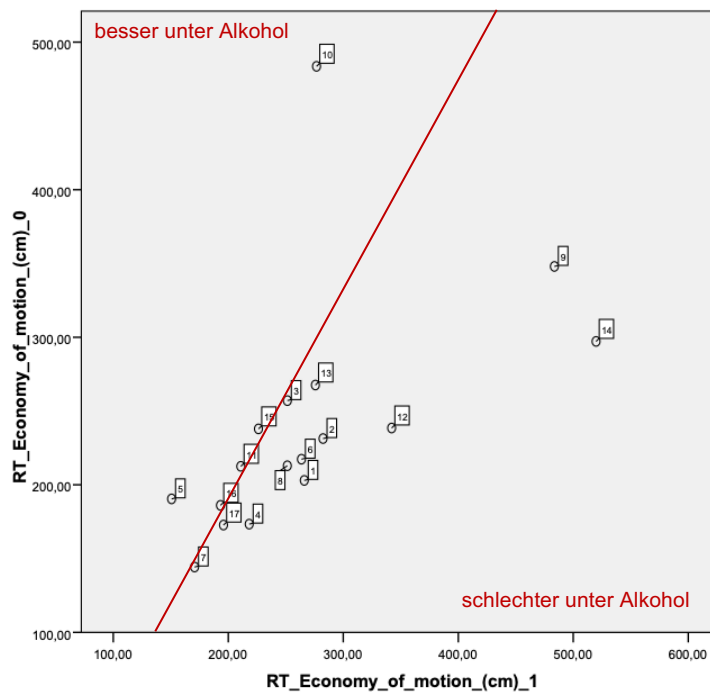


Abbildung 18. Streudiagramm zu „Economy of motion“ der Übung Ring Tower mit (_1) und ohne (_0) Intervention.

Der Vorzeichentest, dessen Ergebnisse in den Tabellen 11 und 12 zu sehen sind, ergab, dass insgesamt zwölf Probanden unter der Intervention mit den Instrumenten eine längere Strecke und fünf Probanden unter dem Einfluss von Alkohol eine kürzere Strecke zurückgelegt haben. Bei einer Teststatistik von $z = 1,455$ und einem Signifikanzniveau von $> 1,96$ zeigt sich hier jedoch kein signifikanter Leistungsunterschied unter Alkoholeinfluss.

Tabelle 11: Häufigkeiten der Differenzen von „Economy of motion“ der Übung Ring Tower

		N
RT_Economy_of_motion_(cm)	Negative Differenzen ^a	5
_1 -	Positive Differenzen ^b	12
RT_Economy_of_motion_(cm)	Bindungen ^c	0
_0	Gesamt	17

a. $RT_Economy_of_motion_ (cm)_1 < RT_Economy_of_motion_ (cm)_0$

b. $RT_Economy_of_motion_ (cm)_1 > RT_Economy_of_motion_ (cm)_0$

c. $RT_Economy_of_motion_ (cm)_1 = RT_Economy_of_motion_ (cm)_0$

Tabelle 12: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Economy of motion“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	12,000 ^a
Standardfehler	2,062
Standardisierte Teststatistik	1,455
Asymptotische Sig. (zweiseitiger Test)	,146
Genauere Sig. (zweiseitiger Test)	,143

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.1.6 Instrument collisions

Das Streudiagramm zu „Instrument collisions“ ist in Abbildung 19 zu sehen. Auffallend ist, dass die Werte in diesem Diagramm im Vergleich zu den vorherigen Streudiagrammen breiter gestreut sind und sich somit eine deutlichere Tendenz darstellt, ob der Proband unter Alkoholeinfluss häufiger oder seltener eine Kollision der Instrumente zeigte. Proband 9 hat ohne Alkoholeinfluss zwölf Kollisionen und alkoholisiert nur zwei Kollisionen gezeigt. Bei Proband 14 verhält sich dies umgekehrt; nüchtern traten bei ihm zwei Kollisionen auf und unter Alkoholeinfluss zeigte er zwölf „Instrument collisions“.

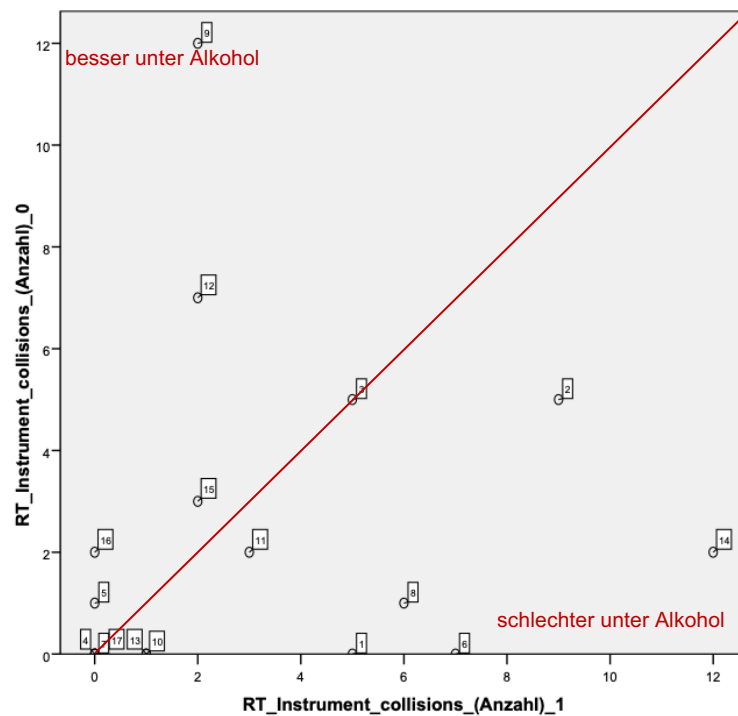


Abbildung 19. Streudiagramm zu „Instrument collisions“ der Übung Ring Tower mit (_1) und ohne (_0) Intervention.

Aus den Tabellen 13 und 14 geht hervor, dass insgesamt 8 Probanden unter der Intervention mehr Kollisionen gezeigt haben als ohne Intervention. Vier Probanden haben hingegen weniger Instrument Collisions gezeigt und vier Probanden haben vor wie auch unter der Intervention eine identische Leistung gehabt. Bei einem z-Wert von $z = ,555$ zeigt sich diese Auswertung nicht signifikant und die Nullhypothese kann nicht verworfen werden.

Tabelle 13: Häufigkeiten der Differenzen von „Instrument collisions“ der Übung Ring Tower

		N
RT_Instrument_collisions_	Negative Differenzen ^a	5
(Anzahl)_1 -	Positive Differenzen ^b	8
RT_Instrument_collisions_	Bindungen ^c	4
(Anzahl)_0	Gesamt	17

a. $RT_Instrument_collisions_ (Anzahl)_1 < RT_Instrument_collisions_ (Anzahl)_0$

b. $RT_Instrument_collisions_ (Anzahl)_1 > RT_Instrument_collisions_ (Anzahl)_0$

c. $RT_Instrument_collisions_ (Anzahl)_1 = RT_Instrument_collisions_ (Anzahl)_0$

Tabelle 14: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instrument collisions“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	8,000 ^a
Standardfehler	1,803
Standardisierte Teststatistik	,555
Asymptotische Sig. (zweiseitiger Test)	,579
Genauere Sig. (zweiseitiger Test)	,581

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.1.7 Instruments out of view

In Abbildung 20 sieht man die Streuung der einzelnen Messwerte von „Instruments out of view“. Die meisten Messwerte liegen hier nah beieinander und zeigen nur eine geringe Streubreite. Beispielsweise hat Proband 10 unter der Intervention eine sichtbar geringere Strecke in Zentimetern zurücklegen müssen, in denen die Instrumente außer Sichtweite waren, als vor der Intervention.

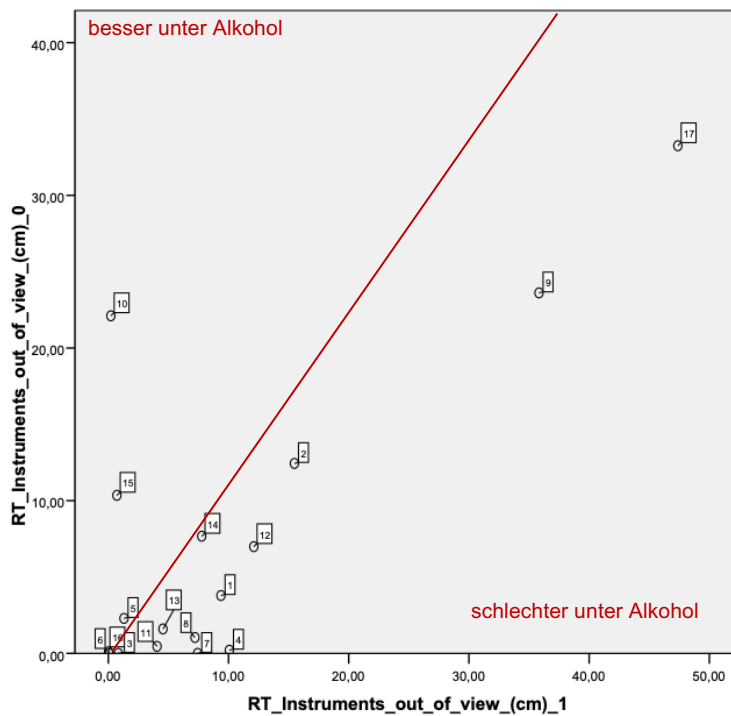


Abbildung 20. Streudiagramm zu „Instruments out of view“ der Übung Ring Tower mit (_1) und ohne (_0) Intervention.

Wie in Tabelle 15 abgebildet ist, haben 13 Probanden unter der Intervention eine längere Strecke der Instrumente außerhalb ihrer Sichtweite zurückgelegt und 4 Probanden haben darunter mit ihren Instrumenten eine kürzere Strecke zurückgelegt. Der Wert für die standardisierte Teststatistik ist in Tabelle 16 zu sehen und entspricht mit $z = 1,940$ etwa dem kritischen Wert von 1,96. Alkoholisiert wurde demnach statistisch signifikant eine größere Strecke der Instrumente außerhalb der Sichtweite zurückgelegt und sich somit verschlechtert.

Tabelle 15: Häufigkeiten der Differenzen von „Instruments out of view“ der Übung Ring Tower

		N
RT_Instruments_out_of_view_(cm)_1 -	Negative Differenzen ^a	4
RT_Instruments_out_of_view_(cm)_0	Positive Differenzen ^b	13
	Bindungen ^c	0
	Gesamt	17

a. $RT_Instruments_out_of_view_cm_1 < RT_Instruments_out_of_view_cm_0$

b. $RT_Instruments_out_of_view_cm_1 > RT_Instruments_out_of_view_cm_0$

c. $RT_Instruments_out_of_view_cm_1 = RT_Instruments_out_of_view_cm_0$

Tabelle 16: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instruments out of view“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	13,000 ^a
Standardfehler	2,062
Standardisierte Teststatistik	1,940
Asymptotische Sig. (zweiseitiger Test)	,052
Genauere Sig. (zweiseitiger Test)	,049

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.1.8 Drops

Die Abbildung 21 gibt die Streuung der Messwerte zur Metrik „Drops“ wieder. Man kann daraus ablesen, dass vier Probanden (1, 14, 16, 17) unter der Intervention zwei „Drops“ mehr zeigten. Weitere sieben Probanden (2, 3, 5, 6, 8, 10, 13) zeigten unter dem Einfluss von Alkohol einen „Drop“ mehr als ohne Alkohol. Zwei Testpersonen (4,11) haben in beiden Fällen 0 „Drops“ gezeigt und vier Probanden (7, 9, 12, 15) haben unter der Intervention einen Drop weniger gezeigt.

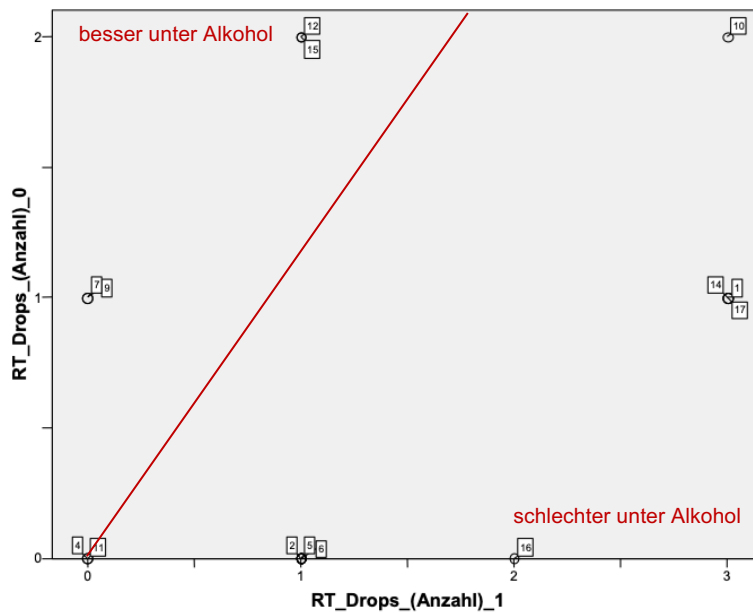


Abbildung 21. Streudiagramm zu „Drops“ der Übung Ring Tower mit (₁) und ohne (₀) Intervention.

In Tabelle 17 ist zu sehen, dass elf Probanden unter der Intervention häufiger etwas fallen lassen haben als ohne die Beeinflussung durch Alkohol. Vier Versuchspersonen haben unter diesem Einfluss seltener etwas fallen lassen und bei zwei Personen zeigte sich kein Unterschied (= Bindungen). Der z-Wert als Ausmaß der standardisierten Teststatistik ist in Tabelle 18 festgehalten. Mit einem Wert von 1,549 liegt dieser unter dem kritischen Wert von 1,96. Eine signifikante Verbesserung oder Verschlechterung zeigt sich hier unter Alkoholeinfluss nicht.

Tabelle 17: Häufigkeiten der Differenzen von „Drops“ der Übung Ring Tower

		N
RT_Drops_(Anzahl)_1 -	Negative Differenzen ^a	4
RT_Drops_(Anzahl)_0	Positive Differenzen ^b	11
	Bindungen ^c	2
	Gesamt	17

a. $RT_Drops_ (Anzahl)_1 < RT_Drops_ (Anzahl)_0$

b. $RT_Drops_ (Anzahl)_1 > RT_Drops_ (Anzahl)_0$

c. $RT_Drops_ (Anzahl)_1 = RT_Drops_ (Anzahl)_0$

Tabelle 18: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Drops“ der Übung Ring Tower

Gesamtzahl	17
Teststatistik	11,000 ^a
Standardfehler	1,936
Standardisierte Teststatistik	1,549
Asymptotische Sig. (zweiseitiger Test)	,121
Genauere Sig. (zweiseitiger Test)	,118

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2 Suture Sponge

In Tabelle 19 sind Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken von der Übung Suture Sponge zu sehen. Die einzelnen Ergebnisse verhalten sich indirekt proportional zur Leistung. Das bedeutet, je kleiner der untersuchte Wert (in Zentimeter, Sekunden oder Anzahl) ist, desto größer ist der Leistungswert. Bei allen Metriken zeigen sich die Mittelwerte unter der Intervention ($_1$) höher als vor der Intervention ($_0$). Das bedeutet, dass sich die Leistung im Mittel unter Alkoholeinfluss verschlechtert hat. Bei Metriken mit Variationskoeffizienten > 1 kann angenommen werden, dass die Standardabweichung über dem Mittelwert liegt. Somit zeigt sich hier eine hohe Streubreite der Messwerte. Zeigen sich die Variationskoeffizienten < 1 , liegt die Standardabweichung somit unter dem Mittelwert und es zeigt sich eine kleinere Streubreite der Werte.

Tabelle 19: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken der Übung Suture Sponge

Metrik	N	Minimum	Maximum	Mittelwert	Standardabweichung	Variationskoeffizient
SS_Time_to_complete_(sec)_0	20	234,61	1303,17	544,62	217,069	0,398
SS_Time_to_complete_(sec)_1	20	310,20	1494,68	683,24	267,661	0,391
SS_Economy_of_motion_(cm)_0	20	377,84	1986,98	742,72	325,861	0,438
SS_Economy_of_motion_(cm)_1	20	487,69	2747,98	941,15	492,212	0,522
SS_Instrument_collisions_(Anzahl)_0	20	8,00	84,00	35,45	18,115	0,511
SS_Instrument_collisions_(Anzahl)_1	20	11,00	175,00	59,55	39,160	0,657
SS_Excessive_Instrument_Force_(sec)_0	20	0,49	85,24	23,08	20,895	0,905
SS_Excessive_Instrument_Force_(sec)_1	20	4,03	75,72	30,45	19,628	0,644
SS_Instruments_out_of_view_(cm)_0	20	1,96	159,12	25,42	33,456	1,316
SS_Instruments_out_of_view_(cm)_1	20	10,29	799,00	71,38	172,658	2,418
SS_Master_workspace_range_(cm)_0	20	5,26	13,27	8,07	1,792	0,222
SS_Master_workspace_range_(cm)_1	20	6,21	11,30	8,38	1,226	0,146
SS_Drops_(Anzahl)_0	20	0,00	3,00	0,40	0,754	1,885
SS_Drops_(Anzahl)_1	20	0,00	9,00	1,85	2,059	1,112
SS_Missed_Targets_(Anzahl)_0	20	13,00	83,00	39,60	17,789	0,449
SS_Missed_Targets_(Anzahl)_1	20	32,00	99,00	57,35	19,784	0,344

4.1.2.1 Time to complete

Abbildung 22 stellt die Streuung der Leistungen bei der Metrik „Time to complete“ der Übung Suture Sponge dar. Die y-Achse gibt die Ergebnisse ohne Intervention und die x-Achse die Ergebnisse unter der Intervention wieder. Ein extremer Ausreißer ist Proband 1, der unter der Intervention eine deutlich geringere Zeit benötigt hat, die Übung zu beenden als vor der Intervention. Ein weiterer Ausreißer präsentiert Proband 9. Dieser hat vor der Intervention eine deutlich geringere Zeit benötigt, die Übung zu beenden.

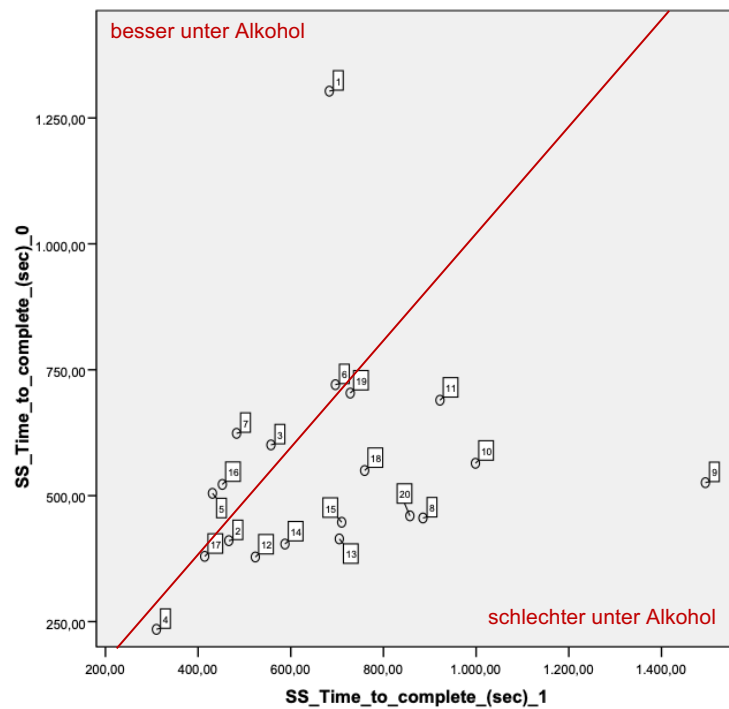


Abbildung 22. Streudiagramm zu „Time to complete“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

Anhand der Tabelle 20 ist zu entnehmen, dass 14 Probanden unter der Intervention in der Kategorie „Time to complete“ eine längere Zeit zum Abschließen der Aufgabe benötigten. 6 Versuchspersonen sind unter der Intervention schneller geworden. Der z-Wert als Ausmaß der standardisierten Teststatistik ist in Tabelle 21 dargestellt. Mit einem Wert von $z = 1,565$ liegt dieser unter dem kritischen Wert von 1,96. Somit ist das Ergebnis nicht signifikant.

Tabelle 20: Häufigkeiten der Differenzen von „Time to complete“ der Übung Suture Sponge

		N
SS_Time_to_complete_(sec)	Negative Differenzen ^a	6
_1 -	Positive Differenzen ^b	14
SS_Time_to_complete_(sec)	Bindungen ^c	0
_0	Gesamt	20

a. $SS_Time_to_complete_1 < SS_Time_to_complete_0$

b. $SS_Time_to_complete_1 > SS_Time_to_complete_0$

c. $SS_Time_to_complete_1 = SS_Time_to_complete_0$

Tabelle 21: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Time to complete“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	14,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	1,565
Asymptotische Sig. (zweiseitiger Test)	,118
Genaue Sig. (zweiseitiger Test)	,115

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2.2 Economy of motion

Das Streudiagramm zu „Economy of motion“ der Übung Suture Sponge ist in Abbildung 23 dargestellt. Es lässt sich erkennen, dass die Werte nah beieinander liegen und eine geringe Streuung vorliegt. Die Probanden 1 und 9 liegen außerhalb dieser konzentrierten Menge und stellen Ausreißer dar.

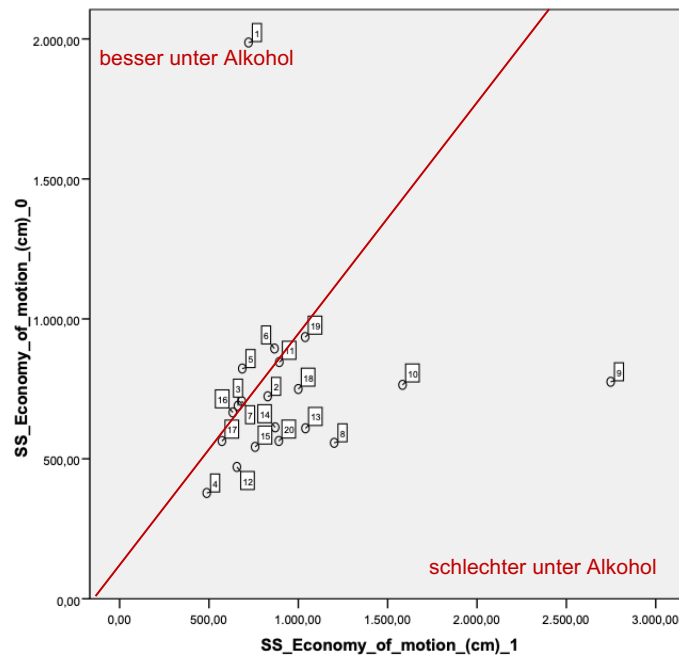


Abbildung 23. Streudiagramm zu „Economy of motion“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

Der Vorzeichentest, dessen Ergebnisse in den Tabellen 22 und 23 zu sehen sind, ergab, dass 14 Probanden unter der Intervention ein größeres Bewegungsausmaß und 6 Probanden ein kleineres Bewegungsausmaß benötigt haben. Bindungen, also kein Leistungsunterschied vor oder unter der Intervention waren in dieser Rubrik nicht zu sehen. Bei einer Teststatistik von $z = 1,565$ zeigte sich keine Signifikanz.

Tabelle 22: Häufigkeiten der Differenzen von „Economy of motion“ der Übung Suture Sponge

		N
SS_Economy_of_motion_(cm)_1 -	Negative Differenzen ^a	6
SS_Economy_of_motion_(cm)_0	Positive Differenzen ^b	14
	Bindungen ^c	0
	Gesamt	20

a. $SS_Economy_of_motion_1 < SS_Economy_of_motion_0$

b. $SS_Economy_of_motion_1 > SS_Economy_of_motion_0$

c. $SS_Economy_of_motion_1 = SS_Economy_of_motion_0$

Tabelle 23: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Economy of motion“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	14,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	1,565
Asymptotische Sig. (zweiseitiger Test)	,118
Genaue Sig. (zweiseitiger Test)	,115

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2.3 Instrument collisions

In der untenstehenden Abbildung 24 ist die Streuung der Messwerte zu „Instrument collisions“ bei Suture Sponge zu sehen. Man erkennt in dem Graphen, dass die Mehrheit der Ergebnisse vor der Intervention (= _0) weniger oder gleich viele Kollisionen im Gegensatz zu unter der Intervention (= _1) gezeigt haben. Vor allem die Probanden 10, 12, 13 und 20 haben unter der Intervention sichtbar mehr als doppelt so viele Kollisionen gezeigt.

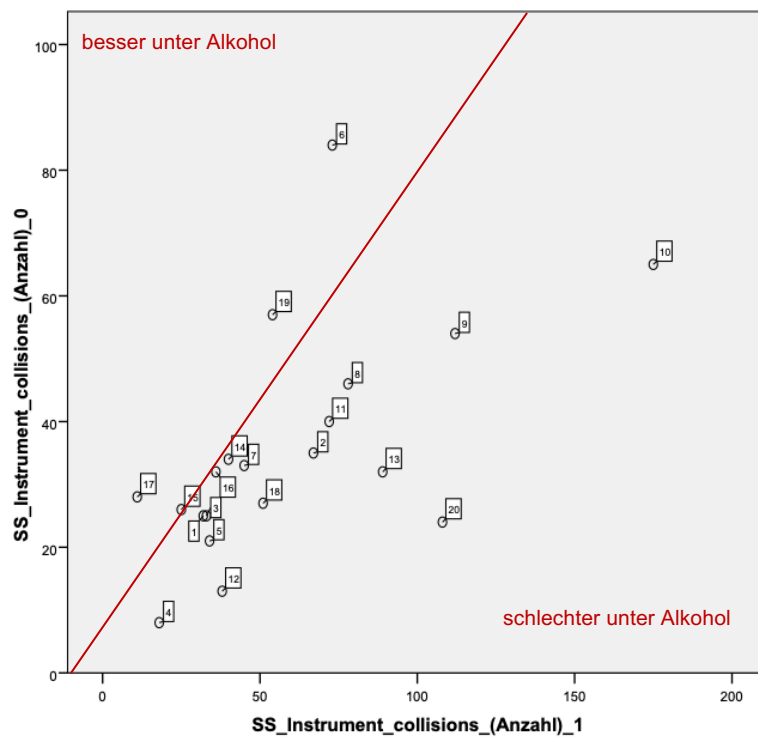


Abbildung 24. Streudiagramm zu „Instrument collisions“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

In Tabelle 24 sieht man, dass 16 Probanden unter der Intervention häufiger Instrumentenkollisionen und 4 Probanden darunter seltener Instrumentenkollisionen gezeigt haben. Die Teststatistik präsentiert sich in Tabelle 25. Bei einem z-Wert von $z = 2,460$ zeigt sich die Auswertung signifikant. Die Alkoholisierung hat in dieser Metrik also zu einer signifikant höheren Anzahl an Instrumentenkollisionen geführt.

Tabelle 24: Häufigkeiten der Differenzen von „Instrument collisions“ der Übung Suture Sponge

		N
SS_Instrument_collisions	Negative Differenzen ^a	4
_(Anzahl)_1 -	Positive Differenzen ^b	16
SS_Instrument_collisions	Bindungen ^c	0
_(Anzahl)_0	Gesamt	20

a. $SS_Instrument_collisions_1 < SS_Instrument_collisions_0$

b. $SS_Instrument_collisions_1 > SS_Instrument_collisions_0$

c. $SS_Instrument_collisions_1 = SS_Instrument_collisions_0$

Tabelle 25: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instrument collisions“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	16,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	2,460
Asymptotische Sig. (zweiseitiger Test)	,014
Genauere Sig. (zweiseitiger Test)	,012

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2.4 Excessive instrument force

Das Streudiagramm der Abbildung 25 zeigt, dass die Messwerte nah beieinander liegen und eine geringe Streubreite aufweisen. Proband 9 zeigte unter Alkoholeinfluss eine länger andauernde (gemessen in Sekunden) Kraftereinwirkung auf die Greifinstrumente. Proband 2 zeigte bei akuter Alkoholisierung eine kürzer andauernde Kraftereinwirkung auf die Greifer.

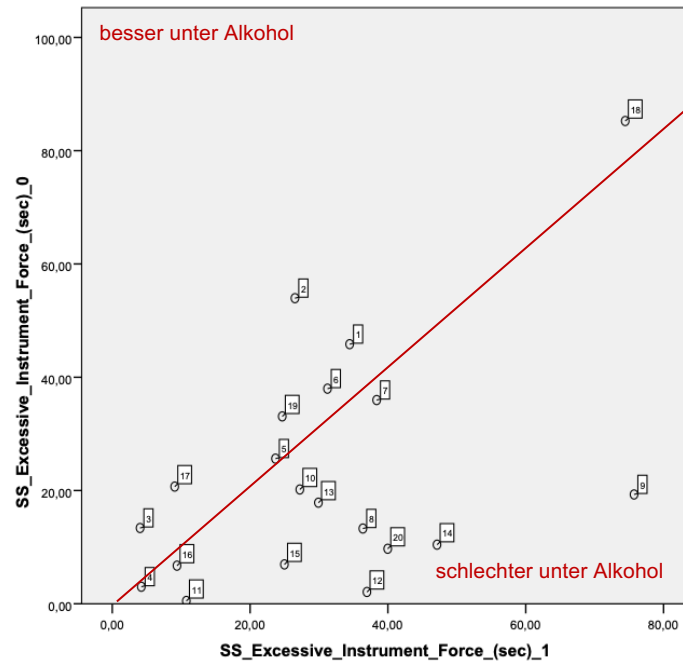


Abbildung 25. Streudiagramm zu „Excessive instrument force“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

Es konnte ermittelt werden, dass 12 Probanden unter akutem Alkoholkonsum eine länger anhaltende Krafteinwirkung zeigten. Acht Probanden haben unter dem Einfluss von Alkohol eine kürzere Dauer bei der Krafteinwirkung dargeboten. Der Wert für die standardisierte Teststatistik liegt mit $z = ,671$ unter dem kritischen Wert von 1,96 und zeigt sich somit nicht signifikant. Die Nullhypothese wird für diese Metrik beibehalten. Oben beschriebene Ergebnisse sind in den Tabellen 26 und 27 dargestellt.

Tabelle 26: Häufigkeiten der Differenzen von „Excessive instrument force“ der Übung Suture Sponge

		N
SS_Excessive_Instrument_Force_(sec)_1	Negative Differenzen ^a	8
(sec)_1 -	Positive Differenzen ^b	12
SS_Excessive_Instrument_Force_(sec)_0	Bindungen ^c	0
(sec)_0	Gesamt	20

a. $SS_Excessive_Instrument_Force_1 < SS_Excessive_Instrument_Force_0$

b. $SS_Excessive_Instrument_Force_1 > SS_Excessive_Instrument_Force_0$

c. $SS_Excessive_Instrument_Force_1 = SS_Excessive_Instrument_Force_0$

Tabelle 27: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Excessive instrument force“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	12,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	,671
Asymptotische Sig. (zweiseitiger Test)	,502
Genauere Sig. (zweiseitiger Test)	,503

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2.5 Instruments out of view

Anhand Abbildung 26 wird ersichtlich, dass die Ergebnisse der Kategorie Instruments out of view bis auf zwei Ausreißer nah beieinander liegen. Bei den Ausreißern handelt es sich um Proband 9, welcher die Instrumente unter der Intervention eine längere Strecke außerhalb der Sichtweite bewegt hat und um Proband 1, der die Instrumente unter der Intervention eine kürzere Strecke außerhalb der Sichtweite bewegt hat.

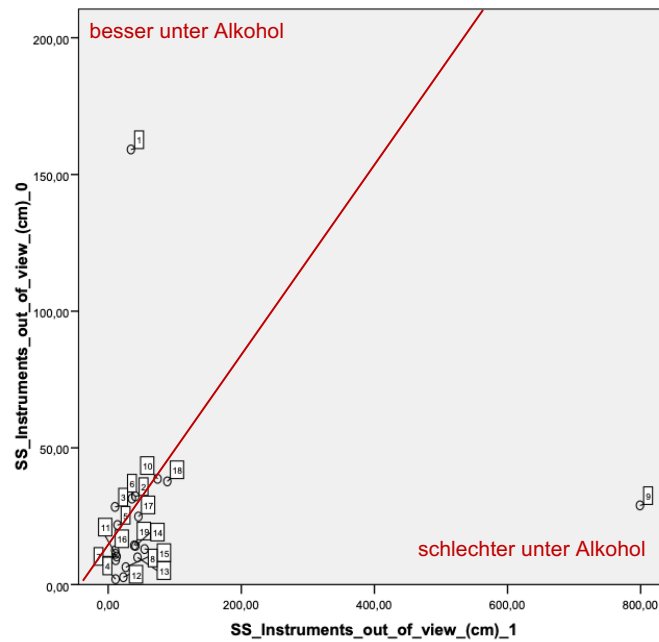


Abbildung 26. Streudiagramm zu „Instruments out of view“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

Der Vorzeichentest zu „Instruments out of view“ der Übung Suture Sponge, dessen Ergebnisse in den Tabellen 28 und 29 zu sehen sind, ergab, dass insgesamt 15 Probanden unter der Intervention eine längere Strecke mit den Instrumenten außerhalb der Sichtweite zurücklegten und 5 Probanden darunter eine kürzere Strecke außerhalb ihrer Sichtweite zurücklegten. Bindungen, also kein Leistungsunterschied vor oder unter der Intervention ergaben sich hier nicht. Bei einer Teststatistik von $z = 2,012$ zeigt sich der Test signifikant. Die Intervention Alkoholeinfluss hat in diesem Fall zu einer längeren Strecke der Instrumente außerhalb des Sichtfeldes und somit zu einer schlechteren Performance geführt.

Tabelle 28: Häufigkeiten der Differenzen von „Instruments out of view“ der Übung Suture Sponge

		N
SS_Instruments_out_of_view	Negative Differenzen ^a	5
_(cm)_1 -	Positive Differenzen ^b	15
SS_Instruments_out_of_view	Bindungen ^c	0
_(cm)_0	Gesamt	20

a. $SS_Instruments_out_of_view_ (cm)_1 < SS_Instruments_out_of_view_ (cm)_0$

b. $SS_Instruments_out_of_view_ (cm)_1 > SS_Instruments_out_of_view_ (cm)_0$

c. $SS_Instruments_out_of_view_ (cm)_1 = SS_Instruments_out_of_view_ (cm)_0$

Tabelle 29: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Instruments out of view“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	15,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	2,012
Asymptotische Sig. (zweiseitiger Test)	,044
Genauere Sig. (zweiseitiger Test)	,041

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2.6 Master workspace range

Anhand der Verteilung in Abbildung 27 erkennt man die Streuung der Leistungen vor und unter Intervention der Rubrik „Master workspace range“. Proband 1 kann als Ausreißer definiert werden. Dieser benötigte unter der Intervention ein geringeres Ausmaß an Arbeitsumfang, gemessen in Zentimetern. Ein weiterer Ausreißer stellt Proband 9 dar, der unter der Intervention ein höheres Ausmaß an Arbeitsumfang benötigte und somit eine schlechtere Leistung geboten hat.

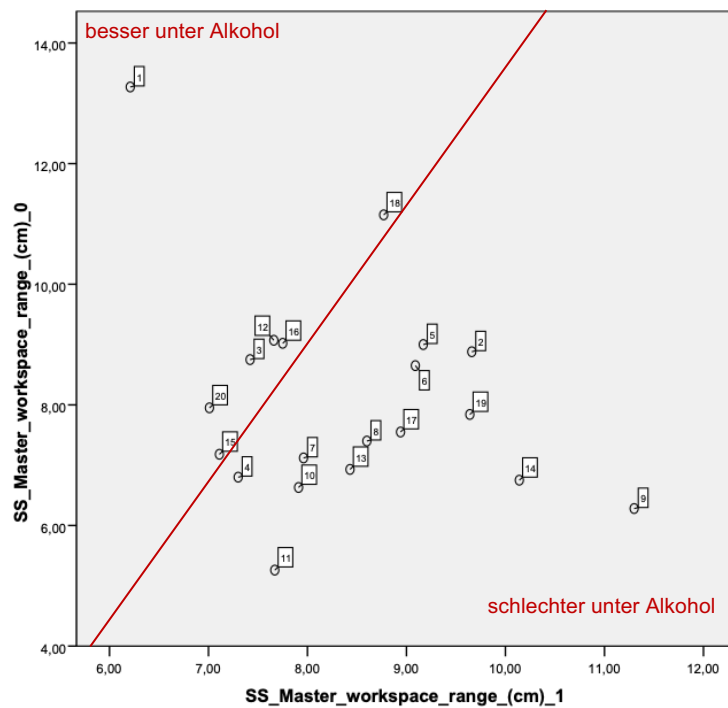


Abbildung 27. Streudiagramm zu „Master workspace range“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

In der Kategorie Master workspace range haben 13 Probanden unter der Intervention ein größeres Ausmaß an Arbeitsumfang benötigt und 7 Probanden haben darunter ein geringeres Ausmaß benötigt. Der Wert für die standardisierte Teststatistik liegt mit $z = 1,118$ unterhalb des kritischen Wertes von 1,96 und zeigt sich somit nicht signifikant. Die Nullhypothese muss hier beibehalten werden. Die beschriebenen Ergebnisse sind in den Tabellen 30 und 31 dargestellt.

Tabelle 30: Häufigkeiten der Differenzen von „Master workspace range“ der Übung Suture Sponge

		N
SS_Master_workspace_range	Negative Differenzen ^a	7
_(cm)_1 -	Positive Differenzen ^b	13
SS_Master_workspace_range	Bindungen ^c	0
_(cm)_0	Gesamt	20

a. $SS_Master_workspace_range_ (cm)_1 < SS_Master_workspace_range_ (cm)_0$

b. $SS_Master_workspace_range_ (cm)_1 > SS_Master_workspace_range_ (cm)_0$

c. $SS_Master_workspace_range_ (cm)_1 = SS_Master_workspace_range_ (cm)_0$

Tabelle 31: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Master workspace range“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	13,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	1,118
Asymptotische Sig. (zweiseitiger Test)	,264
Genauere Sig. (zweiseitiger Test)	,263

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2.7 Drops

Die Abbildung 28 zeigt das Streudiagramm der Metrik „Drops“. Es ist zu sehen, dass die meisten der Probanden unter der Intervention mindestens 1 und maximal 9 „Drops“ mehr aufweisen.

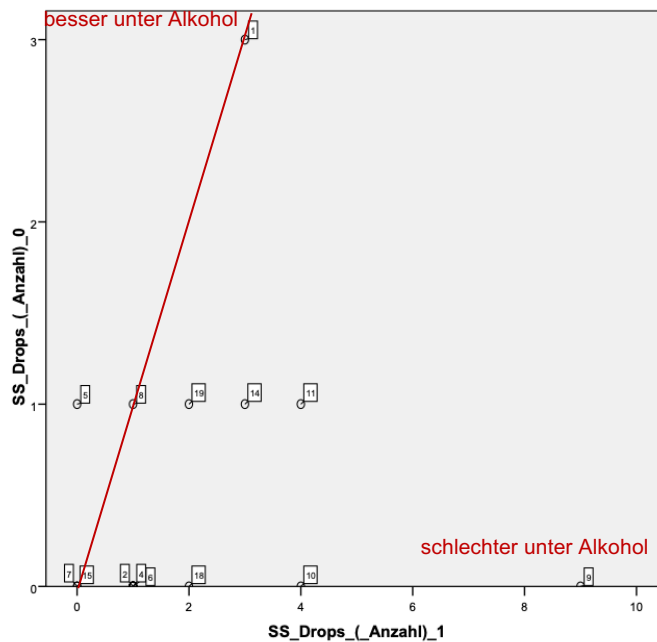


Abbildung 28. Streudiagramm zu „Drops“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

In den folgenden Tabellen 32 und 33 ist zu sehen, dass 15 Versuchspersonen unter der Intervention häufiger „Drops“ zeigten. Ein Versuchsteilnehmer hatte im alkoholisierten Zustand seltener „Drops“. Vier Probanden haben keine Performance-Unterschiede gehabt. Der z-Wert liegt hier bei 3,250 und somit über dem kritischen Wert von 1,96. Dieser Test ist demnach signifikant. Die Intervention hat in diesem Fall signifikant häufiger dazu geführt, dass die Probanden die Nadel fallen lassen haben.

Tabelle 32: Häufigkeiten der Differenzen von „Drops“ der Übung Suture Sponge

		N
SS_Drops_(Anzahl)_1 -	Negative Differenzen ^a	1
SS_Drops_(Anzahl)_0	Positive Differenzen ^b	15
	Bindungen ^c	4
	Gesamt	20

a. $SS_Drops_ (Anzahl)_1 < SS_Drops_ (Anzahl)_0$

b. $SS_Drops_ (Anzahl)_1 > SS_Drops_ (Anzahl)_0$

c. $SS_Drops_ (Anzahl)_1 = SS_Drops_ (Anzahl)_0$

Tabelle 33: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Drops“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	15,000 ^a
Standardfehler	2,000
Standardisierte Teststatistik	3,250
Asymptotische Sig. (zweiseitiger Test)	,001
Genauere Sig. (zweiseitiger Test)	,001

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.2.8 Missed Targets

Das Streudiagramm zu „Missed targets“ ist in Abbildung 29 dargestellt. Die Streuung der Messwerte ist in dieser Metrik im Vergleich zu den anderen Metriken ausgeprägter. Man sieht, dass die Mehrheit der Probanden unter der Intervention eine höhere Anzahl an verpassten Zielen gezeigt hat. Zu erwähnen sind als Ausreißer Proband 1, welcher eine deutlich höhere Anzahl verpasster Ziele ohne den Einfluss von Alkohol gezeigt hat und Proband 9, welcher unter der Intervention eine deutlich höhere Anzahl verpasster Ziele gezeigt hat.

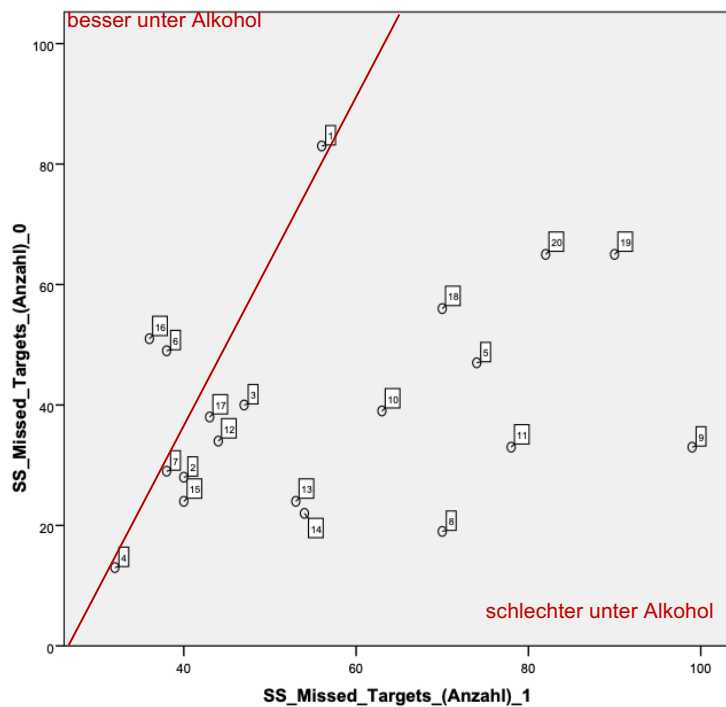


Abbildung 29. Streudiagramm zu „Missed targets“ der Übung Suture Sponge mit (_1) und ohne (_0) Intervention.

Anhand der Tabelle 34 sieht man, dass 17 Probanden unter Einsatz von Alkohol häufiger das Ziel verfehlt haben. Bei 3 Probanden zeigte sich alkoholisiert seltener ein „Missed target“. Bei einem z-Wert von 2,907 ist das Ergebnis signifikant (s. Tabelle 35). In dieser Metrik konnte also nachgewiesen werden, dass die Alkoholisierung zu einer signifikant höheren Anzahl an verfehlten Zielen geführt hat.

Tabelle 34: Häufigkeiten der Differenzen von „Missed targets“ der Übung Suture Sponge

	N
SS_Missed_Targets_(Anzahl)_1 - Negative Differenzen ^a	3
SS_Missed_Targets_(Anzahl)_0 Positive Differenzen ^b	17
Bindungen ^c	0
Gesamt	20

a. $SS_Missed_Targets_Anzahl_1 < SS_Missed_Targets_Anzahl_0$

b. $SS_Missed_Targets_Anzahl_1 > SS_Missed_Targets_Anzahl_0$

c. $SS_Missed_Targets_Anzahl_1 = SS_Missed_Targets_Anzahl_0$

Tabelle 35: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert von „Missed targets“ der Übung Suture Sponge

Gesamtzahl	20
Teststatistik	17,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	2,907
Asymptotische Sig. (zweiseitiger Test)	,004
Genauere Sig. (zweiseitiger Test)	,003

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.3 Stacking Challenge

In Tabelle 36 sind Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken von der Übung Stacking Challenge zu sehen. Bei der Übung Stacking Challenge gibt es nur den Gesamtscore als Auswertungsparameter. Im Gegensatz zu den vorherigen Auswertungsparametern von Ring Tower und Suture Sponge gilt hier, je höher der Score ist, desto besser war die Performance des Probanden. Man erkennt, dass die Mittelwerte der Ergebnisse ohne Intervention und mit Intervention kaum divergent sind. Betrachtet man den Minimal- und Maximalwert der beiden Interventionsgruppen ist zu sehen, dass die Gesamtheit der Messwerte ohne Intervention einen höheren Minimal- als auch Maximalwert erreicht hat. Die Variationskoeffizienten liegen in beiden Gruppen unter 1 und die Standardabweichungen liegen unterhalb des Mittelwertes. Somit zeigt sich bei den Ergebnissen beider Interventionsgruppen eine kleine Streubreite.

Tabelle 36: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der Übung Stacking Challenge

Metrik	N	Minimum	Maximum	Mittelwert	Standardabweichung	Variationskoeffizient
ST_Overall_Score_(Punkte)_0	20	20,02	90,59	57,65	15,504	0,268
ST_Overall_Score_(Punkte)_1	20	0,63	77,06	57,85	15,281	0,264

Dass die Messwerte in beiden Gruppen keine große Streubreite aufweisen, ist in dem Diagramm der Abbildung 30 zu sehen. Als Ausreißer erkennt man die Probanden 1 und 20, welche unter der Intervention eine höhere Punktzahl erreicht haben. Außerdem Proband 13, der ohne Alkoholeinfluss eine größere Punktzahl erlangt hat.

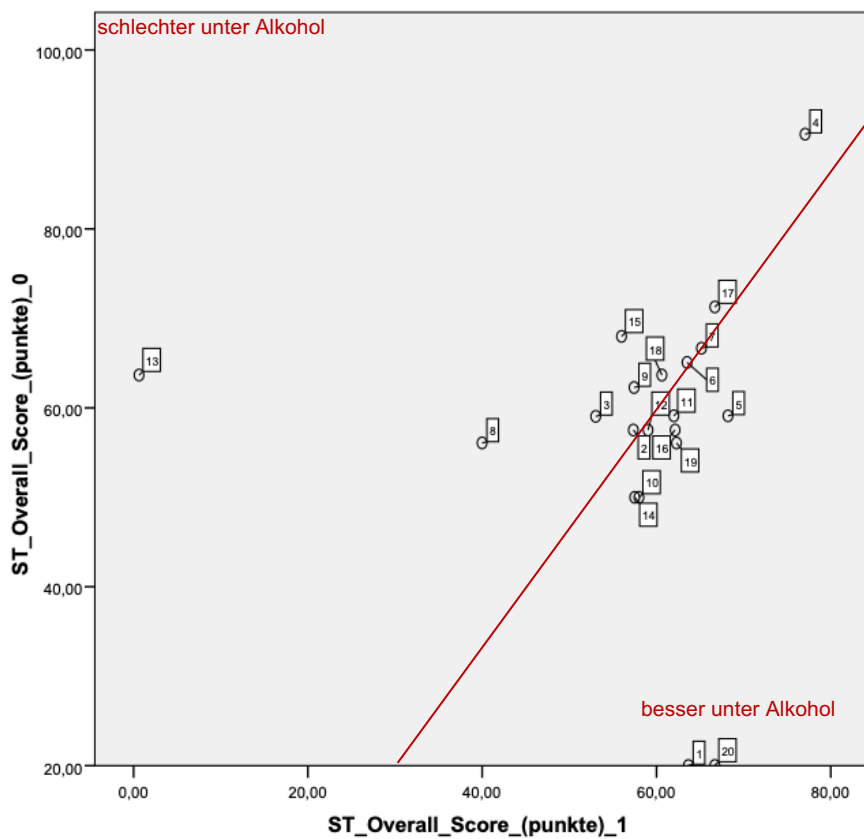


Abbildung 30. Streudiagramm zu der Übung Stacking Challenge mit (_1) und ohne (_0) Intervention.

In untenstehender Tabelle 37 erkennt man, dass insgesamt 9 Probanden unter der Intervention eine höhere Punktzahl erreicht haben. 11 Probanden haben ohne Alkohol eine höhere Punktzahl erlangen können. Bei einer standardisierten Teststatistik von $z = -,224$ liegt bei diesem Test keine statistische Signifikanz vor (s. Tabelle 38).

Tabelle 37: Häufigkeiten der Differenzen der Übung Stacking Challenge

	N
ST_Overall_Score_(punkte)_1 - Negative Differenzen ^a	11
ST_Overall_Score_(punkte)_0 Positive Differenzen ^b	9
Bindungen ^c	0
Gesamt	20

a. $ST_Overall_Score_(\text{punkte})_1 < ST_Overall_Score_(\text{punkte})_0$

b. $ST_Overall_Score_(\text{punkte})_1 > ST_Overall_Score_(\text{punkte})_0$

c. $ST_Overall_Score_(\text{punkte})_1 = ST_Overall_Score_(\text{punkte})_0$

Tabelle 38: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert der Übung Stacking Challenge

Gesamtzahl	20
Teststatistik	9,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	-,224
Asymptotische Sig. (zweiseitiger Test)	,823
Genauere Sig. (zweiseitiger Test)	,824

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.1.4 Vitruvian Operation

In Tabelle 39 sind Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der einzelnen Metriken von der Übung Vitruvian Operation zu sehen. Bei dieser Übung gibt es wie bei der Übung Stacking Challenge nur den Gesamtscore als Auswertungsparameter. Hier gilt also ebenso, je höher der Score ist, desto besser war die Performance des Probanden. Man erkennt, dass die Mittelwerte der Ergebnisse ohne Intervention und mit Intervention kaum divergent sind. Auch die Minimal- und Maximalwerte weichen kaum voneinander ab. Die Variationskoeffizienten liegen in beiden Gruppen unter 1 und die Standardabweichungen liegen unterhalb des Mittelwertes. Somit zeigt sich bei den Ergebnissen beider Interventionsgruppen eine kleine Streubreite.

Tabelle 39: Minima, Maxima, Mittelwerte, Standardabweichungen und Variationskoeffizienten der Übung Vitruvian Operation

Metrik	N	Minimum	Maximum	Mittelwert	Standardabweichung	Variationskoeffizient
VI_Overall_Score_(Punkte)_0	20	6,00	48,00	31,85	10,095	0,316
VI_Overall_Score_(Punkte)_1	20	0,00	49,00	27,30	11,416	0,418

In Abbildung 31 sieht man die Streuung der einzelnen Messwerte von Vitruvian Operation. Die Messwerte weisen keine große Streubreite auf.

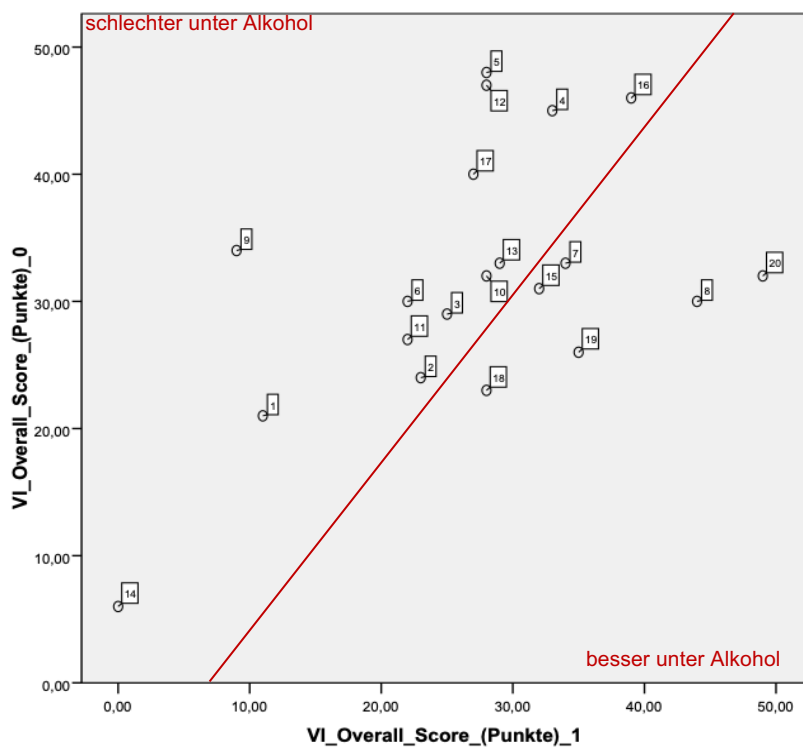


Abbildung 31. Streudiagramm zu der Übung Vitruvian Operation mit (_1) und ohne (_0) Intervention.

Die Ergebnisse des Vorzeichentest zu Vitruvian Operation ist in Tabelle 40 zu sehen. Daraus lässt sich ableiten, dass 14 Probanden ohne Alkoholeinfluss eine höhere Punktzahl erreicht haben als mit Alkohol. Die restlichen 6 Probanden hatten unter Alkoholeinfluss die höhere Punktzahl. Eine Signifikanz liegt bei dieser Testung bei einem z-Wert von -,565 nicht vor (s. Tabelle 41).

Tabelle 40: Häufigkeiten der Differenzen der Übung Vitruvian Operation

		N
VI_Overall_Score_(Punkte)_1 -	Negative Differenzen ^a	14
VI_Overall_Score_(Punkte)_0	Positive Differenzen ^b	6
	Bindungen ^c	0
	Gesamt	20

a. $VI_Overall_Score_ (Punkte)_1 < VI_Overall_Score_ (Punkte)_0$

b. $VI_Overall_Score_ (Punkte)_1 > VI_Overall_Score_ (Punkte)_0$

c. $VI_Overall_Score_ (Punkte)_1 = VI_Overall_Score_ (Punkte)_0$

Tabelle 41: Zusammenfassung des Vorzeichentest mit standardisiertem z-Wert der Übung Vitruvian Operation

Gesamtzahl	20
Teststatistik	6,000 ^a
Standardfehler	2,236
Standardisierte Teststatistik	-1,565
Asymptotische Sig. (zweiseitiger Test)	,118
Genauere Sig. (zweiseitiger Test)	,115

a. Der exakte p-Wert wird anhand der binomialen Verteilung berechnet, weil 25 oder weniger Fälle vorhanden sind.

4.2 Tabellarische Zusammenfassung der Ergebnisse

In untenstehender Tabelle 42 sind zusammenfassend die Metriken mit signifikanter Verschlechterung unter Alkoholeinfluss den Metriken ohne signifikante Verschlechterung gegenübergestellt.

Tabelle 42: Darstellung der Metriken mit signifikanter sowie ohne signifikante Verschlechterung

Metriken mit signifikanter Verschlechterung	Metriken ohne signifikante Verschlechterung
Instruments out of view (RT)	Time to complete exercise (RT)
Instrument collisions (SS)	Instrument collisions (RT)
Instruments out of view (SS)	Wire collisions (RT)
Drops (SS)	Wire contact duration (RT)
Missed Targets (SS)	Economy of motion (RT)
	Drops (RT)
	Tower Knockoffs (RT)
	Time to complete Exercise (SS)
	Excessive Instrument force (SS)
	Economy of motion (SS)
	Master Workspace Range (SS)
	Vitruvian Operation (Gesamtscore)
	Stacking Challenge (Gesamtscore)

4.3 Sekundäre Messparameter

4.3.1 Blutalkoholkonzentration

Die unter der Testung vorhandene Blutalkoholkonzentration der einzelnen Probanden wird in Tabelle 43 dargestellt. Tabelle 44 zeigt Minimal- und Maximalwerte, Mittelwert und Standardabweichung der ermittelten Promillewerte.

Tabelle 43: Blutalkoholkonzentration aller Probanden

Proband	Blutalkoholkonzentration
18001	0,67‰
18002	0,63‰
18003	0,74‰
18004	0,91‰
18005	0,51‰
18006	0,68‰
18007	0,82‰
18008	0,61‰
18009	0,53‰
18010	0,89‰
18011	0,6‰
18012	0,49‰
18013	0,48‰
18014	0,5‰
18015	0,7‰
18016	0,55‰
18017	0,53‰
18018	0,62‰
18019	0,79‰
18020	0,7‰

Tabelle 44: Deskriptive Statistik zu Blutalkoholkonzentration

	N	Minimum	Maximum	Mittelwert	Standardabweichung
BAK	20	,48‰	,91‰	,6475‰	,13154‰
Gültige Werte (Listenweise)	20				

4.3.2 Trinkgewohnheiten

Anhand eines Fragebogens wurden die Trinkgewohnheiten der Probanden ermittelt. Die Ergebnisse und deren Erläuterung sind in Tabelle 45 aufgeführt. Abbildung 32 stellt die Häufigkeitsverteilung der Trinkgewohnheiten dar. Man sieht, dass die meisten Probanden angegeben haben, zwei- bis viermal im Monat Alkohol zu konsumieren. Ein Konsum von viermal in der Woche hat kein Proband angegeben.

Tabelle 45: Auflistung zu den Trinkgewohnheiten der Probanden

Proband	Trinkgewohnheit
18001	1
18002	2
18003	2
18004	1
18005	3
18006	2
18007	2
18008	1
18009	0
18010	3
18011	3
18012	3
18013	0
18014	3
18015	3
18016	2
18017	2
18018	1
18019	2
18020	2

Alkoholkonsum: 0 = nie; 1 = 1x/Monat; 2 = 2-4x/Monat; 3 = 2-3x/Woche; 4 = 4x/Woche

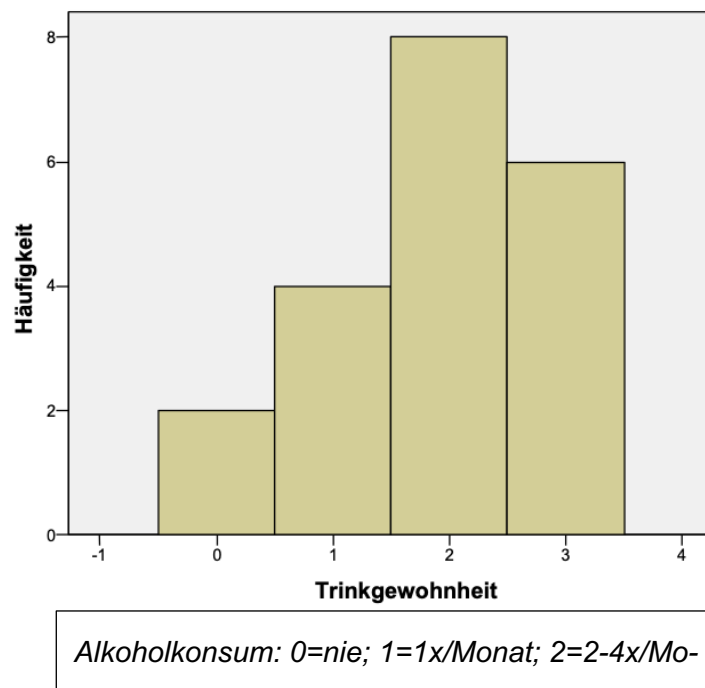


Abbildung 32. Häufigkeitsverteilung zur Trinkgewohnheit.

5 Diskussion

Diese Arbeit ist die erste Arbeit, die einen Zusammenhang zwischen Alkoholkonsum und der operativen Performance an einem Roboter-Simulator systematisch untersucht. Die Arbeit behandelt in einem Zeitalter von Technik, Digitalisierung und steigender Relevanz der Roboter-Chirurgie ein sehr präsenten Thema.

5.1 Zusammenfassung und Interpretation der Ergebnisse

Betrachtet man die Werteverteilung der Streudiagramme, ist in den meisten Fällen eine lineare Regression erkennbar. Bei einigen Metriken zeigt sich also ein Zusammenhang, im Sinne von ähnlichen Leistungen vor und unter der Intervention. Es gibt jedoch auch Diagramme, die sich eher inhomogen und mit mehreren Ausreißern präsentieren. Hier zeigte sich die Streubreite ausgeprägter. Dazu zählen unter anderem die Streudiagramme zu den Metriken „Wire contact duration“ (RT), „Drops“ (RT), „Tower knockoff“ (RT), „Instrument collisions“ (RT), „Drops“ (SS). Um jedoch genau untersuchen zu können, ob sich die Leistung der Probanden unter der Intervention verbessert oder verschlechtert haben oder ob eine gleichbleibende Leistung dargeboten wurde, erfolgte die Analyse mit einem Vorzeichentest. Hier war zu sehen, dass es in der Gesamtheit zu einer Verschlechterung unter der Intervention gekommen ist. Es gab keine Metrik, in der eine Verbesserung unter der Intervention überwogen hat. Insbesondere in der Übung Suture Sponge kam es in den Metriken „Drops“, „Instrument collisions“, „Instruments out of view“, und „Missed targets“ zu einer signifikanten Verschlechterung unter Alkoholeinfluss. Die Übung Suture Sponge benötigt unter anderem ein hohes Maß an Präzision. Hier konnte also nachgewiesen werden, dass die Präzision durch den Einfluss von Alkohol maßgeblich negativ beeinflusst wurde. Auch in der Metrik „Instruments out of view“ der Übung Ring Tower wurde eine signifikante Verschlechterung nachgewiesen. In zahlreichen weiteren Metriken war ebenfalls eine tendenzielle Verschlechterung unter der Intervention zu sehen, welche jedoch keine Signifikanz nachwies (z-Wert >1,96). Dazu zählen „Wire collisions“ Ring Tower (11 verschlechtert/ 6 verbessert), „Economy of motion“ Ring Tower (12 verschlechtert/ 5 verbessert), „Wire contact duration“ Ring Tower (11 verschlechtert/ 6 verbessert), „Drops“ Ring Tower (11 verschlechtert/ 4 verbessert/ 2 gleichbleibend), „Master workspace range“ Suture Sponge (13 verschlechtert/ 7 verbessert), „Time to complete“ Suture Sponge (14 verschlechtert/ 6 verbessert), „Economy of motion“ Suture Sponge (14 verschlechtert/ 6 verbessert) und Vitruvian Operation (14 verschlechtert/ 6 verbessert). Die Ergebnisse der Übung Stacking Challenge sind insofern interessant, als dass sich hier in etwa ein Gleichgewicht von Verschlechterung (11 Probanden) und Verbesserung (9 Probanden) unter der Intervention zeigt.

5.2 Vergleich mit bestehender Literatur

Untersuchungen zu Alkohol und dessen Auswirkung, nicht zuletzt auch auf chirurgische Fähigkeiten sind schon länger Gegenstand wissenschaftlicher Forschung. Im Allgemeinen repräsentieren die Ergebnisse der in dieser Arbeit dargestellten Studien durchweg eine Abnahme der kognitiven und psychomotorischen Leistungen durch den Einfluss von Alkohol akut bzw. in einem zeitlichen Zusammenhang. Doch bis heute gibt es keine einheitliche gesetzliche Regelung über ein Verbot von Alkoholkonsum während, unmittelbar vor oder in einem gewissen zeitlichen Zusammenhang zum Dienstbeginn, wenngleich dies dem Berufsethos des Arztes in Anlehnung an den hippokratischen Eid widerspricht (Güntert, 2018).

Die Untersuchungen, die Kirby et al. (2012) veröffentlichten, sind dieser Studie vom methodischen Vorgehen her sehr ähnlich, jedoch fanden die Untersuchungen nicht an einem Roboter-Simulator statt. Es wurde ebenfalls eine gewisse Menge Alkohol verabreicht, jedoch gemessen in konsumierten Gläsern Wein und nicht in einer genauen Grammangabe um eine gewisse BAK möglichst genau zu erreichen. Kirby et al. argumentierten, dass diese Maßeinheit in der Realität eher eine Projektion auf die eigene Person ermöglicht, als eine festgelegte BAK. Demgegenüber sollte mithilfe der Untersuchung am Roboter-Simulator einerseits überprüft werden, ob Alkoholkonsum Einfluss auf die chirurgische Performance hat. Andererseits könnten mithilfe der genauen Ermittlung der BAK ergänzend Aussagen darüber getroffen werden, inwieweit der Grad der Alkoholisierung eine bessere oder schlechtere Leistung suggeriert. Bei den Studien, die Dorafshar et al. (2002), Kocher et al. (2006) und Van Dyken et al. (2013) publizierten, konnte mithilfe einer Atemalkoholmessung lediglich eine geschätzte BAK (wenn nach gesetzlicher Regelung für eine Atemalkoholkonzentration von 0,25 mg/l eine BAK von 0,5 Promille definiert wurde) objektiviert werden. In der dieser Arbeit zugrunde liegenden Studie wurde die BAK mittels Blutentnahme unmittelbar nach Beendigung der Übungen und anschließender laborchemischer Aufbereitung bestimmt. Diese Methode wurde bisher in keiner der betrachteten Studien verwendet. Gallagher et al. (2011) hingegen forderten im Rahmen ihrer Testung keine vorgegebene BAK oder Trinkmenge/Glasmenge. Die Probanden sollten die Menge an Alkohol zu sich nehmen, ab der sie sich subjektiv betrunken fühlten, was sicherlich ein realitätsnahes und denkbares Szenario darstellt. Um einen Zusammenhang oder eine Wechselbeziehung des Alkoholausmaßes und der Leistung treffen zu können, erweist sich dies jedoch als keine geeignete Methode.

Bei den Ausführungen der Studie von Dorafshar et al. (2002) wurden 28 Probanden zufällig einer Versuchs- oder Kontrollgruppe zugeordnet. Eine Randomisierung wie es hier erfolgt ist

soll verhindern, dass „berücksichtigte Einflussvariablen mit gegebenenfalls nicht berücksichtigten wichtigen Einflüssen korrelieren“ (Stocker & Steinke, 2016, S. 194), um somit trotz der Störgrößen eine höhere statistische Sicherheit zu erlangen. Bei der Arbeit von Kocher et al. als auch bei der Veröffentlichung von Kirby et al. (2012) wurden entweder fünf oder vier Chirurgen getestet. Ist ein Ergebnis oder ein Zusammenhang statistisch signifikant, kann man aus wahrscheinlichkeitstheoretischer Sicht davon ausgehen, dass es sich hier um einen nicht zufällig entstandenen Effekt handelt (Stocker & Steinke, 2016). Bei einer vergleichsweise kleinen Stichprobe wie sie bei Kocher et al. (2006) und bei Kirby et al. (2012) gewählt wurde, ist die Wahrscheinlichkeit, einen zufälligen Effekt aufzuzeigen höher, als in einer größeren Stichprobe. Bei diesen beiden Studien wurden keine Aussagen über eine Signifikanz der Ergebnisse getroffen. In dieser Studie wurden, ähnlich zu Van Dyken et al. (2013) 20 Chirurgen bzw. sich in der chirurgischen Weiterbildung befindliche Ärzte als Kontroll- und als Versuchsgruppe betrachtet. Anhand dieser relativ kleinen Stichprobe mit unterschiedlichen in Abhängigkeit des Ausbildungsstandes operativen Fähigkeiten, sollte das Ergebnis somit hauptsächlich auf die Wirkung des Alkohols zurückzuführen sein. Wie bei Kocher et al. (2006) wurden auch unsere Probanden in ausreichend zeitlichem Abstand zum Studientag an dem Simulator geschult, um somit einen möglicherweise das Ergebnis beeinflussenden Lerneffekt weitestgehend verhindern zu können. Eine weitere Parallele zu dem Studiendesign der Studie am Roboter-Simulator lässt sich zu der Arbeit von Dorafshar et al. (2002) ziehen. Hier wurde eine Verblindung der Studienteilnehmer aufgrund der allgemein in der Bevölkerung bekannten physiologischen und verhaltensbezogenen Effekte von Alkohol als ungeeignet betrachtet. Van Dyken et al. (2013) und Gallagher et al. (2011) konzentrierten sich bei ihrem Versuch vor allem auf die Auswirkungen des Alkohols am Morgen nach dem Konsum. Da alkoholranke Menschen je nach Stadium der Erkrankung häufig versuchen, die Sucht weitestgehend zu verbergen (Challenge-A, 2021), würde der Konsum in den meisten Fällen vermutlich außerhalb der Dienstzeiten stattfinden. Die Überprüfung der chirurgischen Fähigkeit am Tag nach einer Alkoholeinnahme präsentiert somit eine Methode, die in hohem Maß mit der Realität vereinbar wäre. Bei den Untersuchungen von Kirby et al. (2012) wurde die Bearbeitung und das Ergebnis der HNO-Operation am Modell von zwei Chirurgen, die diesem Fachgebiet zugeteilt sind, bewertet. Das bedeutet, hier wurde kein objektiv validiertes Scoringssystem verwendet, sondern eine auf individuellen Erfahrungen und Kompetenzen basierende chirurgische Einschätzung. Jedoch konnte anhand dieser Bewertungsmethode eine ergänzende Aussage über Interaktion und Kommunikation der Probanden mit dem Team getroffen werden. Eine durch Alkohol negativ beeinflusste Kommunikation und Teamarbeit, wie sie bei Kirby et al (2012) beobachtet wurde, repräsentiert einen Risikofaktor für die Patientensicherheit.

Der Studie am Roboter-Simulator standen insgesamt 18 Bewertungsparameter der vier Übungen zur Verfügung. In 16 dieser Metriken war eine Verschlechterung unter akutem Alkoholeinfluss zu verzeichnen. Eine statistische Signifikanz konnte hier bei „Instruments out of view“ (Ring Tower), „Drops“ (Suture Sponge), „Instrument collisions“ (Suture Sponge), „Instruments out of view“ (Suture Sponge) und „Missed targets“ (Suture Sponge) nachgewiesen werden. Diese Metriken geben unter anderem Hinweise auf die Geschicklichkeit und Aufmerksamkeit eines Probanden. Ein Bewertungsparameter, den Kocher et al. (2006), Dorafshar et al. (2002) und Gallagher et al. (2011) in ihren Versuchen betrachteten, war unter anderem die Diathermienutzung. Bei Operationen wird Strom häufig genutzt, um Gewebe zu durchtrennen bzw. um ein Gefäß zu veröden. Diese Anwendung wurde bei den genannten Studien näher betrachtet und stellt verglichen mit unserer Studie ebenfalls ein Indikator für die Aufmerksamkeit und die Geschicklichkeit des Probanden dar. Die Auswertung aller drei Studien zeigte, dass es unter Alkoholeinfluss beziehungsweise in einem zeitlichen Zusammenhang zum Alkoholkonsum zu einer Verlängerung der Diathermienutzung, als auch zu fehlerhafter Nutzung gekommen ist. Bei zwei (Dorafshar, O'Boyle, & McCloy, 2002), (Gallagher, et al., 2011) von drei Studien zeigte sich die Auswertung zur Diathermienutzung, wo sich die Leistung unter Alkoholeinfluss bzw. in einem zeitlichen Zusammenhang zum Alkoholenuss verschlechterte, statistisch signifikant. Zusätzlich wurde die Fehlerzahl untersucht. Ebenso wie die Diathermienutzung kann auch die Fehlerzahl als Maßstab für die Aufmerksamkeit und Geschicklichkeit einer Versuchsperson betrachtet werden. Bei Kocher et al. (2006) zeigte sich diese am Morgen nach dem Konsum geringer als im Zustand der akuten Alkoholisierung. Nach Annahme der Autoren lag dieser Effekt an einer erhöhten Wachsamkeit durch eine bereits geringe Schlafdauer. Hier sollten jedoch auch die Symptome einer akuten Alkoholisierung, wie unter anderem eine veränderte Sinneswahrnehmung, als ursächlich für eine höhere Fehleranzahl mitbedacht werden (Mußhoff, Madea, & Heß, 2014). Bei Gallagher et al. (2011) war bei der Fehleranzahl ebenfalls eine Verschlechterung zu sehen, jedoch zeigte sich nur der Unterschied der Messung um 13 Uhr nachweisbar signifikant. Da hier mehrere Messzeitpunkte betrachtet wurden, kann ein Ermüdungszustand als das Ergebnis beeinflussende Variable nicht ausgeschlossen werden. Ein weiterer Bewertungsparameter, den unter anderem Gallagher et al. (2011) betrachteten, stellte die Bearbeitungszeit dar. Diese ergab eine signifikante Verlängerung unter dem Einfluss von Alkohol. Jedoch wurde neben der Variable Alkohol auch untersucht, ob sich zu verschiedenen Messzeitpunkten nach dem Alkoholkonsum ein bestimmter Effekt zeigt. Man sah, dass die Übung um 9 Uhr am Morgen nach der Alkoholeinnahme sogar signifikant schneller bewältigt wurde als zum Ausgangsmesszeitpunkt. Mithilfe der Studie am

MIMIC dV-Trainer® wurde ausschließlich die Leistung unter akutem Alkoholeinfluss untersucht. In Bezug auf die Zeit konnte beim Betrachten der gewonnenen Ergebnisse bei der entsprechenden Metrik der Übung Suture Sponge eine Tendenz zu einer länger benötigten Bearbeitungszeit unter der Intervention detektiert werden. Bei der Übung Ring Tower hingegen ist die Mehrzahl (neun) der Probanden sogar besser geworden. Es sollte allerdings bedacht werden, dass eine längere Bearbeitungszeit auch qualitative Einbußen im Sinne einer geringeren Präzision und einer höheren Fehlerrate implizieren kann.

Bei den Übungen Vitruvian Operation als auch Stacking Challenge musste die Leistung innerhalb einer vorgegebenen Zeit erbracht werden. Die damit geschaffene Situation kann man in gewissem Maße mit einer notfallmäßigen oder einer nicht planmäßig verlaufenden Operation vergleichen. Auch hier sollte in verhältnismäßig kurzer Zeit eine möglichst präzise Arbeit vollbracht werden. Die Erwartungen an den Chirurgen können in solch einer Situation ergänzend als positiver jedoch auch als negativer Stressor dienen. Die Leistung innerhalb einer Drucksituation wurde unter anderem bei der Arbeit von Van Dyken et al. (2013) bewertet. Hier wurden drei Kriterien als Grundlagen der laparoskopischen Chirurgie unter einer zeitlichen Vorgabe abgefragt. Eine Signifikanz der zeitlichen Differenz zeigte sich jedoch nicht. Ebenfalls vergleichbar mit unseren Erkenntnissen waren hier die Ergebnisse des intrakorporalen Nähens. Es zeigten sich signifikante Unterschiede in der Genauigkeit des Nähens am Morgen nach dem Alkoholkonsum. In Anlehnung an unsere Auswertung zeigten sich in der Übung Suture Sponge, welche die Anforderungen eines präzisen Nahtvorganges simuliert, zahlreiche signifikante Differenzen. Ein Unterschied, der mit hoher Wahrscheinlichkeit nicht zufällig vorliegt, konnte hier bei den Metriken „Drops“, „Instrument collisions“, „Instruments out of view“ und „Missed targets“ nachgewiesen werden. Diese Parameter können unter anderem durch die bereits beschriebene visuelle Einschränkung durch Alkoholkonsum mit konsekutiver Wahrnehmungsstörung von bewegten Objekten und seitlicher Sinneseindrücke bedingt sein (Mußhoff, Madea, & Heß, 2014). Auch ein Verlust der Muskelkoordination zählt zu den Folgen von akutem Alkoholkonsum und kann sich in einem Verlust der Feinmotorik bemerkbar machen (Mußhoff, Madea, & Heß, 2014), was mitunter eine höhere Anzahl von verfehlten Zielen verursachen kann. Auch die von Weiss et al. (2013) beschriebenen durch Alkohol induzierten Dysfunktionen wie Störungen der motorischen Regulation und einer Beeinträchtigung der Aufmerksamkeit kann diese signifikante Verschlechterung der beschriebenen Parameter entscheidend beeinflussen.

Anhand der Auswertung des Vorzeichentests zu der Übung Stacking Challenge sieht man, dass sich etwa genauso viele Probanden unter der Intervention verschlechtert als auch verbessert haben. Bei dieser Übung wird verlangt, dass man unter Zeitdruck einen möglichst hohen Turm mit Würfeln und Dominosteinen baut. Anhand der Höhe des Turmes wird dann die Punktzahl zugeordnet. Nach Ablauf der Zeit wird ausschließlich die Höhe des Turmes bewertet und dementsprechend die Punktzahl vergeben. Sollte der Turm nach Ablauf der Zeit durch einen Zusammenstoß mit einem Greifer oder ähnlichem umgestürzt sein, gibt es 0 Punkte beziehungsweise die Punktzahl für den eventuell noch stehenden Teil des Turmes. Die Tatsache, dass sich bei der Auswertung zu Stacking Challenge eine ausgewogene Leistung der Probanden zeigt, könnte darin begründet sein, dass Alkohol durch eine Enthemmung und Kritikminderung eine erhöhte Risikobereitschaft hervorruft (Mußhoff, Madea, & Heß, 2014). Weiss et al. (2013) beschrieben, dass sich mithilfe neurophysiologischer Studien eine besonders hohe Sensibilität im Bereich des inferioren medialen präfrontalen Cortex und des orbitofrontalen Cortex auf alkoholtoxische Effekte zeigte. Bei womöglich langanhaltendem Konsum im Rahmen einer Alkoholabhängigkeit kann es in diesem Bereich zu Läsionen kommen, was Betroffene besonders risikohafte Entscheidungen ohne Berücksichtigung langfristiger Konsequenzen treffen lässt. Im Rahmen der dieser Arbeit zugrundeliegenden Studie wurde jedoch die Auswirkung von akutem Alkoholkonsum untersucht, daher kann zu diesem Langzeit-Effekt kein erklärender Zusammenhang geschaffen werden.

Korrespondierend zu der Auswertung dieser Arbeit wurde bei der Studie von Dorafshar et al. (2002) ebenfalls die Bewegungsökonomie betrachtet. Es konnte bei den Untersuchungen von Dorafshar et al. ein statistisch signifikanter Unterschied in dieser Kategorie festgestellt werden. Bei den Ergebnissen der Studie am MIMIC dV-Trainer® konnte in der Metrik „Economy of motion“ der Übungen Ring Tower und Suture Sponge kein statistisch signifikanter Unterschied detektiert werden. Jedoch sah man, dass sich jeweils etwa zwei Drittel der Probanden unter der Intervention verschlechtert und ein Drittel verbessert haben. Um eine gute Bewegungsökonomie zu gewährleisten, sind unbeeinträchtigte motorische, kognitive als auch visuelle Funktionen notwendig. Schon geringe Mengen Alkohol beeinflussen das Sehvermögen und führen unter anderem zu einer Herabsetzung der Geschwindigkeit des Tiefenschärfesehvorgangs. Dies führt dazu, dass die Fixation bewegter Objekte und somit die Orientierung im Raum herabgesetzt ist (Mußhoff, Madea, & Heß, 2014). Koordinierte Blick- und Führungsbewegungen sind eingeschränkt, was die Economy of motion maßgeblich beeinflusst.

Nach dem veröffentlichten Artikel von Geuenich et al (2011) konnte anhand einer Online-Befragung aufgezeigt werden, dass viele Ärzte beruflichen Stress mit dem Konsum von Alkohol

oder Medikamenten kompensieren. Bei einer ergänzenden Datenerhebung mittels Befragung zur Trinkgewohnheit gaben fast die Hälfte unserer Probanden an, zwei- bis viermal im Monat Alkohol zu konsumieren. Keiner der Probanden hat angegeben, viermal die Woche oder häufiger Alkohol zu trinken und zwei Probanden haben angegeben, nie Alkohol zu trinken. Anhand dieser Daten konnte ausgeschlossen werden, dass ein Proband mit einer möglicherweise vorliegenden erhöhten Toleranzgrenze an dem Versuch teilnimmt. Die Repräsentativität dieser Daten wird maßgeblich durch eine wahrheitsgetreue Beantwortung der Probanden bedingt. Dies kann jedoch durch die nicht anonymisierte Befragung beeinflusst worden sein.

5.3 Praktische Implikationen und Bedeutung in der Chirurgie

Eine verzögerte Bearbeitungszeit, wie sie zum Beispiel bei der Metrik Wire contact duration mit einer tendenziellen Verlängerung unter Alkoholeinfluss nachgewiesen wurde, kann in vielen klinischen Situationen schwerwiegende Konsequenzen haben. Auch Einbußen in der Feinmotorik, die sich unter anderem in einer statistisch signifikanten Verschlechterung in den Übungen Suture Sponge und Ring Tower präsentierten, können sich demnach negativ auf den klinischen Alltag eines Chirurgen und damit einhergehend auch auf die Patientensicherheit auswirken.

Es gibt bisher keine gesetzliche Regelung zum Umgang mit Alkohol bei Ärzten. Dabei wäre es nach den Erkenntnissen dieser Studie gegebenenfalls sinnvoll, hier eine gesetzliche Regulierung zum Konsum von Alkohol im oder unmittelbar in zeitlichem Zusammenhang zum Dienst einzuführen. In der Luftfahrt oder auch in der Atomindustrie gibt es bereits klare und transparente Richtlinien zum Alkoholkonsum, diese könnte in der Ärzteschaft, als auch in bei den Patienten zu einer höheren Sicherheit beitragen (Kirby, Kapoor, Das-Purkayastha, & Harries, 2012). Allerdings muss bedacht werden, dass die Überwachung einer solchen gesetzlichen Regulierung im klinischen Alltag sicher eine Herausforderung darstellt und eine Einführung durch vielfältige rechtliche Grauzonen limitiert sein könnte. Eine beispielhafte Möglichkeit zur Regelung des Alkoholkonsums zeigt die Luftfahrtindustrie, in der seit Februar 2021 stichprobenartige Testungen von einigen Akteuren auf Alkohol und Drogen (Drägerwerk AG & Co. KgaA, kein Datum) stattfinden. Denkbar wäre die Umsetzung auch im klinischen Alltag mit einer nicht invasiven stichprobenartigen Atemalkoholtestung vor Operationsbeginn. Ein Alkoholkonsum in zeitlichem Zusammenhang zum Dienstbeginn, der sich nachweislich beziehungsweise durch eine den Alkoholkonsum bedingte schlechtere Schlafqualität und einem „Katereffekt“ ebenfalls negativ auf die Leistung auswirkt (Dorafshar, O'Boyle, & McCloy, 2002), kann damit in den meisten Fällen, in denen ein Atemalkoholwert häufig nicht mehr nachweis-

bar ist, nicht identifiziert werden. In solchen Fällen eine anwendbare und mit dem Persönlichkeitsrecht vereinbare Vorgehensweise zur Überwachung zu realisieren, würde zweifellos eine schwierige bis unmögliche Aufgabe darstellen.

5.4 Limitierung und Ausblick

Während der Datenerhebungsphase, die in etwa zwölf Monate dauerte, wurde die Software des MIMIC dV-Trainer® aktualisiert und die Übung Ring Tower aus unbekanntem Gründen in der neuen Software nicht mehr zur Verfügung gestellt. Deshalb konnte die Übung Ring Tower nur von 17 Probanden absolviert und ausgewertet werden, während die anderen Übungen von 20 Probanden durchgeführt worden. Da sich hier bei der Auswertung in den einzelnen Metriken deutliche Tendenzen zur Verschlechterung unter der Intervention jedoch ohne statistische Signifikanz gezeigt haben, lässt dies die Vermutung nahe, dass bei einer vollständigen Probandenzahl gegebenenfalls eine Signifikanz nachgewiesen worden wäre. Unsere Stichprobengröße orientierte sich an der statistischen Vorgabe auf Grundlage der Ergebnisse der ersten zehn Probanden. Außerdem wurde die Probandenanzahl maßgeblich vom Erfolg der Rekrutierung potentieller Probanden limitiert. Es gestaltete sich teilweise äußerst schwierig, an dem Experiment und dem Simulator interessierte Assistenz- und Fachärzte im Rahmen des klinischen Dienstmodells und der hohen Arbeitsbelastung zu gewinnen. Bei einer größeren vorhandenen Probandenanzahl wären einige Ergebnisse unter Umständen noch deutlicher ausgefallen und die Erkenntnisse wären eher auf die Realität anwendbar. Weiterhin wäre es bei einer größeren Anzahl von Versuchspersonen möglich gewesen, eine Randomisierung vorzunehmen. Dadurch könnte einem möglicherweise ergebnisbeeinflussenden Lerneffekt entgegengewirkt werden, anstatt diesen mithilfe der durchgeführten Vorstudie nicht gänzlich ausschließen zu können. Statistisch gesehen stellt dies eine hochwertigere Methode dar, um eine Allgemeingültigkeit anhand der Ergebnisse ableiten zu können.

Die Ergebnisse die dieser Arbeit zugrunde liegen, sind überwiegend mithilfe validierter und objektiver Instrumente generiert worden. Darunter zählen unter anderem das Auswertungssystem der Software des MIMIC dV-Trainer® und die Atem- und Blutalkoholmessung. Jedoch zählen Kommunikation und Teamarbeit im klinischen Alltag und insbesondere im Operationssaal eine ausschlaggebende Rolle, da Missverständnisse oder gar eine fehlende Kommunikation im Team ein genauso großes Risiko für die Patientensicherheit darstellen, wie eine kognitive oder motorische Einschränkung des Operateurs. Diese Fähigkeiten eines Chirurgen, eine verständliche Kommunikation zu bieten und im Team zu interagieren, wie es während einer Operation notwendig ist, wurde in unserem Versuch nicht untersucht. Dahingehend könnten weitere Untersuchungen wie es zum Beispiel Kirby et al. (2012) getan haben erfolgen,

um eine umfassende Aussage über die Wirkung von Alkoholkonsum nicht nur auf die manuell chirurgische Performance, sondern auch auf die interpersonelle und rhetorische Leistung eines Chirurgen treffen zu können.

Die meisten in dieser Arbeit vorgestellten Studien zu dem Thema Alkohol und chirurgische Leistung haben überwiegend die Ergebnisse am Morgen nach dem Alkoholkonsum untersucht. Diesen Aspekt haben wir innerhalb unserer Studie nicht getestet, sodass dazu keine Beurteilung erfolgen konnte. Grundsätzlich kann eine schlechte chirurgische Leistung am Morgen nach dem Konsum durch einen Katereffekt, eine mutmaßlich schlechte Schlafqualität (Dorafshar, O'Boyle, & McCloy, 2002) oder einer vorliegenden Restalkoholisierung schwerwiegende Konsequenzen im klinischen Alltag bedeuten. Außerdem kommt dieser Zustand womöglich häufiger im klinischen Alltag vor, als eine akute Alkoholisierung. Dahingehend wäre eine weitere ergänzende Testung sinnvoll gewesen, um umfangreichere Informationen liefern zu können.

Die erhobenen Daten über die subjektiven Trinkgewohnheiten der Probanden bieten eine Hilfestellung für die Einschätzung einer eventuell vorliegenden erhöhten Toleranzgrenze und einem Gewöhnungseffekt bei regelmäßigem Konsum. Unter Umständen wäre bei einer mutmaßlich erhöhten Toleranzgrenze und einem Gewöhnungseffekt eine größere Alkoholmenge notwendig gewesen, um die akut psychomotorischen und kognitiven Folgen des Alkohols hervorgerufen. Da diese Trinkgewohnheiten jedoch nicht anonymisiert erfragt wurden, kann dies dazu geführt haben, dass die Probanden womöglich nicht wahrheitsgetreu geantwortet haben und somit die verabreichte Trinkmenge nicht angepasst wurde.

Es wurden validierte Instrumente verwendet, um sowohl die Blutalkoholkonzentration (laborchemische Bestimmung), als auch die chirurgische Leistung objektiv zu messen. Weiterhin wurde ein möglicher Lerneffekt im Rahmen der vorhandenen Probandenanzahl weitestgehend minimiert. Die Erkenntnisse dieser Studie erbrachten teilweise signifikante Unterschiede bei den Messungen unter Alkoholkonsum und der Großteil der Leistungen zeigte eine tendenzielle Verschlechterung. Allerdings bieten die oben aufgeführten Kritikpunkte dieser Arbeit hilfreiche Ansätze für weiterführende Untersuchungen und könnten als Basis für nachfolgende Forschungsfragen nützlich sein.

6 Zusammenfassung

Ärzte zeigen nachweislich einen deutlich höheren stressbedingten Substanzmissbrauch als die Normalbevölkerung (Geuenich, 2011). Demnach ist die Gefahr, von einem alkoholisierten Chirurgen behandelt zu werden nicht auszuschließen. Allgemein wird angenommen, dass die chirurgische Leistung durch Alkoholkonsum abnimmt, was bisher nicht für die komplexen Aufgaben der Roboterchirurgie erfasst wurde. Im Rahmen dieser Studie wurde die chirurgische Leistung an einem validierten Operationssimulator, dem MIMIC dV-Trainer®, unter dem Einfluss von Alkohol gemessen. Dazu wurden im Rahmen der Datenerhebungsphase 20 Chirurgen und chirurgisch arbeitende Weiterbildungsassistenten rekrutiert. Diese haben eine Auswahl von vier Übungen, die unter anderem technische, manuelle sowie kognitive Fähigkeiten überprüfen, absolviert. Im Anschluss erhielten die Probanden eine mithilfe von Geschlecht und Gewicht ermittelte Menge Ethanol, um etwa eine Blutalkoholkonzentration von 0,8 Promille zu erreichen. In regelmäßigen Atemalkoholmessungen wurden anschließend die vorhandenen Alkoholwerte kontrolliert. Anschließend wurden die initial nüchtern absolvierten Übungen erneut durchgeführt. Die gewonnenen Ergebnisse wurden mithilfe verschiedener statistischer Testverfahren untersucht. In den Metriken „Instruments out of view“ (RT), „Drops“ (SS), „Instrument collisions“ (SS), „Instruments out of view“ (SS) und „Missed targets“ (SS) wurde eine signifikante Verschlechterung unter der Intervention nachgewiesen. Die signifikanten Ergebnisse wurden durch den sogenannten Vorzeichentest ermittelt, wobei ein Signifikanzniveau von $z = 1,96$ festgelegt wurde. Der Einfluss von Alkohol hat also dazu geführt, dass sich wichtige motorische Fähigkeiten eines Chirurgen am Robotik-Simulator verschlechtert haben. Auf Basis der gewonnenen Resultate ergibt sich demnach ein potenziell erhöhtes Risiko für Patienten, wenn ein Chirurg alkoholisiert in der Roboter-Chirurgie arbeitet. Die Erkenntnisse dieser Arbeit könnten dazu dienen, die Sicherheitsmaßnahmen in einem Operationssaal oder gar generell im medizinischen Bereich zu überdenken und ein Alkoholverbot für Ärzte und Chirurgen auszusprechen, was eventuell zu einer höheren Patientensicherheit beitragen würde.

7 Literaturverzeichnis

- Agha R. A., Fowler, A. J. (Februar 2015) The Role and Validity in surgical Simulation. *International surgery*: 350-357. DOI: 10.9738/INTSURG-D-14-00004.1
- Arden, C. J.-D. (2018) *Aktuelles aus Klinik und Praxis der Urologie*. Heidelberg: Springer.
- Bendel, P. D. (2018) *Gabler Wirtschaftslexikon - Operationsroboter* [E-Book]. Wiesbaden: Springer Gabler. DOI: <https://wirtschaftslexikon.gabler.de/definition/operationsroboter-100561/version-346084>
- Bundeszentrale für gesundheitliche Aufklärung. Alkohol am Steuer. [Internet]. [zitiert am 18.10.2021]. URL: <https://www.kenn-dein-limit.de/alkoholverzicht/alkohol-am-steuer/>
- Bundeszentrale für gesundheitliche Aufklärung. Kenn dein Limit. [Internet]. [zitiert am 05.06.2020]. URL: <https://www.kenn-dein-limit.info/wie-wirkt-alkohol.html>
- Challenge-A* (2021). Alkoholsmus-Symptome. [Internet]. [zitiert am 18.10.2021]. URL: <https://www.challenge-a.com/blog/alkoholismus-symptome/>
- Dorafshar A. H., O'Boyle D. J., McCloy, R. F. (29. Juli 2002). Effects of a moderate dose of alcohol on simulated laparoscopic surgical performance. *Surgical Endoscopy*: 1753-1758. DOI: 10.1007/s00464-001-9052-3
- Drägerwerk AG & Co. KgaA. Sicherstellen der Flugtauglichkeit von Piloten und Flugbegleitern. [Internet]. [zitiert am 18.10.2021]. URL: https://www.draeger.com/de_de/Safety/Drugs-Alcohol-Testing-At-Work/Aviation
- Feussner H., Wilhelm D. (März 2016). Minimally invasive surgery and robotic surgery: surgery 4.0?. *Der Chirurg*: 189-194.
- Güntert, A. Pschyrembel online (2018). Hippokratischer Eid [Internet]. 05/2018. [zitiert am 18.10.2021]. URL: <https://www.pschyrembel.de/Hippokratischer%20Eid/K09UA>
- Gallagher A. G., Boyle E., Toner P., Neary P. C., Andersen D. K., Satava R. M., Seymour N. E. (April 2011). Persistent next-day effects of excessive alcohol consumption on laparoscopic surgical performance. *Archives of surgery (Chicago, Ill. : 1960)*: 419-426.
- Geuenich K. (Februar 2011). Alkohol- und Medikamentenmissbrauch bei Ärzten - Gefährliche Stressbewältigung. *Deutsches Ärzteblatt*(Heft 2): 73-74.

-
- Graw P. D. (2015). Wie viel Alkohol hat Ihr Patient wirklich intus? MMW Fortschritte der Medizin: 46.
- Kenney P. A., Wszolek M. F., Gould J. J., Libertino J. A., Moinzadeh A. (10. April 2009). Face, Content and Construct Validity of dV-Trainer, a Novel Virtual Reality Simulator for Robotic Surgery. Urology: 1288-1292.
- Kirby G., Kapoor K., Das-Purkayastha P., Harries M. (März 2012). The effect of alcohol on surgical skills. Annals of the Royal College of Surgeons of England: 90-93.
- Kocher H., Warwick J., Al-Ghnaniem R., Patel A. (März 2006). Surgical dexterity after a 'night out on the town'. ANZ journey of surgical: 110-112.
- Lendvay T. S., Brand T.C., White L., Kowalewski T., Jonnadula S., Mercer L. D., Khorsand D., Andros J., Hannaford B., Satava R. M. (11. April 2013). Virtual reality robotic surgery warm-up improves task performance in a dry laboratory environment: a prospective randomized controlled study. Journal of the American College of Surgeons: 1181-1192. DOI: 10.1016/j.jamcollsurg.2013.02.012
- MacCraith E., Forde J.C., Davis N.F. (22. Februar 2019). Robotic simulation training for urological trainees: a comprehensive review on cost, merits and challenges. Journal of Robotic Surgery: 371-377.
- Mimic Technologies. (2016). Exercise Instructions. Mimic Technologies, Inc.: Seattle, Washington, Vereinigte Staaten.
- Mimic Technologies. (2018). dV Trainer by mimic User Guide Msim Rev 3.6. Mimic Technologies, Inc.: Seattle, Washington, Vereinigte Staaten.
- Mimic Technologies, Inc. (2018). Mimic. [Internet]. 2018 [zitiert am 08.06.2020] URL: <https://mimicsimulation.com/dv-trainer/>
- Mußhoff F., Madea B., Heß C. (2014). Toxikologie. In: Madea B., Hrsg. Rechtsmedizin. 3.Auflage. Bonn: Springer. 707-729.
- Planing P. (2021). Statistik Grundlagen [E-Book]. Stuttgart: Planing P. DOI: <https://statistikgrundlagen.de/ebook/chapter/korrelation/>

-
- Seymour N. E., Gallagher A. G., Roman S. A., O'Brian M. K., Bansal V. K., Andersen D. K., Satava R. M. (Oktober 2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of Surgery*: 458–464.
- Stocker T. C., Steinke I. (2016). *Statistik- Grundlagen und Methodik*. Mannheim: De Gruyter Oldenbourg. 137-146
- Universität Zürich (11.12.2020). Methodenberatung [Internet]. 12/2020. [zitiert am 18.10.2021] URL:
https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/unterschiede/zentral/vo_rzeichen.html
- Van Dyken I., Szlabick R. E., Sticca R. P. (24.09.2013). Effect of alcohol on surgical dexterity after a night of moderate alcohol intake. *American journal of surgery*: 964-969.
- Virtanen M. E., Jokela M., Nyberg S.T., Madsen I.E.H., Lallukka T., Ahola K., Alfredsson L., Batty G.D., Bjorner J.B., Borritz M., Burr H., Casini A., Clays E., De Bacquer D., Dragano N., Erbel R., Ferrie J.E., Fransson E.I. . . .Kivimäki M. (13.01.2015). Long working hours and alcohol use: systematic review and meta-analysis of published studies and unpublished individual participant data. *The british medical journal*: 350. DOI: 10.1136/bmj.g7772
- Weiss E. M., Singewald E., Ruepp B., Marksteiner J. (19.07.2013). Alkohol induzierte kognitive Dysfunktion. *Wiener Medizinische Wochenschrift*: 9-14. DOI: 10.1007/s10354-013-0226-0
- Wengel A. (08.05.2020). Alkohol [Internet]. 05/2020. [zitiert am 05.06.2020]. URL:
<https://www.planet-wissen.de/gesellschaft/rauschkittel/alkohol/index.html>
- Wilson M. S., Middlebrook A., Sutton C., Stone R. (Dezember 1997). MIST VR: A virtual reality trainer for laparoscopic surgery assesses performance. *Annals of the Royal college of Surgeons of England*: 403-404.

8 Anhang

8.1 Tabelle zu den Leistungsmetriken

Time to complete Exercise	Diese Metrik erfasst die Gesamtzeit, die ein Proband vom Start bis zur Beendigung der Übung benötigt. Je weniger Zeit in Sekunden benötigt wird, desto höher ist die Punktzahl.
Drops	Wenn während der Übung ein von den Probanden geführter Ring oder eine Nadel unbeabsichtigt in einen nicht dafür vorgesehenen Bereich fällt, erfolgt ein Punkteabzug.
Economy of motion	Diese Metrik misst die Gesamtstrecke (in Zentimetern), die alle Instrumente während einer Übung zurücklegen. Je weniger Zentimeter insgesamt an Strecke zurückgelegt werden, desto höher ist die Punktzahl, die zum Gesamtscore gezählt wird. Somit bewertet diese Metrik die Effizienz innerhalb der Übung.
Instrument collisions	Der Parameter gibt wieder, wie oft die Instrumente aneinanderstoßen. Die Kollisionen werden addiert. Je mehr Zusammenstöße stattgefunden haben, desto weniger Punkte gibt es.
Wire collisions	Die Metrik gibt wieder, wie häufig man fälschlicherweise bei der Übung Ring Tower den Draht mit dem Ring berührt.
Wire contact duration	Hier wird die Gesamtzeit, die durch die fälschlicherweise berührten Kontakte mit dem Draht entstanden ist, bewertet.
Tower knockoff	Diese Metrik misst die Anzahl, wie häufig ein Drahturm durch einen der Greifer abgerissen wird.
Excessive instrument force	Gibt die Gesamtzeit (in Sekunden) wieder, in der eine übermäßige Instrumentenkraft angewendet wird, welche über der vorgeschriebenen Grenzkraft liegt. Kräfte auf ein Instrument können durch Kollisionen und durch Handlungen wie dem Zurückziehen des Gewebes, dem Einstechen einer Nadel oder dem Ziehen an einer

	Naht entstehen. Je kürzer diese Kraftanwendung dauert, desto mehr Punkte werden vergeben.
Instruments out of view	Diese Metrik erfasst die Gesamtstrecke (in Zentimetern), die von Instrumenten außerhalb des Sichtfeldes des Benutzers zurücklegt wird. Je kleiner diese Gesamtstrecke ist, desto mehr Punkte erhält der Anwender dafür.
Master Workspace Range	Durch ein definiertes kugelförmiges Volumen sollen alle Positionen der beiden Mastergriffe während der Übung widergespiegelt werden. Innerhalb dieser gedachten Kugel wird die Strecke in Zentimetern gemessen, die ein Proband darin mit den Greifern zurücklegt. Je kleiner die Gesamtstrecke ist, desto höher ist die Punktzahl.
Missed Targets	Jedes verpasste Nadelziel wird bewertet. Je weniger verpasste Ziele, desto mehr Punkte erhält man.

Anmerkung. Die Inhalte der Tabelle basieren auf eigens gemachten Erfahrungen am MIMIC dV-Trainer® als auch auf Basis von Mimic Technologies (2018)


8.2 Bewertungsschlüssel der einzelnen Leistungsmetriken

A. Ring Tower

Metriken, die bei dieser Übung in die Bewertung einfließen und die dazugehörigen Bewertungsschlüssel:

Time to Complete Exercise

Total time the user spends on the exercise.

Pass: 

Value: 96.51 sec **Score:** 130.49


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	138.84	100
Worst	278.2	0

Abbildung A1. Bewertungsschlüssel zu „Time to complete“ der Übung Ring Tower.

Instrument Collisions

Total number of instrument-on-instrument collisions exceeding a minimum force threshold.

Pass: 

Value: 0 **Score:** 200


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	2	100
Worst	4	0

Abbildung A2. Bewertungsschlüssel zu „Instrument collisions“ der Übung Ring Tower.

Wire Collisions

Number of times a ring collides with a wire.

Pass: 

Value: 16 **Score:** 105.88


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	17	100
Worst	27	0

Abbildung A3. Bewertungsschlüssel zu „Wire collisions“ der Übung Ring Tower.

Wire Contact Duration

Total time a ring is in contact with a wire.

Pass: 

Value: 22.08 sec **Score:** 167.52


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	67.99	100
Worst	213.39	0

Abbildung A4. Bewertungsschlüssel zu „Wire contact duration“ der Übung Ring Tower.

Economy of Motion

Total distance traveled by all instruments.

Pass: 

Value: 195.82 cm **Score:** 72.99


Baseline Values

	Raw Value	Score Value
Best	133.89	200
Proficiency	182.65	100
Worst	231.41	0

Abbildung A5. Bewertungsschlüssel zu „Economy of motion“ der Übung Ring Tower.

Instruments Out Of View

Total distance traveled by instruments outside the user's field of view.

Pass: 

Value: 47.38 cm **Score:** 33.64


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	20.25	100
Worst	61.13	0

Abbildung A6. Bewertungsschlüssel zu „Instruments out of view“ der Übung Ring Tower.

Drops

Number of times any object is dropped in an inappropriate region of the scene.

Pass: 

Value: 3 **Score:** 50


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	2	100
Worst	4	0

Abbildung A7. Bewertungsschlüssel zu „Drops“ der Übung Ring Tower.

Tower Knockoffs

Number of times towers get knocked off.

Pass: 

Value: 0 **Score:** 200

Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	0	100
Worst	5	0

Abbildung A8. Bewertungsschlüssel zu „Tower Knockoffs“ der Übung Ring Tower.

B. Suture Sponge

Metriken die bei dieser Übung in die Bewertung einfließen und die dazugehörigen Bewertungsschlüssel:

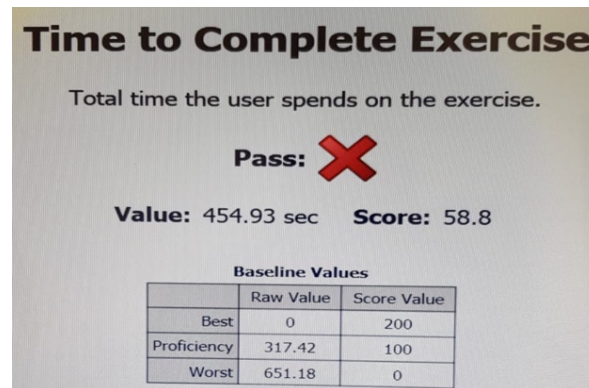


Abbildung B1. Bewertungsschlüssel zu „Time to complete“ der Übung Suture Sponge.

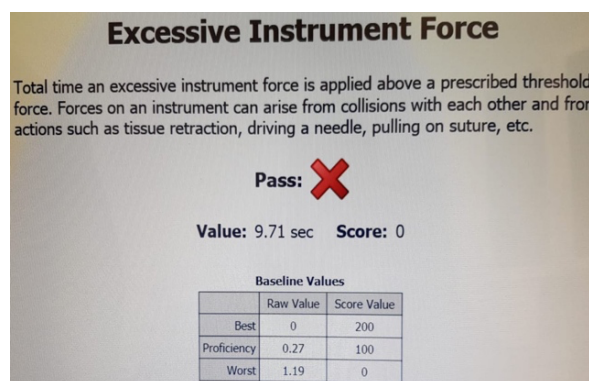



Abbildung B2. Bewertungsschlüssel zu „Excessive Instrument Force“ der Übung Suture Sponge.

Economy of Motion

Total distance traveled by all instruments.

Pass: 

Value: 563.77 cm **Score:** 52.26


Baseline Values

	Raw Value	Score Value
Best	59.62	200
Proficiency	400.86	100
Worst	742.1	0

Abbildung **B3**. Bewertungsschlüssel zu „Economy of motion“ der Übung Suture Sponge.

Instruments Out Of View

Total distance traveled by instruments outside the user's field of view.

Pass: 

Value: 8.9 cm **Score:** 84.08


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	6.17	100
Worst	23.33	0

Abbildung **B4**. Bewertungsschlüssel zu „Instruments out of view“ der Übung Suture Sponge.

Master Workspace Range

The larger of the two radii of motion of the user's working volume on master grips.

Pass: 

Value: 7.95 cm **Score:** 22.87


Baseline Values

	Raw Value	Score Value
Best	2.74	40
Proficiency	8.82	20
Worst	14.9	0

Abbildung **B5**. Bewertungsschlüssel zu „Master Workspace Range“ der Übung Suture Sponge.

Instrument Collisions

Total number of instrument-on-instrument collisions exceeding a minimum force threshold.

Pass: 

Value: 24 **Score:** 0


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	3	100
Worst	11	0

Abbildung **B6**. Bewertungsschlüssel zu „Instrument collisions“ der Übung Suture Sponge.

Drops

Number of times any object is dropped in an inappropriate region of the scene.

Pass: 

Value: 0 Score: 200


Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	1	100
Worst	3	0

Abbildung **B7**. Bewertungsschlüssel zu „Drops“ der Übung Suture Sponge.

Missed Targets

Number of missed targets.

Pass: 

Value: 65 Score: 0

Baseline Values

	Raw Value	Score Value
Best	0	200
Proficiency	15	100
Worst	42	0

Abbildung **B8**. Bewertungsschlüssel zu „Missed Targets“ der Übung Suture Sponge.

C. Vitruvian Operation

Die Bewertung dieser Übung basiert nur auf der erreichten Punktzahl innerhalb einer vorgegebenen Zeit mit folgender Bewertungsmatrix:

Points
Number of points made in a game.

Pass: ✓

Value: 49 **Score:** 49

Baseline Values

	Raw Value	Score Value
Best	60	60
Proficiency	30	30
Worst	0	0

Abbildung C1. Bewertungsschlüssel zu der Übung Vitruvin Operation.

D. Stacking Challenge

Die Bewertung dieser Übung basiert nur auf der erreichten Punktzahl innerhalb einer vorgegebenen Zeit mit folgender Bewertungsmatrix:

Points
Number of points made in a game.

Pass: ✗

Value: 66.7 **Score:** 66.7

Baseline Values

	Raw Value	Score Value
Best	228	228
Proficiency	114	114
Worst	0	0

Abbildung D1. Bewertungsschlüssel zu der Übung Stacking Challenge.

8.3 Bestimmung der benötigten Alkoholmenge

	männlich	weiblich
Beginn: Alkoholaufnahme	17:15	17:00
Beginn: Experiment	18:00	18:00
Körpergewicht [kg]	98	63
Alkoholgehalt Getränk [Vol%]	38%	38%
angestrebte BAK bei Beginn Experiment	0,8	0,8
Volumen alkoholische Getränk [mL]	257	148
Ende: Experiment	18:30	18:30
Erwartete BAK Experimentende	0,73	0,73
Proband nüchtern gg.	23:20	23:20

Abbildung E. Exemplarische Darstellung der Excel-Datei für die Bestimmung der benötigten Alkoholmenge.

8.4 Flowchart zum geplanten Studienablauf

